

**Міністерство освіти і науки України**  
**Національний університет «Запорізька політехніка»**

Кваліфікаційна наукова  
праця на правах рукопису

**ВЕДМЕДЄВ Станіслав Русланович**

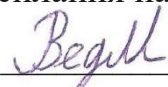
УДК 004.82:633.854.78

**ДИСЕРТАЦІЯ**  
**ПІДТРИМКА ПРИЙНЯТТЯ РІШЕНЬ ПРИ ВИРОЩУВАННІ**  
**СОНЯШНИКУ НА ОСНОВІ ЦИФРОВОЇ ТЕХНОЛОГІЇ**  
**ФЕНОТИПУВАННЯ**

124 системний аналіз

Подається на здобуття наукового ступеня доктор філософії.

Дисертація містить результати власних досліджень. Використання ідей,  
результатів і текстів інших авторів мають посилання на відповідне джерело

 С. Р. Ведмедєв

Науковий керівник: **ТЕРЕЩЕНКО Еліна Валентинівна,**  
кандидат фізико-математичних наук, доцент

Запоріжжя – 2026

## АНОТАЦІЯ

**Ведмедєв С. Р. Підтримка прийняття рішень при вирощуванні соняшнику на основі цифрової технології фенотипування.** – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 124 – «Системний аналіз» – Національний університет «Запорізька політехніка», Запоріжжя, 2026.

Дисертаційне дослідження присвячено вирішенню науково-практичної проблеми підтримки прийняття рішень при вирощуванні соняшнику на основі цифрової технології фенотипування. Вирішення цієї проблеми пов'язане з подоланням розбіжності між стрімким зростанням обсягів первинних цифрових агротехнологічних даних та низьким рівнем їх системної типізації й логічної структурованості для автоматизованого прийняття рішень.

**Об'єктом дослідження** є процеси прийняття рішень при вирощуванні соняшника в умовах цифрової трансформації аграрної галузі.

**Предметом дослідження** є методи формалізації знань на основі онтологічного моделювання, алгоритми цифрового фенотипування рослинних об'єктів за зображеннями та цифрові технології інтеграції даних для моделювання взаємодії «генотип — середовище».

**Мета дослідження** полягає у розробці цифрової технології фенотипування рослинного об'єкта на основі онтологічної моделі та формування датасетів, що забезпечує формалізоване представлення, ефективну інтеграцію фенотипових даних для прийняття рішень при вирощуванні соняшника.

Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатка.

У вступі обґрунтовано актуальність теми дисертаційного дослідження, наведено мету, об'єкт та предмет дослідження, описано методи дослідження, сформульовано наукову новизну та практичне значення одержаних результатів, подано відомості про апробацію та публікації за темою дисертації.

У першому розділі «Процеси цифрової трансформації аграрної галузі» виконано комплексний системний аналіз сучасних процесів цифрової

трансформації аграрної галузі. Проведено аналіз сучасних підходів до цифрового фенотипування рослин, побудови цифрових двійників біологічних об'єктів, застосування методів комп'ютерного зору, онтологічного моделювання та систем підтримки прийняття рішень в аграрній сфері. Встановлено обмеження існуючих рішень, пов'язані з фрагментарністю даних, відсутністю єдиного семантичного середовища та недостатньою інтеграцією результатів фенотипування з механізмами інтелектуального аналізу.

У другому розділі «Концептуальні засади та інструментарій побудови цифрових моделей рослини соняшнику» представлено теоретико-методологічне обґрунтування побудови цифрової моделі рослини соняшника, основою якої є взаємодія «генотип — середовище». Основним інструментом формалізації знань та побудови логічного каркаса цифрової моделі обрано предметну онтологію. Застосування онтологічного моделювання та апарату семантичних правил дозволяє побудувати структуровану систему знань, здатну до автоматизованого логічного виведення на рівні екземплярів і виявлення прихованих закономірностей.

У третьому розділі «Онтологічне моделювання вирощування соняшника» виконано комплексний системний аналіз, формалізацію та оцінку структурно-семантичної якості розробленої бази знань онтології соняшнику HELIANTUS.

У четвертому розділі «Цифрова технологія фенотипування та формування datasets» удосконалено цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets), яка забезпечує семантично узгоджене формування наборів даних відповідно до онтологічної моделі HELIANTUS та їх подальшу інтеграцію до бази знань. Розроблено метод цифрового фенотипування рослинних об'єктів на основі зображень та виконано програмно-алгоритмічну реалізацію розробленого методу, проведено комплексні експериментальні дослідження, а також здійснено валідацію та оцінку ефективності створеного інструментарію при обробці реальних біологічних об'єктів.

Для досягнення мети необхідно було виконати такі **завдання**.

1. Провести системний аналіз процесів цифрової трансформації аграрної галузі та визначити вимоги до формування інформаційно-аналітичного середовища вирощування соняшнику, що забезпечує інтеграцію різнорідних даних, їх семантичне представлення та використання у системах підтримки прийняття рішень.

2. Удосконалити цифрову модель соняшнику, яка ґрунтується на моделі взаємодії «генотип — середовище».

3. На основі розробленої цифрової моделі соняшнику вдосконалити онтологічну модель, яка інтегрує параметри генотипу, селекції та агротехнологій вирощування соняшнику з урахуванням кліматичних показників та екологічних аспектів.

4. Розробити метод цифрового фенотипування рослинних об'єктів на основі методів комп'ютерного зору для автоматизованого виділення та кількісної оцінки морфологічних і фізіологічних ознак соняшнику.

5. Удосконалити цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets), яка забезпечує автоматизований збір, стандартизацію та збереження фенотипових ознак на основі онтологічної моделі.

### **Наукова новизна роботи**

1. Удосконалено структуру цифрової моделі рослини соняшнику на основі чотирирівневої архітектури Digital Phenotyping System, яка забезпечує інтеграцію процесів збору, збереження, обробки та аналізу різнорідних даних. Запропонована архітектура поєднує результати лабораторного та польового фенотипування з параметрами навколишнього середовища, агротехнологічними факторами та джерелами дистанційного моніторингу, що створює основу для побудови цифрового двійника рослини та реалізації предиктивного моделювання її розвитку.

2. Удосконалено предметну онтологічну модель агросистеми вирощування соняшнику HELIANTUS, яка забезпечує семантичну інтеграцію генетичних,

фенотипових, агрокліматичних та технологічних даних у межах єдиного інформаційного простору. Запропонований підхід передбачає трансформацію результатів цифрового фенотипування у формалізовані знання на основі концепцій TBox та ABox, що створює можливість застосування механізмів логічного виведення для автоматизованого аналізу причинно-наслідкових зв'язків у системі «генотип — середовище».

3. Вперше розроблено метод цифрового фенотипування рослинних об'єктів на основі зображень, який реалізує оптимізацію параметрів реєстрації даних. Метод забезпечує автоматизоване виділення та високоточне оцінювання морфологічних, структурних і колірних характеристик насіння соняшнику за кількісними дескрипторами.

4. Удосконалено цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets), яка забезпечує семантично узгоджене формування наборів даних відповідно до онтологічної моделі HELIANTUS та їх подальшу інтеграцію до бази знань.

**Практична значимість** отриманих результатів полягає у створенні алгоритмічного, програмного та семантичного інструментарію, який забезпечує підтримку прийняття рішень при вирощуванні соняшнику на основі цифрової технології фенотипування.

Створено інформаційне середовище прийняття рішень на основі механізмів автоматизованого логічного виведення в онтології HELIANTUS. Практична реалізація онтологічної моделі агросистеми вирощування соняшнику дозволила об'єднати різноманітні дані (генетичні, фенотипові, агрокліматичні, технологічні) в єдину базу знань. Це дає змогу виявляти причинно-наслідкові зв'язки в системі «генотип — середовище» та приймати обґрунтовані рішення щодо оптимізації технологій вирощування соняшнику.

На основі розробленого методу обробки цифрових зображень створено програмний комплекс (мовою Python), який автоматизує виділення, сегментацію та аналіз морфометричних і колірних характеристик насіння соняшнику. Це

забезпечує селекціонерів та агрономів об'єктивними кількісними дескрипторами для прийняття рішень щодо якості насінневого матеріалу, ідентифікації зразків та відбору ліній незалежно від умов зйомки й освітлення.

Впровадження чотирирівневої архітектури Digital Phenotyping System забезпечує практичну можливість консолідації даних лабораторного, польового фенотипування та дистанційного моніторингу з факторами довкілля, що може стати технологічною основою для побудови «цифрових двійників» рослини. Створена структура даних є основою для предиктивного моделювання розвитку культури та інтелектуальної підтримки прийняття рішень на різних етапах вегетації.

**Ключові слова:** підтримка прийняття рішень, інтелектуальна система, математичне моделювання, задача оптимізації, граф, онтологія, цифрова модель, цифровий двійник, комп'ютерний зір, цифрове фенотипування, цифровізація, аграрний сектор, десикація, кліматичні показники, екологічні аспекти.

## ABSTRACT

Vedmedev S. R. Decision support in sunflower cultivation based on digital phenotyping technology. – Qualification scientific work in the form of a manuscript.

Dissertation for the degree of Doctor of Philosophy in the specialty 124 – «System Analysis» – National University “Zaporizhzhya Polytechnic”, Zaporizhzhia, 2026

The dissertation research is devoted to solving the scientific and practical problem of decision support in sunflower cultivation based on digital phenotyping technology. Solving this problem is associated with overcoming the discrepancy between the rapid growth of the volume of primary digital agrotechnological data and the low level of their systematic typing and logical structuring for automated decision-making.

**The object** of the study is the decision-making processes in sunflower cultivation in the context of digital transformation of the agricultural industry.

**The subject** of the study is methods of formalizing knowledge based on ontological modeling, algorithms for digital phenotyping of plant objects based on images, and digital data integration technologies for modeling the "genotype - environment" interaction.

**The purpose** of the study is to develop a digital technology for phenotyping a plant object based on an ontological model and forming datasets, which provides a formalized representation, effective integration of phenotypic data for decision-making in sunflower cultivation.

The dissertation consists of an introduction, four chapters, conclusions, a list of references, and an appendix.

In the introduction, the relevance of the dissertation research topic is substantiated; the aim, object, and subject of the study are presented; research methods are described; the scientific novelty and practical significance of the obtained results are formulated; and information on approbation and publications related to the dissertation is provided.

The first chapter, “Processes of Digital Transformation in the Agricultural Sector” presents a comprehensive systems analysis of modern processes of digital transformation in agriculture. It examines current approaches to digital plant phenotyping, the development of digital twins of biological objects, the application of computer vision methods, ontological modeling, and decision support systems in agriculture. Limitations of existing solutions are identified, related to fragmented data, the absence of a unified semantic environment, and insufficient integration of phenotyping results with intelligent data analysis mechanisms.

In the second chapter, “Conceptual Foundations and Tools for Building Digital Models of the Sunflower Plant” a theoretical and methodological framework for constructing a digital model of the sunflower plant is presented, based on the interaction between “genotype–environment.” The primary instrument for knowledge formalization and the construction of the logical framework of the digital model is a domain ontology. The use of ontological modeling and semantic rule systems enables the development of a structured knowledge system capable of automated logical inference at the instance level and the detection of hidden patterns.

In the third chapter, “Ontological Modeling of Sunflower Cultivation” a comprehensive system analysis, formalization, and assessment of the structural and semantic quality of the developed sunflower ontology knowledge base HELIANTUS are carried out.

In the fourth chapter, “Digital Technology for Phenotyping and Dataset Formation” an improved digital phenotyping technology is developed, including the formation of integrated datasets, which ensures semantically consistent data generation in accordance with the HELIANTUS ontological model and their subsequent integration into the knowledge base. A method of digital phenotyping of plant objects based on images is developed, along with its software and algorithmic implementation. Comprehensive experimental studies are conducted, and validation and evaluation of the effectiveness of the developed tools are performed using real biological objects.

To achieve the goal, it was necessary to perform the following **tasks**.

1. To conduct a systems analysis of digital transformation processes in the agricultural sector and define requirements for the formation of an information-analytical environment for sunflower cultivation, ensuring the integration of heterogeneous data, their semantic representation, and their use in decision support systems.

2. To improve the digital model of the sunflower plant based on the “genotype–environment” interaction model.

3. Based on the developed digital sunflower model, to enhance the ontological model that integrates parameters of genotype, breeding, and agronomic cultivation technologies of sunflower, taking into account climatic indicators and environmental aspects.

4. To develop a method of digital phenotyping of plant objects based on computer vision techniques for the automated extraction and quantitative assessment of morphological and physiological traits of sunflower.

5. To improve digital phenotyping technology with the formation of integrated datasets, ensuring automated collection, standardization, and storage of phenotypic traits based on the ontological model.

### **Scientific novelty of the work.**

1. The structure of the digital sunflower plant model has been improved based on a four-level Digital Phenotyping System architecture, which ensures the integration of processes for data collection, storage, processing, and analysis of heterogeneous data. The proposed architecture combines results of laboratory and field phenotyping with environmental parameters, agronomic factors, and remote monitoring sources, creating a foundation for constructing a digital twin of the plant and enabling predictive modeling of its development.

2. The domain ontological model of the sunflower cultivation agro-system HELIANTUS has been improved, ensuring semantic integration of genetic, phenotypic, agroclimatic, and technological data within a unified information space. The proposed approach involves the transformation of digital phenotyping results into

formalized knowledge based on TBox and ABox concepts, enabling the application of logical inference mechanisms for automated analysis of cause-and-effect relationships in the “genotype–environment” system.

3. For the first time, a method of digital phenotyping of plant objects based on images has been developed, which implements optimization of data acquisition parameters. The method enables automated extraction and high-precision assessment of morphological, structural, and color characteristics of sunflower seeds using quantitative descriptors.

4. The digital phenotyping technology has been improved through the formation of integrated datasets, ensuring semantically consistent dataset generation in accordance with the HELIANTUS ontological model and their subsequent integration into the knowledge base.

**Practical Significance of the obtained results** lies in the creation of algorithmic, software, and semantic tools that provide decision support for sunflower cultivation based on digital phenotyping technology.

An information decision-support environment has been developed based on automated logical inference mechanisms within the HELIANTUS ontology. The practical implementation of the ontological model of the sunflower agro-system has enabled the integration of heterogeneous data (genetic, phenotypic, agroclimatic, and technological) into a unified knowledge base. This makes it possible to identify cause-and-effect relationships in the “genotype–environment” system and to make informed decisions regarding the optimization of sunflower cultivation technologies.

Based on the developed image processing method, a software system (implemented in Python) has been created that automates the extraction, segmentation, and analysis of morphometric and color characteristics of sunflower seeds. This provides breeders and agronomists with objective quantitative descriptors for decision-making regarding seed quality, sample identification, and line selection, regardless of imaging conditions and lighting.

The implementation of the four-level Digital Phenotyping System architecture enables the practical integration of laboratory phenotyping, field phenotyping, and remote monitoring data with environmental factors. This can serve as a technological foundation for constructing plant digital twins. The resulting data structure provides a basis for predictive modeling of crop development and intelligent decision support at different stages of vegetation.

**Keywords:** decision support, intelligent system, mathematical modelling, optimization problem, graph, ontology, digital model, digital twin, computer vision, digital phenotyping, digitalization, agricultural sector, desiccation, climate indicators, environmental aspects.

## СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

### Статті у наукових фахових виданнях України:

1. Ведмедєв, С. (2024). Розробка програмного забезпечення для збирання фенотипічних даних насіння соняшнику. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 4, 53–60, <https://doi.org/10.32782/IT/2024-4-7> ISSN: 2786-5088 (Online) <https://journals.politehnica.dp.ua/index.php/it/article/view/707/634>
2. Ведмедєв С.Р., Терещенко Е.В. (2025) Цифрова модель рослини соняшника для фенотипування в задачах селекції. *Інформаційні технології та суспільство*. 3 (18), 32-39. <https://doi.org/10.32689/maup.it.2025.3.4> <https://journals.maup.com.ua/index.php/it/article/view/5153/5474> - сторінка статті ISSN (Online): 2786-5479
3. Ведмедєв С. Терещенко Е. (2026) Фенотипування на основі зображень: ідентифікація оптимальних параметрів отримання зображень для аналізу морфології насіння. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 2, 45-54.

### Матеріали конференцій Scopus

4. Bakurova A., Vedmedeva K., Vedmedev S., Tereschenko E. (2023) Ontological Model of Helianthus Cultivation in Ukrainian Conditions. *Proceedings of the 7th International Conference on Computational Linguistics and Intelligent Systems. Volume II: Computational Linguistics Workshop*, Kharkiv, Ukraine, April 20-21. CoLinS 2023.130-140. ISSN 1613-0073 <https://ceur-ws.org/Vol-3396/paper10.pdf>

### Статті у іноземних фахових журналах

5. Aliiev E., Vedmedeva K., Machova T., Vedmedev S. (2024) Study of the distribution of phenotypic characteristics of sunflower seeds in a head of different genotypes. *Bulgarian Journal of Crop Science*, 62 (4), 73-89. <https://doi.org/10.61308/JJXX8922>.

### Статті в Українських фахових журналах Б спеціальностей агрономія та агроінженерія

6. Махова Т. В., Ведмедєв С. Р., Поляков О. І. (2024) Створення баз даних фенотипових ознак та добір ліній соняшнику особливого призначення. *Науково-технічний бюлетень Інституту олійних культур НААН. Запоріжжя*, 35. 51-62. <https://doi.org/10.36710/IOC-2023-35-05>  
[http://www.bulletin.imk.zp.ua/pdf/2023/35/Makhova\\_35.pdf](http://www.bulletin.imk.zp.ua/pdf/2023/35/Makhova_35.pdf).

### Матеріали конференцій

7. Ведмедєв С.Р., Терещенко Е.В. (2022) Створення цифрової моделі соняшника. *Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій: Тези доповідей XI Міжнародної науково-практичної конференції (12-14 грудня 2022 р., м.Запоріжжя)*. Запоріжжя: НУ «Запорізька політехніка», 136-137.
8. Бакурова А.В., Ведмедєв С.Р., Терещенко Е.В. (2023) Предметна онтологія «Селекція соняшнику». *Економічна кібернетика: інструменти, моделі і методи інтелектуальної підтримки прийняття рішень: збірник наукових праць за матеріалами Всеукраїнської інтернет-конференції, м. Дніпро, 1-2 березня 2023 р. Дніпро : УДУНТ,46-51*

9. Бакурова А. В. Ведмедєв С.Р., Терещенко Е. В. (2023) Розробка компетентносних запитів для онтології HELIANTHUS. *Тиждень науки. Тези доповідей науково-технічної конференції, Запоріжжя, 24-28 квітня 2023 р.* Запоріжжя, 11-12.
10. Бакурова А. В. Ведмедєв С.Р., Терещенко Е. В. (2023) Проблема формування запитів до онтології. *Комбінаторні конфігурації та їхні застосування: Матеріали XXV Міжнародного науково-практичного семінару імені А. Я. Петренюка, Запоріжжя – Кропивницький, 14-16 червня 2023 року*) Запоріжжя : НУ «Запорізька політехніка», 95-100.
11. Ведмедєв С. Р. (2023) Розробка системи фенотипування насіння соняшнику для селекції. *Збірник тез Міжнародної наукової інтернет-конференції «Олійні культури: сьогодення та перспективи».* (21 березня 2023 р.). Запоріжжя. ІОК НААН, 13-14.
12. Ведмедєв С. Р. (2024) Розробка програмного забезпечення для збирання фенотипічних даних насіння соняшнику. *I (VII) міжнародна науково-практична конференція здобувачів вищої освіти і молодих учених «Інформаційні технології: теорія і практика».* Тези доповідей (Дніпро 20 – 22 березня 2024). М-во освіти і науки України, Нац. техн. ун-т «Дніпровська політехніка». Дніпро: Свідлер А.Л., 347-350.
13. Ведмедєв С. Р. (2025) Створення онтології фенотипічних ознак соняшника (*Helianthus phenotype: seed & head*). *Матеріали XXVII Міжнародного науково-практичного семінару «Комбінаторні конфігурації та їхні застосування», присвяченого 125-річчю Національного університету «Запорізька політехніка» (Запоріжжя-Кропивницький-Київ, 4-6 червня 2025 р.),* 149-150. <https://zp.edu.ua/wp-content/uploads/2025/11/cca-2025-proc.pdf>
14. Ведмедєв С. Р. (2025) Створення програмного забезпечення для розпізнавання різнокольорових насінин на різних фонах. *Тиждень науки-2025. Факультет комп'ютерних наук і технологій. Тези доповідей наук.-*

*практ. конференції, Запоріжжя, 14–18 квітня 2025 р. Запоріжжя: НУ «Запорізька політехніка», 399-400.*  
<https://zr.edu.ua/science/conferences/year-2025/conference-2025-3-1/>

15. Ведмедева К.В., Ведмедєв С.Р. (2026) Можливості використання фенотипування в селекції рослин та біотехнології. *Актуальні питання біотехнології, екології та природокористування: матеріали Міжнар. наук. конф., 16–17 квітня 2026 р.* Держ. біотехнол. ун-т. Харків: ДБТУ, <http://btu.kharkov.ua/nauka/konferentsiyi/>
16. Ведмедєв С., Терещенко Е. (2026) Метод цифрового фенотипування рослинних об'єктів на основі зображень. *Комбінаторні конфігурації та їхні застосування: матеріали XXVIII Міжнар. наук.-практ. семінару (Кропивницький-Запоріжжя-Київ, 7–9 травня 2026 р.)* Кропивницький: ПП «Ексклюзив-Систем», 75–79.

## ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ.....	18
ВСТУП.....	20
1 ПРОЦЕСИ ЦИФРОВОЇ ТРАНСФОРМАЦІЇ АГРАРНОЇ ГАЛУЗІ.....	30
1.1 Системний аналіз сучасного стану цифровізації аграрних досліджень	30
1.2 Сучасний стан аграрної галузі в Україні .....	52
1.3 SWOT-аналіз впровадження цифрових та онтологічних технологій у селекції соняшнику в умовах Agriculture 4.0 .....	56
1.4 Висновки до розділу 1 .....	63
2 КОНЦЕПТУАЛЬНІ ЗАСАДИ ТА ІНСТРУМЕНТАРІЙ ПОБУДОВИ ЦИФРОВИХ МОДЕЛЕЙ РОСЛИНИ СОНЯШНИКУ .....	66
2.1 Концепція цифрової моделі рослини та платформені рішення.....	66
2.2 Структура та етапи формування цифрової моделі рослини соняшнику....	75
2.3 Структура та етапи формування цифрової моделі фенотипу соняшнику кондитерського напрямку використання.....	80
2.4 Онтології: принципи, аксіоматика, класи, властивості, обмеження .....	95
2.5 Метрики оцінки якості онтології.....	105
2.6 Висновки до розділу 2 .....	114
3 ОНТОЛОГІЧНЕ МОДЕЛЮВАННЯ ВИРОЩУВАННЯ СОНЯШНИКА.....	116
3.1 Формалізація вимог до онтологічної моделі та детермінація інформаційних потреб користувача .....	116
3.2 Створення архітектури онтологічної моделі.....	126
3.3 Семантична оцінка якості онтологічної моделі.....	149
3.4 Оцінка якості структури онтологічної моделі .....	156
3.5 Висновки до розділу 3 .....	171
4 ЦИФРОВА ТЕХНОЛОГІЯ ФЕНОТИПУВАННЯ ТА ФОРМУВАННЯ DATASETS.....	174
4.1 Цифрова технологія фенотипування із формуванням інтегрованих наборів даних (datasets) .....	175
4.2 Математична постановка задачі цифрового фенотипування рослинних об'єктів.....	178

4.3 Створення датасету фотозображень насінини соняшнику .....	179
4.4 Метод цифрового фенотипування рослинних об'єктів на основі зображень	180
4.5 Визначення оптимального набору параметрів зйомки .....	183
4.6 Розробка програмного забезпечення вивчення морфологічних ознак насінини соняшнику за фотозображеннями.....	190
4.7 Результати експериментів застосування МЦФЗ.....	197
4.8 Висновки до розділу 4 .....	204
ВИСНОВКИ.....	206
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	209
ДОДАТОК А Інструментальні засоби розробки .....	236
ДОДАТОК Б Акти впровадження .....	240
ДОДАТОК В Фрагменти коду.....	243

## ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

**БПЛА** — безпілотний літальний апарат.

**ДСТУ** — Державний стандарт України

**ІСППР** — інтелектуальна система підтримки прийняття рішень

**МЦФЗ** — метод цифрового фенотипування рослинних об'єктів на основі зображень

**ФСМР** — функціонально-структурні моделі рослин

**ШІ** — штучний інтелект

**АBox** (*Assertion Box*) — бокс тверджень; компонента онтології, яка містить знання про конкретні екземпляри класів, їхні властивості та зв'язки між ними

**AI** (*Artificial Intelligence*) — штучний інтелект

**API** (*Application Programming Interface*) — інтерфейс прикладного програмування

**FAIR** (*Findable, Accessible, Interoperable, Reusable*) — міжнародні принципи управління науковими даними: легкознаходжуваність, доступність, сумісність, придатність для повторного використання

**НPK** (*Nitrogen, Phosphorus, Potassium*) — комплекс макроелементів (азот, фосфор, калій) у мінеральному живленні рослин

**OntoGraf** — спеціалізований плагін у середовищі Protégé для інтерактивної візуалізації та аналізу онтологічних структур

**OWL** (*Web Ontology Language*) — мова веб-онтологій; семантичний стандарт W3C для формального представлення знань

**Reasoners** — машини логічного виведення

**RDF** (*Resource Description Framework*) — базовий семантичний стандарт W3C для формального представлення знань у вигляді графової структури атомарних тверджень

**RGB** (*Red, Green, Blue*) — адитивна колірна модель, що описує кодування цифрових зображень через червоний, зелений та синій канали

**SPARQL** (*SPARQL Protocol and RDF Query Language*) — семантична мова запитів до даних, представлених у форматах RDF або OWL

**SWOT** (*Strengths, Weaknesses, Opportunities, Threats*) — метод системного аналізу зовнішніх і внутрішніх факторів ризику

**SWRL** (*Semantic Web Rule Language*) — мова семантичних веб-правил; підмножина правил логічного виведення, що інтегрується з OWL-онтологіями

**TBox** (*Terminological Box*) — термінологічний бокс, компонента онтології, яка визначає її концептуальну структуру (класи, підкласи, аксіоми та властивості)

**W3C** (**World Wide Web Consortium**) — міжнародний консорціум із стандартизації вебтехнологій

## ВСТУП

**Актуальність теми дисертації.** Сучасний розвиток аграрного виробництва характеризується активним впровадженням концепції Agriculture 4.0, яка передбачає широке використання цифрових технологій, засобів моніторингу, штучного інтелекту та аналітики даних для підвищення ефективності виробництва сільськогосподарської продукції. Одним із пріоритетних напрямів цифрової трансформації аграрної галузі є селекція рослин, де результативність прийняття рішень значною мірою залежить від повноти, достовірності та інтегрованості інформації про генотипові, фенотипові, агротехнологічні, кліматичні та екологічні характеристики досліджуваних об'єктів.

Особливої актуальності зазначені питання набувають у селекції соняшнику, яка потребує оброблення значних обсягів різномірних даних, отриманих із польових експериментів, систем моніторингу та результатів фенотипування. Незважаючи на стрімке зростання обсягів даних, накопичення відбувається у розрізних інформаційних ресурсах, набори даних відрізняються структурою, форматами подання та рівнем деталізації, що ускладнює їх комплексне використання в цифровій селекції. У сучасних дослідженнях наголошується на необхідності інтеграції фенотипічних даних та створення спеціалізованих цифрових ресурсів для прискорення селекційного процесу (Tarasov, 2019).

У зв'язку з цим виникає науково-практична суперечність між потребою ефективного використання великих обсягів цифрових агротехнологічних даних у процесах вирощування та селекції соняшнику і недостатнім рівнем їх структурованості, стандартизації та семантичної сумісності для автоматизованого опрацювання. Одним із найбільш перспективних підходів до подолання цієї суперечності є застосування онтологічного моделювання, яке забезпечує формалізоване представлення предметної області та інтеграцію різномірних джерел знань у єдиному інформаційному просторі (Chornyj & Dosyn,

2025). Поєднання онтологічних підходів із методами штучного інтелекту та аналізу даних створює передумови для побудови інтелектуальних систем підтримки прийняття рішень, здатних здійснювати комплексний аналіз інформації та формувати рекомендації для цифрової селекції соняшнику (Zhang et al., 2025).

У зв'язку з цим актуальним науковим завданням є розроблення онтологічної моделі соняшнику та цифрової технології формування семантично сумісних наборів даних як інформаційної основи інтелектуальної системи підтримки прийняття рішень в умовах Agriculture 4.0.

**Зв'язок роботи з науковими програмами, планами, темами.** Дисертаційну роботу виконано на кафедрі системного аналізу та обчислювальної математики Національного університету «Запорізька політехніка» у безпосередньому зв'язку з планами науково-дослідних робіт установи: науково-дослідна робота з номером державної реєстрації 0121U113264, 2021–2024 рр. «Розвиток методів дослідження складних соціально-економічних систем на основі інтелектуальних технологій» та науково-дослідна робота 2024-2027 рр. 05014 «Прийняття ефективних рішень на основі інтелектуальних технологій та відкритих даних».

Дисертаційне дослідження присвячено вирішенню науково-практичної проблеми підтримки прийняття рішень при вирощуванні соняшнику на основі цифрової технології фенотипування. Вирішення цієї проблеми пов'язане з подоланням розбіжності між стрімким зростанням обсягів первинних цифрових агротехнологічних даних та низьким рівнем їх системної типізації й логічної структурованості для автоматизованого прийняття рішень.

**Об'єктом дослідження** є процеси прийняття рішень при вирощуванні соняшника в умовах цифрової трансформації аграрної галузі.

**Предметом дослідження** є методи формалізації знань на основі онтологічного моделювання, алгоритми цифрового фенотипування рослинних

об'єктів за зображеннями та цифрові технології інтеграції даних для моделювання взаємодії «генотип — середовище».

**Мета дослідження** полягає у розробці цифрової технології фенотипування рослинного об'єкта на основі онтологічної моделі та формування датасетів, що забезпечує формалізоване представлення, ефективну інтеграцію фенотипових даних для прийняття рішень при вирощуванні соняшника.

Для досягнення мети необхідно було виконати такі **завдання**.

1. Провести системний аналіз процесів цифрової трансформації аграрної галузі та визначити вимоги до формування інформаційно-аналітичного середовища вирощування соняшнику, що забезпечує інтеграцію різномірних даних, їх семантичне представлення та використання у системах підтримки прийняття рішень.

2. Удосконалити цифрову модель соняшнику, яка ґрунтується на моделі взаємодії «генотип — середовище».

3. На основі розробленої цифрової моделі соняшнику вдосконалити онтологічну модель, яка інтегрує параметри генотипу, селекції та агротехнологій вирощування соняшнику з урахуванням кліматичних показників та екологічних аспектів.

4. Розробити метод цифрового фенотипування рослинних об'єктів на основі методів комп'ютерного зору для автоматизованого виділення та кількісної оцінки морфологічних і фізіологічних ознак соняшнику.

5. Удосконалити цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets), яка забезпечує автоматизований збір, стандартизацію та збереження фенотипових ознак на основі онтологічної моделі.

**Методи дослідження.** Для вирішення поставлених задач у роботі використано:

- **методи системного аналізу** – для визначення архітектури системи цифрового фенотипування; формалізації задачі цифрового фенотипування рослинних об'єктів як оптимізаційної задачі вибору параметрів фотофіксації, що

забезпечує найменшу похибку вимірювання морфологічних ознак; побудови методу цифрового фенотипування рослинних об'єктів на основі зображень;

- **методи онтологічного моделювання** – для концептуалізації предметної області дослідження шляхом визначення таксономічної структури класів, задання семантичних зв'язків між ними та типізації параметрів цифрових двійників рослин;

- **методи дескриптивної логіки** — для побудови логічного каркаса онтологічної бази знань шляхом розмежування термінологічної складової (ТВох, що визначає систему класів, обмежень та аксіом роз'єднаності) і концептуальних тверджень про конкретні об'єкти (АВох, що описує властивості та метрики реальних зразків соняшнику), а також для забезпечення автоматизованого логічного виведення й верифікації структури моделі;

- **методи математичної статистики** – для формулювання та перевірки статистичних гіпотез щодо значущості виявлених переваг певних наборів параметрів зйомки з метою обґрунтування їхнього статусу як «оптимальних» при цифровому фенотипуванні;

- **методи цифрової обробки зображень та алгоритмізації** – для розробки, оптимізації та програмної реалізації (із використанням мови Python) алгоритмів автоматизованого виділення, сегментації та вимірювання морфометричних ознак насіння соняшнику.

### **Наукова новизна одержаних результатів**

1. **Удосконалено** структуру цифрової моделі рослини соняшнику на основі чотирирівневої архітектури Digital Phenotyping System, яка забезпечує інтеграцію процесів збору, збереження, обробки та аналізу різнорідних даних. Запропонована архітектура поєднує результати лабораторного та польового фенотипування з параметрами навколишнього середовища, агротехнологічними факторами та джерелами дистанційного моніторингу, що створює основу для побудови цифрового двійника рослини та реалізації предиктивного моделювання її розвитку.

2. **Удосконалено** предметну онтологічну модель агросистеми вирощування соняшнику HELIANTUS, яка забезпечує семантичну інтеграцію генетичних, фенотипових, агрокліматичних та технологічних даних у межах єдиного інформаційного простору. Запропонований підхід передбачає трансформацію результатів цифрового фенотипування у формалізовані знання на основі концепцій TBox та ABox, що створює можливість застосування механізмів логічного виведення для автоматизованого аналізу причинно-наслідкових зв'язків у системі «генотип — середовище».

3. **Вперше** розроблено метод цифрового фенотипування рослинних об'єктів на основі зображень, який реалізує оптимізацію параметрів реєстрації даних. Метод забезпечує автоматизоване виділення та високоточне оцінювання морфологічних, структурних і колірних характеристик насіння соняшнику за кількісними дескрипторами.

4. **Удосконалено** цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets), яка забезпечує семантично узгоджене формування наборів даних відповідно до онтологічної моделі HELIANTUS та їх подальшу інтеграцію до бази знань.

**Практичне значення одержаних результатів** полягає у створенні алгоритмічного, програмного та семантичного інструментарію, який забезпечує підтримку прийняття рішень при вирощуванні соняшнику на основі цифрової технології фенотипування.

Створено інформаційне середовище прийняття рішень на основі механізмів автоматизованого логічного виведення в онтології HELIANTUS. Практична реалізація онтологічної моделі агросистеми вирощування соняшнику дозволила об'єднати різноманітні дані (генетичні, фенотипові, агрокліматичні, технологічні) в єдину базу знань. Це дає змогу виявляти причинно-наслідкові зв'язки в системі «генотип — середовище» та приймати обґрунтовані рішення щодо оптимізації технологій вирощування соняшнику.

На основі розробленого методу обробки цифрових зображень створено програмний комплекс (мовою Python), який автоматизує виділення, сегментацію та аналіз морфометричних і колірних характеристик насіння соняшнику. Це забезпечує селекціонерів та агрономів об'єктивними кількісними дескрипторами для прийняття рішень щодо якості насіннєвого матеріалу, ідентифікації зразків та відбору ліній незалежно від умов зйомки й освітлення.

Впровадження чотирирівневої архітектури Digital Phenotyping System забезпечує практичну можливість консолідації даних лабораторного, польового фенотипування та дистанційного моніторингу з факторами довкілля, що може стати технологічною основою для побудови «цифрових двійників» рослини. Створена структура даних є основою для предиктивного моделювання розвитку культури та інтелектуальної підтримки прийняття рішень на різних етапах вегетації.

Результати досліджень, зокрема розроблений програмний комплекс, що містить модулі для визначення геометричних (фенотипових) характеристик та кольору насінини соняшника, впроваджено на кафедрі «Системний аналіз та обчислювальна математика» Національного університету «Запорізька політехніка» у межах кафедральної НДР № 05014 «Прийняття ефективних рішень на основі інтелектуальних технологій та відкритих даних». Матеріали дисертації використовуються як лекційні курси та практичні приклади під час проведення лабораторних занять з дисциплін «Теорія прийняття рішень», «Аналіз даних», «Основи системного аналізу» та «Моделювання соціально-економічних систем», а також інтегровані як змістовні задачі у курсове та дипломне проєктування здобувачів першого (бакалаврського) та другого (магістерського) рівнів вищої освіти спеціальності 124 «Системний аналіз» (Акт впровадження від 28.11.2025 р., ДОДАТОК Б).

Розроблену систему цифровізації візуальної інформації для дослідження фенотипу соняшнику впроваджено в лабораторії генетики та генетичних ресурсів Інституту олійних культур Національної академії аграрних наук

України у межах виконання наукового проєкту ДБ05022 (РК 0121U113264) «Розвиток методів дослідження складних соціально-економічних систем на основі інтелектуальних технологій». Матеріали дослідження застосовано у процесах розробки приладу фенотипування насіння соняшнику та для побудови бази даних морфологічних ознак ліній соняшнику для подальшого проведення селекційних доборів та ідентифікації зразків (Акт впровадження № 448 від 20.12.2023 р., ДОДАТОК Б).

Результати дисертаційної роботи, а саме метод цифрового фенотипування об'єктів агросистем на основі аналізу цифрових зображень та онтологічну модель вирощування соняшнику, впроваджено в наукову діяльність Інституту олійних культур Національної академії аграрних наук України (лабораторія генетики та генетичних ресурсів). Розроблені цифрові технології використовуються для фенотипування насіння соняшнику, а також для виконання фундаментальних завдань НААН у межах науково-дослідних робіт: ПНД 17 «Генетичні ресурси рослин» (РК 0120U105673) та ПНД 25 «Формування, використання і збереження генетичних ресурсів рослин» (РК 0126U001156В) з метою формування колекцій олійних культур, виділення та створення джерел і донорів цінних господарських ознак (Акт впровадження від 18.06.2026 р. № 162, ДОДАТОК Б).

### **Особистий внесок здобувача**

Усі основні наукові результати дисертаційної роботи отримано автором особисто. У роботах, опублікованих зі співавторами, здобувачеві належать такі результати: (Ведмедєв та Терещенко, 2025) — запропоновано концептуальний підхід до побудови цифрової моделі рослини соняшника, формалізовано морфологічні параметри для фенотипування та обґрунтовано їх застосування в задачах автоматизації селекційного процесу; (Ведмедєв та Терещенко, 2026) — розроблено метод цифрового фенотипування рослинних об'єктів на основі методів комп'ютерного зору для автоматизованого виділення та кількісної оцінки морфологічних і фізіологічних ознак соняшнику; (Bakurova, Vedmedeva, Vedmedev&Tereschenko, 2023) — розроблено семантичну структуру

онтологічної моделі вирощування соняшнику (HELIANTUS), сформовано концептуальні класи (TBox) та екземпляри (ABox), а також запропоновано систему SWRL-правил для автоматизованого логічного виведення; (Aliiev, Vedmedeva, Machova & Vedmedev, 2024) — розроблено алгоритмічне забезпечення для аналізу просторового розподілу фенотипових характеристик насіння, виконано математичну обробку експериментальних даних та верифікацію результатів для різних генотипів; (Махова, Ведмедєв, Поляков, 2024) — спроектовано схему бази даних фенотипових ознак соняшнику, розроблено програмні модулі інтеграції даних для автоматизованого добору ліній особливого призначення за вимогами Інституту олійних культур НААН.

**Апробація результатів дисертації.** Основні результати дисертаційного дослідження були оприлюднені та апробовані на таких наукових заходах:

7th International Conference on Computational Linguistics and Intelligent Systems, Volume II: Computational Linguistics Workshop (Kharkiv, Ukraine, April 20–21, 2023);

XI Міжнародній науково-практичній конференції «Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій» (м. Запоріжжя, 12–14 грудня 2022 р.);

Всеукраїнській інтернет-конференції «Економічна кібернетика: інструменти, моделі і методи інтелектуальної підтримки прийняття рішень» (м. Дніпро, 1–2 березня 2023 р.);

Науково-технічній конференції «Тиждень науки – 2023» (м. Запоріжжя, 24–28 квітня 2023 р.);

XXV Міжнародному науково-практичному семінарі імені А. Я. Петренюка «Комбінаторні конфігурації та їхні застосування» (Запоріжжя – Кропивницький, 14–16 червня 2023 р.);

Міжнародній науковій інтернет-конференції «Олійні культури: сьогодення та перспективи» (м. Запоріжжя, ІОК НААН, 21 березня 2023 р.);

I (VII) Міжнародній науково-практичній конференції здобувачів вищої освіти і молодих учених «Інформаційні технології: теорія і практика» (м. Дніпро, 20–22 березня 2024 р.);

XXVII Міжнародному науково-практичному семінарі «Комбінаторні конфігурації та їхні застосування», присвяченому 125-річчю Національного університету «Запорізька політехніка» (Запоріжжя – Кропивницький – Київ, 4–6 червня 2025 р.);

Науково-практичній конференції «Тиждень науки – 2025» (Факультет комп'ютерних наук і технологій, м. Запоріжжя, 14–18 квітня 2025 р.);

Міжнародній науковій конференції «Актуальні питання біотехнології, екології та природокористування» (Державний біотехнологічний університет, м. Харків, 16–17 квітня 2026 р.);

XXVIII Міжнародному науково-практичному семінарі «Комбінаторні конфігурації та їхні застосування» (Кропивницький – Запоріжжя – Київ, 7–9 травня 2026 р.).

**Публікації.** За темою дисертації опубліковано 16 наукових праць, з яких 4 статті у наукових фахових виданнях України, 1 стаття у науковому виданні іншої держави, 9 тез наукових доповідей, 1 матеріали міжнародної конференції, що індексовані у Scopus.

**Використання засобів штучного інтелекту.** Під час підготовки тексту дисертації автором було залучено інструментарій генеративного штучного інтелекту на базі розширених мовних моделей (LLM Gemini від Google) для покращення мовного стилю та для генерації рисунку 4.1.1, 4.6.1 (промпт представлено в Додатку А). Усі текстові фрагменти, попередньо сформовані за допомогою ШІ, були ретельно перевірені, змістовно доповнені та відредаговані автором перед включенням до остаточної редакції дисертації.

**Структура та обсяг дисертації.** Дисертація викладена на 247 сторінках комп'ютерного тексту. Включає анотацію, вступ, 4 розділи, які містять 22

таблиці, 34 рисунка, загальні висновки. Список використаних джерел налічує 232 найменування, у тому числі 187 – латиницею. Містить три додатки.

# 1 ПРОЦЕСИ ЦИФРОВОЇ ТРАНСФОРМАЦІЇ АГРАРНОЇ ГАЛУЗІ

## 1.1 Системний аналіз сучасного стану цифровізації аграрних досліджень

Сучасний етап розвитку аграрної галузі характеризується активним впровадженням цифрових технологій, що формують основу концепції Agriculture 4.0 (Pilarczyk et al., 2024). На відміну від традиційних підходів до автоматизації виробництва, Agriculture 4.0 передбачає створення інтегрованого цифрового середовища, у межах якого здійснюється збір, обробка, аналіз та інтерпретація даних для підтримки прийняття рішень на всіх рівнях управління агровиробничими процесами.

Перехід до індустріальної парадигми AgTech 4.0 є ключовим чинником структурної трансформації аграрного сектору. Впровадження наскрізних цифрових платформ дозволяє суттєво підвищити факторну продуктивність підприємств, що, за висновками Васильєвої та Душина (2026), є базовою умовою для забезпечення економічної стійкості, формування вищої доданої вартості та сталого розвитку бізнесу в сучасних турбулентних умовах.

Цифрова трансформація аграрної галузі супроводжується переходом від реактивного управління до проактивного управління, заснованого на використанні інтелектуальних систем, в яких підтримка прийняття рішень щодо оптимізації виробничих процесів відбувається на аналізі великих масивів різномірних даних. У такому контексті особливого значення набувають системи підтримки прийняття рішень (Decision Support Systems, DSS), які інтегрують інформацію про стан посівів, характеристики ґрунтів, кліматичні показники, технічні ресурси та економічні параметри господарства (Zhai et al., 2020).

Однією з фундаментальних складових Agriculture 4.0 є математичне моделювання агровиробничих процесів. Математичні моделі дозволяють описувати взаємозв'язки між природними, технологічними та економічними факторами, прогнозувати розвиток агроєкосистем, оцінювати можливі сценарії управління та визначати оптимальні стратегії використання ресурсів.

Особливу роль у функціонуванні інтелектуальних систем відіграють кліматичні показники, які безпосередньо впливають на продуктивність сільськогосподарського виробництва. У межах концепції Agriculture 4.0 відбувається моніторинг за допомогою сенсорних мереж, метеостанцій, супутникових систем спостереження та безпілотних літальних апаратів, після чого отримані дані використовуються для прогнозування врожайності та формування управлінських рішень (Гнатієнко та ін., 2024).

Важливим напрямом розвитку цифрового сільського господарства є врахування екологічних аспектів агровиробництва. Використання інтелектуальних систем дозволяє мінімізувати надлишкове внесення добрив і засобів захисту рослин, оптимізувати споживання водних ресурсів, контролювати рівень антропогенного навантаження на ґрунти та знижувати викиди парникових газів. Таким чином, цифрова трансформація сприяє не лише підвищенню економічної ефективності аграрних підприємств, а й забезпеченню принципів сталого розвитку та екологічної безпеки (Trusova et al., 2022).

Важливою складовою Agriculture 4.0 є Селекція 4.0 (Breeding 4.0), яка передбачає використання цифрових технологій, геномних даних, високопродуктивного фенотипування та методів штучного інтелекту для прискорення селекційного процесу. У такій системі математичне моделювання та інтелектуальні системи забезпечують прогнозування адаптивності генотипів до різних кліматичних умов, що сприяє створенню нових сортів і гібридів, стійких до посухи, температурних стресів та інших несприятливих факторів середовища.

Селекція сільськогосподарських культур — це цілеспрямований процес відбору та схрещування рослин з бажаними ознаками з метою покращення їхніх продуктивних, якісних та адаптивних властивостей у наступних поколіннях. Селекція є ключовим елементом аграрної науки та основою підвищення продуктивності і стабільності агроecosистем (Hudzenko et al., 2022; Jacobsen et al., 2020).

Цілями селекції рослин є створення нових та покращення існуючих сортів і гібридів, що характеризуються високою врожайністю, підвищеною стійкістю до хвороб, шкідників та несприятливих кліматичних умов (посуха, холод) (García-Coronado et al., 2025), а також покращеною якістю продукції (смак, поживність, вміст вітамінів).

Цифрові методи дозволяють зібрати великі масиви даних та підвищити точність оцінки фенотипу: high-throughput phenotyping із використанням дронів, сенсорів і інфрачервоного сканування; інтеграція моделі взаємодії «генотип — середовище» (G×E) для прогнозування поведінки рослин у різних середовищних умовах; використання цифрових моделей і симуляцій ідеотипів рослин, що дозволяє проєктувати бажані морфологічні та фізіологічні ознаки (Yoosefzadeh-Najafabadi, 2025).

Концепція «Селекції 4.0» (Breeding 4.0) є сучасним етапом розвитку аграрної науки, який базується на повній цифровізації, використанні технологій штучного інтелекту, Big Data та високоточного редагування геному. Це перехід від «мистецтва відбору» до прецизійної інженерії біологічних систем, де головним інструментом селекціонера стає не лише поле, а й потужний обчислювальний кластер.

Ключовими складовими Селекції 4.0 є такі.

1. Високопродуктивне фенотипування (High-Throughput Phenotyping). Раніше дані збирали вручну (лінійкою в полі). Тепер це роблять БПЛА з мультиспектральними камерами, наземні роботи та IoT-сенсори. Це дозволяє отримувати тисячі вимірювань у динаміці, формуючи «цифровий слід» рослини (Kundu et al., 2024).

2. Big Data та Штучний інтелект (AI) Інтеграція «омікських» даних (геноміка, протеоміка, метаболоміка) з кліматичними даними створює гігантські масиви інформації. Алгоритми машинного навчання прогнозують результат схрещування ще до того, як насіння буде висіяне в ґрунт (Zhu et al., 2025).

3. Моделювання взаємодії «генотип — середовище». Це серце

концепції 4.0. Замість того, щоб просто вимірювати врожай, система буде математичну модель того, як конкретний генотип поведеться в тисячах варіантів умов середовища. Це дозволяє створювати «дизайнерські» сорти для конкретних полів (Lucido et al., 2025).

4. Редагування геному (CRISPR/Cas9). Точкова зміна ДНК дозволяє вносити необхідні правки в геном швидше та точніше, ніж традиційні методи селекції, фактично «програмуючи» стійкість до хвороб або посухи (García-Coronado et al., 2025).

5. Інтероперабельність та стандарти (BrAPI) Обмін даними між лабораторіями світу в режимі реального часу. Це перетворює селекцію на глобальну мережу знань, де результати випробувань у Бразилії можуть допомогти селекціонеру в Україні (Selby et al., 2025).

Отже, сучасна селекція поєднує традиційні та інноваційні методи, цифрове фенотипування. Це забезпечує ефективне створення нових сортів, підвищення їх стійкості до стресових факторів та оптимізацію насінництва, що є ключовими завданнями аграрної науки в умовах глобалізації та зміни клімату (Singh et al., 2015, Danzi et al., 2022).

Для селекції окремих об'єктів існують розроблені схеми селекційного процесу та звичайні методики вимірювання рівня ознак.

Фенотип - це сукупність властивостей і ознак організму, що склалися на основі взаємодії генотипу із умовами зовнішнього середовища на конкретному етапі розвитку, що є основою концептуальної моделі взаємодії «генотип — середовище» ( $G \times E$ ). Ознаки організму можна поділити на такі які визначаються зовнішнім виглядом та його складовими, та ознаки які неможна безпосередньо побачити – біохімічні (вміст окремих речовин та їх співвідношення: білки, жири, вуглеводи та їх краси і підкласи). Зазвичай фенотип це саме те що можна побачити зовні: колір, розміри, розташування, форма кожного з органів та їх частин. Ці ознаки є основою для селекції рослин (Алієв Е. Б., Ведмедєва К.В., 2024).

Фенотип залишається основним критерієм відбору, але сучасні технології дозволяють оцінювати його точніше та комплексніше, інтегруючи дані молекулярних маркерів, цифрового фенотипування та моделювання взаємодії «генотип — середовище» (G×E) (Fan et al., 2021).

Для успішного розвитку селекції при використанні великих масивів даних рекомендовано інвестувати у створення надійних, інтегрованих баз даних, які централізують різноманітні типи даних, пов'язаних з комп'ютерною селекцією рослин. Такі платформи не лише покращать доступність даних, але й покращать спільні зусилля між дослідницькими установами та географічними локаціями.

Хмарні інфраструктури можуть сприяти обміну даними в режимі реального часу та спільному аналізу, дозволяючи комп'ютерним селекціонерам рослин ефективніше використовувати глобальні генетичні ресурси. (Yoosefzadeh-Najafabadi, 2025).

Концепція FAIR-принципів сформувалася як відповідь на зростаючий обсяг наукових даних та потребу у забезпеченні їх довготривалої придатності, прозорості й повторного використання в різних галузях, зокрема біоінформатиці, агрономії, геномиці та екології. Витоки принципів пов'язані з рухом за відкриті дані, що активізувався у 2000-х роках, коли наукові ініціативи, такі як GenBank, Human Genome Project та Protein Data Bank, зіткнулися з проблемами стандартизації форматів, описів і доступу до даних.

Після 2016 року принципи FAIR були швидко інституціоналізовані. Європейська Комісія, створюючи Європейський відкритий науковий простір (EOSC), включила FAIR як центральний нормативний компонент. Провідні фінансуючі організації, як-от Horizon Europe, NSF (США) та Wellcome Trust, почали вимагати дотримання FAIR у проектних заявках та планах управління даними. У галузі біології та аграрних наук розвиток FAIR пришвидшили великі міжнародні ініціативи — ELIXIR, CGIAR Big Data Platform, AgBioData Consortium, які створили інфраструктури і стандарти для метаданих, онтологій

та репозиторіїв, що забезпечують інтероперабельність даних (Wang J.L. et al., 2024).

У 2020-х роках акцент у розвитку FAIR змістився від декларативних принципів до їхнього технічного забезпечення: створення API-орієнтованих репозиторіїв, машинно-читаємих метаданих, стандартів онтологічного опису, схем перевірки FAIRness (FAIR Maturity Models), а також інструментів автоматизованої оцінки FAIR-рівня даних. У сучасних дослідженнях, зокрема в фенотипуванні та моделюванні взаємодії «генотип — середовище» (G×E), FAIR-принципи стали передумовою інтеграції багатомодальних даних — генотипових, фенотипових, екологічних, експериментальних метаданих і польових вимірювань (Jacobsen et al., 2020).

Однак, з розвитком технологій, особливо з експоненціальним зростанням передових систем, таких як штучний інтелект (ШІ) та системи цифрових двійників, виникло кілька обмежень у спектрі принципів FAIR. Ці обмеження в сферах інтерпретованості, справедливості, відтворюваності, машинної дії та конфіденційності стають дедалі критичнішими, оскільки системи прийняття рішень на основі даних стають більш автономними та ефективними. (Bashir, 2025).

Сучасна аграрна наука характеризується інтенсивною інтеграцією цифрових технологій, що забезпечує підвищення ефективності дослідницьких та селекційних процесів. Інтегрована цифрова інфраструктура селекції — це комплекс апаратних, програмних та інформаційних ресурсів, що забезпечує стандартизоване збирання, зберігання, обробку, інтеграцію та аналіз генотипових, фенотипових та екологічних даних. До ключових тенденцій інтеграції цифрових технологій в аграрну галузь належать автоматизація польових експериментів, використання сенсорних систем і дистанційного моніторингу, а також комплексний збір та аналіз багатоджерельних даних, які формують основу для переходу до цифрових селекційних платформ.

Автоматизація польових експериментів. Одним із базових джерел, що заклало фундамент автоматизації у польових дослідженнях, є робота Pierce та Nowak (Pierce and Nowak, 1999). У ній обґрунтовано концепцію прецизійного землеробства та механізми використання ранніх GPS-технологій для підвищення точності експериментальних вимірювань. Автори показують, що інтегровані системи навігації стали основою автоматизованого збору польових даних.

Дослідження Shanahan з співавторами (2001) демонструє перехід до напівавтоматичних систем фіксації просторової мінливості урожайності. У статті запропонована методологія використання вбудованих сенсорів у комбайнах для прямої автоматичної реєстрації врожайності. Це стало першим кроком до створення польових систем реального часу.

У роботі Auernhammer (2001) детально описано концепцію ISOBUS як універсального стандарту комунікації між сільськогосподарськими машинами. Стаття показує, як інтеграція стандарту ISO 11783 створила умови для уніфікації автоматизованого обладнання в польових експериментах. Це поклало початок розвитку автоматизованих платформ з порівнянною міжмашинною сумісністю.

Дослідження Iqbal з співавторами (2020) представляє автономну роботизовану платформу для фенотипування високої пропускнуої здатності. Роботизована система здатна працювати у повністю автономному режимі, збираючи морфологічні та спектральні дані рослин. Це демонструє перехід до безперервного автоматизованого фенотипування у польових умовах.

Використання сенсорних систем і дистанційного моніторингу. Робота Mulla (2013) узагальнює розвиток дистанційного зондування у прецизійному землеробстві протягом 25 років. Автор аналізує роль спектральних індексів, супутникових даних та методів обробки зображень у оцінюванні стану посівів. Стаття також визначає ключові виклики, серед яких — роздільна здатність та точність алгоритмів.

У дослідженні Hatfield та співавт. (2008) розглянуто фізіологічні основи спектральних індексів і принципи їх використання для діагностики стресів.

Автори пропонують стандартизовані підходи до інтерпретації спектральних даних, що стало основою сучасних сенсорних систем. Робота демонструє причинно-наслідковий зв'язок між фізіологією рослин і дистанційними вимірюваннями.

Стаття Jones & Vaughan (2010) систематизує основи дистанційного моніторингу рослин у агрономії, включно з порівнянням різних сенсорних платформ. Автори аналізують наземні та аерокосмічні сенсори, підкреслюючи їх сильні та слабкі сторони. Це джерело стало одним з перших комплексних описів мультисенсорних систем у сільському господарстві.

Сучасне дослідження Di Tocco та співавт. (2023) представляє мультисенсорну платформу для моніторингу рослин у режимі реального часу. Інтеграція мультиспектральних камер, LIDAR, мікрокліматичних сенсорів дає змогу комплексно оцінювати фізіологічний стан рослин. Автори демонструють, що мультисенсорна інтеграція значно підвищує точність діагностики.

Комплексний збір та аналіз багатоджерельних даних. Робота Cooreg зі співавторами (2014) є фундаментом сучасних цифрових моделей продуктивності рослин. Автори інтегрують генотипові, фенотипові та екологічні дані в єдину систему прогнозування. Дослідження визначає аналітичні основи для моделювання G×E-взаємодій.

У статті François Tardieu et al. (2017) запропоновано системний підхід до організації даних для цифрової фенотипізації. Автор підкреслює необхідність структурованих протоколів та уніфікації форматів даних у польових експериментах. Публікація пропонує рамкову модель для інтеграції багатоджерельної інформації.

Робота Rincet зі співавторами (2018) демонструє можливість поєднання генотипових маркерів, даних середовища та просторової статистики для покращення прогнозування селекційної цінності. Автори застосовують сучасні статистичні моделі для багаторівневої інтеграції інформації. Робота вважається одним із ключових зразків аналітичних моделей у цифровій селекції.

Сучасне дослідження (Murphy et al., 2024) показує використання машинного навчання для інтеграції даних із сенсорних платформ, супутників та систем високопродуктивного фенотипування. Автори демонструють, що поєднання гетерогенних наборів даних підвищує точність прогнозування селекційної цінності.

Цифровізація селекційних процесів суттєво трансформує управління генетичною та фенотипічною інформацією, насамперед завдяки появі інфраструктур, здатних інтегрувати генотипічні, фенотипічні та допоміжні метадані в єдиному інформаційному просторі. Платформи на кшталт BreedBase забезпечують уніфіковані сховища даних і функціонал для їх аналітичної обробки, що робить можливим швидке виявлення статистичних зв'язків між генотипом та фенотипом, пришвидшуючи селекцію перспективних ліній і гібридів. Така інтегрованість суттєво підвищує точність прогнозування агрономічно важливих ознак, що відзначено в сучасних дослідженнях інфраструктур для breeding informatics (Morales et al., 2022).

Впровадження цифрових систем також сприяє стандартизації протоколів збору, валідації та архівації даних, що є особливо важливим для мультисайтових експериментів. Інформаційні системи фенотипування, як-от PHIS, забезпечують суворе дотримання метаданих, контроль версій, комплексний опис експериментальних умов та автоматизоване відстеження технологічних операцій. Це зменшує частку людських помилок, підвищує реплікованість польових досліджень і дає можливість порівнювати дані між різними лабораторіями та селекційними програмами. Ефективність таких підходів підтверджено в роботах, присвячених стандартизації та управлінню великими масивами фенотипових даних (Neveu et al., 2019).

Ключовим наслідком цифровізації є можливість застосування методів геномної селекції та обчислювальних предиктивних моделей, що істотно скорочує тривалість селекційного циклу. Системи, орієнтовані на підтримку breeding pipelines, як-от Breeding Management System (IBP), дозволяють

об'єднувати результати польових випробувань, генотипування, геномного прогнозування та середовищних моделей в єдиний робочий процес. Додатково, репозиторії типу PGP (Plant Genomics and Phenomics Research Data Repository) забезпечують відкритий доступ до великих масивів фенотипових та геномних даних, полегшуючи колаборацію, мета-аналіз та глобальне масштабування нових гібридів. Доведено, що такі системи пришвидшують прийняття селекційних рішень та підвищують їх точність, оскільки забезпечують принцип FAIR-даних — знаходжуваність, доступність, інтеоперабельність та повторне використання (Arend et al., 2016).

Серед найбільш важливих і поширених цифрових платформ та інструментів, що вже використовуються в міжнародних селекційних програмах, можна виокремити такі:

BreedBase (зокрема Cassavabase, Sweetpotatobase) — екосистема баз даних для управління генотиповими, фенотиповими та польовими даними з підтримкою Crop Ontology та повною інтеграцією інструментів аналізу (США (Університет Корнелла / Cornell University, Інститут Бойса Томпсона / Boyce Thompson Institute)) (Lyons, 2021);

PHIS (Phenotyping Hybrid Information System) — інтегрована інформаційна система польового фенотипування, розроблена INRAE, що забезпечує збір, управління й аналітику різних потоків фенотипових даних (Франція, INRAE – Національний науково-дослідний інститут сільського господарства, продовольства та навколишнього середовища) (Neveu et al., 2019);

Breeding Management System (BMS, IBP) — комплекс інструментів для планування експериментів, збору польових даних, їх цифрової стандартизації та подальшого аналізу в рамках селекційних програм (Міжнародний консорціум (координується Integrated Breeding Platform (IBP), яка є частиною системи CGIAR. Розробка здійснювалася за участі різних центрів, зокрема з Нідерландів та Великої Британії) ('The Breeding Management System,' 2019);

FieldBook — мобільний застосунок для стандартизованого польового збору фенотипових даних, який інтегрується з платформами BMS та BreedBase (США, Університет штату Канзас / Kansas State University). (Rife and Poland, 2026);

AgroFIMS — система структурованого цифрового опису агрономічних і селекційних експериментів на основі стандартів MIAPPE і Crop Ontology (Міжнародний консорціум (Alliance of Bioversity International and CIAT, частина системи CGIAR, з ключовим внеском із Франції). (AgroFIMS: Agricultural Field Information Management System, 2024.);

PlantCV та ImageBreed — набори інструментів для автоматизованої обробки зображень і високопродуктивного фенотипування, що дозволяють інтегрувати дані з камер і сенсорів у єдині робочі процеси (США (Donald Danforth Plant Science Center, Університет штату Оклахома, США (Університет Корнелла / Cornell University, Boyce Thompson Institute) (PlantCV PlantCV: Open-Source Image Analysis for Plant Phenotyping, 2024);

DSSAT та APSIM — провідні моделі симуляції росту й розвитку сільськогосподарських культур, що широко використовуються як компоненти предиктивних селекційних платформ для попереднього оцінювання реакції генотипів на варіювання середовища (США (Ініційовано під егідою Проєкту IBSNAT, штаб-квартира була в Університеті Гаваїв; зараз підтримується DSSAT Foundation, Австралія (Agricultural Production Systems Research Unit – APSRUCSIRO) (ImageBreed: Image-Based Phenotyping Tools, 2023);

AlphaSimR, BGLR, BGGE — статистичні та симуляційні R-пакети для побудови геномних прогнозних моделей, генерування *in silico* селекційних циклів і моделювання G×E у селекційних програмах (Міжнародна розробка (AlphaSimR – Австралія, BGLR/BGGE – вчені з різних установ, зокрема США та Мексики). (Faux, A. et al. 2016; Holzworth, D. et al., 2024).

Ініціативи цифрових двійників рослин у межах проєктів Horizon Europe та Cybele — формують нове покоління систем, здатних симулювати розвиток

фенотипу в різних агроекологічних сценаріях на основі багаторівневих моделей росту. Програми Horizon Europe та проєкт Cybele є загальноєвропейськими ініціативами, що об'єднують консорціуми з багатьох країн-членів ЄС (Horizon Europe: Digital Twins for Crops Initiative, 2023).

Фенотипування рослин — це комплексне вимірювання морфологічних, фізіологічних, біохімічних та екологічних ознак, які відображають прояв генотипу в конкретних умовах середовища. Фенотипування рослин у сучасному розумінні – це комплексна оцінка складних ознак рослин, таких як ріст, розвиток, толерантність, стійкість, архітектура, фізіологія, екологія, врожайність, а також базове вимірювання окремих кількісних параметрів, які формують основу для більш складних ознак (Watt et al., 2020; Abebe et al., 2023). Фенотип рослини включає ці складні ознаки, і прикладами їх прямих вимірювальних параметрів є морфологія коренів (Clark et al., 2011; Walter et al., 2009; Kumar et al., 2014; Flavel et al., 2012), біомаса (Golzarian et al., 2011; Clark et al., 2011), характеристики листя (Golzarian et al., 2011; Arvidsson et al., 2011), характеристики плодів (Brewer et al., 2006; Monforte et al., 2013), ознаки, пов'язані з урожайністю (Duan L. F. et al., 2011), фотосинтетичну ефективність (Bauriegel et al., 2011), а також реакцію на біотичний та абіотичний стрес (Rao and Laxman, 2013; Balachandran et al., 1997). Окремі дослідники навіть виділяють галузь науки феноміку, яка забезпечує високопродуктивну кількісну оцінку ознак рослин і тварин на системному рівні. Феноміка, або результат фенотипування рослинних об'єктів за останнє десятиріччя отримала достатньо широке застосування, зокрема і у селекції.

Сучасна селекція й агробіологічні дослідження потребують стандартизованих, масштабованих та надійних методів фенотипування, щоб коректно співвідносити фенотип із генотипом та умовами вирощування (SpringerLink. High-Throughput Plant Phenotyping..., 2022)

Традиційні та лабораторні методи фенотипування, що використовувалися десятиліттями, передбачають ручні вимірювання морфометричних показників

(наприклад, висота рослини, довжина листків, маса насіння), візуальну оцінку стану рослин, а також фізіологічні або біохімічні аналізи (вміст пігментів, білків, реакція на стрес тощо). Такі підходи зручні для обмеженої кількості зразків і є відносно недорогими. Проте вони мають суттєві недоліки: висока трудомісткість, суб'єктивність оцінки, обмеження масштабів дослідження та низьку продуктивність (Gosseau et al., 2019).

У відповідь на ці обмеження з'явилися методи фенотипування в контрольованих умовах теплиць або ростових камер. Фенотипування рослин стало основним напрямом досліджень у селекції рослин (Neumann et al., 2015).

Тепличне фенотипування дозволяє стандартизувати фактори середовища — освітлення, температуру, вологість — і тим самим зменшити змінність, зумовлену зовнішніми факторами. Це дає змогу простежити розвиток рослин, їх реакції на умови росту чи стреси, а також порівнювати різні генотипи в однакових умовах. Водночас такі умови можуть не відтворювати повною мірою природні агроекологічні — і отже, результати не завжди корелюють з польовою продуктивністю (Gosseau et al., 2019).

Революцію у фенотипуванні принесли високопродуктивні (high-throughput) методи, поєднані з сенсорними технологіями та обробкою даних автоматизованими системами. Зокрема, автоматизовані тепличні платформи із камерною фотозйомкою дають змогу регулярно фіксувати зміни морфології рослин — зокрема висоти, площі надземної частини, швидкості росту тощо. Так, у класичному дослідженні було представлено плагін для ImageJ, який автоматично обчислює такі параметри, як висота, ширина та проєктована площа надземної частини на основі зображень з бокового та верхнього ракурсу (Choudhury et al., 2019). Такі підходи значно підвищують продуктивність збору даних і зменшують трудовитрати. В останні роки високопродуктивні платформи фенотипування були розгорнуті в ростових камерах або оранжереях (Abebe et al., 2023; Kundu et al., 2024; Liu F. et al., 2025a). Ці платформи використовують робототехніку (Шаптала & Миронова, 2023), точний контроль навколишнього

середовища та технології обробки зображень (апаратне та програмне забезпечення) для оцінки росту та продуктивності рослин.

Однак ці платформи призначені для обмеженого кола видів, охоплюючи невеликі рослини-розетки, такі як арабідопсис (Fang et al., 2016) та первинні зернові культури (Jansen et al., 2009; Tardieu F. et al., 2018; Li Y. W. et al., 2021). Необхідно розробити загальні платформи та рішення, що дозволяють одночасно проводити фенотипну оцінку кількох видів. Зараз вже апробують обладнання для дослідження великих рослин, зокрема соняшника (Gosseau et al., 2019; Balliau et al., 2021).

Зображувальні методи фенотипування (imaging phenotyping) сьогодні є одними з найпотужніших і найуніверсальніших. У межах цих методів використовують різні технології: RGB камери, мультиспектральну й гіперспектральну зйомку, теплове (інфрачервоне) зображення, флуоресцентну візуалізацію, тривимірне 3D-сканування, а також томографічні методи (MRT, PET, CT) для детальних структурних і функціональних досліджень (Li L. et al., 2014; Iyer-Pascuzzi et al., 2010). За допомогою цих технологій можна оцінювати як морфологію (форма, об'єм, структура), так і фізіологічний стан — наприклад, статус водного обміну, фотосинтетичну активність, відповідь на стрес, захворювання тощо (Li L. et al., 2014; Iyer-Pascuzzi et al., 2010).

Сенсорні та роботизовані системи для польового фенотипування відкривають нові можливості для масштабних досліджень у реальних агроєкосистемах. Згідно з останнім оглядом, автономні наземні роботи і мобільні платформи можуть ефективно збирати фенотипічні дані на великих площах, вимірюючи параметри, які раніше було важко або неможливо оцінити вручну — наприклад, дрібномасштабні морфометричні показники, динаміку росту, структуру рослини, а також відповіді на змінні умови середовища (Shi et al., 2025). Нові системи роблять фенотипування більш об'єктивним, менш залежним від людського фактора, і дають змогу охопити великі генетичні популяції. Так створено автоматизовані тепличні системи для високопродуктивного

фенотипування рослин (Hartmann et al., 2011), пристрій та технологія по автоматизації обробки селекційного матеріалу рису (Duan L. F. et al., 2011), метод для вивчення показників кореневої системи заснований на аналізі зміщення частинок ґрунту (Wao et al., 2018). Найбільше за це десятиріччя працювали з фенотипом коренів, оскільки до цього науковцям морфометрія кореня була практично не доступною. Були проведені та розроблені методики з фенотипування коренів з залученням зображень кореневої системи отриманих за допомогою хвиль різної довжини та дво і тривимірних зображень (Iyer-Pascuzzi et al., 2010; Clark et al., 2011), створені платформи для аналізу цих зображень (Yazdanbakhsh and Fisahn, 2009).

Необроблені дані зображень можна обробляти різними способами, і за останні кілька років використовують два основні підходи: 2D-обробка зображень і алгоритми обробки 3D-сітки. На культурі бавовнику було вперше розроблено запропоновано метод 3D реконструкції (Parroki et al., 2012). Потім запропоновані способи вдосконалено для отримання високопродуктивних параметрів характеристик архітектури рослин (Fang et al., 2016).

В світовій селекції технології фенотипування з залученням програмної обробки інформації використано в двох напрямках: вивчення активності фотосинтезу – накопичення органічної речовини та вивчення посухостійкості. Ці два процеси взаємопов'язані. Питання селекції вирішували ймовірнісним шляхом оцінки користі та ризику кожної комбінації алелів після оцінки зображень та врожайності дослідних ділянок на соняшнику (Tardieu F. et al., 2018).

Великі обсяги інформації породжують проблеми оптимізації і збереження, які розглядаються в наукових дослідженнях з різних точок зору (Kayad et al., 2022).

В сучасному аналізі зображень використовують велике різноманіття алгоритмів, зокрема: алгоритми з аналізу хмари точок (Wang Y. S. et al., 2019),

метод гістограми (Pattanaik and Swarnkar, 2018), локальної дискримінантної проєкції (MLDP) (Zhang S. W. et al., 2020).

Результати з фенотипування знімків та інформацій з баз даних інших факторів та подій використовується для створення моделей, та прогнозів з розвитку подій. Для просторово-часового моделювання розробляють середовища за допомогою мови програмування Python (Stoma et al., 2011; Hennig et al., 2017). Для роботи з мікро об'єктами – клітинами створено інструмент CellProfiler для розпізнання окремих клітин та подальшого машинного навчання (Dordea et al., 2016; Macedo and Cunha, 2016), метод визначення концентрації нуклеїнових кислот у зразку шляхом поділу його на краплі з нерівномірним розподілом за об'ємом (Yen et al., 2019), залучення більшої кількості діапазонів хвиль та зйомку з кількох ракурсів (стереозйомку) (Zhang P. et al., 2014).

Для кількісної оцінки складних ознак пов'язаних з ростом, врожайністю та пристосуванням до стресу і для фенотипування рослин у контрольованих екологічних системах (у ростових камерах або в теплиці) або в польових умовах (Arvidsson et al., 2011; White et al., 2012; Walter et al., 2012; Berger et al., 2010; Yagaliyeva et al., 2025) використовують методи візуалізації. Сучасні методи візуалізації мають високу роздільну здатність і дозволяють візуалізувати багатовимірні та багатопараметричні дані. Мета візуалізації — кількісно виміряти фенотип через взаємодію між світлом і рослинами, наприклад, відбиті фотони, поглинені фотони або фотони, що проходять. Зображення на різних довжинах хвиль використовується для різних аспектів фенотипування рослин. Видиме зображення в основному використовується для вимірювання аспектів архітектури рослин, таких як прогнозована біомаса на основі зображень, площа листя, колір, динаміка росту, сила розсади, морфологія насіння, архітектура коренів, оцінка тяжкості захворювань листя, врожайність, кількість і розподіл плодів. Флуоресцентне зображення використовувалося для виявлення захворювань при генетичній стійкості до захворювань.

У контрольованому середовищі (у ростовій камері або в теплиці) видиме зображення використовувалося в першу чергу для аналізу біомаси пагонів (Golzarian et al., 2011; Arvidsson et al., 2011), ознак врожайності (Lingfeng Duan et al., 2011), ознак волоті (Ikeda et al., 2010), швидкості вбирання та проростання (Dias et al., 2011), морфологія листків (Hoyos-Villegas et al., 2014), сила проростків (Fiorani and Schurr, 2013; Walter et al., 2012), довжина колеоптилі та біомаса в стадії виростання (Richards et al., 2010), морфологія насіння (Joosen et al., 2012; Chern et al., 2007) та структура коре Ця система використовувала кілька кутів огляду (зазвичай два види збоку та вид зверху) для вилучення а математичне співвідношення між цими трьома видимими зображеннями для біомаси пагона або площі листя.

Під час використання видимого зображення витягу з динаміки морфології росту рослин, корневих систем або особливостей поверхні насіння використовується серія стандартних алгоритмів попередньої обробки та сегментації аналізу зображень, таких як алгоритм вододілу, метод сегментації кольорів на основі простору RGB, модель сегментації зображення (Chern et al., 2007).

Отже, сучасне фенотипування поєднує класичні методи з новітніми сенсорними, зображувальними та автоматизованими технологіями — від лабораторного та тепличного до польового й роботизованого (Білка та ін., 2025). Такий підхід дає можливість масштабних, об'єктивних, повторюваних і детальних досліджень фенотипів, що критично важливо для селекції, вивчення адаптації, стійкості до стресів, а також фундаментальних досліджень у рослинній біології (Watt et al., 2020).

Калібрування і валідація вимірювань є критично важливими для забезпечення точності та відтворюваності фенотипічних даних. Калібрування передбачає регулярну перевірку та налаштування сенсорів, камер, ваг, лінійок та інших інструментів, щоб їх показники відповідали відомим стандартам або еталонам. Наприклад, фотометричні сенсори калібрують за допомогою

стандартних спектральних еталонів, а системи 3D-сканування — за допомогою калібрувальних мішеней з відомими геометричними параметрами (Dumschott et al., 2023; Papoutsoglou, 2021).

Валідація вимірювань передбачає перевірку отриманих даних шляхом порівняння з контрольними методами або незалежними вимірюваннями. Це може включати ручну перевірку морфометричних параметрів на невеликій підгрупі рослин, повторні вимірювання в різних умовах, або порівняння результатів різних сенсорних систем (Cook et al., 2020). Валідація дозволяє оцінити точність, повторюваність та чутливість методів фенотипування і визначити діапазон їх надійності.

Фенотипічні вимірювання піддаються різноманітним джерелам варіабельності, які можуть впливати на точність та відтворюваність даних: біологічна варіабельність (відмінності між генотипами, розвиток рослин на різних етапах вегетації, вплив мікроекологічних умов), екологічна варіабельність (зміни температури, вологості, освітлення, складу ґрунту, присутність шкідників або симбіотів), технічні помилки (похибки сенсорів, камер, ваг, неідеальна сегментація зображень, шум у сигналу), методологічні джерела (різна підготовка рослин, нестандартні протоколи, помилки при введенні даних або їх обробці), а також суб'єктивність оцінки, що виникає при ручному фенотипуванні (Li J.J. et al., 2025; Gaggion et al., 2025; Wang R.F. et al., 2025). Врахування цих джерел варіабельності та їх контроль через повторювані вимірювання, стандартизовані протоколи та статистичну обробку дозволяє підвищити якість даних та достовірність результатів (Zhang Y and Zhang, 2018).

Стандартизація опису фенотипів забезпечує їх інтеоперабельність та порівнянність між дослідженнями. Основні підходи включають формальні онтології Crop Ontology (CO), Plant Trait Ontology (TO), Phenotype And Trait Ontology (PATO) та Environment Ontology (ENVO) (Abebe et al., 2023; Clarke et al., 2023), що уніфікують терміни та визначення ознак, а також детальну інформацію про методи вимірювань, умови росту, генотипи, сенсорне

обладнання та параметри обробки даних. Використання структурованих файлів (CSV, JSON, ISA Tab) забезпечує інтеграцію фенотипових, генотипових та екологічних даних (Cadle-Davidson et al., 2019).

Для створення цифрових моделей рослин і інтеграції фенотипових, генотипових та екологічних даних у сільському господарстві застосовуються різні спеціалізовані онтології. Вони забезпечують стандартизоване представлення знань, що дозволяє ефективно обмінюватися даними, інтегрувати різні джерела та проводити семантичний аналіз.

Онтологія Crop Ontology надає стандартизовані терміни для агрономічних характеристик культурних рослин (Crop Ontology <https://cropontology.org/>). Вона включає дані про морфологічні та фізіологічні ознаки рослин; показники продуктивності та якості врожаю; методи оцінки фенотипу. Crop Ontology дозволяє інтегрувати фенотипові дані з різних джерел і підтримує сумісність між базами даних у міжнародних селекційних програмах.

Environment Ontology (ENVO) — онтологія для опису екологічних та середовищних умов, що впливають на ріст і розвиток рослин (The Environment Ontology 2020). ENVO включає типи середовищ (поле, теплиця, лабораторні умови); кліматичні та ґрунтові характеристики; антропогенні впливи. ENVO дозволяє інтегрувати середовищні фактори з фенотиповими даними, що є критично важливим для аналізу впливу генотип  $\times$  середовище. (Buttigieg P.L. et al., 2016).

Plant Trait Ontology (TO) — онтологія для опису ознак рослин, включно з морфологічними, фізіологічними та біохімічними характеристиками (Plant Trait Ontology <https://obofoundry.org/ontology/to.html>). Забезпечує стандартизовану термінологію для селекційних та фенотипових досліджень. Використовується для узгодження даних між різними культурами та експериментами.

Phenotype And Trait Ontology (PATO) — онтологія для формалізації фенотипових ознак та їх значень, що може застосовуватися до будь-яких біологічних організмів. Визначає якісні та кількісні фенотипові атрибути

(Phenotype And Trait Ontology <https://obofoundry.org/ontology/pato.html>).

Підтримує інтеграцію даних між Plant Trait Ontology та іншими ресурсами. Використовується для автоматичного аналізу фенотипів у цифрових моделях.

Дотримання міжнародних принципів FAIR сприяє відтворюваності результатів, підвищує точність прогнозів фенотипу та дозволяє моделювати складні взаємодії «генотип – середовище» (G×E), що є ключовим для сучасної селекції (Cadle-Davidson et al., 2019; Zhang Y and Zhang, 2018). Забезпечення того, що дані можна знайти, передбачає наявність унікальних ідентифікаторів (DOI, URI), стандартизованих назв сутностей та структурованих метаданих, які описують умови експерименту, генетичний матеріал, сенсорні платформи, параметри калібрування обладнання та методи обробки даних (Wilkinson et al., 2016). Доступність даних реалізується через відкриті або контрольовані репозиторії, що підтримують машинно-читані протоколи доступу (наприклад, OAI-PMH, HTTPS API) та визначають правила правомірного використання інформації. Інтероперабельність забезпечується застосуванням стандартизованих форматів (CSV, JSON-LD, ISA-Tab), онтологій та контрольованих словників (CO, TO, PATO, ENVO), які дозволяють інтегрувати фенотипові, генотипові та екологічні дані з різних дослідницьких платформ. Придатність до повторного використання гарантується наявністю докладних протоколів вимірювань, описів сенсорних систем, інформації про похибки та калібрування, ліцензійних умов, а також документації щодо попередньої обробки, фільтрації та трансформації даних.

У контексті моделювання G×E принципи FAIR є критично важливими, оскільки забезпечують можливість поєднання даних різних масштабів — від високочастотних сенсорних потоків до польових звітів і багаторічних баз спостережень. Стандартизовані та добре описані дані дозволяють використовувати сучасні статистичні та машинні моделі, включно з багаторівневими, просторово-часовими та нейромережевими підходами, що потребують великої кількості якісних і структурованих даних. Таким чином,

впровадження FAIR-принципів не лише підсилює якість фенотипічних даних, але й формує основу для побудови національних та міжнародних цифрових інфраструктур селекції, орієнтованих на відтворюваність, масштабованість та довгострокове збереження наукового знання. У галузі біології та аграрних наук моделі використовуються для кількісного опису фізіолого-біохімічних процесів, морфологічних характеристик рослин, а також для аналізу взаємодії з абіотичними та біотичними факторами середовища.

Перехід до цифрових і предиктивних селекційних платформ є визначальним трендом сучасної аграрної науки, оскільки такі системи забезпечують інтеграцію генотипових, фенотипових і екологічних даних у єдиному цифровому середовищі. Ці платформи дозволяють моделювати гібридні комбінації, прогнозувати їхню реакцію у різних умовах вирощування, застосовувати алгоритми машинного навчання для відбору перспективних ліній, а також оптимізувати селекційні схеми. Серед найбільш відомих платформ цього класу — BreedBase, Breeding Management System (BMS), FieldBook, AgroFIMS, а також моделювальні системи DSSAT і APSIM, інструменти комп'ютерного фенотипування PlantCV, ImageBreed, PHIS та системи геномного прогнозування AlphaSimR, BGLR, BGGE.

Для селекції соняшнику ці платформи відкривають нові можливості у стандартизованому збиранні фенотипових даних щодо ключових ознак — урожайності, толерантності до посухи, стійкості до хвороб (зокрема *Phoma macdonaldii*, *Alternaria* spp., *Sclerotinia sclerotiorum*), характеристик кошика, індексу площі листя та профілю жирних кислот олії. Інтеграція онтологічних стандартів через Crop Ontology дозволяє формувати формалізовані фенотипові набори даних, сумісні між локаціями та роками випробувань, що особливо важливо для багаторічних програм створення нових гібридів соняшнику.

Використання систем високопродуктивного фенотипування, таких як PlantCV, ImageBreed, PHIS, забезпечує автоматизоване вимірювання морфометричних, структурних та спектральних параметрів рослин соняшнику.

Поєднання аналізу зображень з БПЛА та мультиспектральною зйомкою дозволяє отримувати великі масиви даних про розвиток рослин, їх просторову неоднорідність, фотосинтетичну активність і приховані стресові реакції. Це особливо важливо для соняшнику як культури з високою екологічною пластичністю.

Моделі розвитку рослин DSSAT і APSIM забезпечують можливість оцінки поведінки гібридів соняшнику в широкому спектрі агроекологічних умов ще до проведення польових випробувань (Awais et al., 2020). Такі моделі дозволяють прогнозувати урожайність, реакцію на посуху, тепловий стрес, різні режими удобрення та строки сівби, що критично важливо для України, де спостерігається зростання кліматичних ризиків у Степовій та Лісостеповій зонах.

Інструменти геномного прогнозування — AlphaSimR, BGLR, BGGE — дозволяють моделювати ефект гетерозису, прогнозувати продуктивність гібридів і аналізувати реакцію генотипів на різні середовища (G×E). Для соняшнику, як класичної гібридної культури, це створює можливість значно прискорити цикл створення ліній і гібридів, проводячи частину оцінювання *in silico* (Bancic et al., 2025).

Перспективним напрямом є розвиток концепції цифрових двійників рослин, які інтегрують фізіологічні, генетичні та екологічні моделі. Для соняшнику такі двійники можуть симулювати розвиток рослин за різних технологій вирощування, густоти стояння, водного забезпечення та удобрення, що дозволяє зменшити потребу у великих польових випробуваннях. (Villalobos, F.J., Ritchie, J.T., 2023).

Розглянемо можливості цифрових і предиктивних платформ для соняшнику

Селекція соняшника, як однієї з основних олійних культур світу, має стратегічне значення для продовольчої безпеки та сталого землеробства, зважаючи на його роль у формуванні сівозмін та адаптивний потенціал у різних агрокліматичних умовах. Посилення вимог до врожайності, адаптивності та

біотичної стійкості сортів і гібридів зумовлює необхідність інтеграції сучасних технологічних підходів у селекційний процес. Цифрові селекційні платформи відкривають нові можливості для вдосконалення селекційних програм соняшнику, забезпечуючи стандартизоване збирання, інтеграцію та аналіз багаторівневих даних про генотипи, фенотипи та умови вирощування. Використання BreedBase, BMS, FieldBook та AgroFIMS дає змогу формувати структуровані фенотипові бази даних для ключових ознак соняшнику — продуктивності, стійкості проти біотичних стресів (*Phoma macdonaldii*, *Alternaria* spp., *Sclerotinia sclerotiorum*), а також показників якості олії (Connor and Sadras, 1992; Velasco et al., 2016). Стандартизація ознак через Crop Ontology забезпечує можливість порівнювати результати між роками, локаціями й партнерами, що є критично важливим для багаторічних програм селекції соняшнику (Menda et al., 2025).

Високопродуктивне фенотипування за допомогою PlantCV, ImageBreed та PHIS створює нову якість у вимірюванні морфометричних характеристик соняшнику, включно з тривалістю фаз органогенезу, параметрами листової поверхні, структурою та діаметром кошика, щільністю та однорідністю насіння. Інтеграція безпілотних літальних апаратів, мультиспектральної та теплової зйомки з алгоритмами обробки зображень дозволяє оцінювати однорідність розвитку рослин і виявляти ранні стресові реакції (Pierz et al., 2023). Це особливо важливо для соняшнику, який демонструє високу чутливість до посухи та температурних екстремумів. Використання таких систем дає змогу скоротити витрати на польові оцінки та підвищити точність добору в ранніх поколіннях.

## 1.2 Сучасний стан аграрної галузі в Україні

Системно-аналітичний огляд сучасного стану аграрної галузі України та напряму селекційних досліджень демонструє, що галузь переживає глибоку трансформацію під впливом цифровізації та поширення даноцентричних

підходів до управління складними біологічними системами. Структурні зміни, спричинені кліматичними зрушеннями, економічними коливаннями та безпековими загрозами, істотно впливають на продуктивність і стабільність агросектору, формуючи нові вимоги до інфраструктури збору та обробки даних (Григорук and Співак, 2026; Giuliani and Baron, 2025).

Юридичну та організаційну основу діяльності у сфері насінництва в Україні складають закони України, зокрема:

- Закон України «Про насіння і садивний матеріал» — регламентує виробництво, сертифікацію, обіг і контроль якості насіння та садивного матеріалу (Про насіння і садивний матеріал : Закон України 2002).
- Закони про охорону прав на сорти рослин, карантин рослин, біобезпеку (ГМО) — забезпечують правові гарантії, контроль за захистом сортів і безпекою використання нових технологій. (Про державну систему біобезпеки при створенні 2007).

У 2025 році затверджено нові Методичні вимоги у сфері насінництва, які регламентують порядок збереження сортових та посівних якостей насіння кормових культур і спрямовані на підтримку високої якості насіння. (Про затвердження Методичних вимог у сфері насінництва 2025).

У межах системи насінництва, що включає різні категорії насіння (добазове, базове, сертифіковане), стандарти та технології виробництва регламентуються відповідними ДСТУ. Основні стандарти, що застосовуються:

1. ДСТУ 2240-93 — «Насіння сільськогосподарських культур. Сортові та посівні якості. Технічні умови» (ДСТУ 2240-93, 1994);
2. ДСТУ 4138-2002 — «Насіння сільськогосподарських культур. Методи визначення якості» (ДСТУ 4138-2002, 2002);
3. ДСТУ 7160:2020 — «Насіння овочевих, баштанних, кормових і пряно-ароматичних культур. Сортові та посівні якості. Технічні умови» (ДСТУ 7160:2020, 2020);

4. ДСТУ 6006:2008 — «Насіння овочевих, баштанних культур та кормових коренеплодів. Пакування, маркування, транспортування та зберігання. Технічні умови» (ДСТУ 6006:2008, 2008);

5. ДСТУ 2949-94 — «Насіння сільськогосподарських культур. Терміни та визначення» (ДСТУ 2949-94, 1994).

Ці стандарти забезпечують державну сертифікацію та контроль якості насіння, визначають порядок виробництва та обігу насіння різних категорій, а також встановлюють методи оцінки його посівних і сортових якостей. У сучасному даноцентричному агровиробництві (де використовуються цифрові двійники, ШІ для прогнозування врожайності та супутникові дані) ці закони та ДСТУ виконують роль стандартів вхідних даних. Без чіткого дотримання ДСТУ неможливо побудувати адекватну математичну модель розвитку рослини, оскільки початкові біологічні параметри насіння будуть хаотичними.

Кліматичні зміни, зокрема зміщення ізотермічних зон, зростання частоти посух і температурних стресів, призводять до зміни структури посівних площ, переорієнтації виробників на посухостійкі культури та необхідності широкого застосування польових сенсорних систем і супутникового моніторингу (Zatserkovnyı et al., 2025). Безпекові фактори, зокрема масштабне мінування та втрата частини агропотужностей, актуалізували розвиток автономних цифрових платформ для дистанційного контролю стану посівів та інфраструктури агровиробництва (Wagner et al., 2025).

Сучасна парадигма розвитку аграрної науки в Україні характеризується поступовим переходом від традиційних методів селекції до інтегрованих цифрових екосистем. Цей процес визначається як трансформація від поодиноких експериментів до формування локальних цифрових хабів у структурі провідних науково-дослідних установ, хоча імплементація цифрових платформ усе ще залишається фрагментарною, через значну капіталоємність інфраструктури та дефіцит спеціалізованих кадрів, зокрема біоінформатиків (Колісніченко, 2025). У цьому контексті виникає потреба у модернізації систем збору, зберігання та

інтеграції даних. Сучасні селекційні програми дедалі більше залежать від цифрових репозитаріїв, баз генетичних ресурсів і платформ фенотипування, які забезпечують доступ до якісної інформації щодо сортових характеристик. В Україні такі системи формуються на основі діяльності лише на базі колекцій генетичних ресурсів зокрема Інституту олійних культур Національної академії аграрних наук України (Алієв Е. Б. Ведмедева К.В., 2024). Хоча бази генетичних ресурсів наявні в багатьох установах, які об'єднані у Національний центр генетичних ресурсів рослин України існуючий при Інституті рослинництва ім. В. Я. Юр'єва НААН (Рябчун et al., 2014).

Усі селекційні установи активно вивчають і починають впроваджувати елементи цифрової селекції, зокрема бази даних польових випробувань, електронні паспорти сортів, геопросторові інформаційні системи. Основним драйвером цифровізації вітчизняної селекції виступає інтеграція у глобальні дослідницькі мережі, де ключову роль відіграє співпраця з міжнародними центрами групи CGIAR, такими як CIMMYT та ICARDA. Зокрема, пріоритетним напрямом є впровадження системи Breeding Management System (BMS) на базі Інституту рослинництва ім. В. Я. Юр'єва та Селекційно-генетичного інституту (СГІ), що дає змогу стандартизувати протоколи фенотипування, уніфікувати селекційні цикли та забезпечити повну прозорість генеалогічних даних (Strengthening Wheat Breeding in Ukraine..., 2022). Паралельно з цим, у межах проєктів FAO та USAID, здійснюється цифровізація державних реєстрів та створення «цифрових двійників» (Digital Twins) баз даних генетичних ресурсів, що є критично важливим для забезпечення продовольчої безпеки в умовах воєнного стану (Digitalization of the Seed Sector in Ukraine..., 2023.; Ukraine Agriculture Resilience Initiative, 2024). Додатковим інструментом модернізації виступають програми Horizon Europe, завдяки яким українські інститути залучаються до консорціумів, що забезпечують технічне переоснащення системами високоточного польового фенотипування з використанням БПЛА та мультиспектральних сенсорів.

Практична реалізація таких цифрових рішень чітко відображає спеціалізацію наукових центрів: якщо Інститут рослинництва ім. В. Я. Юр'єва функціонує як центральний вузол цифровізації Національного генбанку через інтеграцію реляційних баз даних, то СГІ–НЦНС фокусується на математичному моделюванні взаємодії «генотип — середовище» за допомогою моделей типу DSSAT для прогнозування адаптивності культур до аридних умов Степу. Своєю чергою, Миронівський інститут пшениці імені В. М. Ремесла активно впроваджує цифрові мобільні системи FieldBook, що дозволяє стандартизувати польові журнали та мінімізувати суб'єктивні помилки при первинному зборі даних (Pichura V. et al., 2023; Pichura V et al., 2025). Попри вказані успіхи, формування цілісної національної Big Data платформи стримується проблемою ізольованості даних (Data Silos), коли кожна установа накопичує інформацію у власних закритих форматах, що унеможлиблює проведення системного аналізу (Pichura V et al., 2025). Відсутність єдиних стандартів метаданих, зокрема недотримання Breeding API (BrAPI), значно ускладнює інтеграцію українських масивів даних у глобальні наукові мережі, а постійні безпекові ризики в прифронтових зонах змушують науковців екстрено переносити критичну інформацію в хмарні середовища AWS та Google Cloud. Таким чином, перехід до концепції «Селекції 4.0» потребує не лише технологічного оновлення, а й фундаментальної зміни підходів до математичного моделювання процесів росту та розвитку рослин (Digitalization of the Seed Sector in Ukraine..., 2023; Ukraine Agriculture Resilience Initiative, 2024).

### **1.3 SWOT-аналіз впровадження цифрових та онтологічних технологій у селекції соняшнику в умовах Agriculture 4.0**

Системно-аналітичний огляд дозволяє виявити як потенціал сучасних цифрових інструментів, так і низку обмежень, пов'язаних із фрагментованістю даних, відсутністю єдиних стандартів, недостатнім рівнем фінансування

цифрової інфраструктури та браком взаємодії між науковими центрами. Актуальним є формування національної екосистеми digital breeding, яка охоплюватиме фенотипування, генотипування, польову аналітику, агрокліматичні дані та онтологічно структуровані інформаційні ресурси для забезпечення інтеграції даних і подальшого використання в інтелектуальних системах аналізу селекційної інформації (Колісніченко, 2025). Важливим теоретико-методологічним підґрунтям для розгортання таких систем є дослідження процесів цифровізації як глобального чинника структурної трансформації та забезпечення економічної стійкості суб'єктів господарювання (Васильєва & Душин, 2026), а також врахування соціо-економічного виміру та людського потенціалу в умовах формування єдиного цифрового простору (Vasyl'yeva et al., 2023). При цьому впровадження новітніх цифрових рішень в аграрному секторі має розглядатися в контексті сталого розвитку та оптимізації факторної продуктивності, де сучасний інструментарій комп'ютерного моделювання виступає драйвером підвищення ефективності (Vasyl'yeva et al., 2025). Крім того, побудова наскрізних галузевих платформ вимагає узгодження технічних параметрів із маркетинговим та ресурсним забезпеченням агропродовольчих підприємств, зокрема в розрізі трансформації їхніх внутрішніх бізнес-процесів (Trusova et al., 2022).

Важливою складовою такої екосистеми є онтологічне представлення знань, яке забезпечує семантичну інтеграцію різнорідних даних, формування графів знань та підтримку автоматизованого конструювання моделей прийняття рішень. Застосування онтологічних підходів дозволяє формалізувати предметну область селекції, забезпечити інтеоперабельність даних та створити основу для масштабованих інтелектуальних систем підтримки прийняття рішень (Теплий & Досин, 2026).

Аналіз відкритої наукової літератури свідчить, що в Україні вже існує низка прикладних досліджень, які застосовують елементи прогнозних та інтелектуальних інформаційних систем, закладаючи методологічний фундамент

для цифрової селекції. Особливе місце тут посідають математичні методи інтеграції інструментів машинного й глибинного навчання для предиктивного моделювання та прогнозування врожайності соняшнику, що дозволяють обробляти просторово-часові масиви даних моніторингу культури (Hnatiienko, 2024).

Подальший розвиток цього напрямку пов'язаний із переходом до мультифакторного аналізу та залученням апарату штучних нейронних мереж, які дають змогу синтезувати складні нелінійні залежності між різномірними агрокліматичними показниками та продуктивністю сільськогосподарських культур (Гнатієнко та ін., 2024). На сучасному етапі такі підходи еволюціонували до систем точного, просторово-орієнтованого прогнозування, які на основі геопросторового аналізу та оптимізованих алгоритмів машинного навчання забезпечують локальну прив'язку прогнозних моделей до конкретних ділянок вирощування соняшнику (Hnatiienko et al., 2025).

Паралельно з цим, у публікації Кириченко та ін. (Кириченко et al., 2020) проведено прогнозування динаміки вегетації гібридів соняшнику в різних ґрунтово-кліматичних умовах українського Степу, що наближається до моделей взаємодії генотипу та середовища «генотип-середовище». Дослідження Чуйко зі співавторами (Чуйко et al., 2025) описує агробіологічну оцінку гібридів соняшнику у Лісостепових і Степових зонах, формуючи докладний фенотиповий масив, який може бути основою для подальшої оцифровки в рамках селекційних платформ.

Водночас у відкритій літературі не виявлено документованих випадків повного використання міжнародних відкритих селекційних платформ (*BreedBase*, *BMS*, *APSIM*, *DSSAT*) для соняшнику саме в українських дослідженнях. Часто подібні розробки залишаються внутрішніми комерційними проектами приватних компаній (наприклад, *Kernel*, *IMC*, *HarvEast*, які активно використовують супутникові індекси NDVI для моніторингу стресових факторів) і не публікуються у наукових журналах.

В умовах України, де соняшник є стратегічною олійною культурою, цифрові платформи надають критично важливі інструменти для адаптивної та кліматично-орієнтованої селекції. Для степових та лісостепових зон, що характеризуються частими посухами й зростанням частоти теплових стресів, моделювання на основі APSIM Initiative (n.d.) та DSSAT Foundation (n.d.) дає змогу прогнозувати реакцію гібридів під різними сценаріями кліматичних змін. Використання систем електронного польового фенотипування FieldBook (PhenoApps, n.d.) дозволяє стандартизувати дані між станціями та роками, а інтеграція BreedBase (Morales et al., 2022) — формувати спільні бази селекційних даних. Це створює підґрунтя для прискорення розробки посухостійких, високоолійних та пластичних гібридів соняшнику.

Створення онтологічної моделі фенотипів соняшнику повинно розглядатися не лише як засіб структуризації даних, а як інформаційна основа для подальшого розвитку інтелектуальних систем аналізу та підтримки прийняття рішень. Онтології забезпечують формалізацію знань предметної області, підтримують механізми семантичного пошуку, виявлення прихованих взаємозв'язків між об'єктами та побудову рекомендаційних сервісів на основі накопиченого досвіду й експертних знань (Oborska et al., 2023). Це створює передумови для формування цифрової платформи селекції соняшнику, орієнтованої на аналіз взаємодії «генотип – середовище» та підтримку прийняття селекційних рішень.

Розглядаючи регіональні загрози для України варто відзначити, що у Південному Степу спостерігається значний водний дефіцит, що підсилює прояви взаємодії «генотип – середовище» ( $G \times E$ ) для ознак посухостійкості, що підтверджується даними Інституту олійних культур НААН. У Лісостеповій зоні відзначаються різкі коливання ураженості фомопсисом, що ускладнює точне моделювання без застосування стандартизованих протоколів фенотипування. Щодо практичного впровадження, то українські аграрні компанії, зокрема Kernel, ІМС та HarvEast, активно використовують супутникові індекси NDVI для

моніторингу стресових факторів та оптимізації розташування дослідних ділянок. Дослідження, проведені на платформі «Миронівський інститут пшениці», демонструють ефективність використання змішаних моделей BLUP для оцінки стабільності генотипів у багаторічних випробуваннях (Demudov et al., 2024).

Попри наявність досліджень із застосуванням елементів цифрового моніторингу та прогнозування, реалізація українських проєктів у глобальних open-source платформах стикається з низкою перешкод.

1. Нестача стандартизованих даних у міжнародних форматах. Багато польових досліджень виконано без інтеграції онтологій та структурованих фенотипових баз, що ускладнює імпорт українських даних до глобальних платформ, таких як BreedBase чи BMS.

2. Відсутність повного циклу цифрової селекції. У багатьох українських проєктах використовується лише частина методів (супутниковий моніторинг, фенотипування, статистичне моделювання), без інтеграції геномної інформації та прогнозних моделей взаємодії «генотип-середовище»  $G \times E$ .

3. Обмежені технічні та ресурсні можливості. Для повноцінного використання міжнародних платформ потрібні обчислювальні потужності, стабільний збір польових і геномних даних та висока кваліфікація персоналу. Часто такі умови не повністю реалізовані в українських установах.

4. Відмінності агроєкологічних умов. Дані міжнародних платформ і моделей орієнтовані на інші кліматичні та ґрунтові умови; пряме застосування без адаптації до степових і лісостепових зон України може бути некоректним.

5. Ліцензійні та організаційні бар'єри. Деякі платформи та бази даних мають обмеження щодо внесення зовнішніх даних або публікації результатів, що перешкоджає інтеграції українських проєктів у світову інфраструктуру.

В Україні вже здійснюється часткове застосування цифрових підходів у селекції соняшнику. Зокрема: прогнозування динаміки вегетації та врожайності гібридів на основі супутникових і кліматичних даних; польові агробіологічні оцінки та фенотипування морфометричних і продуктивних ознак; оцінка

стійкості до фомопсису та інших хвороб. Однак повної інтеграції в глобальні open-source платформи наразі не зафіксовано, що підкреслює необхідність розвитку власної національної цифрової інфраструктури.

Для комплексної оцінки передумов, можливостей та обмежень цифрової трансформації селекції соняшнику в таблиці 1.3.1 наведено SWOT-аналіз створення онтологічно керованого цифрового середовища даних в умовах Agriculture 4.0.

Таким чином, проведений аналіз та SWOT-аналіз дозволили встановити, що основною проблемою цифрової трансформації селекції соняшнику в Україні є не відсутність даних, а недостатній рівень їх структурованості, стандартизації та семантичної сумісності. Наявні фенотипові, генотипові, агрокліматичні та виробничі дані залишаються фрагментованими, що обмежує можливість їх ефективного використання у цифрових моделях та інтелектуальних системах аналізу. Вирішення цієї проблеми потребує розроблення онтологічної моделі соняшнику та цифрової технології формування семантично сумісних наборів даних, які забезпечать основу для створення бази знань цифрової селекції та функціонування інтелектуальної системи підтримки прийняття рішень щодо аналізу, інтеграції та використання селекційної інформації.

Таблиця 1.3.1. - SWOT-аналіз передумов створення онтологічно керованого цифрового середовища даних для селекції соняшнику в умовах Agriculture 4.0

<b>Компонент та SWOT</b>	<b>Ключові положення</b>	<b>Наукові уточнення та приклади для України</b>
<b>S — Strengths (Сильні сторони)</b>	Наявність значних обсягів накопичених селекційних та агротехнологічних даних	Багаторічні результати польових досліджень, фенотипування та випробування гібридів створюють інформаційну основу для формування цифрових моделей соняшнику
	Наявність різномірних джерел цифрової інформації (фенотипові, генотипові,	Забезпечує можливість комплексного опису об'єкта дослідження та аналізу взаємодії «генотип–середовище» (G×E)

Компонент та SWOT	Ключові положення	Наукові уточнення та приклади для України
	агрокліматичні, ґрунтові та технологічні дані)	
	Розвиток технологій Agriculture 4.0	Використання дистанційного зондування, IoT, комп'ютерного зору та автоматизованого моніторингу забезпечує отримання нових цифрових даних
	Наявність сучасних методів онтологічного моделювання та інженерії знань	Створює можливість формалізації понять предметної області, побудови семантичних зв'язків та формування бази знань для інтелектуальних систем підтримки прийняття рішень.
<b>W — Weaknesses (Слабкі сторони)</b>	Фрагментованість та неоднорідність існуючих цифрових даних	Дані зберігаються у різних форматах, структурах та інформаційних системах, що ускладнює їх інтеграцію
	Відсутність формалізованої онтологічної моделі вирощування соняшнику	Обмежує можливість семантичного узгодження даних та їх машинної інтерпретації
	Недостатній рівень стандартизації опису фенотипових ознак	Ускладнює формування сумісних datasets для подальшого використання у моделях машинного навчання
	Переважаючі локальні підходи до накопичення селекційної інформації	Знижує можливості повторного використання даних та інтеграції у цифрові платформи
<b>O — Opportunities (Можливості)</b>	Розроблення цифрової та онтологічної моделей соняшнику	Поєднання структурованих фенотипових, генотипових та середовищних даних створює основу для використання статистичних і машинних моделей оцінювання взаємодії «генотип–середовище» (G×E).
	Створення технології формування семантично сумісних наборів даних	Забезпечує автоматизоване формування структурованих datasets на основі онтологічного опису, семантичне узгодження різномірних джерел інформації та їх використання в інтелектуальних системах підтримки прийняття рішень.

Компонент та SWOT	Ключові положення	Наукові уточнення та приклади для України
	Формування бази знань цифрової селекції соняшнику	Дані фенотипування дозволяють формувати лінії з підвищеною стійкістю до посухи, теплових хвиль та нових рас вовчка, що поширюються на Півдні України.
	Інтеграція з міжнародними цифровими платформами селекції	Забезпечує можливість використання стандартизованих підходів та обміну даними
	Застосування методів штучного інтелекту та машинного навчання	Дозволяє використовувати накопичені знання для прогнозування продуктивності та адаптивності генотипів
<b>T — Threats (Загрози)</b>	Експоненційне накопичення цифрових аграрних даних без їх семантичної структуризації	Може призвести до накопичення великих, але малоприматних для аналізу інформаційних масивів
	Відсутність єдиних стандартів представлення селекційних даних	Ускладнює інтеграцію українських даних у міжнародні цифрові екосистеми
	Кліматична нестабільність та зміна агроекологічних умов	Підвищує складність формування універсальних моделей опису взаємодії «генотип–середовище»
	Обмеженість ресурсів для цифрової трансформації селекційних установ	Уповільнює впровадження сучасних технологій збору, зберігання та аналізу даних
	Втрата або недоступність частини експериментальних даних	Зменшує повноту цифрових моделей та обмежує можливості формування репрезентативних баз знань

#### 1.4 Висновки до розділу 1

У першому розділі виконано комплексний системний аналіз сучасних процесів цифрової трансформації аграрної галузі. За результатами проведеного теоретико-методологічного огляду сформульовано такі висновки:

1. Обґрунтовано необхідність трансформації процесів прийняття рішень у сфері вирощування та селекції соняшнику (*Helianthus*) на основі переходу до даноцентричних (*data-driven*) методів. Установлено, що сучасний

аграрний сектор функціонує в умовах високої стохастичності, невизначеності та ризиків біотичного стресу. Показано, що традиційні емпіричні підходи до управління не дозволяють повною мірою врахувати динамічні параметри цифрового фенотипування рослин, що зумовлює потребу в розробці інтелектуальних систем підтримки прийняття рішень (ІСППР).

2. Визначено нормативно-технологічні обмеження, які можуть бути використані як стандартизовані параметри вхідних даних для системного моделювання. На основі аналізу чинного законодавства України («Про насіння і садивний матеріал», «Про охорону прав на сорти рослин») та нормативних вимог державних стандартів (ДСТУ 2240-93, ДСТУ 4138-2002, ДСТУ 7160:2020) виокремлено жорсткі критерії сортових і посівних якостей насіння. Ці показники визначено як базові обмеження для побудови математичних моделей оптимізації процесів вирощування та селекції.

3. Встановлено обмеження існуючих моделей знань та інформаційних систем у контексті їх адаптації до задач цифрової селекції соняшнику. Встановлено, що існуючі глобальні онтологічні системи не забезпечують необхідного рівня деталізації морфометричних і колірних параметрів рослин, які отримують сучасними методами комп'ютерного зору. Крім того, вони не відображають специфічні причинно-наслідкові зв'язки та управлінські рішення, що притаманні вітчизняним системам землеробства, зокрема для специфічних агроекологічних умов Південної та Східної України. Це обґрунтовує необхідність розробки адаптованої онтологічної моделі предметної області.

4. Визначено та науково обґрунтовано інструментальний каркас дослідження. Обґрунтовано застосування методів онтологічного моделювання. Встановлено, що обробка таких семантичних мереж та проведення імітаційного моделювання для великих національних баз даних потребує залучення високопродуктивних обчислювальних ресурсів.

5. Виявлено ключову науково-практичну суперечність предметної області, яка полягає у суперечності між стрімким зростанням обсягів цифрових

агротехнологічних даних та недостатнім рівнем їх структурованості, стандартизації і семантичної сумісності для використання в інтелектуальних системах підтримки прийняття рішень. Відповідно, основними завданнями подальших досліджень є:

- удосконалення цифрової та онтологічної моделей соняшнику, що інтегрують характеристики генотипу, фенотипу, селекційного процесу та агротехнологій;
- розроблення методу цифрового фенотипування на основі технологій комп'ютерного зору для отримання стандартизованих фенотипових характеристик;
- розроблення цифрової технології формування семантично сумісних наборів даних на основі онтологічного представлення знань про соняшник для наповнення бази знань та інформаційного забезпечення інтелектуальної системи підтримки прийняття рішень у цифровій селекції.

## 2 КОНЦЕПТУАЛЬНІ ЗАСАДИ ТА ІНСТРУМЕНТАРІЙ ПОБУДОВИ ЦИФРОВИХ МОДЕЛЕЙ РОСЛИНИ СОНЯШНИКУ

У другому розділі представлено теоретико-методологічне обґрунтування та практичний інструментарій розробки цифрової моделі рослини соняшника на основі методів формалізації знань та алгоритмів цифрового фенотипування. Розглянуто комплексний підхід до моделювання взаємодії «генотип — середовище», що охоплює інтеграцію морфологічних, біохімічних та агрофізичних характеристик агроecosистеми, з ефектом деталізації динамічного впливу факторів зовнішнього середовища, просторово-географічного детермінізму та агротехнологій на формування цільових ознак продуктивності культури. Основним інструментом формалізації знань та побудови логічного каркаса цифрової моделі обрано предметну онтологію, що забезпечує семантичний стандарт термінології, дозволяє побудувати структуровану систему знань (TBox), здатну до автоматизованого логічного виведення (Reasoning) на рівні екземплярів (ABox) і виявлення прихованих селекційних та фітосанітарних закономірностей.

### 2.1 Концепція цифрової моделі рослини та платформені рішення

Математичне моделювання забезпечує формалізоване відображення об'єктів, процесів або явищ із метою їх аналізу, прогнозування та оптимізації. Інструментальною реалізацією для моделювання в межах цього дослідження визначено цифрову модель рослини, що дозволяє об'єднати у єдиний цифровий простір знань дані про складні процеси взаємодії в системі «генотип — середовище» (рис.2.1.1). Цифрова модель рослини (англ. digital plant model, virtual plant) — це інтегрована комп'ютерна репрезентація рослинного організму,

яка поєднує морфологічні, фізіологічні, біохімічні та генетичні параметри рослини з даними про умови навколишнього середовища.

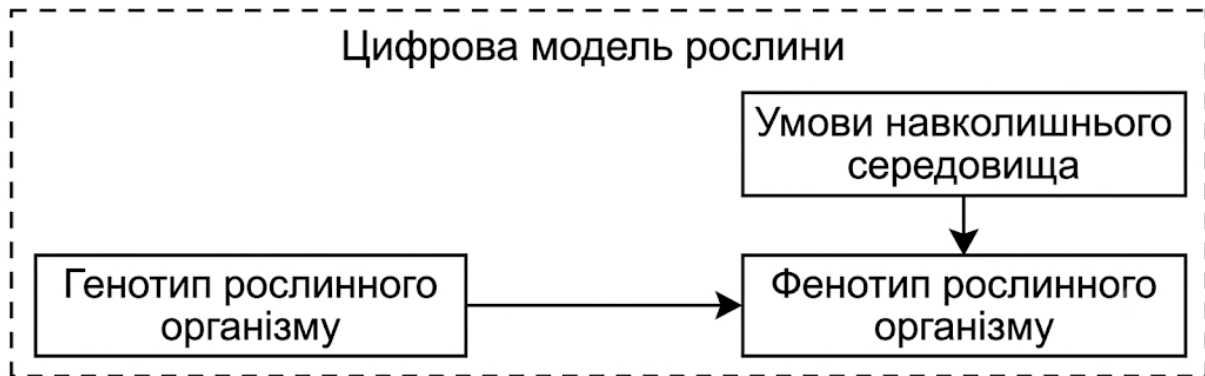


Рисунок 2.1.1 – Концепція цифрової моделі рослини, яка відображає взаємодію в системі «генотип — середовище» (розроблено автором)

Така модель ґрунтується на багаторівневій структурі даних і алгоритмах моделювання процесів росту, розвитку, адаптивних реакцій та загальної продуктивності рослини в межах різних сценаріїв середовища. Цифрова модель виступає основою для симуляційних досліджень *in silico*, оскільки дозволяє відтворювати динамічну поведінку біологічної системи віртуально, істотно зменшуючи потребу в проведенні тривалих і ресурсомістких польових експериментів. У сучасних селекційних дослідженнях цифрові моделі рослини виконують низку критично важливих функцій, що безпосередньо впливають на ефективність і точність процесу створення нових сортів та гібридів. З науково-методологічної точки зору цифрова модель об'єднує чотири фундаментальні кластери даних та підходів.

Генотипічні дані: локуси кількісних ознак (QTL), SNP-профілі, молекулярно-генетичні маркери, гени-ефектори.

Фенотипічні дані високої роздільної здатності: морфологічні ознаки органів, фізіологічні криві, спектральні індекси, масиви даних сенсорних та роботизованих платформ.

Екологічні та агротехнічні параметри: фізико-хімічні характеристики ґрунту, температурні режими, вологість, параметри мікроклімату та технологічні операції.

Механістичні та машинні моделі: біофізичні моделі росту, моделі донорно-акцепторних зв'язків (джерело–стік), моделі водного балансу агроєкосистеми, а також ML-алгоритми прогнозування цільових ознак (Bin Nasir and Li, 2024; Saritha and Thangaraja, 2023, Integrated Breeding Platform. 2019).

Цифрові моделі рослини дозволяють проводити високоточне прогнозування продуктивності та стабільності гібридів. Комп'ютерна симуляція росту та розвитку рослини за множиною змінних сценаріїв середовища забезпечує можливість оцінювання потенційної врожайності, стабільності та адаптивності генотипів без необхідності проведення тривалих польових випробувань. Це істотно прискорює селекційний цикл та дає змогу оперативно порівнювати перспективні лінії та мінімізує фінансові й часові витрати на експериментальні дослідження. Наприклад, у сучасних дослідженнях феноміки рослин особлива увага приділяється переходу від статичних описів до динамічного просторово-часового моделювання ознак та інтеграції зворотного зв'язку з навколишнім середовищем. У цьому контексті актуальними є детерміністичні, ймовірнісні та генеративні підходи моделювання, які розглядаються у фундаментальних працях останніх років, зокрема у роботах Мохамеда Деббага (Debbagh M. et al., 2025). Автори підкреслюють, що класичні експериментальні детерміністичні підходи мають обмеження за умов високої мінливості клімату, через що виникає гостра потреба у впровадженні умовних генеративних моделей та динамічних фреймворків, здатних враховувати стохастичну невизначеність середовища, що розвивається. Сучасний етап розвитку цифрового фенотипування базується на інтеграції методів регресії та моделей представлення на основі глибоких нейронних мереж для розв'язання задач прогнозного розпізнавання образів у 2D та 3D просторах. Застосування функціонально-структурних моделей рослин (FSPM) у синергії з генеративно-

змагальними чи дифузійними нейромережами дозволяє не просто фіксувати поточний габітус, а симулювати архітектуру органів (листоків, стебла, суцвіть) у динаміці вегетації. Перспективним вектором розвитку таких систем, що безпосередньо корелює з завданнями цього дослідження, є конвергенція знань, специфічних для предметної області (онтологічних моделей і правил), із методами, керованими великими масивами даних (Data-Driven). Такий гібридний підхід відкриває можливості для створення масштабованих цифрових репозиторіїв та успішного впровадження інтелектуальних моделей росту рослин у реальні застосування систем підтримки прийняття рішень (СППР) для точного землеробства та превізуалізації результатів селекції. Цифрові моделі ефективно застосовуються для моделювання взаємодії «генотип — середовище» ( $G \times E$ ). Вони створюють математичний інструментарій для аналізу того, як конкретні алелі, локуси кількісних ознак (QTL) або їхні комбінації визначають фенотипічну реакцію рослини в різних екологічних умовах, включаючи критичні абіотичні стреси — посуху, екстремально високі температури чи дефіцит азотного живлення (Smet et al., 2023). Такий підхід дозволяє ідентифікувати генотипи з підвищеною екологічною пластичністю та формує підґрунтя для розроблення кліматично орієнтованих селекційних стратегій. Імітаційне моделювання просторової архітектоники рослини, параметрів фотосинтетичної активності та архітектури кореневої системи забезпечує можливість випробування різних агротехнічних рішень у віртуальному середовищі. Це дозволяє оптимізувати густоту стояння, режими поливу, структуру внесення мінеральних добрив та інші технологічні фактори, що безпосередньо впливають на кінцевий фенотип та його стабільність (Grobler and Celano, 2025). Цифрова модель рослини виконує функцію інтеграційної платформи між високошвидкісними фенотипічними системами та алгоритмами геномного прогнозування. Поєднання симуляційних моделей з методами GBLUP, Random Forest, градієнтним бустингом та глибинними нейромережами забезпечує підвищення точності оцінювання селекційної цінності, дозволяє формувати

багатопараметричні предиктивні моделі та покращує загальну ефективність прийняття селекційних рішень (Sinha et al., 2025). Завдяки можливості проведення масштабних *in silico* експериментів (комп'ютерного моделювання для аналізу, прогнозування та дизайну біоструктур), цифрова модель суттєво прискорює ключові етапи селекційного процесу. Це дозволяє здійснювати одночасне оцінювання тисяч генотипів у великій кількості сценаріїв розвитку, мінімізуючи час на первинний відбір, оптимізуючи вибір батьківських пар і забезпечуючи оперативне формування поколінь із прогнозовано високим потенціалом продуктивності та стійкості (Louarn and Song, 2020). У контексті розвитку інтегрованих цифрових платформ для аграрної науки цифрова модель рослини стає центральним елементом концепції *predictive breeding* — селекції, що спирається на інтелектуальний аналіз даних та прогнозування. Вона формує основу для розробки цифрових двійників культур (*crop digital twins*). Цифровий двійник рослини (*digital twin of a plant*) — це комп'ютерно реалізована динамічна модель, яка безперервно відображає морфологічний стан реального біологічного об'єкта, підтримує двосторонній зв'язок з фізичним середовищем і оновлюється на основі даних сенсорного моніторингу та зовнішніх впливів (Ziamtsov I., Navlakha S. 2020; Andres F. et.al.2021; Mildenhall B. Et al. 2021; Господаренко, Г. М. та ін. 2021; Wang, J., et al.2022; Зозуля О. Л. Та ін, 2023).

За рахунок використання оптичних, спектральних, глибинних і екологічних сенсорів модель акумулює дані про ключові ознаки росту (геометрія, орієнтація листя, щільність надземної маси, стан генеративних органів) та виявляє аномальні відхилення у розвитку (Атлас морфологічних ознак...2018; Andres F. et.al.2021). У практиці фенотипування для просторового моделювання широко застосовуються *mesh*-моделі, які відтворюють зовнішню поверхню об'єкта на основі полігональної трикутної сітки, та *voxel*-моделі, що формують об'ємне представлення на основі регулярної тривимірної дискретизації простору (Intelligent Monitoring... 2022).

Цифровий двійник виконує функцію відтвореної цифрової репрезентації, що може бути збережена у форматі багатосарової бази даних, масштабована, використана в симуляційних експериментах або передана в інші дослідницькі чи виробничі середовища (Aliyev E.B. 2019; Wang, J., et al. 2022,). Зокрема, метод NeRF (Neural Radiance Fields) ефективно використовується для реконструкції складних геометрій рослин та дрібних структур у статичному лабораторному середовищі, що є критично корисним для селекційної роботи за умов належної фіксації зразків (Atkinson, N.J., Lilley, C.J., Urwin, P.E. 2013; Zhang, Z., et al. 2020, Семенова, І. Г. 2017). Сучасні методи 3D-реконструкції на основі NeRF, згорткових нейронних мереж (CNN) та 3D Гаусового розсіювання (3D Gaussian Splatting) детально розглянуто в фундаментальних роботах (Gao Y., Qi X. 2024, Li Q. et al. 2025, AgEval: A 2023). Важливим напрямом є розробка моделей, що враховують розвиток рослини в часі, тобто темпоральних цифрових двійників. У роботі (GrowSplat Tsafaris S. A. 2016) запропоновано підхід на основі 3D Gaussian Splatting, що дозволяє створювати послідовності 3D-моделей для кожного моменту часу та виконувати їх просторово-часове вирівнювання (alignment). Це забезпечує точне відстеження мікроморфологічних змін і динамічний аналіз темпів росту органів. У зв'язку з цим постає гостра необхідність у створенні уніфікованих інформаційно-аналітичних систем, які б інтегрували високопродуктивні фенотипічні дані з параметрами зовнішнього середовища, зокрема, ґрунтовими характеристиками, кліматичними умовами та агротехнічними прийомами. Формування такої єдиної семантичної бази знань є критично важливим для підвищення точності селекційної оцінки, забезпечення репродуктивності результатів між різними дослідними платформами, сезонами та географічними зонами, а також для об'єктивного аналізу складних кількісних ознак у процесі генетичного поліпшення соняшника. Концептуальним основам розробки предметної онтології вирощування соняшника присвячено фундаментальну працю (Martin-Brualla R. 2021).

У процесі генетичного поліпшення сільськогосподарських культур, зокрема соняшника (*Helianthus annuus L.*), перед селекціонером постає комплекс завдань, що визначають загальну ефективність селекційного циклу. Першочерговим є розроблення або вдосконалення аналітичних методик і протоколів кількісного оцінювання ознак, які мають суттєве господарське значення. Наступним етапом є застосування цих методик для репрезентативного фенотипування селекційного генофонду з метою встановлення спадкової стабільності та генетичної детермінованості відповідних ознак. Заключним кроком у селекційному процесі є формування нових генотипічних конструкцій — ліній, сортів або гібридів, які характеризуються стабільним і високим рівнем експресії цільових фенотипових ознак.

На сучасному етапі розвитку аграрної науки провідну роль у цьому процесі відіграє цифрове фенотипування, що виступає технологією високоточної кількісної оцінки морфологічних, фізіологічних та біохімічних ознак рослин. Традиційно оцінка складних ознак, таких як стійкість до абіотичних (посуха, висока температура) та біотичних (збудники хвороб, шкідники, зокрема вовчок соняшниковий) факторів, здійснювалася переважно за допомогою суб'єктивних бальних шкал, які ґрунтувалися на візуальних спостереженнях дослідника. У багатьох випадках така оцінка супроводжувалася умовною кількісною інтерпретацією, наприклад, у вигляді відсотка уражених рослин або ступеня деградації листової поверхні. Водночас обмежена кількість повторень, обумовлена ресурсними обмеженнями, додатково знижувала статистичну достовірність отриманих результатів. У зв'язку з цим фенотипування набуло особливого поширення насамперед у дослідженнях стресостійкості, де дозволило досягти істотного прогресу завдяки більш об'єктивному, відтворюваному та кількісному підходу до оцінки ознак (Liu S. et al., 2020; Adebola T. et al., 2025). Паралельно з розвитком фенотипових платформ активно розвиваються молекулярно-генетичні технології, зокрема методи картування геному та аналізу асоціацій генотип-ознака (GWAS). Завдяки цим підходам було

ідентифіковано численні локуси, пов'язані з такими ознаками, як посухостійкість; найглибші дослідження дозволили локалізувати відповідні гени на конкретних ділянках геному (Zhu H., Huang L., Li Y. 2024; Кендзьора, Н. З. 2020). Однак наступним кроком у розумінні природи цих ознак є вивчення функціональної ролі відповідних генів у формуванні кінцевого фенотипу, що потребує залучення фізіолого-біохімічних досліджень (Примак І. Д., 2017). Такі роботи наразі мають поодинокий характер через їх високу вартість, трудомісткість і технологічну складність як на фенотиповому, так і на молекулярному рівні.

Розвиток цифрових моделей рослин протягом останніх двох десятиліть відбувався у тісному взаємозв'язку з еволюцією інструментальних платформ, що забезпечують інтеграцію алгоритмічних підходів, симуляцій зростання, 3D-реконструкцій та зв'язку із фенотиповими базами даних.

Найбільш впливовими серед таких середовищ стали OpenAlea, GroIMP, L-Py та спеціалізована модель MAppleT, кожна з яких зробила вагомий внесок у загальну методологію цифрового фенотипування та проектування віртуальних рослин.

OpenAlea (INRIA, Франція) — це модульна платформа з відкритим кодом, створена для аналізу рослинної архітектури, моделювання ростових процесів та інтеграції різнорідних даних. Ключовими особливостями системи є компонентна структура (workflow-oriented), що полегшує інтеграцію математичних моделей; наявність інструментів для 3D-реконструкції, підтримка L-систем та симуляцій екологічних взаємодій у мультидисциплінарних задачах (від морфометричного аналізу до фізіологічного моделювання). OpenAlea стала фактичною базою для розвитку низки спеціалізованих бібліотек, зокрема L-Py.

L-Py — це Python-орієнтований модуль над платформою OpenAlea, що надає високоефективний інтерфейс для побудови L-систем і швидких симуляцій архітектури рослини. Його переваги зумовлені повною інтеграцією з екосистемою Python і можливістю паралельної обробки великих масивів

фенотипових даних; підтримкою параметризованих L-систем, що дозволяють моделювати конкретні генотипи або умови середовища, а також зручністю застосування у швидкій прототипізації моделей. L-Py значною мірою сприяв стандартизації підходів до семантичного опису цифрових рослин.

GroIMP (Growth Grammar-related Interactive Modelling Platform) — середовище, оптимізоване для моделювання складних 3D-структур із використанням розширених граматики розвитку — eXtended L-систем (XL). Її ключові інновації полягають у наявності потужного інструментарію 3D-рендерингу та симуляції світлових потоків методом трасування променів (ray-tracing), що важливо для досліджень взаємодії «рослина–світло». Платформа підтримує розширені граматики для опису розвитку архітектури та можливості моделювання міжвидової ресурсної конкуренції (competition modelling).

MAppleT — спеціалізована модель віртуального дерева яблуні, створена для аналізу архітектурного розвитку, продуктивності та реакції на агротехнічні операції. Включає детальну параметризацію органів та габітусу культури, симуляцію багаторічного росту й плодоношення, а також можливість інтеграції з даними польових фенотипувань та LIDAR-зйомок. MAppleT є класичним прикладом вузькоспеціалізованих платформ, які демонструють, як точні архітектурні моделі можуть слугувати практичним інструментом для селекції та садівництва.

Загальні тренди еволюції платформ та порівняльний аналіз розвитку систем OpenAlea, GroIMP, L-Py та MAppleT дозволяють виділити кілька ключових тенденцій.

Перехід до відкритих екосистем: Домінування Python-орієнтованих інструментів, що полегшують інтеграцію з реляційними та семантичними фенотиповими БД.

Зростання біологічного реалізму: Глибоке включення моделей фізіології, спектральних характеристик світлових умов та ресурсної конкуренції.

Спрямованість до інтеграції: Об'єднання систем з високопродуктивним цифровим фенотипуванням, методами машинного зору та концепцією динамічних цифрових двійників рослин.

Розвиток доменно-специфічних моделей: Створення систем (за аналогією з MAppleT), які виступають технологічними прототипами для побудови онтологічно узгоджених систем фенотипових БД для нових цільових культур, зокрема для соняшника.

## **2.2 Структура та етапи формування цифрової моделі рослини соняшнику**

Рослина є складною біологічною системою, що характеризується багаторівневою структурою, динамічними процесами росту та розвитку, а також взаємодією з зовнішнім середовищем. З погляду системного аналізу, рослина розглядається як інтегрована система, яка складається з взаємопов'язаних підсистем:

Морфологічна підсистема — органи рослини (корінь, стебло, листки, квітка, плоди) та їх топологічна і функціональна організація.

Фізіологічна підсистема — процеси фотосинтезу, транспірації, метаболізму, росту та диференціації тканин.

Генетична підсистема — генотипічна детермінація розвитку та реакцій на фактори середовища (взаємодія «геном-середовище»).

Екологічна підсистема — взаємодія з середовищем, включно з ресурсами (світло, вода, поживні речовини), стресовими факторами та іншими організмами.

Системний підхід до рослини передбачає математичне та комп'ютерне моделювання як цілісної системи, де всі підсистеми взаємодіють динамічно та нелінійно; ідентифікацію ключових факторів, що визначають продуктивність та адаптивні властивості; прогнозування поведінки системи під впливом змін середовища або агротехнічних факторів; оптимізацію та контроль процесів розвитку рослини, наприклад, у селекції або точному землеробстві.

Рослина як об'єкт системного аналізу є основою для створення цифрових моделей та цифрових двійників, оскільки її складна структура і функції добре піддаються формалізації через FSPM, AI-моделі, математичні симуляції; підсистеми можуть бути представлені як компоненти з параметрами, що змінюються в часі; можливе інтегрування генотипічних, фенотипічних та екологічних даних для прогнозування росту, розвитку та адаптаційних реакцій.

За допомогою системного підходу вирішуються такі ключові завдання.

Проводиться виявлення критичних факторів росту та розвитку. Проводиться аналіз взаємодії між морфологічними та фізіологічними процесами. Відбувається прогнозування продуктивності та реакції на стресові фактори. Моделювання застосовується для підтримки прийняття рішень в селекційних та агротехнічних завданнях. Відбувається інтеграція даних з цифрових моделей, сенсорних систем та фенотипових баз.

Структурні моделі (також відомі як архітектурні або морфологічні) зосереджуються на детальному описі тривимірної (3D) геометрії та топології рослини, тобто її архітектури (de Reffye et al., 1998).

Основний фокус цих моделей є представлення рослини як системи взаємопов'язаних елементарних одиниць (наприклад, міжвузлів, листків, бруньок).

Функціональні моделі (також відомі як процесно-орієнтовані або фізіологічні моделі) зосереджуються на динаміці метаболічних та фізіологічних процесів, які керують ростом і розвитком рослини, часто ігноруючи або сильно спрощуючи її 3D архітектуру (Guo et al., 2006).

Функціонально-структурні моделі рослин (ФСМР, англ. Functional-Structural Plant Models, FSPM) є найкомплекснішим підходом, що забезпечує експліцитну інтеграцію 3D архітектури рослини з описом фізіологічних процесів, які керують її ростом та розвитком (Sievänen et al., 2014; Guo et al., 2006).

ФСМР вимагають детального опису трьох взаємопов'язаних підсистем.

Перша підсистема, яка потребує опису, це модель архітектури, структурний компонент. Топологія та геометрія визначає, як органи (модулі) пов'язані між собою (топологія) і де вони розташовані у 3D просторі (геометрія). Часто використовується формалізм L-систем (як основа) або об'єктно-орієнтовані графові структури, де кожен орган (наприклад, листок, міжвузля, брунька) представлений окремим об'єктом із власними атрибутами та станом (Prusinkiewicz, 2004). Моделювання процесів морфогенезу – ініціація та розвиток нових модулів (наприклад, проростання бруньок, формування нового листка), що часто залежить від ресурсного балансу та гормональних сигналів.

Друга підсистема, яка потребує опису, це модель мікросередовища. Найважливішим компонентом є розрахунок впливу радіації, світла. Розраховується перехоплення світла (прямого та розсіяного) на поверхні кожного індивідуального листка або іншого органу, з урахуванням затінення сусідніми органами. Використовуються методи, такі як трасування променів (ray tracing) або радіозиті (radiosity), для точного визначення локальної освітленості (Godin and Sinoquet, 2005). Інші фактори це локальна температура, вологість та швидкість вітру навколо органів.

Третя підсистема, яка потребує опису, це функціональний компонент, модель функції та розподілу ресурсів. Локальні фізіологічні процеси (фотосинтез, транспірація) моделюються на рівні індивідуального модуля, використовуючи локальні умови мікросередовища. Ресурси (наприклад, асимільований вуглець) агрегуються з усіх активних модулів і потім розподіляються по всій рослині. Розподіл керується попитом (потребами ростучих меристем) та пропозицією (швидкістю виробництва асимілятів), а також анатомічними обмеженнями (транспортні шляхи) (Hu et al., 2003).

Розвиток цифрових моделей рослин протягом останніх двох десятиліть відбувався у тісному взаємозв'язку з еволюцією інструментальних платформ, що забезпечують інтеграцію алгоритмічних підходів, симуляцій зростання, 3D-реконструкцію та зв'язок із фенотиповими даними. Найбільш впливовими серед

таких середовищ стали OpenAlea, GroIMP, L-Py та спеціалізована модель MAppleT, кожна з яких зробила власний внесок у загальну методологію цифрового фенотипування та віртуальних рослин.

OpenAlea (INRIA, Франція) — це модульна платформа з відкритим кодом, створена для аналізу рослинної архітектури, моделювання ростових процесів та інтеграції різномірних даних.

Ключові особливості:

- компонентна структура (workflow-oriented), що полегшує створення складних моделей;
- інструменти для 3D-реконструкції, L-систем та симуляцій екологічних взаємодій;
- підтримка мультидисциплінарних задач: від морфометричного аналізу до фізіологічного моделювання.

OpenAlea стала фактичною основою для розвитку низки спеціалізованих бібліотек, зокрема L-Py.

GroIMP (Growth Grammar-related Interactive Modelling Platform) — середовище, оптимізоване для моделювання складних 3D-структур із використанням eXtended L-систем (XL).

Її ключові інновації:

- потужний 3D-рендеринг та симуляція світлових потоків (ray-tracing), що робить GroIMP надзвичайно цінною для досліджень взаємодії рослина–світло;
- розширені граматики для опису розвитку архітектури;
- можливість реалізації моделей взаємодії між рослинами (competition modelling).

GroIMP широко використовується в біофізичних моделях та при створенні віртуальних експериментів.

L-Py — це Python-орієнтований модуль над OpenAlea, що надає доступний інтерфейс для побудови L-систем і швидких симуляцій архітектури рослини.

Переваги L-Py зумовлені:

- інтеграцією з екосистемою Python і можливістю обробки великих масивів фенотипових даних;
- підтримкою параметризованих L-систем, що дозволяють моделювати конкретні генотипи або умови середовища;
- зручністю застосування у навчанні та швидкій прототипізації моделей.

L-Py значною мірою сприяв стандартизації підходів до опису цифрових рослин.

MAppleT — спеціалізована модель віртуального дерева яблуні, створена для аналізу архітектурного розвитку, продуктивності та реакції на агротехнічні операції.

Основні елементи:

- детальна параметризація органів та габітусу яблуні;
- симуляція багаторічного росту та плодоношення;
- можливість інтеграції з даними польових фенотипувань та LIDAR-зйомок.

MAppleT є прикладом вузькоспеціалізованих платформ, які демонструють, як точні архітектурні моделі можуть слугувати практичним інструментом для селекції та садівництва.

Загальні тренди еволюції платформ

Порівняльний аналіз розвитку OpenAlea, GroIMP, L-Py та MAppleT дозволяє виділити кілька ключових тенденцій:

- Перехід до відкритих екосистем та Python-орієнтованих інструментів, що полегшують інтеграцію з фенотиповими БД.
- Зростання рівня біологічного реалізму, включно з моделями фізіології, світлових умов, ресурсної конкуренції.
- Спрямованість до інтеграції з високопродуктивним фенотипуванням та цифровими двійниками рослин.

- Розвиток доменно-специфічних моделей, як MAppleT, які можуть слугувати прототипами для створення онтологічно узгоджених систем фенотипових БД.

### **2.3 Структура та етапи формування цифрової моделі фенотипу соняшнику кондитерського напрямку використання**

Центральним вектором формування цифрової моделі рослини соняшнику є формалізація взаємодії «генетика — середовище». Цифровий образ рослини розглядається як результат взаємодії спадкових властивостей організму та агрокліматичного комплексу вирощування. Таке представлення дає змогу математично та логічно оцінити норму реакції генотипу, виявити рівень його адаптивної пластичності та спрогнозувати динаміку розвитку органів соняшника (зокрема, прогнозування дати цвітіння чи фітосанітарних ризиків) на основі інтеграції даних комп'ютерного зору та ґрунтових аналізів.

Визначимо склад та етапи створення цифрової моделі рослини соняшника, які представлено в роботах автора (Ведмедєв, С & Терещенко, Е. 2025; Vakurova A., Vedmedeva K., Vedmedev S., Tereschenko E. 2023).

У сільськогосподарському виробництві та селекції соняшник (*Helianthus annuus* L.) за господарським призначенням та напрямом використання насіння поділяють на кілька основних типів. Олійний соняшник є основним промисловим типом, насіння якого містить до 54% жиру і використовується для масового виробництва рослинної олії. Кондитерський соняшник вирізняється великим смугастим насінням із низьким вмістом олії та високим вмістом білка, що робить його затребуваним у харчовій промисловості. Спеціалізовані типи охоплюють межові гібриди подвійного використання, силосні сорти для відгодівлі худоби та декоративні рослини для флористики.

Цифрова модель рослини соняшнику включає сукупність формалізованих ознак, що відображають морфо-біологічний стан рослини та її господарську цінність.

Морфометричні характеристики насінини включають фізичні параметри (довжина, ширина, товщина), маса 1000 насінин, форма та виповненість, що є визначальними для кондитерської індустрії.

Архітектоніка кошика (суцвіття) визначається геометричною формою, просторовою орієнтацією (кут нахилу), діаметром та ступенем покриття насінням.

Вегетаційні та архітектурні ознаки рослини включають лінійні розміри (висота), забарвлення органів тощо. Біохімічний профіль визначається кількісним вмістом олії, білків та лушпинністю, що інтегруються в цифрову модель як атрибутивні дані.

При селекційному відборі соняшника важливими є ознаки насінини та кошика, а саме розмір насіння, натура, лушпинність, щільність лушпиння, обрушуваність, відсоток крупного насіння, форма, розміри, забарвлення насіння, вміст білку та олії. В роботі (Bakurova A., et al. 2023b) було визначено перелік характеристик для соняшнику за ознаками насінини та кошика, який базується на вимогах Державної методики України «Методика визначення однорідності та стабільності (ВОС)» та дослідженнях Інституту олійних культур НААН України (Aliyev E.V., 2019). Дослідниками Інституту олійних культур було розширено набір характеристик й введені нові градації для деяких з них. Запропоновано нові градації і вимір ознак у кількісному варіанті для співвідношення площ темного та світлого забарвлень та сили руйнування оболонки насінини. В роботі (Bakurova A., et al. 2023b) зафіксовано такий перелік характеристик для цифрової моделі рослини соняшнику для фенотипування: довжина насінини (мм, см), ширина насінини (мм, см), товщина насінини (мм, см), форма насінини (категоріальна), товщина відносно ширини (тонка, середня, товста), основний колір насінини (категоріальна), смуги на краю насінини (відсутні, слабкі,

виражені), колір смуг (категоріальна), співвідношення площ темного та світлого забарвлень ( $< 1/4$ ,  $1/4-3/4$ ,  $> 3/4$ ), плямистість перикарпію (відсутня, наявна), вага насінини (г), вміст олії у насінні (%), вміст білку у насінні (%), лушпинність (%), вміст олеїнової кислоти в олії (%), вміст пальмітинової кислоти в олії (%), вміст лінолевої кислоти в олії (%), вміст ліноленової кислоти в олії (%), сила руйнування оболонки насінини (Н), положення кошику (кут у градусах або категоріальна), розмір кошику (см), форма кошику зі сторони сім'янок (категоріальна), кількість насінин у кошику (шт.), кількість спіралей насінин у кошику (шт.), кількість насінин у кожній спіралі (шт.), вага насінин у кошику (г). Визначено джерело отримання даних та метод зберігання.

Процес побудови цифрової моделі рослини соняшнику реалізується через послідовність етапів отримання, зберігання, обробки інформації, та на заключному етапі - формування цифрової моделі фенотипу рослини.

Інтегрування етапів в єдину послідовність визначає послідовний тип системи цифрового фенотипування соняшника Digital phenotyping system, загальна структурна схема якої представлена на рис.2.2.1.

Представлена структурна схема Digital phenotyping system відображає процес перетворення даних, що отримані в польових експериментах та лабораторних умовах щодо показників рослинних об'єктів у формалізовану цифрову модель.

На першому етапі відбувається реєстрація первинних даних (Data Collection gadget). На другому етапі формується сховище та систематизація (DATA Base) даних. На цьому етапі необхідно забезпечити принципи FAIR, а саме забезпечення цілісності та довготривалого зберігання, структурування даних за реляційними або графовими моделями, прив'язка сирих даних до унікальних ідентифікаторів генотипів та часових міток (Timestamping).

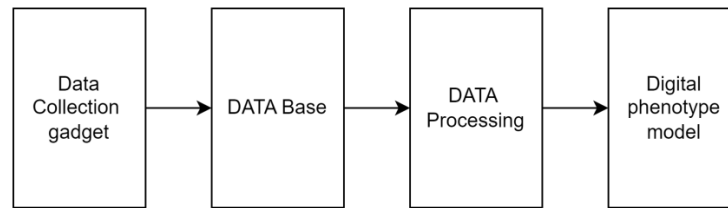


Рисунок 2.2.1 – Загальна структурна схема Digital phenotyping system (розроблено автором)

Модуль інтелектуальної обробки (DATA Processing) для вилучення корисних ознак, а саме обчислення площі, периметра, довжини та геометричних індексів, біометрична корекція щодо усунення шумів та артефактів зйомки для отримання точних кількісних значень.

Результуюча цифрова модель фенотипу (Digital phenotype model) є кінцевим етапом, результатом якого є цифровий образ об'єкта, що містить формалізований опис ознак соняшнику, враховані фактори середовища, набори даних (datasets) для аналізу взаємодії генотипів та часових міток та підтримки селекційних рішень.

Особливий акцент у дослідженні зроблено на фенотипуванні за ознаками насінини та кошика, оскільки саме вони є найбільш інформативними характеристиками адаптивності та продуктивності кондитерського соняшнику.

Для опису фенотипу насіння розроблені спеціальні методики. Державна методика України має назву «Методика визначення однорідності та стабільності (ВОС)» (2003). Європейська методика має назву: «International union for the protection of new varieties of plants (UPOV)» (UPOV TG/81/7 Sunflower Guidelines, 2011). Європейська методика має на одну градацію більше. Фізичні розміри насіння складають від 5 до 25 мм довжини, та від 3 до 7 мм ширини. Товщина насіння менша за ширину. У документах опису існують чотири градації форми

насіння. Забарвлення насіння соняшнику описують за допомогою чотирьох або п'яти ознак з градаціями.

Digital phenotyping system містить три базові підсистеми фенотипу насінини та кошика рослини соняшнику кондитерського, а саме підсистема збору візуальних характерних ознак насінини; підсистема збору характеристик насіння за біохімічними аналізами; підсистема збору характеристик кошику, стебла та кореня рослини.

Підсистема збору візуальних характерних ознак насінини призначена для створення колекції фотозображень насіння різних генотипів (Таблиця 2.2.1). В цій системі будуть відбуватися процедури агрегування інформації щодо розмірів, забарвлень та форми насіння.

Підсистема збору характеристик насіння за біохімічними аналізами призначена для агрегування інформації щодо результатів лабораторних аналізів на вміст білку, олеїнової кислоти, лінолевої кислоти, пальмітинової кислоти, ліноленової кислоти (Таблиця 2.2.2). Додатково ця підсистема долучає до агрегування фізичні результати оцінки сили луцення на відповідному обладнанні з автоматичною фіксацією результатів експерименту.

Кошик характеризується за розміром, формою та нахилом. За нахил кошика відповідають 12 генів з адитивним ефектом, кожен з яких змінює нахил кошика на 15 градусів. Всього реєструють 9 типів положення кошика відносно стебла від горизонтального(0\*) до вертикального (90\*), напівоберненого (135\*) та оберненого донизу (180\*) (UPOV TG/81/7 Sunflower Guidelines, 2011).

Таблиця 2.2.1 – Параметри цифрової моделі фенотипу насінини  
(сформовано автором за даними Інституту олійних культур НААН України)

Характеристики	Тип	Джерело даних
Розмір	Довжина Ширина Товщина	фото
Форма	Видовжена Вузько -яйцеподібна Широко -яйцеподібна Округла	фото
Форма за товщиною відносно ширини	тонка (<0,5 мм) середня (0,51-0,7мм) товста (> 0,71 мм)	фото
Забарвлення	біле білувато-сіре сіре світло-коричневе темно-коричневе чорне пурпурове	фото
смугастість по краях	відсутня або виражена дуже слабо слабо виражена сильно виражена	фото
Смугастість між боковими краями	відсутня або виражена дуже слабо слабо виражена сильно виражена	фото
Забарвлення смужок	біле сіре коричневе чорне	фото
Співвідношення площ темного та світлого забарвлень	< 1/4 1/4<3/4 >3/4	фото
Плямистість перикарпію	«Відсутність» / «Наявність»	фото

Таблиця 2.2.2. – Біохімічні та фізичні характеристики насіння (сформовано автором за даними Інституту олійних культур НААН України)

Характеристики	Джерело даних
Вага насінини, г	Лабораторні аналізи
вміст олії у насінні, %	Лабораторні аналізи
Вміст білку у насінні, %	Лабораторні аналізи
Лушпинність, %	Лабораторні аналізи
Вміст олеїнової кислоти в олії, %	Лабораторні аналізи
Вміст пальмитинової кислоти в олії, %	Лабораторні аналізи
Вміст лінолевої кислоти в олії, %	Лабораторні аналізи
Вміст ліноленової кислоти в олії, %	Лабораторні аналізи
Сила руйнування оболонки насінини	Лабораторні аналізи

Підсистема збору характеристик кошику рослини (Таблиця 2.2.3) призначена для створення колекції фотозображень кошиків різних генотипів та агрегування інформації щодо розмірів кошиків, кількості насінин, кількості виповнених насінин, куту нахилу кошику з фотозображень.

Матеріалом дослідження є насіння соняшнику колекції Інституту олійних культур НААН України. Опис колекції здійснювався за безпосередніми візуальними спостереженнями та за допомогою аналізу зображень фотографування у полі та у лабораторії. Вимірювання сили руйнування здійснювали за допомогою стаціонарного приладу (Алієв Е. Б. Ведмедева К.В., 2024), інформація з якого має прямий канал передачі до ноутбука. Біохімічні показники отримували з даних лабораторії біохімії проведених за ДСТУ.

Для оцінки генотипу та впливу зовнішнього середовища дуже важливими є такі показники як висота рослин і площа та колір листя. Ці ознаки є ключовими індикаторами сили росту, фотосинтетичного потенціалу та стійкості гібридів до вилягання. Для оцінки можливості успішного виживання рослини в складних умовах необхідно враховувати довжину головного кореня та об'єм ґрунту охоплений коренями рослини.

Таблиця 2.2.3. – Характеристики кошика (сформовано автором за даними Інституту олійних культур НААН України)

Характеристики	Тип	Джерело даних
Кошик: положення	горизонтальне вертикальне напівобернене донизу з прямим стеблом напівобернене донизу з зігнутим стеблом обернене донизу зі злегка зігнутим стеблом обернене донизу з прямим стеблом обернене донизу з сильно зігнутим стеблом дуже обернене	фото
Кошик: розмір	малий 3 (5,1 - 10,0 см) середній 5 (10,1 - 20,0 см) великий 7 (20,1 - 40,0 см.)	Фото вимірювання лінійкою
Кошик: за формою з боку сім'янок	сильно увігнута 1 плескатий 3 злегка випуклий 4 сильно випуклий 5 деформований 6	фото

Таблиця 2.2.4 містить морфометричні та архітектурні характеристики стебла і листкового апарату рослини соняшника, сформована відповідно до вимог Державної методики України (ВОС) та сучасних методів дистанційного зондування (БПЛА/комп'ютерного зору).

Для повної дескрипції підземної архітектоніки агроєкосистеми соняшника, яка визначає його посухостійкість та адаптивну пластичність, цифрову модель необхідно розширити підсистемою параметрів кореневої системи. Таблиця 2.2.5 містить морфологічні та структурно-функціональні характеристики кореневої системи рослини соняшника.

Оскільки корінь соняшника є стрижневим і може проникати на глибину понад 1.5–2 метри, його кількісна оцінка є критично важливою для моделювання взаємодії «генотип — середовище» за умов дефіциту вологи.

Таблиця 2.2.4 - Морфометричні та архітектурні характеристики стебла і листкового апарату рослини соняшника (сформовано автором за даними Інституту олійних культур НААН України)

Категорія ознак	Характеристика	Тип даних / Градації	Джерело даних
Параметри стебла	Висота рослини	Кількісний (см)	Польові вимірювання / LiDAR зйомка / БПЛА
	Діаметр стебла (біля основи та під кошиком)	Кількісний (мм)	Польові обміри цифровим штангенциркулем
	Кількість міжвузлів	Кількісний (шт.)	Візуальний підрахунок / Фотофіксація
	Інтенсивність антоціанового забарвлення стебла	Категоріальний (відсутнє або дуже слабе, слабе, середнє, сильне)	Фотозображення (аналіз у колірних просторах RGB/HSV)
	Ступінь галуження стебла	Категоріальний (відсутнє, лише на верхівці, по всій довжині)	Фотографія рослини в повний зріст
	Опушення стебла (у верхній третині)	Категоріальний (слабе, середнє, сильне)	Візуальна оцінка / Макрофотозйомка
Архітектоніка листя	Загальна кількість листків на рослині	Кількісний (шт.)	Візуальний підрахунок / Аналіз 3D-моделей (Voxel/Mesh)
	Просторове розташування (орієнтація листків)	Категоріальний (горизонтальне, підняте, поникле)	Фотозображення / 3D Gaussian Splatting
	Розмір листкової пластинки (середній)	Категоріальний (малий, середній, великий) або Кількісний (довжина/ширина в см)	Фотофіксація листка на каліброваному фоні
Архітектоніка листя	Площа листкової поверхні рослини (LAI)	Кількісний (м <sup>2</sup> /рослину або індекс)	Розрахунок за алгоритмами OpenCV (сегментація зеленої маси) / БПЛА
	Форма верхівки листка	Категоріальний (вузькозагострена, середньозагострена, тупа)	Комп'ютерний аналіз контурів за фото
	Інтенсивність зеленого забарвлення листя	Категоріальний (світло-зелене, зелене, темно-зелене)	Спектральні камери / Колірні індекси фотозображень
	Пухирчастість (текстура поверхні листка)	Категоріальний (відсутня або дуже слабка, слабка, середня, сильна)	Макрофотозйомка (аналіз текстурних ознак)
	Форма поперечного розрізу листка	Категоріальний (пласка, увігнута, випукла)	3D-реконструкція органу
	Зубчастість краю листкової пластинки	Категоріальний (дрібна, середня, крупна)	Цифрова обробка контурів зображення

Таблиця 2.2.5. Морфологічні та структурно-функціональні характеристики кореневої системи рослини соняшника (сформовано автором за даними Інституту олійних культур НААН України)

Категорія	Характеристика	Тип даних / Градації	Джерело даних
Архітектоніка головного кореня	Довжина головного (стрижневого) кореня	Кількісний (см, м)	Польові розкопки (моноліти) / Цифрові різотрони / Симуляційне моделювання
	Діаметр головного кореня біля кореневої шийки	Кількісний (мм)	Вимірювання цифровим штангенциркулем
	Ступінь звивистості стрижневого кореня	Категоріальний (прямий, слабкодеформований, сильнодеформований)	Цифрова обробка фотографій відмитих кореневих систем
Структура бічного розгалуження	Загальна кількість бічних коренів першого порядку	Кількісний (шт.)	Комп'ютерний зір (OpenCV: скелетизація та підрахунок точок розгалуження)
Категорія	Характеристика	Тип даних / Градації	Джерело даних
	Щільність розгалуження (кількість бічних коренів на 10 см стрижня)	Кількісний (шт./10 см)	Розрахунок за алгоритмами обробки зображень
Категорія	Характеристика	Тип даних / Градації	Джерело даних
Просторові та об'ємні метрики	Загальний об'єм ґрунту, охопленій кореневою системою	Кількісний (см <sup>3</sup> або м <sup>3</sup> )	3D-реконструкція (Voxel/Mesh) / Томографія / Георадари
	Загальна довжина кореневої системи (Root Length Density — RLD)	Кількісний (см/см <sup>3</sup> ґрунту)	Лабораторний аналіз вимитих зразків (сканери WinRHIZO)
	Сира та суха біомаса кореневої системи	Кількісний (г)	Лабораторне зважування (деструктивний метод)
Глибинний розподіл	Глибина ефективного поглинання вологи	Кількісний (см)	Розрахунковий метод на основі вологоємних сенсорів ґрунту на різних глибинах
	Архітектура кореневої системи за шарами ґрунту (0–30 см, 30–60 см, 60–100 см)	Пропорційний розподіл біомаси (%)	Метод ґрунтових кернів / Симуляційні моделі росту коренів

У польових умовах селекції пряме ручне фенотипування коренів є деструктивним і трудомістким, тому в сучасних системах дані отримують за допомогою методів викопування монолітів (*Shovelomics*), ризотронів (у контрольованому середовищі) або неінвазивних методів (георадари, томографія, симуляційне моделювання FSPM).

Далі визначимо параметри впливу зовнішнього середовища, які потребують визначення при фенотипуванні в контексті селекційних досліджень.

Значна частина успіху у реалізації генотипу залежить від складу та стану ґрунту, що добре описано у роботах з ґрунтознавства, агрохімії, агрономії та інших науках (Іванова, І. Є., та ін. 2026). Головними факторами для рослин є вміст головних поживних елементів у вигляді доступних сполук азоту, фосфору, калію, а також наявність відповідної органічної складової – гумусу та мікроелементів (Аверченко В. І., Самойленко Н. М., 2018). Основні характеристики ґрунту представлено в таблиці 2.2.6.

Основні кліматичні характеристики, які фіксуються для фенотипування рослин у селекційних дослідженнях представлено в таблиці 2.2.7. Джерелами кліматичних даних можуть бути локальні метеостанції, мобільні агрокліматичні сенсори, супутникові платформи (Copernicus, NASA POWER), дані з агрометеорологічних сервісів (Meteoblue, NOAA, Weather API). В умовах України збільшується посушливість клімату, що доведено Семеновою І.Г. (2017).

Частина кліматичних умов і властивостей ґрунту можна цілеспрямовано модифікувати за допомогою агротехнічних прийомів. Зокрема, йдеться про здійснення додаткового зрошення, внесення мінеральних та органічних добрив, глибоке рихлення ґрунту.

Таблиця 2.2.6. Характеристики ґрунту, які фіксуються для задач фенотипування соняшнику (сформовано автором за даними Інституту олійних культур НААН України)

Категорія	Характеристика	Одиниця вимірювання / Тип даних	Джерело даних
Фізичні властивості	Тип ґрунту	Класифікаційна категорія	Лабораторний аналіз / польова оцінка
	Структура ґрунту	Якісна оцінка	Візуальна оцінка / сенсори
	Щільність	г/см <sup>3</sup>	Лабораторія / польові сенсори
	Пористість	%	Розрахунок
	Вологість	%	Вологоємні сенсори
	Температура	°С	Ґрунтові термометри
	Кам'янистість	% або кількість каміння на м <sup>2</sup>	Візуальна оцінка / сенсори
Хімічні властивості	pH	Шкала кислотності (0–14)	Лабораторія / pH-метри
	Електропровідність	dS/m	Сенсори / лабораторія
	Вміст гумусу	%	Лабораторія
	Вміст органічної речовини	%	Лабораторія
	Азот (N)	мг/кг	Лабораторія
	Фосфор (P)	мг/кг	Лабораторія
	Калій (K)	мг/кг	Лабораторія
	Мікроелементи (Fe, Mn, Zn, B тощо)	мг/кг	Лабораторія
Біологічні властивості	Мікробіологічна активність	Кількість колоній / активність	Лабораторія
	Наявність патогенів	Виявлено / Не виявлено	Лабораторія
Агротехнічні фактори	Історія обробітку поля	Текст / запис	База агрономічних даних
	Система удобрення	Типи добрив, дози, дати внесення	Агрономічна документація
	Зрошення	Обсяг, частота	Сенсори / агродокументація
	Попередники	Назви культур	План сівозміни / документація

Таблиця 2.2.7. Кліматичні характеристики, які фіксуються для фенотипування рослин у селекційних дослідженнях (сформовано автором за даними Інституту олійних культур НААН України)

Категорія	Характеристика	Одиниця вимірювання / Тип даних	Джерело даних
Температурні умови	Середньодобова температура	°С	Метеостанції / датчики
	Максимальна денна температура	°С	Метеостанції / логери
	Мінімальна нічна температура	°С	Метеостанції / логери
	Температурні стреси (кількість днів >30°С або <5°С)	Кількість днів	Метеорологічні архіви
	Сума активних температур за період вегетації рослин	°С	Метеостанції / логери
Опади	Загальна кількість опадів	мм	Метеостанції / дощоміри
	Інтенсивність опадів	мм/год або мм/доба	Метеостанції
	Посушливі періоди	Кількість днів без опадів	Аналіз метеоданих
	Сума опадів за період вегетації рослини	мм	Метеостанції / дощоміри
Сонячне випромінювання	Сонячна радіація (інсоляція)	МДж/м <sup>2</sup> /день або Вт/м <sup>2</sup>	Радіометри / супутникові дані
	Тривалість світлового дня	Години	Астрономічні розрахунки
	Кількість похмурих / сонячних днів	Днів	Супутникові знімки / архіви
Вологість повітря	Середня відносна вологість	%	Метеостанції / гігрометри
	Мінімальна / максимальна	%	Метеостанції / логери
Вітер	Середня швидкість вітру	м/с	Анемометри / метеостанції
	Максимальні пориви вітру	м/с	Метеостанції
	Напрямок вітру	Кут / сектор (N, NE, E, тощо)	Метеостанції
Інші показники	Атмосферний тиск	гПа або мм рт. ст.	Барометри
	Випаровуваність (Evapotranspiration, ET0)	мм/доба	Розрахунок за формулами FAO
	Кількість градобоїв / буревіїв / заморозків	Події (факт фіксації)	Метеозведення / польові спостереження
Кліматичні індекси	Сума активних температур (>10°С)	°С	Агрономічні розрахунки
	Індекс аридності (засушливості)	Безрозмірна величина	Розрахунок за метеоданими
	Сприятливі дні для вегетації	кількість	Агрономічні моделі

Біотичні фактори, обумовлені дією кліматичних чинників і навколишнього середовища — зокрема шкідники та патогенні організми — також чинять істотний вплив на розвиток рослин. Їхня шкодочинність може змінюватися під впливом агротехнічних прийомів і засобів хімічного чи біологічного походження, що застосовуються в агрономічній практиці. Отже, агрономічні заходи відіграють ключову роль у реалізації потенціалу генотипу та його адаптації до конкретних кліматичних умов регіону. Тому врахування агрономічного впливу є необхідною умовою як у процесі фенотипування рослин, так і при побудові цифрових моделей їх розвитку. Дослідження, що проводяться в умовах контрольованого середовища, як правило, передбачають моделювання лише тих факторів, які створені людиною. Відтак, виникає необхідність забезпечення рослини всіма основними ресурсами: водою, світлом, теплом, макро- і мікроелементами. Включення впливу агрономічних чинників, насамперед агротехнічних прийомів (табл. 2.2.8), сприяє підвищенню точності аналізу фенотипових ознак. Інформація збирається вручну, за допомогою агроІТ-систем або автоматично з дронів, сенсорів чи машинної телеметрії.

Digital phenotyping system містить такі базові підсистеми фенотипу рослини соняшнику: підсистема збору візуальних характерних ознак насінини; підсистема збору характеристик насіння за біохімічними аналізами, підсистема збору характеристик кошику, стебла та кореня рослини, а також три підсистеми характеристик ґрунту, кліматичних характеристик та агротехнічних (рис.2.2.2). Агрегування та обробка інформації, яка буде отримана в процесі збору та обробки даних в цих підсистемах стане базою для створення цифрової моделі фенотипу рослини з певним генотипом.

Таблиця 2.2.8. Агротехнічні характеристики, які фіксуються та враховуються при фенотипуванні рослин (сформовано автором за даними Інституту олійних культур НААН України)

Категорія	Характеристика	Одиниця вимірювання / Тип даних	Джерело даних	
Обробіток ґрунту	Тип основного обробітку (оранка, глибоке рихлення, mini-till, no-till)	Кваліфікатор (назва методу)	Польовий журнал, агроплан	
	Глибина основного обробітку	см	Агроагрегати, агрономічна карта	
Сівба	Тип передпосівного обробітку	Кваліфікатор	Польові записи	
	Дата сівби	Дата	Журнал агротехнологій	
	Густота стояння рослин	тис. рослин/га	Польові обміри / дрони	
	Глибина загортання насіння	см	Польові виміри	
	Схема сівби	міжряддя (см)	Сівалки, документація	
Добрива	Вид добрив	НРК, органічні, мікроелементи	Агроплани, лабораторні аналізи	
Категорія	Характеристика	Одиниця вимірювання / Тип даних	Джерело даних	
	Доза внесення	кг/га або л/га	Польові журнали	
	Спосіб внесення (в основне, припосівне, листове, фертигація)	Кваліфікатор	Агрокартографія	
	Кількість внесень	Кількість разів	План-графік	
	Зрошення / волога	Спосіб зрошення (краплинне, дощування, борозенкове)	Кваліфікатор	План зрошення
		Об'єм поданої води	мм або м <sup>3</sup> /га	Системи зрошення
	Кількість поливів	Разів	Агроніма	
Захист рослин	Вид ЗЗР (гербіциди, фунгіциди, інсектициди)	Назва + діюча речовина	Журнал обробок, AgroScout	
	Дата та фаза внесення	Дата / фенологічна фаза	Скаутинг, агрокалендар	
	Доза внесення	л/га або кг/га	Сертифікат, обприскувач	
	Спосіб обробки	Наземний, авіаційний, дрон	Польові журнали	
Інші агроприйоми	Мульчування	Наявність / тип	Польові спостереження	
	Інокуляція / стимуляція	Препарати / методика	Журнал агрозаходів	
	Суміжні культури / сівозміна	Попередник, наступна культура	Агрономічний план	

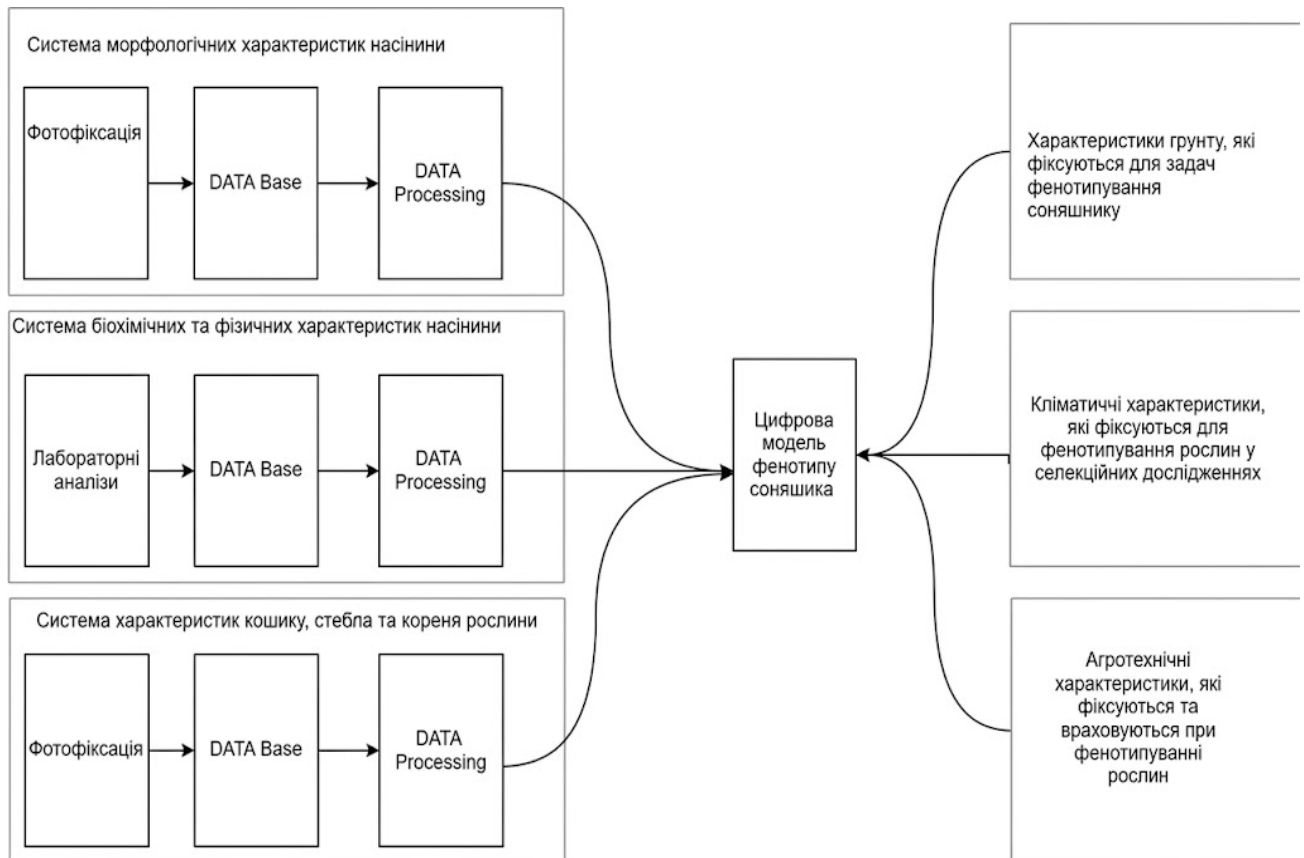


Рисунок. 2.2.2 - Структурна схема системи цифрової моделі рослини соняшнику *Helianthus* (створено автором)

## 2.4 Онтології: принципи, аксіоматика, класи, властивості, обмеження

Онтологічне моделювання є формалізованим підходом до структурного і семантичного опису предметної області, що дозволяє формалізувати знання про об'єкти, їхні властивості та взаємозв'язки. Онтологія містить базис для моделювання певної галузі знань та визначає відношення та домовленість про подання теоретичних основ цієї галузі знань. Відповідно до визначення, яке запропонував Т. Gruber, онтології – це формальні, явні специфікації спільних концептуалізацій (Gruber Т. 1991). Класичне визначення онтології є впорядкованою трійкою виду:

$$O = (T, R, F),$$

де  $T$  – скінчена непуста множина термінів (концептів, понять, класів) галузі знань, яку представляє онтологія  $O$ ;

$R$  – скінчена множина відношень між поняттями галузі знань;

$F$  – скінчена множина функцій інтерпретації (аксіоматизація), заданих на концептах та/або відношеннях онтології  $O$ .

Основні елементи онтології включають класи, екземпляри, властивості та аксіоми, кожен з яких виконує специфічну роль у побудові моделі знань.

Клас (Class) визначає множину об'єктів, що мають спільні характеристики. Клас формалізується як сукупність атрибутів та взаємозв'язків, які визначають об'єкти, що до нього належать.

Екземпляр (Instance / Individual) є конкретним представником класу — реальною чи симульованою сутністю, яка підпорядковується правилам свого класу.

Властивості (Properties / Attributes / Relations) описують ознаки класів та екземплярів і встановлюють зв'язки між об'єктами. В онтологічному моделюванні виділяють два типи властивостей. Datatype properties описують якісні та кількісні характеристики (числові значення, текстові атрибути). Object properties описують взаємозв'язки між екземплярами чи класами. Властивості дозволяють будувати семантичну мережу знань та забезпечують логічне виведення нових фактів.

Аксіоми (Axioms) визначають логічні правила та обмеження, які застосовуються до класів, екземплярів та властивостей, забезпечуючи узгодженість та коректність онтології. Аксіоми класу описують загальні правила для всіх екземплярів класу. Обмеження властивостей визначають допустимі значення та типи зв'язків. Логічні аксіоми дозволяють робити виведення нових знань.

Методи онтологічної інженерії базуються на принципах, запропонованих Т. Грубером (Gruber T., 1991), метою яких є обмін знаннями та взаємодія між програмами на основі спільної концептуалізації.

1. Ясність (Clarity). Онтологія повинна ефективно передавати передбачуване значення визначених термінів. Необхідність фіксування аксіом, які визначають інтерпретації введених термінів.
2. Узгодженість (Coherence). Онтологія є узгодженою, якщо визначальні аксіоми логічно узгоджені, дозволені висновки, включно для понять, що представлені природньою мовою, узгоджуються з визначеннями. Якщо висновки, які можна вивести з аксіом, суперечить визначенню або прикладу, наведеному природною мовою, то онтологія є неузгодженою.
3. Розширюваність (Extendibility). Онтологія визначає такий словник для низки передбачуваних завдань, який може бути розширений та спеціалізований за рахунок нових понять, без змін існуючих визначень.
4. Мінімізація упередженості кодування (Minimal encoding bias). Концептуалізація повинна бути визначена на рівні знань без залежності від конкретного кодування, оскільки агенти обміну знаннями можуть бути реалізовані в різних системах представлення.
5. Мінімізація семантичної пов'язаності та внутрішньої взаємозалежності компонентів (Minimal ontological commitment). Архітектура предметної онтології повинна проєктуватися з урахуванням критерію мінімально необхідної семантичної пов'язаності компонентів, яка є строго достатньою для забезпечення коректного обміну знаннями та інтеграції різнорідних даних. Рівень внутрішньої взаємозалежності концептів мінімізується шляхом формалізації найбільш гнучкої логічної теорії та специфікації виключно тих сутностей і зв'язків, без яких неможлива однозначна інтерпретація та передача інформації в межах досліджуваної системи. Такий підхід забезпечує слабку зчепленість (low coupling) модулів бази знань, спрощує її подальше масштабування та виключає надлишкові логічні обмеження при виведенні рішень.

Методології створення онтологій визначають організаційні, процесні та технологічні задачі, що впливають на якість онтологій та її застосування. У

межах сучасної інженерії знань прийнято виділяти три фундаментальні методологічні парадигми, кожна з яких орієнтована на вирішення специфічного класу завдань при побудові інтелектуальних систем.

Методологія METHONTOLOGY, розроблена Fernández-López et al. (2002) на основі стандартів життєвого циклу програмного забезпечення IEEE 1074, є найбільш деталізованим підходом до інженерії знань, що забезпечує сувору формалізацію на всіх етапах розробки. Процес розпочинається з етапу специфікації (Specification), на якому чітко визначаються мета, обсяг та необхідний ступінь формалізації майбутньої системи, після чого здійснюється концептуалізація (Conceptualization) для створення «проміжної моделі» (Intermediate Representation), що структурує знання у формі словника термінів, таксономій, діаграм бінарних зв'язків та таблиць атрибутів і аксіом. Наступна фаза формалізації (Formalization) забезпечує трансформацію цієї проміжної моделі у формальну структуру, яка на етапі реалізації (Implementation) втілюється у програмний код мовою онтологій OWL, що в подальшому дозволяє здійснювати безперервний супровід (Maintenance) для коригування та оновлення системи. Життєвий цикл онтології також включає паралельні підтримуючі процеси, зокрема інтеграцію вже існуючих моделей для економії ресурсів та етап оцінювання (Evaluation), в межах якого проводиться комплексний технічний аудит моделі для підтвердження її якості та цілісності.

Комплексна інженерна стратегія розробки онтологій On-To-Knowledge (ОТКМ) була розроблена в однойменному європейському проєкті та орієнтована на побудову систем управління знаннями (Knowledge Management Systems) в межах організацій (Castillo-Barrera, F. Et al. 2013). Методологія акцентує увагу на прагматичному аспекті використання знань, де ключовим критерієм успішності є здатність моделі вирішувати конкретні прикладні запити користувачів у межах корпоративної структури. Методологія пропонує п'ять послідовних етапів, які можуть повторюватися в ітераційному режимі.

Техніко-економічне обґрунтування (Feasibility Study). На цьому етапі проводиться аналіз проблемних зон організації, сценарії використання (Use Cases), цільова аудиторія. Проводиться оцінка економічної та технічної доцільності розробки онтологічного рішення.

Підготовчий етап (Kick-off) включає формування специфікації вимог до онтології (ORSD). Проводиться визначення предметної області та рівня деталізації (granularity). Визначаються Компетентнісні запитання (Competency Questions) як перелік запитань природною мовою, відповіді на які мають отримати користувачі системи.

Етап 1. Удосконалення та розробка (Refinement) включає вилучення знань (Knowledge Elicitation) через збір термінів та концептів із документів та через інтерв'ю з експертами. Концептуалізація включає побудову таксономії (ієрархії класів), визначення атрибутів та зв'язків (object/data properties). Формалізація включає опис побудованої моделі мовами представлення знань (переважно OWL 2 або RDF/S).

Етап 2. Оцінювання (Evaluation) включає верифікацію, тобто перевірку логічної несуперечливості онтології, та валідацію, тобто перевірку можливості отримати відповіді на компетентнісні запитання, сформульовані на етапі Kick-off. Аналіз відповідності вимогам кінцевих користувачів.

Експлуатація та супровід (Maintenance). Процеси оновлення онтології при зміні знань у предметній області. Управління версіями та інтеграція нових концептів без порушення існуючих логічних зв'язків.

Методологія DILIGENT (*Distributed, Loosely-controlled and evolVnG Engineering of oNTologies*) це ще одна фундаментальна стратегія в інженерії знань, яка, на відміну від ОТКМ, фокусується на децентралізованій та розподіленій розробці онтологій (Pinto, H., Staab, S., & Tempich, C. 2004). DILIGENT підтримує децентралізовані робочі процеси, залучення експертів у предметній області та управління еволюцією онтологій.

Життєвий цикл DILIGENT складається з п'яти ітераційних кроків:

Етап перший Побудова (Build) на якому створюється базова "ядерна" онтологія (core ontology).

Другий етап Локальна адаптація (Local Adaptation). Адаптація базової "ядерної" онтології (core ontology) до прикладних задачах через створення "локальних версій".

Аналіз (Analysis). Проводиться збір та аналіз проведених локальних змін для виявлення патернів, конфліктів, корисних доповнень, які з'явилися в процесі реальної експлуатації.

Ревізія (Revision). На основі аналізу приймається рішення про оновлення "ядерної" онтології. Конфлікти розв'язуються шляхом консенсусу між учасниками.

Локальне оновлення (Local Update). Нова версія "ядерної" онтології розсилається всім учасникам, які потім інтегрують свої специфічні локальні знання вже в оновлену структуру.

На сучасному етапі розвитку інженерії знань спостерігається тенденція до гібридизації методів, де класичні підходи доповнюються інструментами для роботи з великими даними та хмарними сервісами спільної розробки.

Середовище розробки онтологій — це програмний комплекс, що забезпечує автоматизацію етапів побудови, редагування, візуалізації та логічної перевірки онтологічних моделей.

Сучасні середовища розробки онтологій (Ontology Development Environments) класифікуються за рівнем підтримки стандартів (OWL, RDF), можливостями автоматичного логічного виведення (Reasoning) та підтримкою колективної роботи.

Середовище Protégé, розроблене Стенфордським центром біомедичних досліджень (Stanford BMIR), де-факто є світовим стандартом в інженерії знань. Protégé базується на відкритій архітектурі та забезпечує повну підтримку рекомендацій консорціуму W3C, зокрема мов OWL 2 та RDF(S). Створення онтологій у Protégé (Liu S. et al.,2020) дозволяє розробнику оперувати такими

поняттями, як класи (Classes), властивості об'єктів (Object Properties), властивості типу даних (Data Properties) та екземпляри (Individuals). Protégé надає інтерфейс прикладних програм (API), а також інструменти для графічних інтерфейсів користувача (GUI). Користувачі також можуть створювати онтології за допомогою програмування на Java, використовуючи API. Protégé підтримує множинне успадкування та нові перевірки узгодженості даних, а також має надійну масштабованість.

Завдяки модульній системі плагінів, Protégé дозволяє розширювати функціонал для візуалізації (OntoGraf), імпорту структурованих даних (Cellfie для Excel) та написання правил (SWRL). Важливою особливістю є інтеграція з машинами логічного виведення (Reasoners), такими як HermiT та Pellet, що дозволяє в автоматичному режимі перевіряти онтологію на несуперечливість (consistency check) та виводити приховані зв'язки між концептами (Flügel, S., Glauber, M., Neuhaus, F., & Hastings, J. 2025). OntoGraf – це плагін для Protégé, інструменту візуалізації онтологій (Sen, C., Summers, J., & Mosko, G. 2013). OntoGraf представляє Клас та Індивід як вузли, а їхній зв'язок виражається як ребра. OntoGraf забезпечує автоматичні макети, такі як Сітка, Радіальна, Пружинна та Дерево. Крім того, користувач може довільно розташовувати всі вузли. Вузли та ребра можна фільтрувати, щоб візуалізувати лише ті типи ребер та вузли, які потрібні користувачеві. Крім того, вузли OntoGraf можна зробити невидимими. Нарешті, користувач може зберегти графік у форматі растрового зображення або зберегти графік у файлі та завантажити його пізніше. Інструмент візуалізації був створений шляхом додавання кількох функцій для ефективної візуалізації доповненої онтології на основі OntoGraf (Alkhamash, E. 2022).

WebProtégé є хмарною реалізацією середовища Protégé, що спеціально оптимізована для групової роботи та розподіленого збору знань. Система надає розвинені інструменти взаємодії в командах, механізми обговорення конкретних концептів, коментування та багаторівневе відстеження змін. WebProtégé є оптимальним інструментом для реалізації методології DILIGENT, оскільки

забезпечує необхідну прозорість та консенсус між територіально розподіленими експертами.

TopBraid Composer — це професійне комерційне середовище, побудоване на базі IDE Eclipse, орієнтоване на створення масштабних графів знань (Enterprise Knowledge Graphs). Платформа вирізняється потужними інструментами для інтеграції з реляційними базами даних (стандарти RDB2RDF) та підтримкою складних запитів мовою SPARQL. Найкраще корелює з методологією On-To-Knowledge (OTKM), де онтологія розглядається як центральний компонент корпоративної інформаційної інфраструктури.

NeOn Toolkit спеціалізоване середовище, розроблене для підтримки життєвого циклу мережових онтологій (Ontology Networks). Підтримує специфічні операції методології NeOn, такі як реінжиніринг неонтологічних ресурсів, модульна декомпозиція та складне відображення (mapping) між гетерогенними моделями знань.

Fluent Editor представляє альтернативний підхід до інженерії знань, що базується на використанні контрольованої природної мови (Controlled Natural Language, CNL), інструмент автоматично транслює природномовні конструкції у формальний OWL-код, зберігаючи математичну строгість моделі.

Компаративна характеристика середовищ розробки представлена у таблиці 2.3.1.

Таблиця 2.3.1 - Компаративна характеристика середовищ розробки

Параметр	Protégé	WebProtégé	TopBraid	NeOn Toolkit	Fluent Editor
Доступність	Free / Open Source	Cloud / Free	Commercial	Free / Open Source	Commercial / Free trial
Основна мета	Наукові дослідження	Колективна робота	Enterprise-рішення	Мережі онтологій	Моделювання на CNL
Логічний вивід	Повний (Reasoners)	Базовий	Повний (SPARQL/SHACL)	Високий	Середній
Складність	Середня	Низька	Висока	Висока	Мінімальна

Методика побудови онтології засобами Protégé запропонована в роботі Литвин та ін. (2009) складається з таких кроків:

Крок 1. Визначення галузі й масштабу онтології. Розробка питань компетентності для перевірки відповідності онтології заданій предметній галузі. Питання компетентності дозволяють оцінити повноту поданої інформації і рівень її деталізації.

Крок 2. Можливість використання наявних онтологій.

Крок 3. Визначення термінів і їх властивостей, які несуть основну інформацію про задану предметну галузь.

Крок 4. Визначення класів і їх ієрархії. Є декілька підходів для побудови ієрархії класів: зверху-вниз, знизу-вгору та комбінований процес.

Крок 5. Визначення властивостей класів. Після визначення певної кількості класів необхідно описати внутрішню структуру понять. Під час кроку 3 були вибрані класи зі створеного списку термінів. Більшість термінів, які залишаться, ймовірно, будуть властивостями цих класів. Усі підкласи класу успадковують властивість цього класу.

Крок 6. Визначення фацетів властивостей. Властивості можуть мати різні фацети, що описують тип і коефіцієнт (потужність) значення властивості, діапазон та інші характеристики, які вона може мати.

Крок 7. Створення екземплярів. Останній крок – це створення окремих екземплярів класів в ієрархії. Для визначення окремого екземпляра необхідно:

- 1) вибрати клас;
- 2) створити окремий екземпляр цього класу;
- 3) увести значення слотів.

Для формалізації відношень між сутностями використовуються об'єктні властивості (Object Properties), що дозволяють вийти за межі простої таксономії:

- Part-of (Частина-чогось): використовується для опису структури ґрунту (компоненти clay, sand як частина mechanical\_composition) або технологічних карт.

- **Has-trait (Має-ознаку):** пов'язує *genotype* з його фенотипічними характеристиками або якісними показниками насіння.
- **Observed-in (Спостерігається-в-ньому):** застосовується для фіксації появи шкідників (*Pests*) або хвороб (*Diseases*) у конкретних агроекологічних умовах.
- **Measured-as (Вимірюється-як):** визначає метрики для кількісних показників, таких як концентрація NPK у ґрунті або параметри сонячної радіації.

Для забезпечення валідності бази знань в онтологію впроваджуються логічні обмеження для забезпечення несуперечності у вигляді аксіом:

- **Domain & Range:** чітко визначають типи суб'єктів та об'єктів для кожної властивості (наприклад, властивість «має рівень гумусу» може мати доменом лише клас *soil*).
- **Disjointness (Неперетинність):** гарантує, що об'єкт не може одночасно належати до взаємовиключних класів (наприклад, сорт не може бути гібридом).
- **Cardinality (Потужність):** обмежує кількість можливих зв'язків (наприклад, одна ділянка може мати лише один основний тип механічного складу ґрунту).

Механізми логічного виведення (*Reasoning*).

Використання формальної логіки дозволяє автоматизованим ризонерам (*Reasoners*) виявляти приховані зв'язки. Наприклад, на основі даних про *climatic\_condition* та *soil*, система може автоматично класифікувати придатність конкретного *genotype* для певної території або прогнозувати ризик розвитку *Diseases*. Це перетворює онтологію зі статичного сховища даних на динамічне середовище генерації нових знань для підтримки прийняття рішень.

## 2.5 Метрики оцінки якості онтології

Модель, що використовуватиметься для визначення метрик, базується на праці Tartir, S. et al. (2005). Вона формально визначає структури схеми та бази знань (БЗ). Ця модель буде застосована для визначення метрик.

Структура онтології— це кортеж із шести елементів

$$O = (C, P, A, H_c, \text{prop}, \text{att}),$$

що складається з двох неперетинних множин  $C$  та  $P$ , елементи яких називаються концептами та відношеннями відповідно; ієрархії концептів  $H_c$ , де  $H_c$  є орієнтованим транзитивним відношенням  $H_c \subseteq C \times C$ , яке також називається таксономією концептів. Запис  $H_c(C_1, C_2)$  означає, що  $C_1$  є підконцептом  $C_2$ ; функції  $\text{prop}: P \rightarrow C \times C$  або  $P(C_1, C_2)$ , яка пов'язує концепти нетаксономічно (функція  $\text{dom}: P \rightarrow C$ , де  $\text{dom}(P) = \prod_1(\text{rel}(P))$ , визначає область визначення (домен);  $\text{range}: P \rightarrow C$ , де  $\text{range}(P) = \prod_2(\text{rel}(P))$ , визначає область її значень (діапазон). Специфічним видом відношень є атрибути  $A$ . Функція  $\text{att}: A \rightarrow C$  пов'язує концепти з літеральними значеннями  $\text{rang}(A) = \text{string}$ .

Системний підхід до верифікації онтологій виконує роль регулятора якості на всіх етапах функціонування бази знань. Зокрема, метричний аналіз забезпечує превентивне виявлення архітектурних ризиків, підвищує рівень довіри користувача до результатів роботи системи та дозволяє здійснювати селекцію моделей на основі кількісних критеріїв придатності, мінімізуючи суб'єктивізм при виборі цільової онтології.

Якість онтологій оцінюється по трьом напрямкам: схема, база знань та метрики класів. Застосування метричного підходу забезпечує об'єктивний аналіз специфічних характеристик моделі, оскільки її архітектура та семантична топологія безпосередньо детермінуються особливостями конкретної предметної області. У Tartir S. зі співавторами (2005) розглядаються дві пов'язані категорії:

метрики схеми та метрики екземплярів. Перша категорія оцінює дизайн онтології та його потенціал для багатого представлення знань. Друга категорія оцінює розміщення даних екземплярів в онтології та ефективне використання онтології для представлення знань, змодельованих в онтології.

Метрики схеми стосуються дизайну онтології та вказують на багатство, ширину, глибину та успадкування схеми онтології.

Метрика багатства зв'язків (Relationship Richness, RR) є інтегральним показником, що відображає семантичну різноманітність та топологічну структуру онтології. Ця метрика дозволяє оцінити ступінь використання нетаксономічних відношень порівняно із загальною кількістю зв'язків у моделі. Математично багатство зв'язків схеми визначається як відношення потужності множини нетаксономічних відношень  $|P|$  до сукупної кількості всіх зв'язків, включаючи зв'язки успадкування (ієрархію підкласів  $|SC|$ ):

$$RR = \frac{|P|}{|SC| + |P|}.$$

Результатом обчислення є нормований коефіцієнт у діапазоні  $[0, 1]$ , який характеризує тип та архітектурну спрямованість онтології. Значення RR близьке до 0 свідчить про те, що модель є переважно таксономічною. У такій системі домінують зв'язки типу «клас-підклас» (IS-A), що характерно для базових класифікаторів або словників з низьким рівнем функціональної взаємодії між концептами. Значення RR близьке до 1 вказує на високу семантичну насиченість онтології. Це означає, що більшість зв'язків описують специфічні для предметної області залежності (наприклад, функціональні, причинно-наслідкові або просторові відношення), що суттєво розширює когнітивні можливості системи та точність логічного виведення.

Метрика атрибутивного багатства (Attribute Richness, AR) визначається середньою кількістю дефініцій атрибутів (слотів) у розрахунку на один клас. Цей

показник є індикатором як архітектурної якості проектування онтології, так і потенційної інформаційної ємності описів екземплярів. Висока концентрація атрибутів свідчить про значний рівень деталізації концептуальної схеми та підвищену семантичну виразність моделі знань.

Математична формалізація показника атрибутивного багатства AR базується на розрахунку математичного очікування кількості дефініцій атрибутів (слотів), що припадають на один концепт системи. Розрахунок індикатора здійснюється як відношення сукупної кількості атрибутів усіх класів  $|att|$  до потужності множини класів  $|C|$ :

$$AR = \frac{|att|}{|C|}.$$

Отримане значення є кількісною мірою семантичної деталізації концептуальної схеми. Високе значення AR свідчить про високу когнітивну точність моделі. Це вказує на те, що архітектура онтології передбачає глибоку формалізацію властивостей кожного об'єкта, забезпечуючи високу інформаційну насиченість бази знань. Низьке значення AR може бути ознакою високого рівня абстракції моделі або недостатньої деталізації властивостей концептів, що обмежує можливості системи щодо точного представлення даних на рівні екземплярів.

Метрика багатства успадкування (Inheritance Richness,  $IR_S$ ) характеризує архітектурну глибину онтології та особливості ієрархічного розподілу знань. Вона є ключовим показником ступеня деталізації категорій та підкатегорій, дозволяючи диференціювати моделі за типом їхньої вертикальної чи горизонтальної орієнтації.

Математично багатство успадкування схеми  $IR_S$  визначається як середнє число кількості підкласів, що припадають на один клас. Ця метрика обчислюється як відношення суми всіх прямих нащадків (підкласів) для кожного

класу до загальної потужності множини класів

$$IR_s = \frac{\sum_{i=1}^{|C|} |H^C(C_1, C_i)|}{|C|},$$

де  $|H^C(C_1, C_i)|$  — кількість безпосередніх підкласів  $C_1$  (direct subclasses) для класу  $C_i$ .

Показник  $IR_s$  дозволяє класифікувати онтології за типом їхньої когнітивної структури. Горизонтальна (планарна) онтологія має незначну кількість рівнів ієрархії при великій кількості підкласів у кожного батьківського концепту. Це свідчить про широке охоплення предметної області при низькому рівні специфікації окремих вузлів. Характеризується високим значенням  $IR_s$ . Вертикальна (глибинна) онтологія має розвинену систему вкладених рівнів успадкування з невеликою кількістю прямих нащадків у кожного класу. Така структура вказує на високу ступінь спеціалізації знань та детальну формалізацію ієрархічних зв'язків. Характеризується низьким значенням  $IR_s$ .

Метрики екземплярів відображають спосіб розміщення даних в онтології. Розміщення даних екземплярів та розподіл даних можуть вказувати на ефективність дизайну онтології та обсяг знань, представлених онтологією. Метрики екземплярів групуються у дві категорії: метрики бази знань, які описують базу знань в цілому, та метрики класів, які описують спосіб використання кожного класу, визначеного в схемі, в базі знань.

Розглянемо метрики бази знань. Метрика багатство класів відображає як екземпляри розподіляються між класами. Кількість класів, які мають екземпляри, порівнюється із загальною кількістю класів, що дає загальне уявлення про те, скільки екземплярів пов'язані з класами, визначеними в схемі. Формально, багатство класів  $CR$  бази знань визначається як відношення кількості класів, що використовуються в базі  $|C'|$ , до кількості класів, визначених в онтологічній схемі  $|C|$ .

$$CR = \frac{|C'|}{|C|}.$$

Результатом буде відсоток, що вказує на те, наскільки багата класами база знань. Таким чином, якщо база знань має дуже низький CR, то база знань не має даних, які б ілюстрували всі знання в схемі. З іншого боку, база знань, яка має дуже високий CR (близько 100%), означатиме, що дані в бази знань представляють більшу частину знань у схемі.

Метрика середня сукупність (Average Population, P), середній розподіл екземплярів по всіх класах, є показником кількості екземплярів порівняно з кількістю класів. Він може бути корисним, якщо розробник онтології не впевнений, чи було вилучено достатньо екземплярів порівняно з кількістю класів.

Формально, середня популяція P класів у БЗ визначається як кількість екземплярів бази знань |I|, поділена на кількість класів, визначених в онтологічній схемі |C|.

$$P = \frac{|I|}{|C|}.$$

Результат буде дійсним числом, яке показує, наскільки добре виконано процес вилучення даних для заповнення бази знань. Наприклад, якщо середня кількість екземплярів на клас низька, то разом із попередньою метрикою це число означатиме, що екземплярів, вилучених у бази знань, може бути недостатньо для представлення всіх знань у схемі. Майте на увазі, що деякі класи схеми можуть мати дуже низьку або дуже високу кількість залежно від характеру того, що вони представляють.

Метрика зв'язності (Cohesion, Coh) є кількістю окремих пов'язаних компонентів у екземплярах, якщо екземпляри та зв'язки між ними розглядати як

граф, де вузли представляють екземпляри, а ребра представляють зв'язки між ними. Метрика зв'язності  $\text{Coh}$  бази знань визначається як кількість окремих зв'язних компонентів  $|\text{SCC}|$  графа, що представляє базу знань.

$$\text{Coh} = |\text{SCC}|.$$

Результатом буде ціле число, що представляє кількість окремих компонентів. Наприклад, більш корисну пропускну здатність алгоритмів семантичного асоціативного виведення можна очікувати від онтології з  $\text{Coh} = 1$ , де всі дані в базі знань пов'язані.

Розглянемо метрики класу.

Метрика важливості класу (Class Importance,  $\text{Imp}$ ) дозволяє оцінити ступінь концентрації екземплярів у межах конкретної гілки онтології. Цей показник характеризує розподіл фактичних даних за класами та допомагає ідентифікувати «фокальні області» (focal areas) моделі — сегменти, які є найбільш затребуваними або детально описаними в поточному контексті експлуатації системи. Формально важливість класу  $C_i$  визначається як відношення кількості екземплярів, що належать до піддерева з коренем у  $C_i$  (включаючи екземпляри самого класу та всіх його нащадків), до загальної потужності множини екземплярів у базі знань  $|I|$ :

$$\text{Imp} = \frac{|C_i|}{|I|},$$

де  $|C_i|$  — сукупна кількість екземплярів у піддереві концепту  $C_i$ ;

$|I|$  — загальна кількість екземплярів в у всій базі знань.

Високе значення  $\text{Imp}$  вказує на те, що дана частина онтології є центральною для поточної предметної області. Суттєві розбіжності у показниках

важливості між різними гілками можуть свідчити про «перекоси» у зборі даних або про наявність надлишкових сегментів схеми, які майже не використовуються на практиці. Аналіз  $Imp$  дозволяє інформувати користувача про те, наскільки база знань готова до вирішення конкретних прикладних задач. Якщо цільовий клас має низький  $Imp$ , система може потребувати додаткового етапу вилучення та інтеграції даних.

Метрика повноти (Fullness,  $F$ ) детально описує метрику середньої популяції бази знань, згадану вище. Повнота ( $F$ ) класу  $C_i$  визначається як фактична кількість екземплярів, що належать до піддерева з коренем у  $C_i(C_i(I))$  порівняно з очікуваною кількістю екземплярів, що належать до піддерева з коренем у  $C_i(C_i'(I))$ .

$$F = \frac{|C_i(I)|}{C_i'(I)}.$$

Результатом формули буде відсоток, що представляє фактичне покриття екземплярів порівняно з очікуваним покриттям. У більшості випадків цей показник є показником того, наскільки добре виконався процес вилучення екземплярів. Наприклад, база знань, де більшість класів мають низький  $F$ , вимагатиме більшого вилучення даних. З іншого боку, база знань, де більшість класів майже заповнені, вказуватиме на те, що вона точніше відображає знання, закодовані в схемі.

Метрика багатство успадкування (Inheritance Richness,  $IR_C$ ) дозволяє оцінити інтенсивність розгалуження та ступінь деталізації знань у межах конкретного сегмента онтології. На відміну від загального показника  $IR_S$ , ця метрика фокусується на локальній топології піддерева. Нехай  $C'$  — множина всіх концептів (вузлів), що входять до піддерева з коренем у класі  $C_i$ . Тоді багатство успадкування класу  $C_i$  визначається як середня кількість підкласів, що припадає на один вузол даного піддерева

$$IR_c = \frac{\sum_{i=1}^{|C'|} |H^C(C_1, C_i)|}{|C'|},$$

де  $|H^C(C_1, C_i)|$  — кількість безпосередніх підкласів  $C_i$  (direct subclasses) для класу  $C_1$ .

Високе значення  $IR_c$  при низькій загальній глибині піддерева вказує на «планарну» спеціалізацію (широкий перелік об'єктів одного рівня). Низьке значення  $IR_c$  зазвичай притаманне глибоким ієрархічним структурам з послідовним уточненням властивостей. Значні аномалії в  $IR_c$  для окремих класів можуть сигналізувати про нерівномірну деталізацію онтології. Це критично для роботи Reasoners, оскільки вузли з надмірним розгалуженням або надлишковою глибиною можуть експоненціально збільшувати час логічного виведення.

Метрика багатство зв'язків (Relationship Richness,  $RR_c$  відображає, скільки властивостей у кожному класі в схемі фактично використовується на рівні екземплярів. Формально, багатство зв'язків класу  $C_i$  визначається як відношення кількості відношень, що фактично використовуються екземплярами  $I_i$  (які належать до класу  $C_i$ ), до загальної кількості відношень, задекларованих для цього класу в схемі:

$$RR_c = \frac{|P(I_i, I_j): I_i \in C_i (I)|}{|P(C_i, C_j)|}.$$

Високе значення  $RR_c$  свідчить про те, що база знань повноцінно використовує семантичний потенціал схеми. Низьке значення  $RR_c$  вказує на «порожню» або надлишкову схему, де задекларовані зв'язки не підкріплені фактичними даними, що може свідчити про недосконалість алгоритмів вилучення даних.

Метрика зв'язності (Connectivity,  $C_n$ ) призначена для оцінки інтенсивності

взаємодії екземплярів певного класу з іншими об'єктами бази знань. Вона працює в тандемі з метрикою важливості  $Imp$ , дозволяючи виділити архітектурні центри системи. Зв'язність класу  $C_n$  визначається як кількість індивідів інших класів  $I_j$ , з якими пов'язані індивіди даного класу через об'єктні властивості:

$$C_n = |I_j, P(I_i, I_j) \wedge I_i \in C_i(I)|.$$

Результатом є ціле число, що відображає «популярність» або центральність класу. Класи з високим показником  $C_n$  відіграють роль семантичних хабів (центральної вузлів), через які проходить найбільша кількість інформаційних потоків. Це допомагає розробнику зрозуміти топологічну структуру онтології та виділити найбільш значущі концептуальні області.

Метрика читабельності (Readability,  $Rd$ ) визначає ступінь насиченості онтологічної схеми когнітивно-орієнтованими дескрипторами (мітками, анотаціями, коментарями), що призначені для інтерпретації людиною-оператором. Даний показник є критичним для інтелектуальних систем, які передбачають високий рівень інтерактивності, візуалізацію результатів семантичного пошуку або верифікацію логічних висновків кінцевим користувачем. Читабельність класу  $C_i$  визначається як сума атрибутів анотацій, таких як коментарі (`rdfs:comment`) та мітки (`rdfs:label`), закріплених за класом:

$$Rd(C_i) = |\{a \in A(C_i) | type(a) = rdfs:comment\}| + |\{a \in A(C_i) | type(a) = rdfs:label\}|.$$

Високий рівень  $Rd$  гарантує, що результати логічного виведення або пошуку в онтології будуть інтерпретовані користувачем коректно.

## 2.6 Висновки до розділу 2

У другому розділі представлено теоретико-методологічне обґрунтування побудови цифрової моделі рослини соняшника, основою якої є взаємодія «генотип — середовище». Основним інструментом формалізації знань та побудови логічного каркаса цифрової моделі обрано предметну онтологію. Застосування онтологічного моделювання та апарату семантичних правил дозволяє побудувати структуровану систему знань, здатну до автоматизованого логічного виведення на рівні екземплярів і виявлення прихованих закономірностей.

1. Удосконалено структуру та архітектурні етапи побудови цифрової моделі рослини соняшника з урахуванням взаємодії «генотип — середовище» шляхом розробки послідовної чотирирівневої системи інтеграції різномірних даних (Digital phenotyping system), яка охоплює процеси апаратної реєстрації, FAIR-сумісного збереження, інтелектуальної обробки та предиктивного моделювання.

Структура моделі поєднує підсистеми збору візуальних, біохімічних та фізичних характеристик насіння й кошика з параметричними матрицями стану ґрунту, кліматичних умов та агротехнічних заходів. При цьому для кожної з шести виділених підсистем чітко визначено й типізовано джерела отримання даних, а також методи їхнього агрегування й довготривалого структурованого зберігання. Запропонована етапність трансформації інформації забезпечує послідовний перехід від реєстрації необроблених мультисенсорних даних до генерації формалізованого цифрового образу фенотипу рослини, що дозволяє кількісно оцінювати адаптивну пластичність та прогнозувати динаміку розвитку культури в мінливих умовах агроценозу.

2. За результатами аналізу структури та етапів формування цифрової моделі рослини, для перетворення розрізнених результатів по фенотипуванню у структуровані знання, обґрунтовано необхідність застосування онтологічного підходу. Доведено, що впровадження онтологічного підходу є системною

умовою для перетворення декомпонованих результатів по фенотипуванню, отриманих з різних джерел, у формалізоване середовище знань, яке формалізує взаємодію «генотип — середовище». Визначено концептуальний базис та критерії метричного аналізу онтологій, що створює формалізоване підґрунтя для забезпечення семантичної сумісності, дотримання принципів FAIR та математико-логічного моделювання досліджуваних процесів у межах парадигми «Селекції 4.0».

Основні наукові результати розділу опубліковано в працях автора (Ведмедєв та Терещенко, 2022, 2025).

## 3 ОНТОЛОГІЧНЕ МОДЕЛЮВАННЯ ВИРОЩУВАННЯ СОНЯШНИКА

### 3.1 Формалізація вимог до онтологічної моделі та детермінація інформаційних потреб користувача

Ціллю створення онтології є структуроване представлення знань предметної галузі вирощування соняшнику в Україні, в якій міститься багатовимірна інформація про дослідження в аграрних науках національних експертів із селекції, генетики та агротехнології вирощування соняшника, а також враховано вимоги державних стандартів України. Практична цінність, яка розглядається при створенні онтології, полягає у розширенні можливостей створення наборів даних, отримання нових знань, швидкого консультування практиків та інтеграції у світову наукову спільноту. За наявності великих колекцій наборів даних генобанків, зібраних на національному та глобальному рівнях, та їх сільськогосподарської цінності існує проблема систематичного застосування генетичних ресурсів через відсутність фенотипної інформації про окремі зразки, яка необхідна для оцінки мінливості на рівні ознаки. Саме ця інформація є міститься в базах даних ІОК НААН (Aliiev E., Vedmedeva K., Machova T., Vedmedev S. 2024).

У межах розробки цифрової моделі соняшника предметну онтологію формалізовано як впорядковану п'ятірку (кортеж) виду:

$$O = (C, R, I, A, SWRL) \quad (3.1)$$

де:

- $C$  (*Concepts*) — множина класів предметної області соняшника;
- $R$  (*Relations*) — множина об'єктних та числових властивостей, що визначають зв'язки та метрики;
- $I$  (*Individuals*) — множина екземплярів або індивідів (ABox), що формують цифрові паспорти конкретних полів та зразків;
- $A$  (*Axioms*) — множина аксіом обмежувального характеру (еквівалентність, диз'юнктивність класів);

- SWRL — підмножина семантичних правил логічного виведення, що детермінують сценарну логіку прийняття рішень та моделювання взаємодії "генотип — середовище".

З позицій теорії описових логік (Description Logics) формалізований кортеж предметної онтології соняшника (3.1) структурно розділяється на два взаємопов'язані компоненти — схему знань  $TBox$  (Terminological Box) та простір фактів  $ABox$  (Assertion Box), що детермінуються як:

$$TBox = \langle C, R_{schema}, A \rangle,$$

$$ABox = \langle I, R_{facts} \rangle,$$

де

$TBox$  (Термінологічний бокс) — визначає концептуальні знання предметної області, поєднує впорядковану множину концептів  $C$ , задекларовані для них правила зв'язків  $R_{schema}$  та загальносистемні обмежувальні аксіоми  $A$ .  $TBox$  формує жорстку метамодель системи, яка є незмінною при масштабуванні бази знань.

$ABox$  (бокс тверджень) — визначає екстенціональні знання (конкретні факти та емпіричні дані). Він акумулює в собі множину індивідів  $I$  та сукупність реальних предикатних зв'язків і числових значень між ними  $R_{facts}$ .  $ABox$  виступає простором цифрових паспортів генотипів, експериментальних зразків рослин та конкретних полів.

Підмножина правил SWRL виконує роль динамічного інтерфейсу логічного виведення, який на основі структурних правил  $TBox$  аналізує первинні факти  $ABox$  та генерує нові знання у вигляді додаткових семантичних зв'язків.

Подальше викладення базується на методиці побудови онтології засобами Protégé (Литвин та ін., 2009), яка розглянута у 2.3.

Перший крок по визначенню галузі й масштабу онтології та розробці

питань компетентності проводилося із залученням експертів Інституту олійних культур НААН України, що є пріоритетним, порівняно з автоматичним методом побудови онтології згідно (Осієвський С.В., Третяк В.Ф. 2022). Із залученням експертів Інституту олійних культур НААН України було проведено перевірку можливості використання наявних онтологій (Крок 2). За результатами аналізу існуючих онтологій, представлених в п.1.2., було прийнято рішення щодо побудови власної онтологічної моделі, адаптованої до запитів вітчизняного агротехнологічного сектору. Вибір на користь розробки оригінальної онтології зумовлений необхідністю первинного впорядкування та формалізації унікальних знань вітчизняних фахівців, зокрема експертів Інституту олійних культур НААН України, чий досвід відображає специфіку вирощування соняшнику в конкретних агрокліматичних зонах України. Існуючі світові онтології, попри їхню масштабність, часто мають занадто узагальнену структуру, яка не дозволяє повною мірою інтегрувати локальні технологічні регламенти, специфіку місцевого генофонду та особливості біологічної стійкості гібридів до регіональних рас патогенів. Таким чином, створення власної моделі є необхідним кроком для забезпечення автентичності бази знань та її адаптації до практичних потреб українського агросектору, що неможливо реалізувати шляхом простої імплементації наявних глобальних рішень.

Експертами Інституту олійних культур НААН України були визначені терміни та їхні властивості, які несуть основну інформацію про задану предметну галузь (Крок 3).

Експертами Інституту олійних культур НААН України були визначені цільові показники та фактори впливу на продуктивність соняшнику (рис.3.1.1). Формування врожайності (продуктивності) соняшнику обумовлене функціонуванням складної багатофакторної системи, де результативність залежить від інтегрального впливу чинників упродовж календарного року та їхнього динамічного розподілу за фазами вегетаційного періоду. Ключовими детермінантами цієї системи виступають якісні характеристики генотипу (сорту

чи гібрида) та його адаптивна здатність до конкретних умов навколишнього середовища. Поряд із біологічним потенціалом рослини, критичне значення мають ґрунтово-кліматичні умови, рівень фітосанітарного тиску з боку хвороб і шкідників, а також застосовані агротехнології та ефективність інструментального контролю за ключовими параметрами середовища на кожному етапі органогенезу.

На продуктивність соняшнику впливає система факторів за рік та вегетаційний період та розподіл факторів за періодом вегетації та календарного року, а також якості генотипу рослини соняшнику (сорт, гібрид) та пристосованість до умов навколишнього середовища, а також ґрунтові умови, кліматичні умови, хвороби, шкідники, технології, контроль факторів.

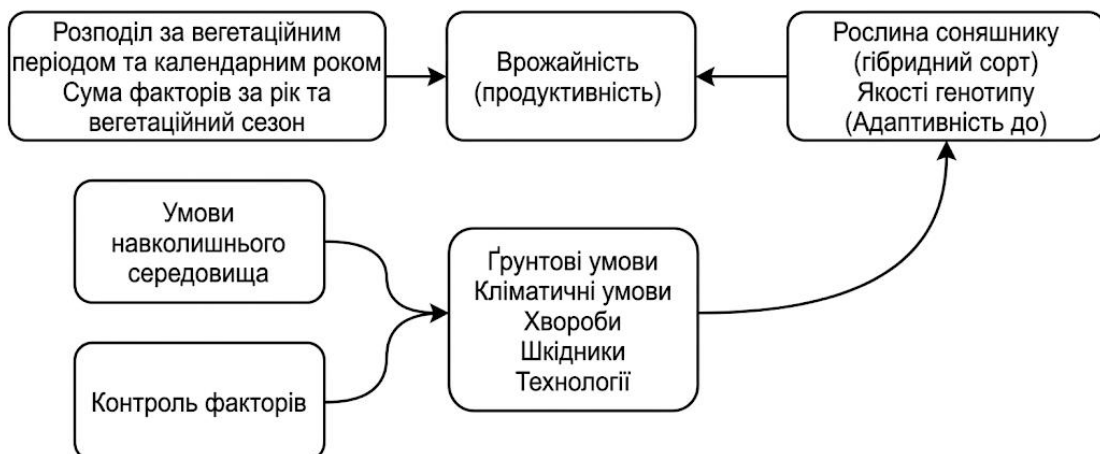


Рисунок 3.1.1 - Цільові показники та фактори впливу на продуктивність соняшнику (розроблено автором)

До складу онтологічної моделі мають бути включені визначені в Розділі 2 складові цифрової моделі фенотипу рослини, параметри ґрунтово-кліматичних умов, ентомологічні та фітопатологічні чинники, а також регламенти застосування пестицидів і добрив. Інформаційну повноту термінологічного словника (табл. 3.1.1) забезпечено залученням нормативної документації та

галузових реєстрів засобів захисту рослин (Міністерство захисту довкілля та природних ресурсів України, 2023). Таблиця 3.1.1 містить структуровану класифікацію ключових факторів впливу на агротехнологічний цикл соняшнику, розподілених за функціональними категоріями.

Таблиця 3.2.1. - Систематизація факторів впливу на продуктивність соняшнику (сформовано автором за даними Інституту олійних культур НААН України)

Ґрунтові умови	Кліматичні умови	Хвороби	Шкідники	Технологія
Механічний склад: пісок, глина, гумус.	Температурний режим	Несправжня борошниста роса (P. halstedii)	Вусач соняшниковий (Agapanthia dahli)	Сівозміна
Хімічний склад: макро- (NPK) та мікроелементи (B, Zn, Mg, Se).	Опади (кількість та інтенсивність)	Склеротиніоз (біла гниль) (S. sclerotiorum)	Дротяники (Athous niger, Agriotes sputator)	Типи технології: класична, no-till, інтенсивна.
Вологоємність: агрегатний стан, колоїдність.	Сонячна радіація	Фомопсис та Фомоз (D. helianthi, P. macdonaldii)	Мідляк піщаний (Opatrum sabulosum)	Гербицидний захист: Clearfield, Express.
Біота: бактерії, черв'яки, симбіонти.	Вітер: швидкість та напрямок	Іржа соняшнику (Puccinia heliantha)	Шипоноска (Mordellistena parvula)	Зрошення (наявність та тип)
Глибина профілю: 1.0 – 1.5 м.	Географія: висота над рівнем моря, широта.	Вугільна та Ризопусна гнилі	Луговий метелик та Сонцевик	Технічне оснащення: наявне обладнання.
		Альтернاریоз та Фузаріоз	Мухи-строкатокрилки	
		Бактеріальні та вірусні хвороби	Соняшникова вогнівка	
		Вовчок соняшниковий (O. cumana)	Листогризучі совки	

Для формалізації вимог до онтології та визначення її функціональних меж було визначено стратегічні напрями, що представлені в таблиці 3.1.2.

Таблиця 3.1.2 Стратегічні функціональні напрями онтології HELIANTUS (розроблено автором за даними Інституту олійних культур НААН України)

Функціональний напрям	Цільове призначення запитань	Рівень реалізації в HELIANTUS
Глобальний інтеграційний домен управління продуктивністю	Синтез усіх факторів для прийняття стратегічного рішення щодо максимізації врожаю. Цільові функції (оптимізація), сценарії «що-якщо», валідація рекомендацій, оцінка ризиків.	Семантично формалізовано. Забезпечено через дедуктивне виведення знань. Перспективне розширення
Фенотиповий	Ідентифікація зв'язків між морфологією та продуктивністю за допомогою комп'ютерного зору.	Апробовано. Реалізовано через формалізацію цифрових фенотипів. Перспективне розширення
Темпоральний	Моделювання критичних стадій вегетації та прогнозування дат технологічних операцій.	Семантично формалізовано. Забезпечено через дедуктивне виведення знань. Перспективне розширення
Генотипів та біологічних властивостей	Опис внутрішнього потенціалу та генетично детермінованих ознак рослини.	Інтегровано в модель. Перспективне розширення
Екологічний	Забезпечення сталості агроecosистеми та управління сівозміною.	Перспективне розширення
Економічний	Оцінка економічної ефективності агротехнологічних рішень, розрахунок собівартості та максимізація маржинального прибутку.	Перспективне розширення.

Запитання компетентності дозволяють визначити структуру класів представлення інформації з необхідним рівнем її деталізації. В таблиці 3.1.3 представлений фрагмент множини компетентностних запитань глобального інтеграційного домену управління продуктивністю, де зібрано запитання щодо теми «Що мені слід зробити, щоб отримати найкращий урожай соняшнику на моєму полі?».

Таблиця 3.1.3. - Компетентні запитання за темою «Що слід зробити, щоб отримати найкращий урожай соняшнику на цьому полі?» (розроблено автором за даними Інституту олійних культур НААН України)

Компетентнісіні запитання (неформальні)	Компетентнісіні запитання (формальні)
Яким чином можна визначити стан родючості ґрунту для культури соняшнику?	Якщо існує час $t$ ТА існує аналіз ґрунту $y$ (на глибині 1–1,5 м) для будь-якого зразка ґрунту $x$ ТА визначено хімічний склад (NPK або мікроелементи (бор, цинк, магній, селен тощо)) і вологоутримуючу здатність (агрегатний стан, колоїдність) АБО живий компонент ґрунту (бактерії, черв'яки, симбіонти тощо) АБО механічний склад ґрунтів (пісок, глина, гумус тощо), тоді приймається рішення $d$ щодо внесення (ТАК / НІ) та часу $t$
Якими є оптимальні норми та терміни внесення добрив у розрахунку на один гектар посівів соняшнику?	Якщо існує рішення $d$ та визначено добриво $f$ для будь-якого соняшнику $s$ , при цьому наявні результати аналізу ґрунту $y$ (на глибині 1–1,5 м) для кожного зразка $x$ , що містять дані про хімічний склад (NPK або мікроелементи: бор, цинк, магній, селен тощо) та живий компонент ґрунту (бактерії, черв'яки, симбіонти тощо), тоді визначається кількість підстилкового гною, норма NPK або обсяг карбаміду (сечовини).
Чи потрібно поливати соняшник?	Якщо визначено часовий параметр $t$ та наявні результати аналізу ґрунту $y$ (на глибині 1,0–1,5 м), що містять дані про його вологоутримуючу здатність (агрегатний стан, колоїдність) або гранулометричний (механічний) склад ґрунту (вміст піску, глини, гумусу тощо), то здійснюється детермінація вибору оптимальної технології зрошення $w$ .
Яким чином визначаються оптимальні терміни та вибір засобів (препаратів) для захисту посівів соняшнику від ентомологічних шкідників?	Якщо ідентифіковано шкідника $z$ , що належить до виду $x$ , та визначено генотип рослини $G$ , при цьому зафіксовано часовий параметр $t$ (фазу вегетації або календарний термін), то для застосування відповідного пестициду $p$ детермінується технологія агрохімічної обробки
Які існують регламенти та методи терапії фітопатологій (хвороб) соняшнику залежно від стадії вегетації та типу збудника?	Якщо ідентифіковано фітопатологію (хворобу) $w$ та визначено специфічний генотип рослини $G$ , при цьому обрано пестицид $p$ , що є ефективним проти даної хвороби в межах часового регламенту $t$ та відповідної технології агрохімічного захисту, то приймається рішення про використання пестициду $p$ для лікування або купірування хвороби $w$ .
Які існують стратегії та методи контролю популяції вівчак соняшникового ( <i>Orobanche cumana</i> Wallr) у структурі посівів соняшнику?	Якщо наявний паразит <i>Orobanche cumana</i> Wallr (OCW) та генотип $G$ із ознакою стійкості до нього, то обробка не проводиться; АБО якщо наявний паразит <i>Orobanche cumana</i> Wallr, а генотип $G$ має ознаку стійкості до нього та резистентність до гербіциду $g$ у час $t$ , і при цьому передбачена технологія агрохімічної обробки гербіцидом $g$ , то проводиться обробка гербіцидом $g$ ; АБО якщо наявний паразит <i>Orobanche cumana</i> Wallr, генотип $G$ має ознаку стійкості до нього та визначено час $t$ , то технологія сівозміни розглядається як стійка до <i>Orobanche cumana</i> Wallr.

В таблиці 3.1.4. представлено формалізацію інформаційних потреб користувача у формі компетентнісних запитань вибору генотипу (сорт/гібрида) соняшнику.

Таблиця 3.1.4. Компетентні запитання щодо вибору генотипу (сорт/гібрида) соняшнику (розроблено автором за даними Інституту олійних культур НААН України)

Компетентнісні запитання (неформальні)	Компетентнісні запитання (формальні)
Який гібрид найкраще підійде для кліматичних умов мого регіону?	Якщо існує Регіон R із заданою сумою активних температур S та прогнозованою тривалістю вегетаційного періоду, ТА існує Генотип G, що належить до відповідної групи стиглості (ранньостиглий, середньостиглий тощо), тоді приймається рішення d про придатність Генотипу G для вирощування в даній зоні.
Як обрати гібрид залежно від цільового призначення (максимальний збір олії чи валовий урожай насіння)?	Якщо встановлено цільовий пріоритет O (вміст олії) АБО Y (врожайність насіння), та існує Генотип G із зафіксованими генетичними показниками потенційної продуктивності P <sub>o</sub> або P <sub>y</sub> , то здійснюється ранжування та вибір Генотипу G, чії параметри максимізують цільову функцію.
Чи забезпечить обраний гібрид стійкість до місцевих рас вовчка соняшникового?	Якщо ідентифіковано наявність раси r паразита <i>Orobanche cymana</i> Wallr. у ґрунті, та існує Генотип G, що володіє домінантною ознакою стійкості до раси r, тоді формується позитивне рішення d щодо вибору даного генотипу для мінімізації біологічних ризиків.
Який тип гібрида обрати для технологій Clearfield або Express (сумісність із гербіцидами)?	Якщо в технологічній карті закладено використання гербіциду типу H (імідазоліони або трибенурон-метил), та існує Генотип G, що має генетичну резистентність до діючої речовини гербіциду H, тоді детермінується можливість застосування даної технологічної пари «Гібрид — Гербіцид».
Чи здатний даний сорт витримати тривалу засуху в період цвітіння?	Якщо умови навколишнього середовища E характеризуються високим індексом посушливості в часовий інтервал t (фаза органогенезу), та існує Генотип G з високим показником адаптивності та посухостійкості, тоді приймається рішення про доцільність його використання як стресостійкого компонента системи.
Як впливає густина посіву на вибір архітектури гібрида?	Якщо визначено заплановану густоту стояння рослин D на момент збирання, та існує Генотип G з відповідними морфологічними ознаками (нахил кошика, висота стебла, стійкість до вилягання), то здійснюється вибір G, що забезпечує оптимальну площу живлення при заданому D.

В таблиці 3.1.5. представлено компетентнісні запитання щодо цифрового моніторингу та оцінки стану посівів.

Таблиця 3.1.5. Формалізація запитів щодо оцінки стану посівів (розроблено автором за даними Інституту олійних культур НААН України)

Компетентнісні запитання (неформальні)	Компетентнісні запитання (формальні)
Чи відповідає поточний розвиток соняшнику календарним нормам?	Якщо визначено Час $t$ та отримано дані морфометрії (висота, кількість листків) для Генотипу $G$ , ТА існує еталонна крива росту для даної кліматичної зони, то здійснюється детермінація відхилення $\Delta$ та оцінка темпів органогенезу.
Як за допомогою цифрових індексів ідентифікувати дефіцит азоту?	Якщо ідентифіковано спектральний індекс NDVI (або GNDVI) нижче порогового значення для фази $t$ , та виключено вплив хвороб, то формується рішення $d$ про наявність дефіциту азотного живлення та розраховується норма підживлення.
Який прогноз врожаю на основі автоматичного розпізнавання параметрів кошика?	Якщо за допомогою методів комп'ютерного зору розпізнано об'єкт «Кошик соняшнику» та визначено його діаметр $D$ у фазі R5.5 (цвітіння), то за допомогою регресійної моделі в онтології обчислюється прогнозована маса насіння з одиниці площі.
Чи існує загроза вилягання посівів при посиленні вітру?	Якщо встановлено параметри архітекtonіки рослини (висота стебла, індекс міцності) та отримано метеопрогноз щодо швидкості вітру $V$ , то обчислюється коефіцієнт стійкості та приймається рішення про ризики втрати врожаю.
Коли настане оптимальний момент для десикації (підсушування) за даними сенсорів?	Якщо зафіксовано фазу R9 (фізіологічна зрілість) за цифровими ознаками (колір тильного боку кошика) та встановлено вологість насіння $W < 30 \%$ , то детермінується час $t$ для проведення десикації.

Експертами було окремо наголошено на критичній важливості такого фактора, як вчасне виконання агротехнологічних процесів. Важливо підкреслити, що детермінація поняття «вчасно» може здійснюватися за сукупністю різних показників, а саме: температурою ґрунту, рівнем сонячної активності, настанням конкретних фаз вегетації, термінами профілактики ураження шкідниками та хворобами, а також сумарною кількістю опадів за визначений період. Перша група запитань стосується тривалості та настання фаз розвитку соняшнику (Growth Stages). Друга група запитань стосується

технологічних вікон (Operational Windows) та планування польових робіт. Третя група запитань стосується ризиків та навколишнього середовища (Temporal Risks) та дозволяє проаналізувати вплив часу на стійкість системи.

Таблиця 3.1.6 - Формалізація темпоральних запитів (розроблено автором за даними Інституту олійних культур НААН України)

Компетентнісні запитання (неформальні)	Компетентнісні запитання (формальні)
Група 1: Фази розвитку (Growth Stages)	
Коли настає фаза цвітіння для гібрида X у конкретній зоні?	hasStartTime для FloweringPhase відносно Hybrid(X) та Region(Z)
Яка середня тривалість періоду «сходи — бутонізація» для ранньостиглих сортів?	Average(Duration) між Emergence та Budding для EarlyMaturityGenotype
Скільки днів триває наливання насіння при поточній сумі температур?	Duration фази SeedFilling як функція від AccumulatedThermalUnits
Яка дата прогнозованого дозрівання на основі дати посіву?	PredictedMaturityDate = SowingDate + sum PhaseDurations
Група 2: Технологічні вікна (Operational Windows)	
Який оптимальний інтервал для посіву при t ґрунту 10-12°C?	TimeInterval де SoilTemp in [10, 12] для операції Sowing
Коли найефективніше проводити обробку фунгіцидами (хвороба Y)?	OptimalTreatmentTime = InfectionDate + IncubationPeriod - Delta
Скільки триває «вікно» дії гербіциду після опадів?	Duration для EffectivenessWindow після PrecipitationEvent
Який крайній термін десикації для вчасного збирання?	LatestDate для Desiccation відносно PlannedHarvestDate
Група 3: Часові ризики (Temporal Risks)	
Чи були заморозки у період формування кошика за останні 5 років?	Exists(FrostEvent) де Date in HeadDevelopmentPeriod
Як довго тривала посуха під час цвітіння у регіоні Z?	Sum(Duration) для DroughtEvent що перетинає FloweringPhase
Коли активність шкідників досягає порогу шкодочинності?	TimeInterval де PestDensity > EconomicThreshold

На Кроці 1-3 проведено формалізацію предметної області та її термінологічного базису шляхом аналізу експертних знань фахівців Інституту олійних культур НААН України, а також аналізу державних стандартів і технічних каталогів. Сформовано репрезентативний термінологічний апарат, який охоплює екологічні чинники, біологічні властивості генотипів,

агротехнологічні регламенти та параметри фенотипування. Системний підхід дозволив структурувати предметну галузь за функціональними напрямками, включно з глобальним інтеграційним доменом управління продуктивністю, що забезпечує прийняття оптимізаційних рішень. Ключовим результатом цього етапу стала розробка системи компетентнісних запитань у неформальній та формальній версіях, що є підґрунтям для створення такої архітектури онтологічної моделі, яка дозволить отримати необхідні відповіді на ці запитання. Такий підхід забезпечив створення логічного фундаменту для подальшої концептуалізації онтології, трансформуючи агрономічні знання у формалізовані правила та аксіоми для систем підтримки прийняття рішень.

### 3.2 Створення архітектури онтологічної моделі

На наступних кроках побудови онтології за визначеною на попередньому етапі інформації закладаються структури, які забезпечують уніфіковане трактування термінології всіма учасниками роботи з онтологією, а саме експертами, розробниками, користувачами. У редакторі **Protégé** побудовано модель предметної онтології «**HELIANTUS**» (рис.3.2.1).

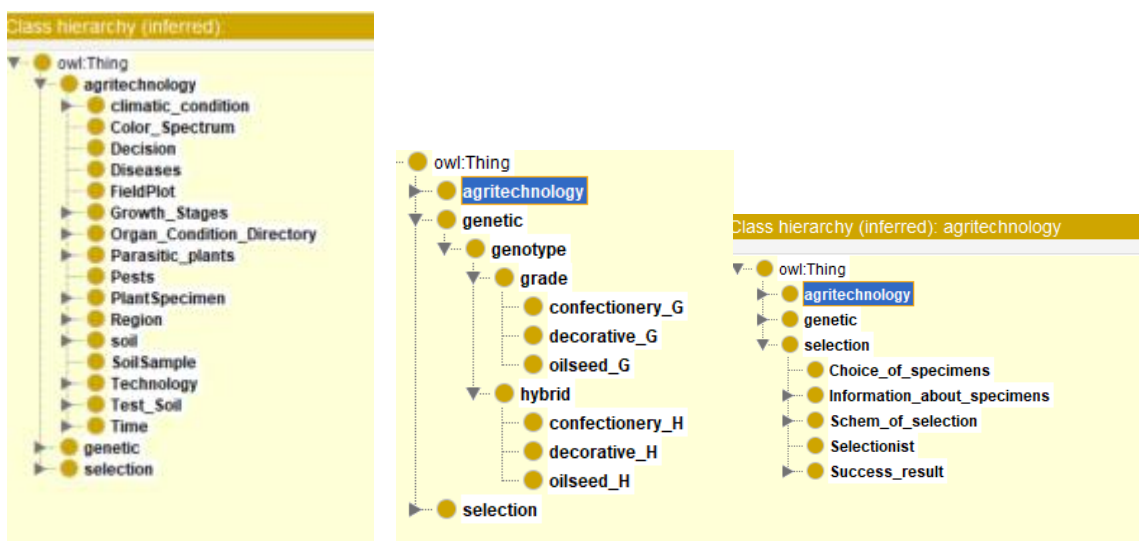


Рисунок 3.2.1 - Архітектура макрокласів Agritechnology, Selection, Genetic

Для визначення корневих класів онтології було використано експертні твердження щодо основних напрямів досліджень культури соняшнику в Інституті олійних культур, а саме: агротехнологія, селекція та генетика.

Верхній рівень розробленої моделі представлений базовим класом owl:Thing, який функціонує як універсальний корінь онтології. Це дозволяє уніфікувати запити до бази знань та забезпечує сумісність онтології з іншими стандартами Semantic Web. Від базового класу owl:Thing походять головні домени знань макрокласи Agritechnology, Selection, Genetic, ієрархічне представлення яких наведено на рис.3.2.2. Кожен із цих доменів має розгалужену систему підкласів (subclasses), що забезпечує вертикальну цілісність моделі. Кожен із цих класів містить підкласи (третій рівень), що відповідають специфічним характеристикам об'єктів певного класу.

Макроклас Agritechnology онтології функціонує як операційне ядро моделі, що формалізує знання про технологічні напрями вирощування соняшнику, які було описано як складові цифрової моделі в Розділі 2. Макроклас Agritechnology забезпечує взаємозв'язок між ресурсним потенціалом господарства (техніка, засоби захисту) та біологічними потребами культури, дозволяючи системам підтримки прийняття рішень формувати оптимальні сценарії агротехнічних заходів на основі аналізу стану ґрунту та динамічних кліматичних факторів. Онтограф макрокласу Agritechnology представлено на рис. 3.2.2. Об'єктні властивості макрокласу Agritechnology представлено на рис. 3.2.3.

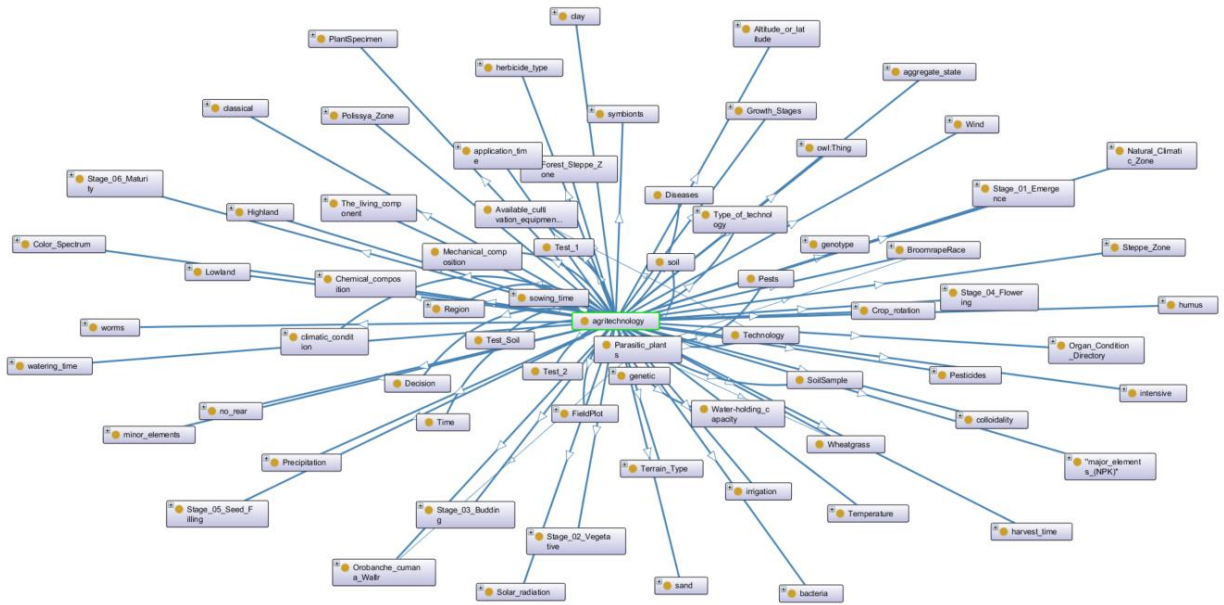


Рисунок. 3.2.2 - Онтограф макрокласу Agritechnology.

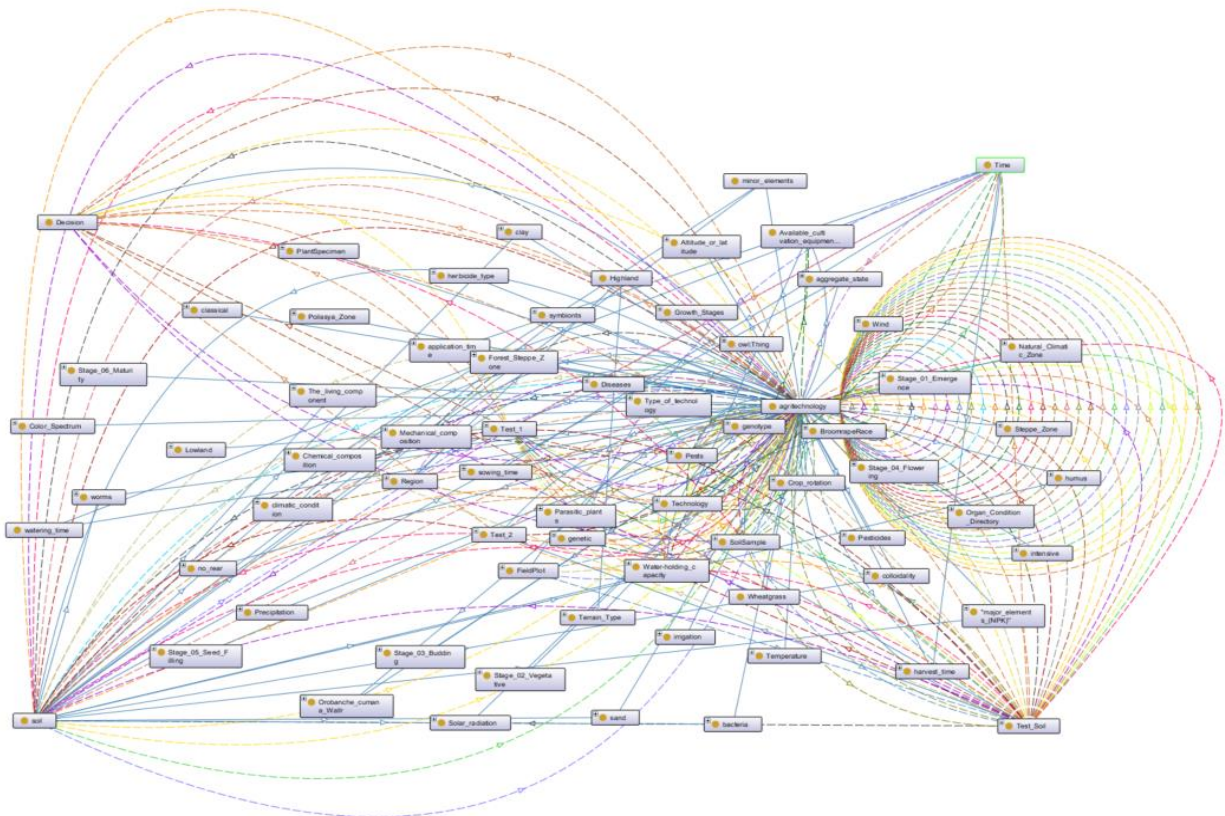


Рисунок. 3.2.3 - Макроклас Agritechnology з об'єктними властивостями.

Введемо два класи PlantSpecimen (індивідуальна рослина на полі) та FieldPlot (польова ділянка), які є ключовими для забезпечення відповідей на

компетентності запитання зі застосуванням системи логічного виведення. На рис.3.2.4 представлено онтограф та об'єктні властивості класів PlantSpecimen та FieldPlot.

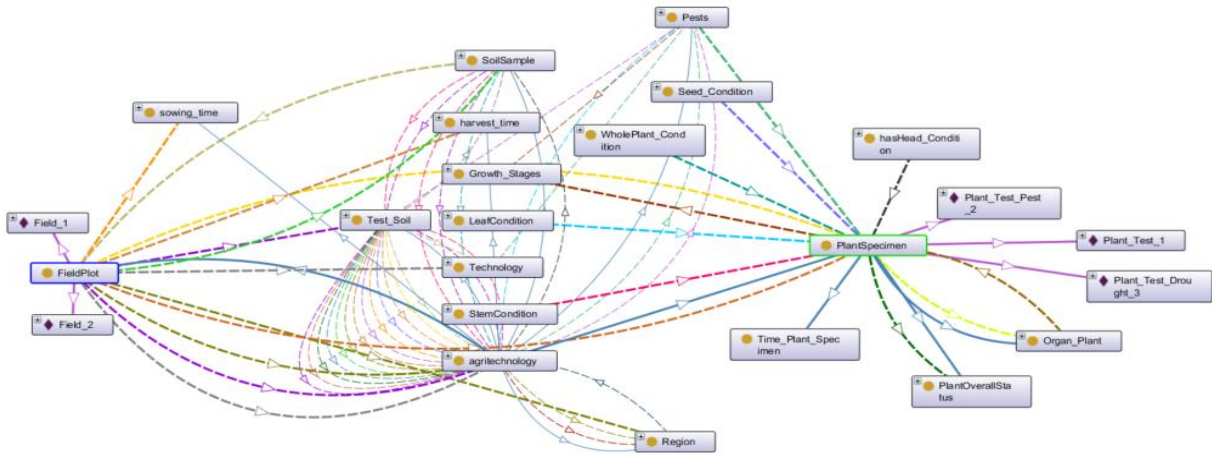


Рисунок. 3.2.4 - Онтограф класів PlantSpecimen та FieldPlot з об'єктними властивостями

На рис.3.2.5 - представлено структурний та семантичний опис класів FieldPlot та PlantSpecimen.



Рисунок. 3.2.5 - Структурний та семантичний опис класів FieldPlot та PlantSpecimen

Клас PlantSpecimen є центральним концептом для цифрового моделювання біологічного стану окремих рослинних організмів безпосередньо в полі, збирає дані про морфологію органів, поточну фазу росту, виявлені стреси та зв'язок із конкретною ділянкою (рис.3.2.6).

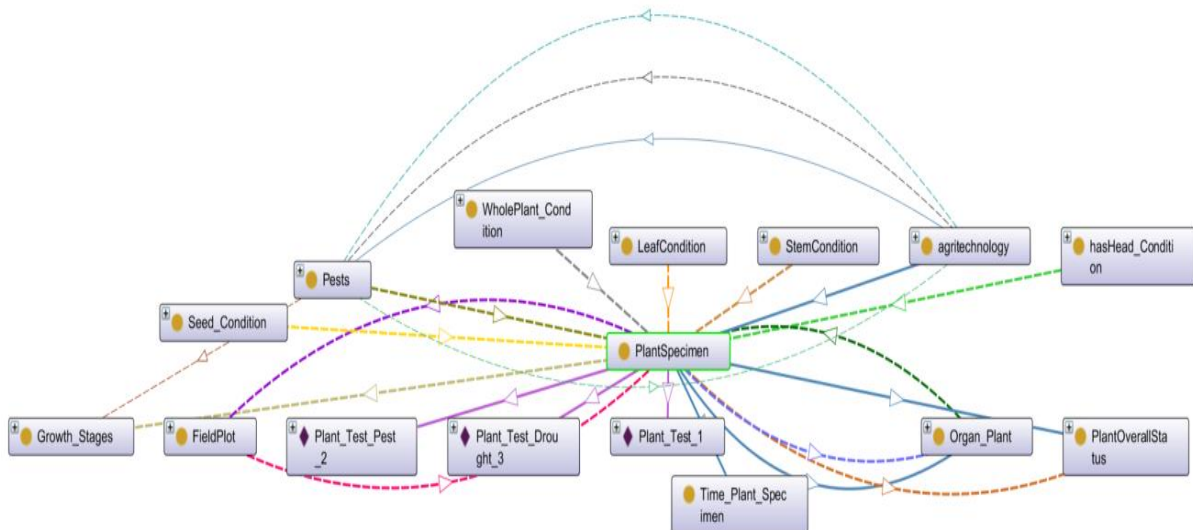


Рисунок 3.2.6 - Онтограф класу PlantSpecimen з об'єктними властивостями

Клас PlantSpecimen інтегрований в онтологію через об'єктні властивості (Object Properties). Властивості, де PlantSpecimen є доменом (Domain), описують характеристики самої рослини та її вихідні зв'язки:

- **hasGrowthStage** — визначає поточну фенологічну фазу рослини,
- **hasOverallStatus** — фіксує інтегральний стан життєдіяльності рослини,
- **hasPart** — реалізує мерологічний зв'язок (частина-ціле), вказуючи на анатомічні органи рослини,
- **isPlantOfField** — задає просторову прив'язку до конкретної польової ділянки (класу FieldPlot), з якої походить цей зразок.

Властивості, де PlantSpecimen є діапазоном (Range), спрямовані на рослину з боку інших сутностей та процесів:

- **attacksOrgan** — пов'язує фітопатогени (гриби, бактерії) або шкідників із рослиною, яку вони уражують;

- `hasHead_Condition`, `hasLeafCondition`, `hasSeed_Condition`, `hasStem_Condition` — специфічні властивості медико-біологічного або морфологічного стану конкретних органів (кошика, листка, насіння, стебла соняшника);
- `hasWholePlant_Condition` — фіксує загальний габітус (зовнішній вигляд) усього рослинного екземпляра;
- `hasSownPlant` — пов'язує поле `FieldPlot` із посіяним на ньому рослинним зразком;
- `isPartof` — зворотна властивість для опису внутрішніх органів, які входять до складу цієї рослини;

Клас `FieldPlot` виступає центральним концептом для моделювання конкретних сільськогосподарських угідь або ділянок поля (рис. 3.2.7). Цей клас консолідує інформацію про технологію вирощування, часові межі польових робіт, посіяні культури та результати лабораторних досліджень (ґрунту і рослин).

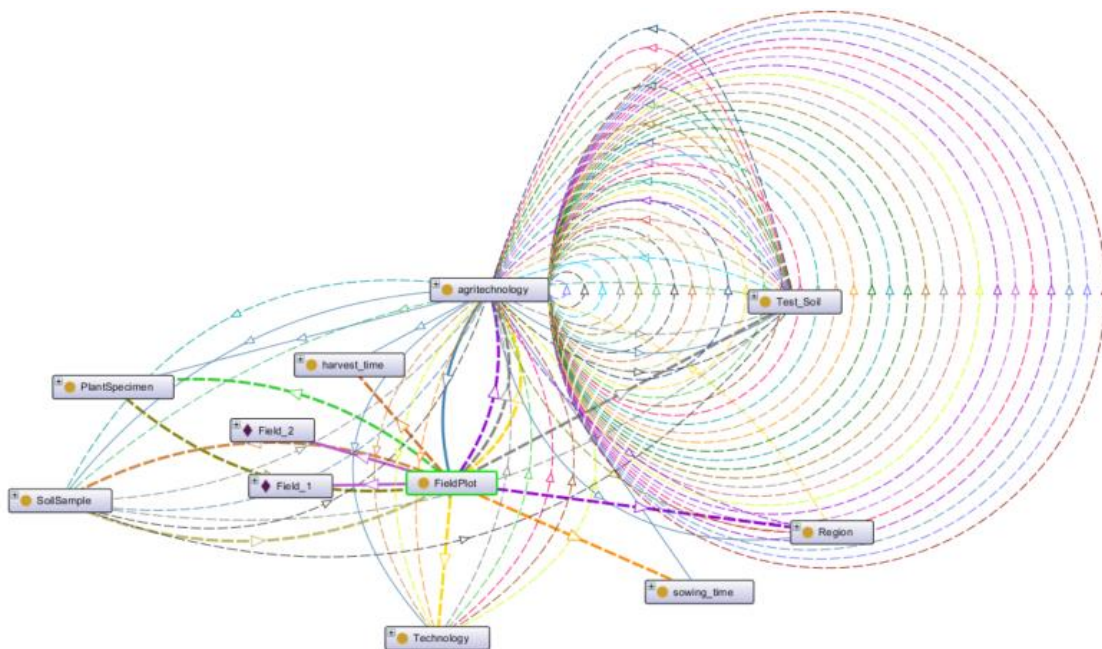


Рисунок 3.2.7 - Онтограф класу `FieldPlot` з об'єктними властивостями

Клас `FieldPlot` активно взаємодіє з іншими концептами онтології через дев'ять об'єктних властивостей (`Object Properties`). Властивості чітко розмежовані за логічним спрямуванням.

Властивості, де `FieldPlot` є доменом (`Domain`), описують характеристики польової ділянки:

- `hasAnalysis` — пов'язує ділянку з результатами проведених аналізів (наприклад, фітопатогенних чи мікробіологічних);
- `hasSoilSample` — вказує на зразки ґрунту, які були відібрані з цієї конкретної ділянки;
- `hasSowingTime` та `hasHarvestTime` — задають часові межі (агротехнічні терміни) посіву та збору врожаю на ділянці;
- `hasSownPlant` — визначає сільськогосподарську культуру (наприклад, конкретний гібрид соняшника), яка засіяна на полі;
- `isLocatedIn` — географічна або адміністративна прив'язка ділянки (наприклад, належність до певної природно-кліматичної зони: Степ, Лісостеп);
- `isUsedTechnology` — вказує на технологію вирощування чи обробітку, що застосовується на цій території.

Властивості, де `FieldPlot` є діапазоном (`Range`), характеризують залежність інших об'єктів від поля:

- `isPlantOfField` — зворотний зв'язок від рослини/культури до польової ділянки, на якій вона росте;
- `isSoilSampleOfField` — зворотний зв'язок від конкретного лабораторного зразка ґрунту до материнського поля, з якого його було взяли.

Клас `Soil` є прямим нащадком макрокласу `Agritechnology`. Клас `Soil` є фундаментальною категорією, що описує фізико-хімічні та біологічні характеристики середовища вирощування сільськогосподарських культур.

Клас Soil інтегрує дані про мінеральний склад (макро- та мікроелементи), гранулометричні показники та стан мікробіоценозу, що дозволяє моделювати оптимальні умови для вегетації культур.

Клас декомпозується на кілька ключових напрямків, що дозволяють комплексно оцінити стан ґрунту.

`Chemical_composition` (Хімічний склад): визначає вміст поживних речовин. `Major_elements_(NPK)`: азот (N), фосфор (P), калій (K). `Minor_elements`: мікроелементи, такі як цинк (zinc), бор (boron), магній (magnesium), селен (selenium). `Mechanical_composition` (Механічний склад): характеризує текстуру та структуру ґрунту через такі компоненти, як: Sand (Пісок); Clay (Глина); Humus (Гумус).

`The_living_component` (Біологічний компонент): описує екосистему мікроорганізмів, зокрема: `Symbionts` (Симбіонти); `Bacteria` (Бактерії).

`Water-holding_capacity` (вологоутримуюча здатність): визначає гідрологічні властивості ґрунту, критичні для планування зрошення.

У контексті агротехнологій клас soil взаємодіє з іншими класами (наприклад, через тестування — `Test_Soil`), забезпечуючи базу даних для прийняття рішень щодо внесення добрив, вибору технологій обробітку та підбору відповідних гібридів рослин. Архітектурне об'єднання доменів Soil та Agritechnology представлено на рис.3.2.8.

Це дозволяє формалізувати прямі та зворотні залежності між агрофізичними параметрами середовища та операційними рішеннями, що є базою для автоматичного логічного виведення у системі підтримки прийняття рішень.

Об'єктні властивості (`Object Properties`) визначають семантичні зв'язки класу soil з іншими компонентами системи агротехнологій.

Клас soil як суб'єкт (`Domain`):

- `hasMechanicalComposition`: пов'язує ґрунт із його механічними характеристиками;

- `hasRecommendedTechnology`: вказує на технології, що пропонуються для цього типу ґрунту;
- `isOptimizedByTechnology`: описує процес покращення властивостей ґрунту під впливом конкретних технологічних операцій.

DL query:

**Query (class expression)**

soil

Execute Add to ontology

**Query results**

Equivalent classes (1 of 1)

- soil

Superclasses (2 of 2)

- agritechnology
- owl:Thing

Direct superclasses (1 of 1)

- agritechnology

Direct subclasses (4 of 4)

- Chemical\_composition
- Mechanical\_composition
- The\_living\_component
- Water-holding\_capacity

Subclasses (15 of 15)

- 'major\_elements\_(NPK)'
- Chemical\_composition
- Mechanical\_composition
- The\_living\_component
- Water-holding\_capacity
- aggregate\_state
- bacteria
- clay
- colloidalilty
- humus
- minor\_elements
- owl:Nothing
- sand
- symbionts
- worms

Instances (9 of 9)

- ◆ Field\_Sample\_01
- ◆ Heavy\_Clay\_Type
- ◆ boron
- ◆ magnesium
- ◆ nitrogen\_N
- ◆ phosphorus\_P
- ◆ potassium\_K
- ◆ selenium
- ◆ zinc

DL query:

**Query (class expression)**

recommendedFor some soil

Execute Add to ontology

**Query results**

Equivalent classes (0 of 0)

Superclasses (2 of 2)

- agritechnology
- owl:Thing

Instances (3 of 3)

- ◆ no\_till\_01
- ◆ no\_till\_02
- ◆ non\_till

Рисунок 3.2.8 - Архітектурне об'єднання доменів Soil та Agritechnology

Клас soil як об'єкт (Range):

- `isMostlySuitableFor`: вказує на придатність певних факторів або культур для даного ґрунту;

- isSoilTest: зв'язок від процедур тестування або аналізів до конкретного об'єкта ґрунту;
- recommendedFor: ключова властивість виводу, де ґрунт є ціллю для рекомендації технології.

Властивості даних (Data Properties) використовуються для кількісної або якісної характеристики фізичних параметрів:

- hasTemperature: фіксує температурний режим ґрунту (Domain: soil).

На рисунку 3.2.9 відображено результати формування об'єктних властивостей та аксіом, які визначають семантичний контекст взаємодії ґрунтових умов із технологічними операціями.

Клас Technology підпорядковується макрокласу Agritechnology. Клас Technology в структурі онтології систематизує знання про виробничі стратегії (класична, інтенсивна, No-till) та необхідні ресурси для їх реалізації, поєднує технічне забезпечення (обладнання для культивування) з агротехнічними заходами (сівозміна, іригація, хімічний захист), забезпечуючи комплексне моделювання процесів вирощування. Онтограф класу Technology представлений на рис. 3.2.10.

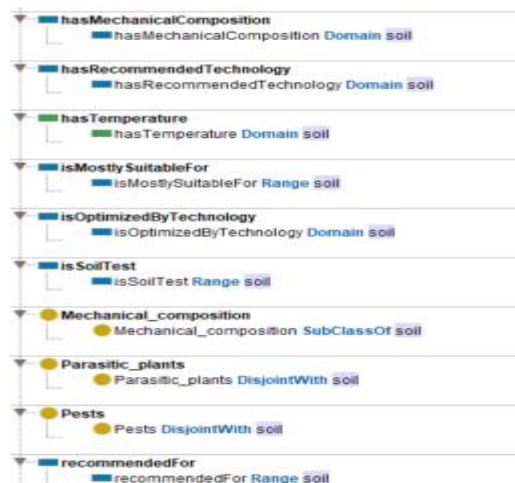


Рисунок 3.2.9 - Специфікація властивостей об'єктів (Object Properties) та аксіом класів, що описують семантичні зв'язки навколо класу soil

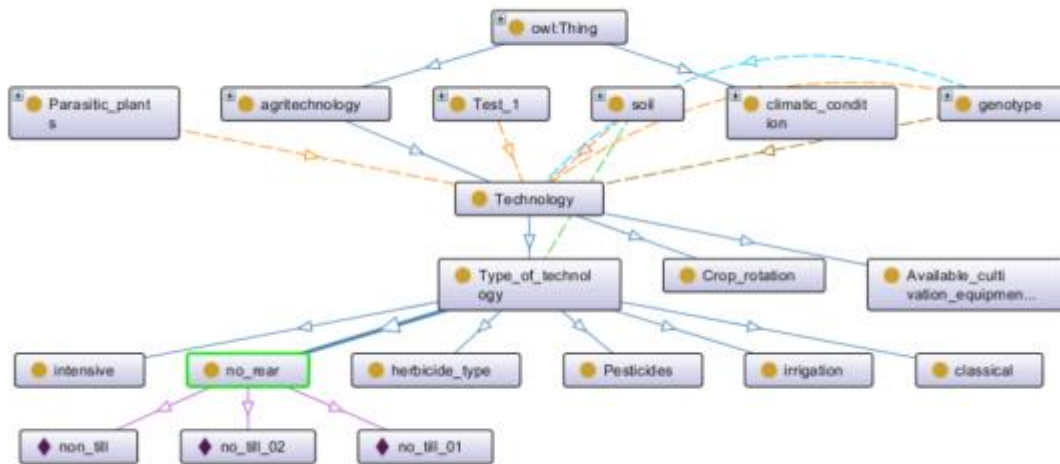


Рисунок 3.2.10 - Онтограф класу Technology

Типологія технологій (Type\_of\_technology) - підклас деталізує конкретні стратегії господарювання, що застосовуються в агровиробництві:

- Classical (Класична): традиційна система обробітку ґрунту та догляду за посівами;
- Intensive (Інтенсивна): технологія, що передбачає високий рівень використання добрив, засобів захисту та механізації для максимізації врожайності;
- No\_rear (система без відвального обробітку / No-till): мінімальний або нульовий обробіток ґрунту для збереження його структури та вологи.

Клас також пов'язаний з інструментами та методами реалізації технологічного процесу:

- Available\_cultivation\_equipmen: наявність та характеристики сільськогосподарської техніки для обробітку;
- Irrigation (Зрошення): підсистема штучного зволоження, що інтегрується в обрану технологію;
- Crop\_rotation (Сівозміна): науково обґрунтоване чергування культур у часі та на полях, що є частиною технологічного регламенту.

Pesticides / Herbicide\_type: хімічні засоби захисту, вибір яких залежить від обраного типу технології.

Наведемо систематизований перелік основних об'єктних властивостей, розподілених за функціональними групами.

Властивості технологічної детермінації (Technology) представлені на рисунку 3.2.11. Ці властивості визначають вимоги та наслідки застосування агротехнологій:

- requiresWeather: Встановлює зв'язок між Technology (Domain) та climatic\_condition (Range) для визначення метеорологічних обмежень операції;
- requiresEquipment: Поєднує Technology з Available\_cultivation\_equipment, описуючи необхідну матеріально-технічну базу;

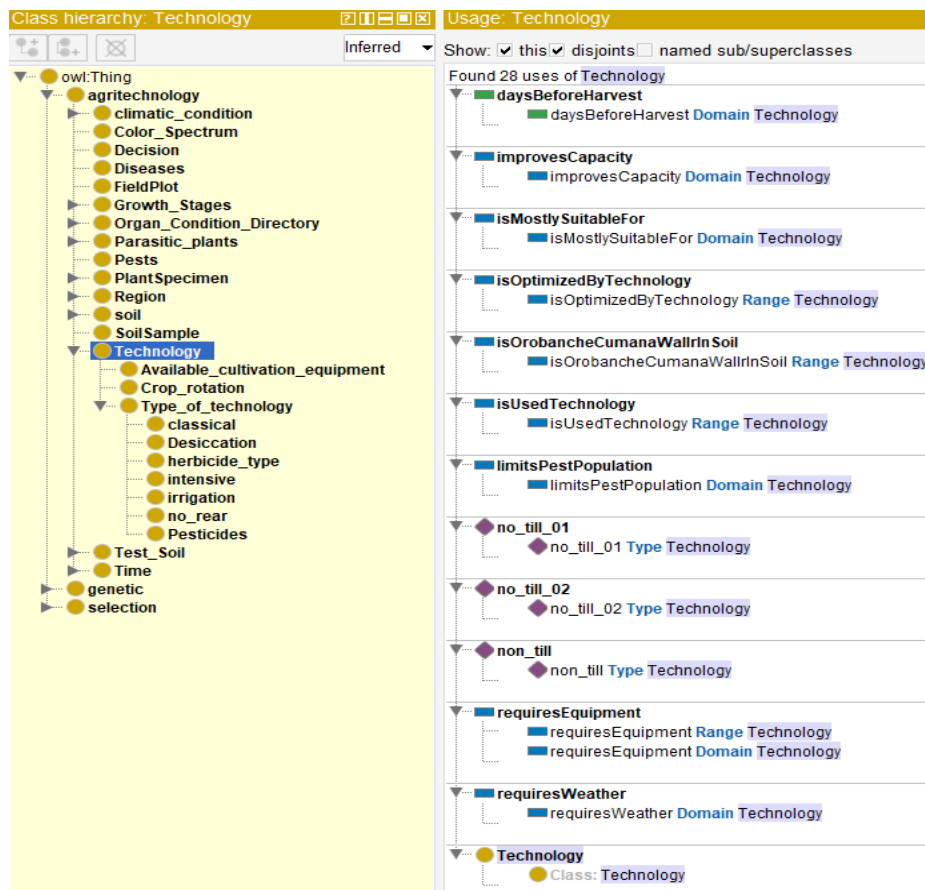


Рисунок 3.2.11 - Клас Technology ієрархічна структура, перелік екземплярів, об'єктні властивості, властивість даних

- `limitsPestPopulation`: Пов'язує `Technology` з класами `Pests` або `Parasitic_plants`, формалізуючи захисну функцію методу;
- `improvesCapacity`: Спрямована від `Technology` до `Water-holding_capacity`, описуючи меліоративний вплив на фізичні властивості ґрунту;
- `recommendedFor`: Ключова властивість виводу, що сполучає `Technology` із `soil` на основі виконання SWRL-правил.

Властивості діагностики та аналізу (`Analytical`).

Використовуються для опису характеристик об'єктів дослідження:

- `hasMechanicalComposition`: З'єднує індивідів класу `soil` із типами механічного складу (`clay`, `sand`, `humus`);
- `hasChemicalComposition`: Пов'язує зразок ґрунту з вмістом макро- та мікроелементів (`NPK`, бор тощо);
- `hasAnalysis`: Загальна властивість, що пов'язує об'єкт `soil` із результатами лабораторних тестів (`Test_Soil`);
- `hasLivingComponent`: Вказує на наявність бактерій чи симбіонтів у структурі ґрунту.

Властивості управління та вибору (`Strategic & Genetic`).

Забезпечують зв'язок між біологічними характеристиками та прийняттям рішень:

- `forGenotypeDependsOn` вказує, що вибір `Technology` (`Range`) детермінований особливостями конкретного генотипу;
- `isOptimizedByTechnology` відображає процес покращення виробничих показників під впливом обраної технології;
- `hasDecision` логічний зв'язок між об'єктом аналізу та прийнятим управлінським рішенням (`Decision`);
- `atTime` часова мітка, що пов'язує рішення або операцію з конкретним періодом (`Time`) — посівом, поливом чи збиранням.

Для формалізації часових параметрів в онтологію введено клас Time (рис. 3.2.12.). Клас Time розміщений в єдиному дереві з такими критичними доменами як soil, Technology та climatic\_condition. Це дозволяє інтегрувати часовий чинник у процес прийняття стратегічних рішень (Decision), забезпечуючи синхронізацію біологічних потреб рослини з технічними можливостями господарства.

Клас Time в структурі онтології виконує роль темпорального регулятора, що забезпечує часову специфікацію агротехнічних заходів та природних факторів.

The image displays two screenshots from an ontology query tool. The left screenshot, labeled 'a)', shows the 'Query results' for the class 'Time'. It lists equivalent classes (Time), superclasses (agritechnology, owl:Thing), direct superclasses (agritechnology), direct subclasses (application\_time, harvest\_time, sowing\_time, watering\_time), subclasses (application\_time, harvest\_time, owl:Nothing, sowing\_time, watering\_time), and instances (14 of 14), including Sowing\_TIME\_UE0100456 and at\_Time. The right screenshot, labeled 'b)', shows the 'Usage: Time' section, which lists 60 uses of the class. It includes relationships such as 'application\_time SubClassOf Time', 'at\_Time Type Time', 'climatic\_condition Time DisjointWith climatic\_condition', 'Decision DisjointWith Time', 'Diseases DisjointWith Time', 'forGenotypeDependsOn Domain Time', 'harvest\_time SubClassOf Time', and 'hasCalendarDate Domain Time'.

a)

б)

Рисунок 3.2.12 - Клас Time а) ієрархічна структура, перелік екземплярів, б) об'єктні властивості та аксіоми класу

Цей клас дозволяє формалізувати знання про сезонну динаміку кліматичних умов, життєві цикли біологічних загроз (шкідників та патогенів), а також визначати часові інтервали реалізації технологічних карт вирощування культур. підкласи якого пов'язані з типовими інформаційними запитам практиків (фермерів) із різних регіонів України, що звертаються за консультаціями до фахівців інституту. Ієрархічна структура класу Time деталізована через функціональні підкласи, що відповідають критичним етапам виробництва:

- `application_time`: регламентує терміни внесення технологічних засобів (добрив, пестицидів);
- `sowing_time`: визначає оптимальне вікно для проведення посівних робіт;
- `watering_time`: фіксує графік іригаційних заходів;
- `harvest_time`: встановлює період збирання врожаю.

Темпоральна інтеграція рішень та операцій реалізована через об'єктні властивості (Object Properties), де клас Time виступає як область значень (Range):

- `atTime`: ключова властивість, що дозволяє пов'язати будь-яку технологічну подію або операцію з конкретним часовим індивідом;
- `isDecision`: специфічний зв'язок, що маркує управлінські рішення часовими мітками, забезпечуючи простежуваність стратегічного планування.

Для внесення конкретних значень (дат, інтервалів або тривалості стадій вегетації) використовується властивість даних(Data Properties):

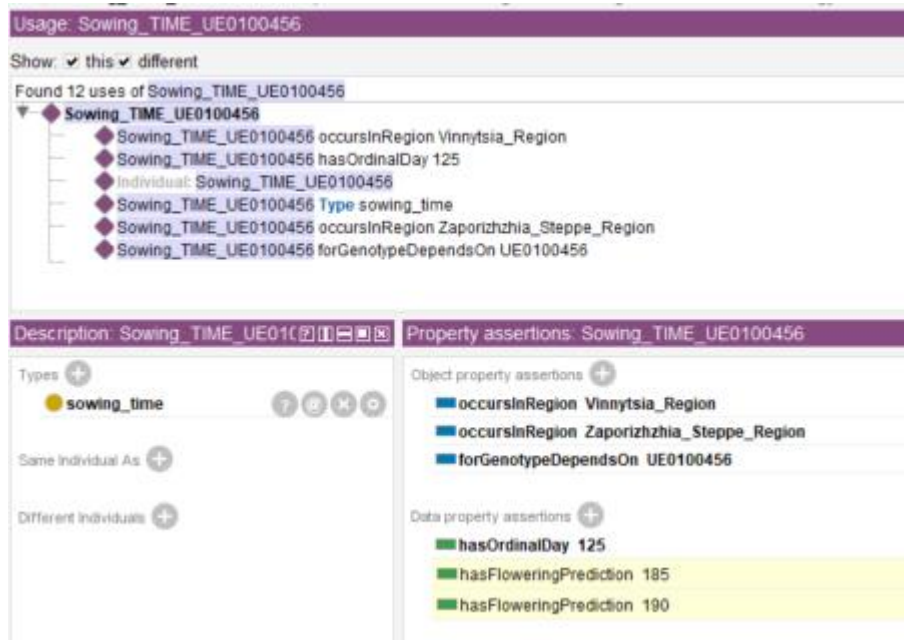
- `in_Time`: де клас Time виступає суб'єктом (Domain) для кількісної або текстової фіксації часових параметрів.

Для забезпечення високої точності роботи ризонера встановлено жорсткі обмеження неперетинності (DisjointWith). Клас Time логічно відділений від усіх матеріальних та процесуальних доменів:

- об'єкти та фактори: `soil`, `climatic_condition`, `Pests`, `Diseases`, `Parasitic_plants`;

- операції та висновки: Technology, Test\_Soil, Decision.

Опис індивідуума підкласу sowing\_time представлено на рисунку 3.2.13.



Рисунк.3.2.13 - Опис індивідуума підкласу sowing\_time

Клас Growth\_Stage містить формалізацію фенологічних фаз розвитку культури відповідно до класифікації Шнейтера та Міллера (Schneider & Miller, 1981) (рис.3.2.14). Онтограф класу Growth\_Stage представлено на рисунку 3.2.15.

Класифікація Шнейтера та Міллера є класичною шкалою макростадій розвитку соняшнику, яка використовується у точній агрономії та селекції для стандартизації фенотипічних спостережень. Нижче наведено зіставлення класу Growth\_Stage з періодизацією Шнейтера-Міллера (табл. 3.2.1).

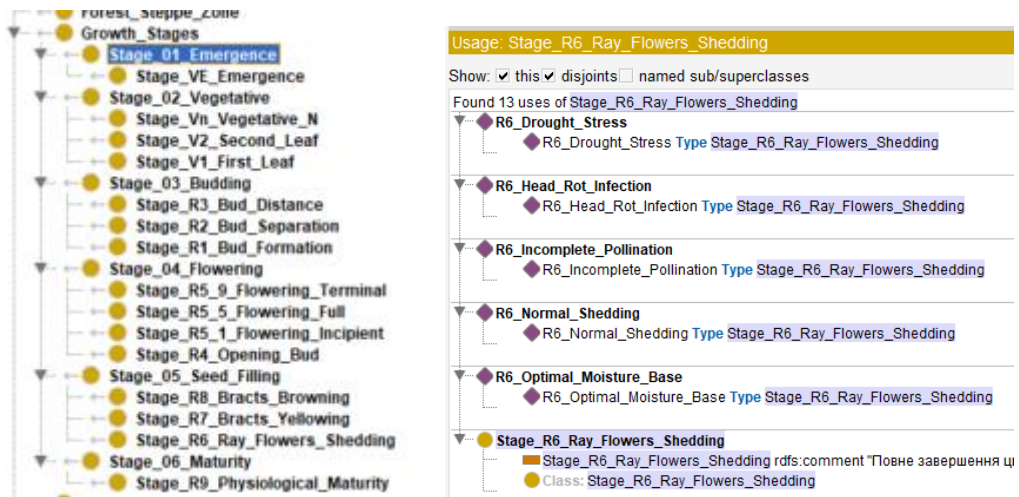


Рисунок.3.2.14 - Ієрархічна структура класу `Growth_Stage` а) та індивідами класу `Stage_R6_Ray_Flowers_Shedding` б)

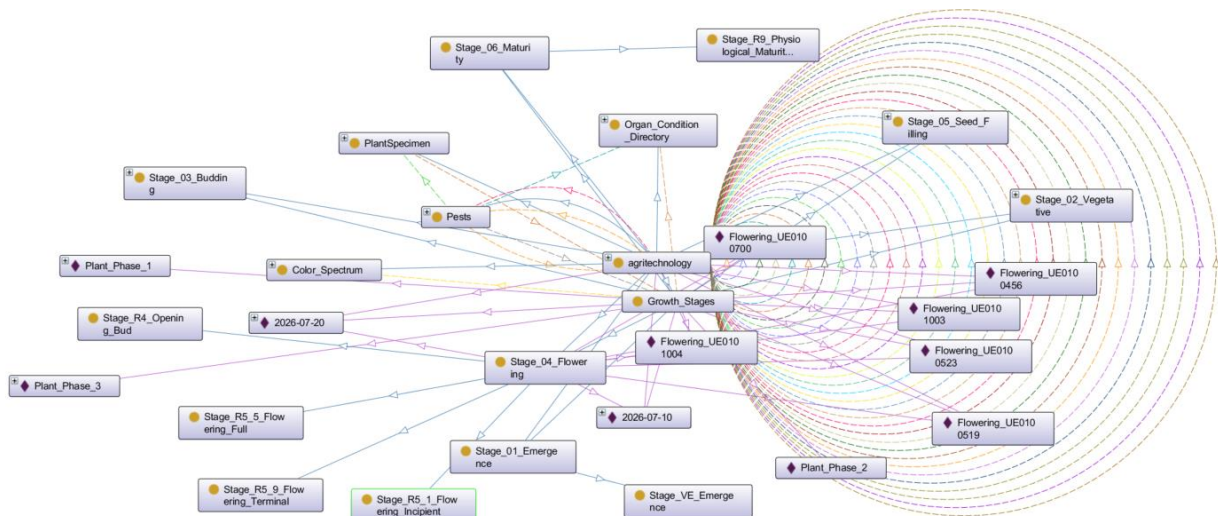


Рисунок.3.2.15 - Онтограф класу `Growth_Stage`

Створено два класи довідники. Клас-довідник `Color_Spectrum` (Спектр кольорів) призначений для формалізації та фіксації оптико-колориметричних параметрів органів соняшнику, що отримуються в результаті попередньої математичної обробки та аналізу цифрових зображень.

Таблиця 3.2.1 - Зіставлення класу Growth\_Stage з періодизацією Шнейтера-Міллера.

Стадія в онтології	Код Шнейтера-Міллера	Опис морфологічних та цифрових ознак
Stage_01_Emergence	VE	Сходи. Поява сім'ядоль над поверхнею ґрунту.
Stage_02_Vegetative	V1 – V(n)	Веgetативні стадії. Рахуються за кількістю справжніх листків завдовжки понад 4 см.
Stage_03_Budding	R1 – R3	Бутонізація (зірковий кошик). Поява мініатюрного бутона, який з часом піднімається над найближчим листям.
Stage_04_Flowering	R5.1 – R5.9	Цвітіння. Рахується у відсотках від загальної кількості квіток на кошику (наприклад, R5.5 — 50% цвітіння).
Stage_05_Seed_Filling	R6 – R8	Наливання та дозрівання насіння. Язичкові квітки в'януть, кошик починає змінювати колір із зеленого на лимонно-жовтий.
Stage_06_Maturity	R9	Фізіологічна зрілість. Тильна сторона кошика стає жовто-бурою, приквітки коричневіють. Вологість насіння природно падає.

Клас Color\_Spectrum виступає в ролі семантичного містка між низькорівневими числовими даними комп'ютерного зору (матрицями пікселів) та високорівневими агрономічними поняттями. Індивіди цього класу (Green, Lemon\_Yellow, Yellow\_Brown, Brown) представляють еталонні колірні маркери, кожен з яких наділений інваріантними числовими межами для каналів колірного простору RGB (зокрема, властивостями даних hasRMin, hasRMax, hasGMin, hasGMax, hasBMin, hasBMax). Така декомпозиція дозволяє алгоритмам комп'ютерного зору здійснювати динамічний запит до бази знань для верифікації кольору без необхідності жорсткого вбудовування числових констант у програмний код розпізнавання.

The screenshot displays an ontology editor interface. On the left, a 'Class hierarchy (inferred)' tree shows 'Color\_Spectrum' as a subclass of 'agritechnology'. The main area is divided into three panels: 'Usage: Color\_Spectrum' showing 15 uses (e.g., Brown Type Color\_Spectrum, Green Type Color\_Spectrum, Lemon\_Yellow Type Color\_Spectrum, Yellow\_Brown Type Color\_Spectrum), 'Annotations: Color\_Spectrum' showing an 'rdfs:comment' 'довідник кольорів', and 'Description: Color\_Spectrum' showing it is a subclass of 'agritechnology' and has instances like 'Brown', 'Green', 'Lemon\_Yellow', 'Plant\_Phase\_1', and 'Yellow\_Brown'.

Рисунок. 3.2.16 - Клас-довідник Color\_Spectrum

Клас-довідник *Organ\_Condition\_Directory* (Довідник станів органів) (рис.3.2.17)

Цей концепт слугує для категоризації та стандартизації якісних морфологічних перетворень архітектоники рослини, які не піддаються прямим лінійним числовим вимірам, проте є критичними тригерами для логічного виведення стадій вегетації за допомогою SWRL-правил.

Індивіди даного класу (*Active\_Growth*, *Bud\_Visible*, *Active\_Flowering*, *Wilting*, *Dry*) описують дискретні якісні стани вегетативних та репродуктивних органів соняшнику. Обмеження діапазону значень (*Range*) об'єктної властивості

hasCondition виключно цим класом забезпечує строге логічне наслідування, що унеможливило некоректну інтерпретацію експертних знань.

The screenshot displays a software interface for class management. On the left, a 'Class hierarchy (inferred)' tree shows 'owl:Thing' as the root, with 'agritechnology' as a direct subclass. Under 'agritechnology', several subclasses are listed, including 'Organ\_Condition\_Directory'. The middle pane, titled 'Usage: Organ\_Condition\_Directory', lists 15 instances of the class: 'Active\_Flowering', 'Active\_Growth', 'Bud\_Visible', 'Dry', 'hasCondition', 'Organ\_Condition\_Directory', and 'Wilting'. The right pane, 'Annotations: Organ\_Condition\_Directory', shows an annotation: 'rdfs:comment [language: uk] довідник станів рослини'. The bottom pane, 'Description: Organ\_Condition\_Directory', shows the class is a subclass of 'agritechnology' and lists its instances: 'Active\_Flowering', 'Active\_Growth', 'Bud\_Visible', 'Dry', and 'Wilting'.

Рисунок.3.2.17 - Клас-довідник Organ\_Condition\_Directory

Клас Genetic є прямим підкласом (subclass) кореневого концепту owl:Thing. Клас Genetic систематизує знання про генетичне різноманіття культури, розділяючи насінневий матеріал на сорти (Grade) та гібриди (Hybrid). Даний клас є центральним елементом для вирішення завдань автоматичного поєднання генотипу з фенотипічними проявами, що дозволяє прогнозувати продуктивність рослин залежно від агроекологічних факторів (грунтових умов та клімату) та обраних технологій вирощування.

Ієрархія властивостей даних (Data Properties) класу Genetic представлена на рис. 3.2.18. Ці показники відповідають характеристикам, що визначені цифровою моделлю в розділі 2.

Клас декомпозується на дві основні категорії, що відображають селекційний статус насіннєвого матеріалу:

- Grade (Сорт): сукупність рослин, створених у результаті селекції, що мають певні спадкові морфологічні, фізіологічні та господарські ознаки;
- Hybrid (Гібрид): результат схрещування генетично відмінних батьківських форм, що характеризується ефектом гетерозису (підвищеною життєздатністю та врожайністю).

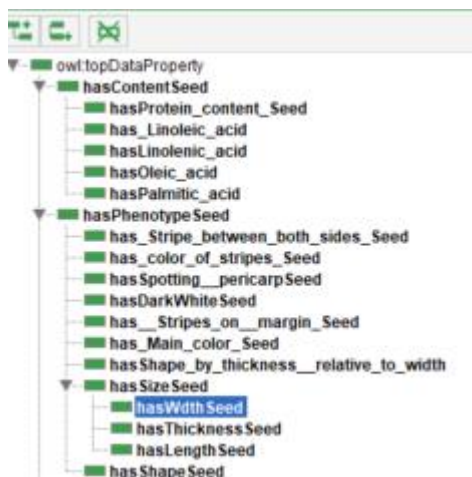


Рисунок.3.2.18 - Ієрархія властивостей даних (Data Properties) класу Genetic

Клас genotype має семантичні зв'язки з:

- Selection (Селекція) та Genetic (Генетика): суміжні класи, що описують походження та методи покращення генотипу;
- Diseases & Pests: для визначення генетичної стійкості конкретного сорту чи гібриду до певних патогенів;
- Climatic\_condition: для оцінки адаптивності генотипу до температурних режимів та кількості опадів.

Клас Pests інтегровано в загальну онтологічну структуру як прямий підклас макрокласу Agritechnology, що дозволяє розглядати біологічні загрози як частину загального технологічного контексту вирощування соняшнику. Процедура логічного виведення (reasoning) підтвердила відсутність логічних суперечностей у визначенні класу (відсутність підкласів, крім системного owl:Nothing).

Визначено 9 специфікованих видів шкідників, характерних для посівів соняшнику (Instances). Серед них Sand tarragon (*Opatrum sabulosum*) та Wireworms (ковалики) — як основні загрози на ранніх етапах вегетації; Sunflower moth (*Homoeosoma nebulella*) та Sunflower spikenose — шкідники, що безпосередньо впливають на формування кошика та якість насіння (рис. 3.2.19). Така деталізація індивідів у межах онтології забезпечує можливість формування точних рекомендацій у системі підтримки прийняття рішень при виявленні конкретних ентомологічних ризиків на основі даних моніторинг

Для реалізації інтелектуальних функцій системи використано логічні правила семантичного виведення (SWRL Rules), які враховують регіональні особливості та технологічні параметри.

На рисунку 3.2.20 представлена ієрархія класів (inferred class hierarchy), отримана після запуску логічного виведення (reasoning), що демонструє уточнену структуру зв'язків між класами Genetic, Time, Technology.

Pests

Execute Add to ontology

**Query results**

Equivalent classes (1 of 1)

- Pests

Superclasses (2 of 2)

- agritechnology
- owl:Thing

Direct superclasses (1 of 1)

- agritechnology

Direct subclasses (1 of 1)

- owl:Nothing

Subclasses (1 of 1)

- owl:Nothing

Instances (9 of 9)

- ◆ ""Sand\_tarragon\_Opatrum\_sabulosum\_(L.,\_1758)""
- ◆ ""Sunflower\_moth\_(sunflower\_moth)\_Homoeosoma\_nebulium\_(Den.\_and\_Schiff.)""
- ◆ ""Thistle\_Vanessa\_cardui\_(Linnaeus,\_1758)""
- ◆ Barbel\_or\_sunflower\_agapanthia\_Agapanthia\_dahli\_Richt
- ◆ Bow\_butterfly\_Margaritia\_sticticalis\_L
- ◆ Fruit\_flies\_Strautia\_longipennis
- ◆ Leaf-eating\_cutworms\_A\_number\_of\_species\_of\_the\_family\_Noctuidae
- ◆ Sunflower\_spikenose\_Mordeisistena\_parvula\_Gyll
- ◆ Wireworms:\_Black\_boletus\_Athous\_niger\_L\_Sowing\_bollard\_Agriotes\_sputator\_L.

Рисунок.3.2.19 - Клас Pests: ієрархічна структура, перелік екземплярів



Рисунок.3.2.20 - Inferred class hierarchy

### 3.3 Семантична оцінка якості онтологічної моделі

Маємо два види оцінок якості онтології: семантичні — з урахуванням питань компетентності, сформульованих експертом, і структурні (кількісні метрики онтографа). Оцінка семантичної якості онтології була проведена на основі можливості отримати відповіді на компетентнісні запитання.

Нижче наведено правила для основних класів онтології — PlantSpecimen та FieldPlot, які дозволяють автоматично переводити візуальні ознаки ураження та дані моніторингу в комплексні діагнози та рекомендації.

Правила для класу PlantSpecimen. Клас PlantSpecimen є основним операндом для SWRL-правил, що діагностують стан рослини.

S14: Успадкування параметрів поля / Перенесення часу посіву на рослину  
 $\text{PlantSpecimen(?p)} \wedge \text{isPlantOfField(?p, ?f)} \wedge \text{hasSowingTime(?f, ?t)} \rightarrow \text{hasSowingTime(?p, ?t)}$

S15: Діагностика абіотичного стресу / В'янення органів через посуху  
 $\text{PlantSpecimen(?p)} \wedge \text{hasPart(?p, ?organ)} \wedge \text{Organ\_Condition\_Directory(?organ)} \wedge \text{hasCondition(?organ, Wilting)} \rightarrow \text{hasOverallStatus(?p, AbioticStress)}$

S16: Верифікація здорового розвитку рослини за станом листка  
 $\text{PlantSpecimen(?p)} \wedge \text{hasPart(?p, ?leaf)} \wedge \text{Leaf\_Plant\_Sample(?leaf)} \wedge \text{hasLeafCondition(?leaf, Healthy\_Leaf)} \wedge \text{hasCondition(?p, Active\_Growth)} \rightarrow \text{hasOverallStatus(?p, Healthy\_Development)}$

S17: Діагностика біотичного стресу / Шкідники та скелетування листка  
 $\text{PlantSpecimen(?p)} \wedge \text{hasPart(?p, ?leaf)} \wedge \text{Leaf\_Plant\_Sample(?leaf)} \wedge \text{hasLeafCondition(?leaf, LeafSkeletonization)} \rightarrow \text{hasOverallStatus(?p, BioticStress)}$

Для перевірки цієї логіки у базі знань створено три тестові індивідууми: Plant\_Test\_1 — базовий зразок, Plant\_Test\_Drought\_3 — зразок для моделювання абіотичного стресу (пов'язаний із правилом S15), Plant\_Test\_Pest\_2 — зразок для моделювання біотичного стресу (пов'язаний із правилом S17), Plant\_Phase6 — зразок для моделювання застосування десикації. Результати роботи резонера для

Plant\_Test\_Pest\_2 та Plant\_Test\_Drought\_3 показано на рис. 3.3.1.

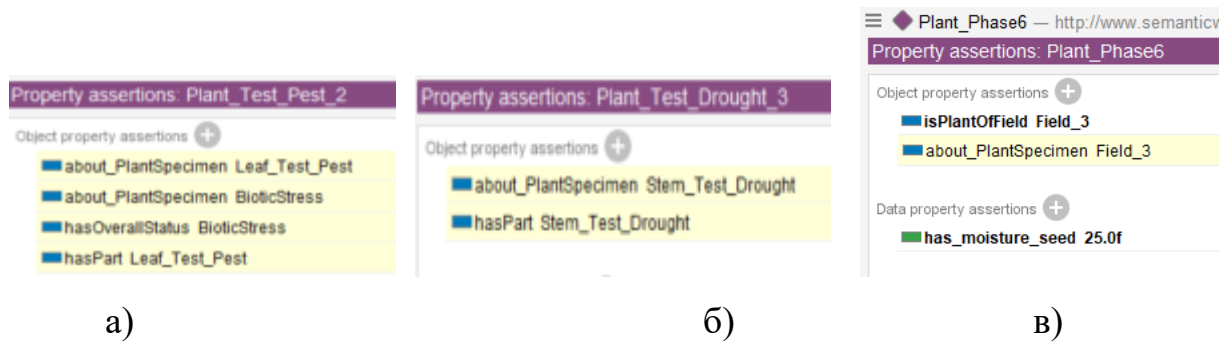


Рисунок 3.3.1 - Результати роботи логичної машини резонера для індивідууму а)Plant\_Test\_Pest\_2, б)Plant\_Test\_Drought\_3, в) Plant\_Phase6 класу PlantSpecimen

Правила для класу FieldPlot (Польова ділянка). Логіка для FieldPlot базується на аналізі зразків ґрунту (SoilSample(?x)), оскільки FieldPlot є єдиним доменом для властивості hasSoilSample. Усі рішення (Decision) автоматично підтягуються до конкретної польової ділянки. У базі знань створено тестові індивідууми Field\_1, Field\_2, Field\_3.

S1: Експертна логіка формування технологічних рекомендацій залежно від складу ґрунту. Правило реалізує диференційований підхід: якщо ґрунт має глинистий механічний склад, генерується рекомендація щодо застосування технології обробки без заднього приводу (no\_rear). S1: soil(?s) ^ hasMechanicalComposition(?s, ?m) ^ clay(?m) ^ no\_rear(?a) -> recommendedFor(?a, ?s).

Правила S2, S3, S4: Прогнозування цвітіння Клас FieldPlot має властивість hasSowingTime, яка веде до класу sowing\_time, що активізує ці правила. Вони додають певну кількість днів до дати посіву залежно від природно-кліматичної зони розташування поля:

- S3 (Степ / Steppe\_Zone): +60 днів. S3: sowing\_time(?s) ^ occursInRegion(?s, ?r) ^ Steppe\_Zone(?r) ^ hasOrdinalDay(?s, ?start) ^ swrlb:add(?predicted, ?start, 60) -> hasFloweringPrediction(?s, ?predicted)

- S2 (Полісся / Polissya\_Zone): +70 днів. S2: sowing\_time(?s) ^ occursInRegion(?s, ?r) ^ Polissya\_Zone(?r) ^ hasOrdinalDay(?s, ?start) ^ swrlb:add(?predicted, ?start, 70) -> hasFloweringPrediction(?s, ?predicted)
- S4 (Лісостеп / Forest\_Steppe\_Zone): +65 днів. S4: sowing\_time(?s) ^ occursInRegion(?s, ?r) ^ Forest\_Steppe\_Zone(?r) ^ hasOrdinalDay(?s, ?start) ^ swrlb:add(?predicted, ?start, 65) -> hasFloweringPrediction(?s, ?predicted)
  - S5: Генетичні характеристики сорту Правило пов'язує генетику сорту з конкретним днем посіву. S5: genetic(?g) ^ isGenotypeForSowing(?g, ?s) ^ hasOrdinalDay(?s, ?val) -> hasSowingDayFromEvent(?g, ?val)
  - Правила глибокого моніторингу (S10, S11, S12) Дозволяють аналізувати шари поля на глибині 1.0–1.5 метри. Якщо зафіксовано відповідні умови, система формує Decision (Рішення) для FieldPlot.
    - S10: Перевірка вологозабезпечення та ущільнення в глибоких шарах SoilSample(?x) ^ Test\_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasSamplingDepth(?y, ?depth) ^ swrlb:greaterThanOrEqual(?depth, 1.0) ^ swrlb:lessThanOrEqual(?depth, 1.5) -> hasFieldStatus(?d, DeepLayerMoistureVerified)
    - S11: Контроль легкого механічного складу та вимивання поживних речовин SoilSample(?x) ^ Test\_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasSamplingDepth(?y, ?depth) ^ swrlb:greaterThanOrEqual(?depth, 1.0) ^ hasSubterraneanTexture(?y, Sandy) -> hasFieldStatus(?d, NutrientLeachingRiskDetected)
    - S12: Діагностика прихованих біотичних загроз — насіння Вовчка SoilSample(?x) ^ Test\_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasSamplingDepth(?y, ?depth) ^ hasBiologicalContamination(?y, Orobanche\_cumana\_Seeds) -> hasFieldStatus(?d, OrobancheDeepInfestationRisk)
- Правила оцінки поля (S6, S7, S8) Фіксують індивідууми, пов'язані з полем.

○ S6: Хімічний склад поля (баланс азоту, фосфору та вологоутримання).  $\text{SoilSample}(?x) \wedge \text{Test\_Soil}(?y) \wedge \text{Time}(?t) \wedge \text{Decision}(?d) \wedge \text{isSoilTest}(?x, ?y) \wedge \text{isDecisionForSample}(?d, ?x) \wedge \text{hasAnalysisTime}(?y, ?t) \wedge \text{hasChemicalFormulas}(?y, \text{True}) \wedge \text{hasMoistureRetention}(?y, \text{True}) \rightarrow \text{hasFieldStatus}(?d, \text{ChemicalAndMoistureAnalysisComplete})$

○ S7: Гідрофізика поля (структура ділянки за співвідношенням «пісок-глина-гумус»).  $\text{SoilSample}(?x) \wedge \text{Test\_Soil}(?y) \wedge \text{Time}(?t) \wedge \text{Decision}(?d) \wedge \text{isSoilTest}(?x, ?y) \wedge \text{isDecisionForSample}(?d, ?x) \wedge \text{hasAnalysisTime}(?y, ?t) \wedge \text{hasSoilTexture}(?y, ?texture) \wedge \text{hasHumusContent}(?y, ?humus) \rightarrow \text{hasFieldStatus}(?d, \text{MechanicalCompositionAnalysisComplete})$ .

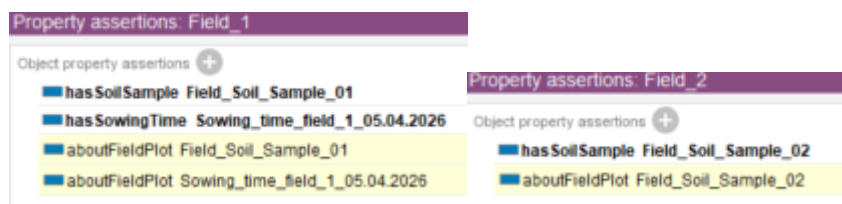
S8: Біологічний потенціал поля (наявність бактерій-фосфатмобілізаторів та мікоризи).  $\text{SoilSample}(?x) \wedge \text{Test\_Soil}(?y) \wedge \text{Time}(?t) \wedge \text{Decision}(?d) \wedge \text{isSoilTest}(?x, ?y) \wedge \text{isDecisionForSample}(?d, ?x) \wedge \text{hasAnalysisTime}(?y, ?t) \wedge \text{hasMicrobialInhabitant}(?x, ?bacteria) \wedge \text{hasSynergisticFungi}(?x, ?fungi) \rightarrow \text{hasFieldStatus}(?d, \text{BiologicalComponentAnalysisComplete})$ .

Створимо правило для вибору технології планової десикації соняшнику

S18:  $\text{PlantSpecimen}(?plant) \wedge \text{hasSownPlant}(?crop, ?plant) \wedge \text{hasGrowthStage}(?crop, \text{R6\_Normal\_Shedding}) \wedge \text{has\_moisture\_seed}(?plant, ?moisture) \wedge \text{swrlb:greaterThanOrEqual}(?moisture, 20.0) \rightarrow \text{isUsedTechnology}(?crop, \text{Desiccation\_Standard})$ .

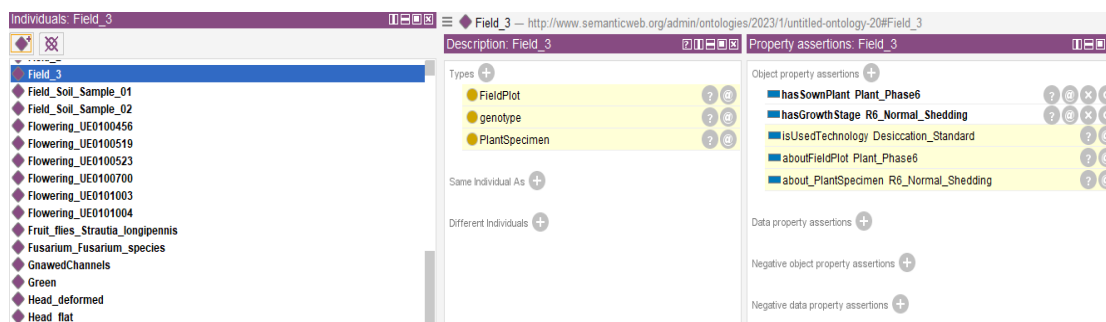
Результати роботи резонера для індивідуумів Field\_1, Field\_2, Field\_3 класу FieldPlot показано на рис. 3.3.2.

У табл. 3.3.1. представлено відповідність між розробленими семантичними SWRL правилами та функціональними напрямками.



а)

б)



в)

Рисунок 3.3.2 - Результати роботи резонера для індивідуумів а) Field\_1, б) Field\_2 та в) Field\_3 класу FieldPlot

Для проведення семантичної верифікації розробленої онтологічної моделі було використано стандартний апарат компетентнісних запитань. Трансформацію неформальних експертних запитів у формальні логічні запити дескрипційної логіки, предикатна форма, SWRL правила у межах бази знань HELIANTUS, деталізовано в табл. 3.3.2.

Таблиця 3.3.1 - Відповідність функціональних напрямів та семантичних SWRL правил

Функціональний напрям	Номер правила
Глобальний інтеграційний домен управління продуктивністю	S1, S6, S7, S8, S10, S11, S12, S18
Фенотиповий	S15, S16, S17
Темпоральний	S2, S3, S4, S14
Генотипів та біологічних властивостей	S5
Екологічний	<i>Перспективне розширення (правила відсутні)</i>
Економічний	<i>Перспективне розширення (правила відсутні)</i>

Таблиця 3.3.2 - Матриця верифікації семантичної якості онтології HELIANTUS на основі компетентнісних запитань та SWRL-правил

Неформальне компетентнісне запитання	Формальне компетентнісне запитання (логічний запит)	Семантичне SWRL-правило виведення
1. Яку технологію обробітку ґрунту доцільно застосувати на ділянках із важким (глинистим) складом?	Знайти технологічні рішення $a \in \text{No\_rear}$ , рекомендовані для зразків ґрунту $s \in \text{soil}$ із механічним складом $m \in \text{clay}$ .	S1: $\text{soil}(?s) \wedge \text{hasMechanicalComposition}(?s, ?m) \wedge \text{clay}(?m) \wedge \text{no\_rear}(?a) \rightarrow \text{recommendedFor}(?a, ?s)$
2. Коли очікується настання фази цвітіння соняшника при посіві у зоні Полісся?	Для події посіву $s \in \text{sowing\_time}$ у регіоні $r \in \text{Polissya\_Zone}$ розрахувати прогнозну дату цвітіння $\text{predicted}$ .	S2: $\text{sowing\_time}(?s) \wedge \text{occursInRegion}(?s, ?r) \wedge \text{Polissya\_Zone}(?r) \wedge \text{hasOrdinalDay}(?s, ?start) \wedge \text{swrlb:add}(?predicted, ?start, 70) \rightarrow \text{hasFloweringPrediction}(?s, ?predicted)$
3. Яка прогнозна дата цвітіння культури для ділянок, розташованих у Степовій зоні?	Для події посіву $s \in \text{sowing\_time}$ у регіоні $r \in \text{Steppe\_Zone}$ розрахувати прогнозну дату цвітіння $\text{predicted}$ .	S3: $\text{sowing\_time}(?s) \wedge \text{occursInRegion}(?s, ?r) \wedge \text{Steppe\_Zone}(?r) \wedge \text{hasOrdinalDay}(?s, ?start) \wedge \text{swrlb:add}(?predicted, ?start, 60) \rightarrow \text{hasFloweringPrediction}(?s, ?predicted)$
4. Як географічне розташування поля в Лісостепу впливає на розрахункову дату цвітіння?	Для події посіву $s \in \text{sowing\_time}$ у регіоні $r \in \text{Forest\_Steppe\_Zone}$ розрахувати прогнозну дату цвітіння $\text{predicted}$ .	S4: $\text{sowing\_time}(?s) \wedge \text{occursInRegion}(?s, ?r) \wedge \text{Forest\_Steppe\_Zone}(?r) \wedge \text{hasOrdinalDay}(?s, ?start) \wedge \text{swrlb:add}(?predicted, ?start, 65) \rightarrow \text{hasFloweringPrediction}(?s, ?predicted)$
5. Яка календарна дата сівби зафіксована для конкретного генотипу соняшника?	Визначити порядковий день сівби $val$ для екземпляра генотипу $g \in \text{genetic}$ , пов'язаного з подією сівби $s$ .	S5: $\text{genetic}(?g) \wedge \text{isGenotypeForSowing}(?g, ?s) \wedge \text{hasOrdinalDay}(?s, ?val) \rightarrow \text{hasSowingDayFromEvent}(?g, ?val)$
6. Чи завершено комплексний хімічний аналіз та оцінку вологоутримання для дослідного поля?	Знайти рішення $d \in \text{Decision}$ для зразків ґрунту $x$ , які мають підтверджені хімічні формули та властивості вологоутримання.	S6: $\text{SoilSample}(?x) \wedge \text{Test\_Soil}(?y) \wedge \text{Time}(?t) \wedge \text{Decision}(?d) \wedge \text{isSoilTest}(?x, ?y) \wedge \text{isDecisionForSample}(?d, ?x) \wedge \text{hasAnalysisTime}(?y, ?t) \wedge \text{hasChemicalFormulas}(?y, \text{True}) \wedge \text{hasMoistureRetention}(?y, \text{True}) \rightarrow \text{hasFieldStatus}(?d, \text{ChemicalAndMoistureAnalysisComplete})$

Неформальне компетентнісне запитання	Формальне компетентнісне запитання (логічний запит)	Семантичне SWRL-правило виведення
7. Яка структура та гідрофізичний стан ґрунтового профілю на аналізованій ділянці?	Ідентифікувати статус рішення <i>d</i> , що базується на аналізі текстури ґрунту <i>texture</i> та вмісту гумусу <i>humus</i> .	S7: SoilSample(?x) ^ Test_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasSoilTexture(?y, ?texture) ^ hasHumusContent(?y, ?humus) -> hasFieldStatus(?d, MechanicalCompositionAnalysisComplete)
8. Чи володіє ґрунт ділянки достатнім корисним біологічним та мікробіологічним потенціалом?	Перевірити наявність у зразку <i>x</i> бактерій-фосфатмобілізаторів <i>bacteria</i> та синергетичних мікоризних грибів <i>fungi</i> .	S8: SoilSample(?x) ^ Test_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasMicrobialInhabitant(?x, ?bacteria) ^ hasSynergisticFungi(?x, ?fungi) -> hasFieldStatus(?d, BiologicalComponentAnalysisComplete)
9. Чи верифіковано оптимальні показники вологозабезпечення у глибоких шарах ґрунту (1.0–1.5 м)?	Виявити рішення <i>d</i> , де глибина відбору зразка <i>depth</i> перебуває в діапазоні [1.0; 1.5] метра.	S10: SoilSample(?x) ^ Test_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasSamplingDepth(?y, ?depth) ^ swrlb:greaterThanOrEqual(?depth, 1.0) ^ swrlb:lessThanOrEqual(?depth, 1.5) -> hasFieldStatus(?d, DeepLayerMoistureVerified)
10. Чи існує ризик вимивання поживних речовин через легкий механічний склад підповерхневих шарів?	Сформулювати статус ризику <i>NutrientLeachingRisk</i> для ділянок із глибиною шару $\geq 1.0$ м та піщаною текстурою <i>Sandy</i> .	S11: SoilSample(?x) ^ Test_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasSamplingDepth(?y, ?depth) ^ swrlb:greaterThanOrEqual(?depth, 1.0) ^ hasSubterraneanTexture(?y, Sandy) -> hasFieldStatus(?d, NutrientLeachingRiskDetected)
11. Чи виявлено приховану фітосанітарну загрозу ураження вовчорком соняшниковим у глибоких шарах?	Зафіксувати біотичну загрозу у формі наявності насіння паразитичної рослини <i>Orobanche cumana</i> у глибоких шарах ґрунту.	S12: SoilSample(?x) ^ Test_Soil(?y) ^ Time(?t) ^ Decision(?d) ^ isSoilTest(?x, ?y) ^ isDecisionForSample(?d, ?x) ^ hasAnalysisTime(?y, ?t) ^ hasSamplingDepth(?y, ?depth) ^ hasBiologicalContamination(?y, Orobanche_cumana_Seeds) -> hasFieldStatus(?d, OrobancheDeepInfestationRisk)
12. Як параметри часу посіву конкретного поля	Здійснити логічне успадкування часових міток <i>t</i> від поля <i>f</i> до	S14: PlantSpecimen(?p) ^ isPlantOfField(?p, ?f) ^ hasSowingTime(?f, ?t) -> hasSowingTime(?p, ?t)

Неформальне компетентнісне запитання	Формальне компетентнісне запитання (логічний запит)	Семантичне SWRL-правило виведення
трансляються на вегетуючі рослини цього поля?	індивіда рослини $p$ \in PlantSpecimen.	
13. Чи ідентифіковано наявність ознак абіотичного стресу (наслідків посухи) за візуальним станом органів?	Визначити інтегральний статус ураження рослини $p$ за умови фіксації ознаки в'янення (Wilting) на її вегетативних органах.	S15: PlantSpecimen(?p) ^ hasPart(?p, ?organ) ^ Organ_Condition_Directory(?organ) ^ hasCondition(?organ, Wilting) -> hasOverallStatus(?p, AbioticStress)
14. Чи відповідає поточний стан вегетації рослини параметрам норми та здорового розвитку?	Верифікувати статус здорового розвитку для рослини $p$ за комбінації здорового листка та стадії активного росту.	S16: PlantSpecimen(?p) ^ hasPart(?p, ?leaf) ^ Leaf_Plant_Sample(?leaf) ^ hasLeafCondition(?leaf, Healthy_Leaf) ^ hasCondition(?p, Active_Growth) -> hasOverallStatus(?p, Healthy_Development)
15. Чи зафіксовано біотичний стрес, спричинений пошкодженням листової поверхні шкідниками?	Діагностувати біотичний стрес для рослини $p$ у разі виявлення ознак деструктивного скелетування листка (LeafSkeletonization).	S17: PlantSpecimen(?p) ^ hasPart(?p, ?leaf) ^ Leaf_Plant_Sample(?leaf) ^ hasLeafCondition(?leaf, LeafSkeletonization) -> hasOverallStatus(?p, BioticStress)
16. Правило автоматичного призначення планової десикації соняшнику на основі цифрового фенотипування у фазі R6.	Якщо рослина належить до поля з фазою завершення цвітіння та осипання язичкових квіток, та вологість насіння $\geq 20\%$ , приймається рішення про застосування стандартної технології десикації.	S18: PlantSpecimen(?plant) ^ hasSownPlant(?crop, ?plant) ^ hasGrowthStage(?crop, R6_Normal_Shedding) ^ has_moisture_seed(?plant, ?moisture) ^ swrlb:greaterThanOrEqualTo(?moisture, 20.0) -> isUsedTechnology(?crop, Desiccation_Standard).

### 3.4 Оцінка якості структури онтологічної моделі

Для забезпечення об'єктивного аналізу структурно-функціональних характеристик розробленої онтології соняшника та оцінки її архітектурної цілісності було застосовано метричний підхід, що базується на методології OntoQA. Оцінювання простору знань проводилося за двома системними

категоріями: метрики схеми (провідні структурні тренди TBox) та метрики екземплярів (ступінь наповненості ABox). (Tartir, S. et al. 2005). Оцінювання проводилося за категоріями *метрики схеми* та *метрики екземплярів*.

Метрика *багатства зв'язків* (Relationship Richness, *RR*)

$$RR = \frac{|P|}{|SC| + |P|} = \frac{68}{269 + 68} = 0,2.$$

Отримане значення свідчить, що модель поєднує строго структуровану ієрархію концептів із помірною щільністю функціональних зв'язків. Опис складних морфологічних та причинно-наслідкових взаємозв'язків між елементами системи забезпечується деталізацією через аксіоми доменів (ObjectPropertyDomain = 114) та діапазонів (ObjectPropertyRange = 110). Це дозволяє системі не лише класифікувати об'єкти, а й описувати складні функціональні та морфологічні взаємозв'язки між ними, що суттєво підвищує точність логічного виведення (Reasoning).

Метрика *атрибутивного багатства* (Attribute Richness, *AR*)

$$AR = \frac{|att|}{|C|} = \frac{53}{139} \approx 0,4.$$

Поточний показник вказує на те, що специфікація параметрів стану системи орієнтована на ключові вузли, а подальше масштабування моделі потребуватиме розширення діапазонів властивостей даних для слабодеталізованих класів.

*Метрика багатства успадкування* (Inheritance Richness, *IR<sub>S</sub>*) дозволяє класифікувати архітектуру як планарну (горизонтально-орієнтовану) онтологію. Це підтверджує широке охоплення предметної області цифрового фенотипування при відносно невисокій вертикальній специфікації (глибині) окремих гілок ієрархії.

Формула для розрахунку  $IR_s$

$$IR_s = \frac{\sum_{i=1}^{|C|} |H^C(C_i, C_i)|}{|C|} = \frac{269}{139} \approx 2,0,$$

де  $|H^C(C_i, C_i)|$ — кількість безпосередніх підкласів  $C_l$  (direct subclasses) для класу  $C_i$ .

Метрика багатство класів ідентифікує відносну кількість концептів схеми (TBox), які мають принаймні один фактичний екземпляр (ABox) у базі знань. Для розрахунку використано кількість заповнених класів

$$CR = \frac{|C'|}{|C|} = \frac{72}{139} \approx 0,5.$$

Отримане значення фіксує, що знання моделі розподілені рівномірно й охоплюють 50% задекларованих концептуальних областей, що підтверджує відповідність розробленої структури TBox практичним завданням цифровізації фенотипових ознак.

Метрика середня сукупність (Average Population,  $P$ )

$$P = \frac{|I|}{|C|} = \frac{145}{139} \approx 1.$$

Результат відображає рівномірне стартове наповнення структури з середнім коефіцієнтом близько одного індивіда на клас, що обґрунтовує необхідність подальшої автоматизованої генерації екземплярів за результатами DATA Processing.

Метрику зв'язності (Cohesion,  $Coh$ ) розраховуємо

$$Coh = \frac{68 (ObjectPropertyAssertion) + 53 (DataPropertyAssertion)}{\text{загальна кількість індивідумів}} = \frac{102}{145} = 0,83.$$

З позицій системного аналізу це свідчить про те, що 83% інформаційного простору онтології HELIANTUS охоплено активними предикатними зв'язками, які безпосередньо задіяні у формуванні логічних висновків. Отримане значення підтверджує достатню щільність семантичної мережі на поточному етапі наповнення бази знань. Для оптимізації процесів дедуктивного виведення за допомогою Reasoners потрібне розширення мережі об'єктних властивостей між індивідами з метою наближення інтеграційного показника до цільового значення 1,0 (абсолютно зв'язний граф знань).

Аналіз метрики важливості термінальних класів (Terminal classes) (*Imp*) представлений в таблиці 3.N.

Дані отримано по результату SPARQL query

```

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
SELECT ?DirectClass (COUNT(?individual) AS ?ExactCount)
WHERE {
    ?individual rdf:type ?DirectClass .
# Відсікаємо загальні системні класи
    FILTER(?DirectClass != owl:NamedIndividual && ?DirectClass !=
    owl:Thing)
# Універсальний фільтр для глибоких (прямих) класів
    FILTER NOT EXISTS {
        ?individual rdf:type ?SubClass .
        SubClass rdfs:subClassOf ?DirectClass .
        FILTER(?SubClass != ?DirectClass)
    }
}
GROUP BY ?DirectClass ORDER BY DESC(?ExactCount)

```

Таблиця 3.4.1. - Важливість термінальних класів (Terminal classes)

№	Прямий клас предметної області (Найглибший рівень специфікації)	Фактична кількість унікальних екземплярів (ExactCount)	Розрахунок відносної важливості (Imp=ExactCount /145)	Важливість концепту у % (Сума = 100%)
1	Diseases (Хвороби)	18	18 / 145	12%
2	Pests (Шкідники)	9	9 / 145	6%
3	Time (Часові мітки операцій)	8	8 / 145	5,5%
4	LeafCondition (Стан листкової поверхні)	8	8 / 145	5,5%
5	HeadPosition (Положення кошика)	8	8 / 145	5,5%
6	Stage_04 Flowering (Фаза цвітіння)	8	8 / 145	5,5%
7	Sowing time (Терміни сівби)	7	7 / 145	4,8%
8	Chemical_composition (Хімічні проби)	7	7 / 145	4,8%
9	Stage_R6 Ray Flowers Shedding	6	6 / 145	4%
10	Head_Shape (Форма кошика соняшника)	6	6 / 145	4%
11	Grade (Товарний сорт)	6	6 / 145	4%
12	Decision (Управлінські рішення)	6	6 / 145	4%
13	Genotype (Генотипи соняшника)	6	6 / 145	4%
14	Forest Steppe Zone (Лісостепова зона)	5	5 / 145	3,4%
15	SoilSample (Проби ґрунту ділянок)	5	5 / 145	3,4%
16	WholePlant_Condition (Загальний стан біомаси)	5	5 / 145	3,4%
17	Seed (Насінневий матеріал)	4	4 / 145	2,7%
18	Polissya_Zone (Зона Полісся)	4	4 / 145	2,7%
19	Color_Spectrum (Колірні моделі детекції)	4	4 / 145	2,7%
20	minor_elements (Вміст мікроелементів)	4	4 / 145	2,7%
21	PlantOverallStatus (Інтегральний статус)	4	4 / 145	2,7%
22	ShapeSeed (Морфологічні ознаки насіння)	4	4 / 145	2,7%
23	Mechanical_composition (Текстура ґрунту)	4	4 / 145	2,7%
24	PlantSpecimen (Зразки рослин)	4	4 / 145	2,7%
25	no gear (Енергоощадний обробіток)	3	3 / 145	2,1%
26	oilseed_G (Олійний напрям використання)	3	3 / 145	2,1%
27	'major_elements_(NPK)' (Макроелементи)	3	3 / 145	2,1%
28	Growth Stages (Загальні стадії росту)	3	3 / 145	2,1%
29	BroomrapeRace (Раси вовчка)	2	2 / 145	1,4%
30	Test Soil (Параметри аналізу ґрунту)	2	2 / 145	1,4%
31	clay (Глинисті фракції)	2	2 / 145	1,4%
32	sand (Піщані фракції)	2	2 / 145	1,4%
33	Leaf_Plant_Sample (Метрики листка)	2	2 / 145	1,4%

№	Прямий клас предметної області (Найглибший рівень специфікації)	Фактична кількість унікальних екземплярів (ExactCount)	Розрахунок відносної важливості (Imp=ExactCount /145)	Важливість концепту у % (Сума = 100%)
34	Steppe Zone (Степова зона)	2	2 / 145	1,4%
35	FieldPlot (Експериментальна ділянка)	2	2 / 145	1,4%
36	Region (Адміністративні регіони)	2	2 / 145	1,4%
37	soil (Загальний тип ґрунту)	2	2 / 145	1,4%
38	confectionery_G (Кондитерський напрямок)	2	2 / 145	1,4%
39	Stem_Plant_Sample (Характеристики стебла)	1	1 / 145	0,7%
40	application_time (Час внесення добрив)	1	1 / 145	0,7%
41	decorative_G (Декоративний напрямок)	1	1 / 145	0,7%
42	Mechanical_composition (Базовий клас)	1	1 / 145	0,7%
43	humus (Коефіцієнт гумусу)	1	1 / 145	0,7%
Ра- зом	Чисті предметні індивіди ABox	145	<b>145 / 145</b>	<b>100,00%</b>

Локальні показники повноти для термінальних класів відповідають межах їхнього прямого наповнення, що доводить структурну стійкість бази знань та її готовність до проведення дедуктивного виведення за допомогою машин логічного виведення (Reasoners).

Метрика повноти (Fullness,  $F$ ) Визначає коефіцієнт інформаційної ємності піддерев онтологічної схеми і характеризує щільність популяції бази знань у розрізі окремих функціональних гілок. Повнота  $F(C_i(I))$  концепту  $C_i(I)$  детермінується як відношення фактичної кількості екземплярів, що належать до піддерева з коренем у  $C_i(I)$  до загальної потужності простору індивідів бази знань  $|I| = 137$ .

Створимо SPARQL query

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
```

```
SELECT ?SubRootClass (COUNT(DISTINCT ?individual) AS ?SubTreeInstances)
```

```
WHERE {
```

```
# Знаходимо всі класи системи
```

```

?SubRootClass rdf:type owl:Class .
FILTER(?SubRootClass != owl:NamedIndividual && ?SubRootClass != owl:Thing)
# Зв'язуємо індивідів як з самим класом, так і з усіма його підкласами будь-якої
глибини (знак це позначає патерн rdfs:subClassOf*)
?subClass rdfs:subClassOf* ?SubRootClass .
?individual rdf:type ?subClass .
}
GROUP BY ?SubRootClass
ORDER BY DESC(?SubTreeInstances)

```

За результатами виконання структурного SPARQL-скрипту встановлено, що домінуючу структурну вагу має підсистема агротехнологічного контексту та біотичних стресів agritechnology Imp = 82,5%. Контур інформаційного супроводу селекції selection акумулює 13,2% емпіричного наповнення системи, фіксуючи параметри вихідного аналізу зразків, що дозволяє оцінити критерії успішності експериментальних досліджень. Найменшу відносну вагу має блок генотипічних характеристик genetic (Imp = 4,4%), що відображає високу ступінь абстракції та типізації базових ліній соняшника: невелика кількість екземплярів-декларацій генотипів компенсується їхньою високою зв'язністю з іншими класами через об'єктні властивості. Таким чином, архітектура онтології HELIANTUS орієнтована на інтенсивний аналіз зовнішніх чинників впливу на культуру, забезпечуючи репрезентативну базу фактів для прийняття рішень у системі підтримки збалансованого агровиробництва.

Метрика багатство успадкування (Inheritance Richness, IR<sub>c</sub>).

SPARQL-запит для розрахунку Inheritance Richness

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
PREFIX owl: <http://www.w3.org/2002/07/owl#>
```

```

SELECT ?SubRootClass (COUNT(DISTINCT ?subClass) AS ?TotalNodesInSubTree)
(COUNT(?directSub) AS ?TotalBranches) (AVG(?directSubCount) AS ?Local_IR)

```

```

WHERE {
  ?SubRootClass rdf:type owl:Class .
  FILTER(?SubRootClass != owl:NamedIndividual && ?SubRootClass != owl:Thing)

  # Знаходимо всі вузли всередині піддерева
  ?subClass rdfs:subClassOf* ?SubRootClass .

  # Для кожного вузла знаходимо кількість підкласів
  {
    SELECT ?subClass (COUNT(?directSub) AS ?directSubCount)
    WHERE {
      ?directSub rdfs:subClassOf ?subClass .
      FILTER(?directSub != ?subClass)
    }
    GROUP BY ?subClass
  }
  ?directSub rdfs:subClassOf ?subClass .
  FILTER(?directSub != ?subClass)
}
GROUP BY ?SubRootClass
ORDER BY DESC(?Local_IR)

```

Отримані значення метрики багатства успадкування в табл. 3.4.2. показують, що різні частини онтології побудовані за різними принципами, залежно від їхнього призначення.

Головний клас *agritechnology* має найвищий показник розгалуження  $IR = 36,2$ . Це пояснюється тим, що він спроектований «у ширину»: під ним знаходиться велика кількість суміжних понять одного рівня (типи технологій, кліматичні умови, географічні регіони). Така структура є оптимальною, оскільки дозволяє описати всі зовнішні фактори та агротехнічні операції, не створюючи занадто довгих і заплутаних ієрархічних ланцюжків «клас-підклас».

Таблиця 3.4.2. Розрахунок локальних показників багатства успадкування для ключових піддерев онтології HELIANTUS

№	Базовий концепт (Корінь піддерева)	Загальна кількість вузлів у піддереві	Загальна кількість відгалужень	Значення метрики багатства успадкування (IR)
1	agritechnology	29	378	36,2
2	selection	5	54	17,05
3	information_about_specimens	6	32	13,97
4	soil	5	11	9,43
5	Phenotype	4	11	9,29
6	Technology	2	3	7,50
7	genetic	4	6	7,09
8	Signs_of_plant	1	7	7,00
9	Type_of_technology	1	9	6,00
10	genotype	3	25	5,86
11	Region	3	30	5,17
12	climatic_condition	1	6	5,00
13	Organ_Condition_Directory	1	5	5,00
14	Time	1	19	4,00
15	Stage_04_Flowering	1	6	4,00
16	Organ_Plant	1	3	4,00
17	Success_result	1	4	4,00
18	Growth_Stages	7	16	3,86
19	PlantSpecimen	2	16	3,57
20	head	2	12	3,00
21	grade	1	3	3,00
22	Stage_02_Vegetative	1	2	2,00
23	Stage_03_Budding	1	1	2,00
24	hybrid	1	2	2,00
25	Natural_Climatic_Zone	1	5	2,00
26	Stage_05_Seed_Filling	1	1	2,00
27	The_living_component	1	2	2,00

№	Базовий концепт (Корінь піддерева)	Загальна кількість вузлів у піддереві	Загальна кількість відгалужень	Значення метрики багатства успадкування (IR)
28	Seed	1	8	2,00
29	Parasitic_plants	2	6	2,00
30	Chemical_composition	1	3	2,00
31	Test_Soil	1	8	2,00
32	Mechanical_composition	4	3	1,50
33	Terrain_Type	1	1	2,00
34	Water-holding_capacity	1	2	2,00
35	Schem_of_selection	1	1	2,00
36	HeadPosition	1	1	1,00
37	Stage_01_Emergence	1	4	1,00
38	Orobanche_cumana_Wallr	1	8	1,00
39	Stage_06_Maturity	1	4	1,00

Навпаки, класи, пов'язані з розвитком рослини та біологічними загрозами (наприклад, Growth\_Stages з IR = 3,86 та Parasitic\_plants з IR = 2,00), побудовані «у глибину». Вони мають низьке значення метрики, що підтверджує логіку послідовного уточнення біологічних ознак — від загального поняття (наприклад, паразитична рослина) до конкретної деталізації (конкретна раса вовчка соняшникового Orobanche cumana). Таке архітектурне рішення є найбільш ефективним для роботи програм логічного виведення (Reasoners). Оскільки сильне розгалуження зосереджене лише на верхніх рівнях онтології, а біологічні класи йдуть чіткими вертикальними ланцюжками, комп'ютер витрачає мінімальну кількість часу та обчислювальних ресурсів на аналіз бази знань і швидко генерує управління рішень.

Метрика багатство зв'язків Relationship Richness,  $RR_c$ .

SPARQL-запит для розрахунку

PREFIX rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>>

PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

```

PREFIX owl: <http://www.w3.org/2002/07/owl#>
SELECT  ?Class  (COUNT(DISTINCT  ?definedProp) AS  ?DefinedProperties)
(COUNT(DISTINCT ?activeProp) AS ?ActiveProperties) (ROWNUM() AS ?Row)
WHERE {
  # Знаходимо предметні класи
  ?Class rdf:type owl:Class .
  FILTER(?Class != owl:NamedIndividual && ?Class != owl:Thing)

  # Визначаємо властивості, закріплені за класом через Domain
  {
    ?definedProp rdfs:domain ?Class .
  }
  UNION
  {
    ?Class rdfs:subClassOf [
      rdf:type owl:Restriction ;
      owl:onProperty ?definedProp
    ] .
  }
}

# Перевіряємо, чи є хоча б один індивід цього класу з цією властивістю
OPTIONAL {
  ?individual rdf:type ?Class .
  ?individual ?definedProp ?value .
  BIND(?definedProp AS ?activeProp)
}
}
GROUP BY ?Class
ORDER BY DESC(?ActiveProperties)

```

Розрахунок метрики багатства зв'язків (RR, табл. 3.4.3.) дозволяє оцінити, наскільки ефективно реальні дані (ABox) використовують логічний потенціал розробленої схеми (TBox).

Таблиця 3.4.3. Оцінка метрики багатства зв'язків для концептів онтології HELIANTUS

№	Клас онтології	Задекларовано властивостей (Pdefined)	Активно властивостей (Pactive)	Значення метрики багатства зв'язків (RR=Pactive/Pdefined)
1	agritechnology	48	16	0,333 (33,3%)
2	Time	8	5	0,625 (62,5%)
3	Test_Soil	21	5	0,238 (23,8%)
4	Growth_Stages	8	5	0,625 (62,5%)
5	genetic	6	4	0,667 (66,7%)
6	Decision	5	2	0,400 (40,0%)
7	soil	5	2	0,400 (40,0%)
8	SoilSample	5	2	0,400 (40,0%)
9	FieldPlot	7	2	0,286 (28,6%)
10	PlantSpecimen	4	1	

Найвищі показники щільності заповнення зв'язків зафіксовано для концептів genetic (RR = 66,7%), Time (62,5%) та Growth\_Stages (62,5%). Це свідчить про те, що більшість задекларованих властивостей для цих класів (наприклад, прив'язка генотипів до фаз вегетації та часових інтервалів) активно використовуються під час моделювання.

Для макро-класу agritechnology показник становить 33,3% (активні 16 властивостей із 48 задекларованих). З позицій системного аналізу це є нормальним значенням для високорівневих концептів. Воно вказує на те, що схема має значний «запас міцності» для подальшого масштабування бази знань. Наявність класів із нульовим значенням активних властивостей (наприклад, selection, Phenotype) відображає поточний етап наповнення онтології. Логічні зв'язки для них повністю спроектовані у схемі, що гарантує цілісність моделі, а їхнє поступове заповнення даними цифрового фенотипування не вимагатиме перебудови загальної архітектури системи. Таким чином, архітектура онтології

HELIANTUS визнається придатною для стабільної роботи машин логічного виведення.

Метрика зв'язності (Connectivity,  $C_n$ ) призначена для оцінки інтенсивності взаємодії екземплярів певного класу з іншими об'єктами бази знань.

SPARQL-запит для пошуку семантичних хабів (Connectivity)

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX owl: <http://www.w3.org/2002/07/owl#>

```

SELECT ?Class (COUNT(DISTINCT ?otherIndividual) AS ?Connectivity)
WHERE {
  # Беремо індивіда, його клас та об'єктний зв'язок до іншого індивіда
  ?individual rdf:type ?Class .
  FILTER(?Class != owl:NamedIndividual && ?Class != owl:Thing)

  {
    ?individual ?property ?otherIndividual .
  }
  UNION
  {
    ?otherIndividual ?property ?individual .
  }
  # Перевіряємо, що зв'язок об'єктний
  FILTER(?property != rdf:type && isURI(?otherIndividual))
  ?otherIndividual rdf:type ?otherClass .
  FILTER(?Class != ?otherClass && ?otherClass != owl:NamedIndividual && ?otherClass
!= owl:Thing)
}
GROUP BY ?Class
ORDER BY DESC(?Connectivity)

```

Оцінка метрики локальної зв'язності та виявлення архітектурних центрів онтології HELIANTUS представлена в табл. 3.4.4.

Головним семантичним хабом системи є клас *agritechnology* Conn = 30, виступає центральним інтегратором, який об'єднує екологічний контекст, параметри ґрунту та технологічні операції в єдину мережу. Другим за значущістю хабом визначено клас *Time* Conn = 20, що вказує на високу темпоральну зв'язність моделі — більшість подій, фаз росту та агротехнічних дій жорстко прив'язані до часових інтервалів. Також чітко простежується логічна сумісність у парі *genetic* та *genotype* (обидва мають Conn = 11). Це доводить, що блоки внутрішнього потенціалу соняшника правильно інтегровані з товарними сортами (*grade*), напрямками використання (*oilseed\_G*) та зовнішніми чинниками вирощування.

Таблиця 3.4.4. Оцінка метрики локальної зв'язності та виявлення архітектурних центрів онтології HELIANTUS

№	Клас (Концепт онтології)	Значення метрики зв'язності (Conn, ціле число)	Статус вузла в топології графа знань
1	<i>agritechnology</i>	30	Головний семантичний хаб (Центр ABox)
2	<i>Time</i>	20	Семантичний хаб другого рівня
3	<i>SoilSample</i>	12	Активний зв'язний вузол
4	<i>sowing_time</i>	11	Активний зв'язний вузол
5	<i>grade</i>	11	Направлений дескриптор
6	<i>genetic</i>	11	Контекстний хаб гілки
7	<i>genotype</i>	11	Контекстний хаб гілки
8	<i>oilseed_G</i>	8	Профільний вузол
9	<i>soil</i>	7	Локальний вузол
10	<i>Region</i>	7	Локальний вузол
11	<i>Technology</i>	5	Локальний вузол
12	<i>Type_of_technology</i>	5	Локальний вузол

№	Клас (Концепт онтології)	Значення метрики зв'язності (Conn, ціле число)	Статус вузла в топології графа знань
13	no_rear	5	Локальний вузол
14	Growth_Stages	5	Локальний вузол
15	PlantSpecimen	4	Локальний вузол
16	Test_Soil	4	Локальний вузол
17	Leaf_Plant_Sample	4	Локальний вузол
18	Decision	4	Управлінський вузол виведення
20	Color_Spectrum	2	Дескриптор комп'ютерного зору
21	BroomrapeRace	2	Вузол фітосанітарного ризику
22	WholePlant_Condition	2	Фенотиповий дескриптор
23	Stage_04_Flowering	2	Темпоральний дескриптор
24	LeafCondition	2	Фенотиповий дескриптор
25	Parasitic_plants	2	Вузол фітосанітарного ризику
26	Orobanche_cumana_Wallr	2	Цільовий вузол паразитарного стресу
27	confectionery_G	2	Профільний вузол

З позицій логічного виведення, наявність стійких хабів із високими цілими значеннями метрики гарантує, що база знань не містить ізольованих концептуальних областей, що дозволяє машинам логічного виведення (Reasoners) будувати безперервні дедуктивні ланцюжки.

Середнє значення метрики читабельності (Readability, Rd)

$$\overline{Rd}(C) = \frac{AnnotationAssertion}{C} = \frac{269}{257} \approx 1.$$

Поточне значення показника свідчить про наявність анотаційної підтримки ключових концептів для забезпечення коректної інтерпретації результатів семантичного аналізу кінцевим користувачем та покращення інтерактивності інтелектуальної системи.

### 3.5 Висновки до розділу 3

У третьому розділі виконано комплексний системний аналіз, формалізацію та оцінку структурно-семантичної якості розробленої бази знань онтології соняшнику HELIANTUS. За результатами проведеного дослідження сформульовано такі наукові та практичні висновки:

1. Удосконалено предметну онтологічну модель агросистеми вирощування соняшнику HELIANTUS, яка забезпечує семантичну інтеграцію генетичних, фенотипових, агрокліматичних та технологічних даних у межах єдиного інформаційного простору. Запропонований підхід передбачає трансформацію результатів цифрового фенотипування у формалізовані знання на основі концепцій TBox та ABox, що створює можливість застосування механізмів логічного виведення для автоматизованого аналізу причинно-наслідкових зв'язків у системі «генотип — середовище».

2. Обґрунтовано архітектурну спроможність та форму схеми онтології TBox. Розрахунок локальних показників багатства успадкування IR виявив диференційовану топологію графа. Головний макро-концепт agritechnology характеризується високим показником розгалуження  $IR = 36,2$ , що підтверджує його планарну архітектуру, орієнтовану на охоплення широкого спектра середовищних чинників без надмірного збільшення глибини вкладеності. Навпаки, блоки біотичних стресів (Parasitic\_plants,  $IR = 2,0$ ; Orobanche\_cumana\_Wallr,  $IR = 1,0$ ) реалізовано як фрагменти строгої вертикальної ієрархії для послідовного логічного уточнення рас вовчка соняшникового. Такий розподіл локальних топологічних метрик оптимізує часові витрати машин логічного виведення, нівелюючи ризик експоненціального зростання обчислювальної складності.

3. Доведено високу інформаційну насиченість та когнітивну точність дефініцій. Завдяки систематизації числових дескрипторів цифрового фенотипування (геометричних параметрів, колірних спектрів та характеристик

органів рослини), онтологія акумулює 53 унікальні властивості даних (*Datatype Properties*) при 139 класах схеми. Отриманий інтегральний показник атрибутивного багатства становить  $AR = 0,4$ , що математично підтверджує достатню щільність параметризованих слотів, необхідних для всебічної цифрової паспортизації експериментальних зразків.

4. Проведено верифікацію репрезентативності та стану заповнення простору фактів (ABox). Шляхом застосування селективних SPARQL-запитів із локалізацією індивідів виключно на рівні термінальних класів усунуто ефект множинного наслідування та сформовано несуперечливий масив обсягом  $|I| = 139$  індивідів. Метрика багатства класів показує, що база знань охоплює 50% концептів таксономії. Аналіз структурної важливості підтвердив раціональну концентрацію інформаційного наповнення навколо цільових зон діагностики фітосанітарного стану: сегмента Diseases (12%) та Pests (6%), що безпосередньо відповідає прикладним задачам дослідження.

5. Математично підтверджено високу системну зв'язність семантичної мережі. Розрахований коефіцієнт інтеграції бази знань становить  $Cohesion = 83\%$  при показнику інтенсивності зв'язків  $Connectivity \approx 2$ , що повністю виключає ризик виникнення ізольованих, логічно відірваних структур даних. Виявлені ключові архітектурні хаби — класи *agritechnology* ( $Conn = 30$ ) та *Time* ( $Conn = 20$ ) — виступають головними семантичними інтеграторами, забезпечуючи безперервність побудови дедуктивних ланцюжків при генерації управлінських рішень.

6. Встановлено належний рівень семантичної інтерпретованості онтології. Показник читабельності ( $Rd\ 1$ ) засвідчує стабільну наявність метаданих та анотаційного супроводу для ключових концептуальних вузлів системи. Поточна насиченість дескрипторами є достатньою для коректної верифікації результатів семантичного пошуку кінцевим користувачем та правильної валідації правил логічного виведення.

Визначено перспективні напрями розвитку та масштабування розробленої моделі. Відкрита архітектура онтології HELIANTUS забезпечує можливість її подальшого розширення. Наступним етапом еволюції системи є автоматизація наповнення простору фактів (ABox) безпосередньо з відкритих репозиторіїв наукових даних.

Основні наукові результати розділу опубліковано в працях автора (Бакурова та ін., 2023; Bakurova, Vedmedeva, Vedmedev & Tereschenko, 2023; Bakurova, Vedmedeva, Vedmedev, Tereschenko & Shyrokograd, 2023)

## 4 ЦИФРОВА ТЕХНОЛОГІЯ ФЕНОТИПУВАННЯ ТА ФОРМУВАННЯ DATASETS

Матеріал розділу присвячено побудові цифрової технології фенотипування із формуванням датасетів, яка забезпечує автоматизований збір та збереження фенотипових ознак на основі онтологічної моделі. Традиційні методи оцінки фізичних характеристик рослин упродовж їхнього вегетативного розвитку передбачають ручне спостереження та фіксацію, що характеризується суб'єктивністю оцінок, низькою точністю, відсутністю стандартизації та потребує значних часових витрат, здійснюється як у польових умовах, так і в лабораторіях (Bakurova, A. et al. 2023a). Ручних вимірювань уникають через неточності, що виникають внаслідок людської помилки, на яку може впливати суб'єктивне сприйняття, індивідуальний досвід та інші особисті фактори. Застосування методів цифрової оцінки фізичних характеристик рослин допомагає мінімізувати такі помилки, забезпечуючи стандартизовані та об'єктивні вимірювання. На застосування методів цифрової оцінки у фенотипуванні рослин впливають фактори дослідницького середовища, системи візуалізації, цільового організму, вибору ознак, архітектури моделі тощо (Li Z., et al. 2020). Не існує універсального робочого процесу, який можна застосувати до всіх типів сільськогосподарських культур без відповідної адаптації стратегій отримання та обробки зображень (Xie C., Yang C. 2020). Кожне завдання фенотипування вимагає вибору відповідних алгоритмів, адаптованих до конкретних характеристик цільового об'єкта або проблеми.

В розділі 4 розроблено цифрову технологію фенотипування насіння соняшника. Насіння як об'єкт фенотипування демонструє унікальні характеристики з точки зору форми, розміру, забарвлення поверхні. Наприклад, насіння соняшнику демонструє значну фенотипічну різноманітність. Тільки за забарвленням перикарпію було виявлено до 12 груп (Gorohivets N. A., Vedmedeva E. V. 2016). Поширені кольори включають чорний або темно-сірий з темними смугами або без них, але є також більш чіткі фенотипи, такі як повністю біле

насіння або насіння з контрастними білими, чорними або сірими смугами (Vedmedeva K., Machova T. 2022). Деякому насінню бракує чіткого шару перикарпії, і натомість воно має коричневі або жовтуваті смуги безпосередньо на насінневі оболонці. При застосуванні до інших культур питання стає ще складнішим, оскільки різноманітність забарвлення та морфології насіння часто більша (Manianga A., Bose C., Sreeraj S. 2024).

Для досягнення надійних наукових результатів експериментальні умови повинні забезпечувати високу роздільну здатність та відтворюваність вимірюваних ознак об'єкта фенотипування. Розробка математично обґрунтованого методу цифрового фенотипування рослинних об'єктів дозволить забезпечити необхідну точність та відтворюваність результатів, що є принципово важливим для прийняття рішень у селекції.

Експериментальне впровадження проводилося на генетичному матеріалі Інституту олійних культур НААН України.

#### **4.1 Цифрова технологія фенотипування із формуванням інтегрованих наборів даних (datasets)**

У межах дослідження розроблено та алгоритмічно реалізовано цифрову технологію фенотипування об'єктів агросистем (на прикладі насіння соняшнику), яка інтегрує методи класичного детермінованого комп'ютерного зору, проміжні структуровані репозиторії табличного типу та інструменти онтологічного інжинірингу знань.

Удосконалено цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets), яка забезпечує автоматизований збір, стандартизацію та збереження фенотипових ознак біологічних об'єктів у контрольованих лабораторних умовах.

На відміну від існуючих технологій, де накопичення експериментальних даних відбувається у вигляді розрізнених статичних таблиць чи файлових архівів

без урахування їхньої подальшої логічної структури (що блокує можливість їх безпосереднього машинного аналізу рушіями логічного виведення), удосконалена технологія реалізована у вигляді цілісної інтегрованої системи, структуровану декомпозицію якої представлено на рис. 4.1.1

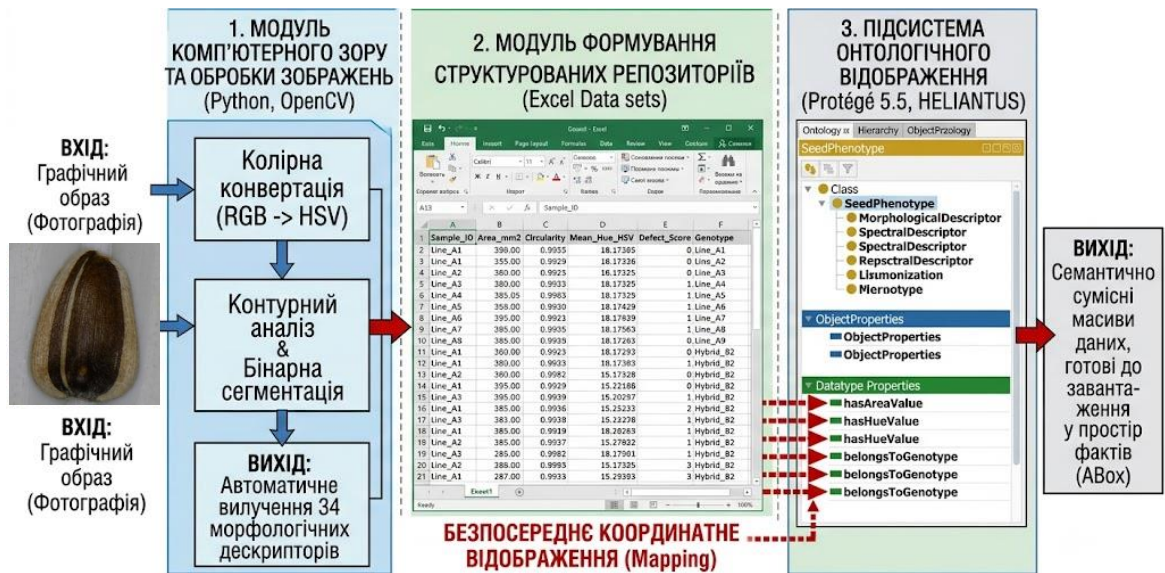


Рис. 4.1.1 - Структурно-функціональна архітектура цифрової технології фенотипування та формування семантично сумісних наборів даних (Datasets) (створено автором із застосуванням ШІ gemini, промт в додатку А).

Компонентна архітектура технології. Структурно розроблена цифрова технологія фенотипування складається з трьох взаємопов'язаних модулів, які забезпечують наскрізний інформаційний потік від первинного цифрового сигналу до логічно впорядкованого масиву знань:

Модуль первинної обробки зображень та вилучення дескрипторів. Цей компонент базується на методах класичного комп'ютерного зору (реалізованого на мові Python із використанням бібліотеки OpenCV). Він забезпечує автоматичний перехід від двовимірного графічного образу насінини (матриці пікселів) до числових характеристик. Модуль виконує колірну конвертацію у простір HSV для нівелювання завад освітлення, реалізує алгоритми контурного

аналізу та бінарної сегментації, автоматично вилучаючи 34 морфологічні ознаки (метричні параметри, ексцентриситет форми, колірні градієнти перикарпію) без залучення людського фактора.

Модуль синтаксичної стандартизації та формування структурованих репозиторіїв. Даний модуль відповідає за генерацію сумісних датасетів. Параметризовані результати обробки графічних образів експортуються у проміжні файли табличного типу (Excel-паспорти зразків). Ключовою особливістю модуля є те, що кожна колонка згенерованої таблиці репрезентує фіксований числовий або текстовий вимір, синтаксично уніфікований під стандарти збереження типів даних.

Підсистема онтологічного відображення. Компонент забезпечує семантичну сумісність між емпіричними вимірами та базою знань. Він реалізує логічне проектування структури репозиторіїв (Excel-датасетів) на розроблену семантичну схему онтології соняшнику HELIANTUS. Координати колонок таблиць жорстко закріплені за відповідними властивостями даних (Datatype Properties) простору TBox, що створює передумови для безпосередньої трансляції структури сформованих наборів даних у простір фактів (Assertion Box — ABox).

Логіка функціонування та системний ефект. Впровадження дворівневої схеми збереження даних «табличний репозиторій — онтологічний простір» дозволяє вирішити ключову проблему системного аналізу в агросфері — несуперечливе об'єднання кількісних експериментальних вимірів із якісними експертними знаннями предметної області.

Функціонування технології за рахунок безпосереднього координатного відображення емпіричних метрик на семантичну схему знань дає можливість здійснювати наскрізну синтаксичну та семантичну стандартизацію первинних параметрів безпосередньо на етапі їх генерації кодом комп'ютерного зору. Оскільки структура репозиторіїв виступає точним інформаційним інтерфейсом

(зліпком) онтології, усувається необхідність у повторній реструктуризації чи ручній переробці файлів при наповненні бази знань актуальними екземплярами.

Це дозволяє забезпечувати динамічне оперативне наповнення простору фактів (ABox) верифікованими цифровими паспортами конкретних селекційних ліній та гібридів, а також гарантувати високу обчислювальну готовність накопичених лабораторних даних для машин логічного виведення (Reasoners). Сформований таким чином граф знань стає основою для автоматизованого побудови дедуктивних ланцюжків у системі підтримки прийняття рішень (СППР) при моделюванні складних причинно-наслідкових зв'язків у взаємодії «генотип — середовищний стрес — керований агротехнологічний вплив».

#### **4.2 Математична постановка задачі цифрового фенотипування рослинних об'єктів**

Введемо математичне формулювання задачі цифрового фенотипування рослинних об'єктів як оптимізаційну задачу вибору параметрів зйомки об'єкту фенотипування, що дає найменшу похибку вимірювання морфологічних ознак.

Нехай задано тривимірний простір пошуку  $X$ , який є декартовим добутком множин параметрів зйомки:

$$X = L \times F \times C$$

де:

$L = \{l_1, l_2, \dots, l_n\}$  — множина значень інтенсивності освітлення (наприклад, у люксах або рівнях потужності освітлювача);

$F = \{f_1, f_2, \dots, f_m\}$  — дискретна множина варіантів колірної фону;

$C = \{c_1, c_2, \dots, c_k\}$  — множина колірних режимів (монохромні або спектральні канали).

Потужність множини  $|X| = n \times m \times k$ .

Отже, множина  $X = \{(l_j, f_j, c_j)\}, j = \overline{1, |X|}$ .

Визначимо цільову функцію  $E(l, f, c)$  як функцію похибки оцінки морфологічної ознаки  $R$ :

$$E(l, f, c) = \frac{|R - R_{\text{ref}}|}{R_{\text{ref}}} \rightarrow \min, \quad (4.1)$$

де:

$R$ — значення морфологічної ознаки, обчислене за цифровим зображенням, отриманим при параметрах  $(l, f, c) \in X$ ;

$R_{\text{ref}}$  — еталонне значення характеристики, виміряне контактним методом.

Необхідно знайти такий вектор оптимальних параметрів  $x^* = (l^*, f^*, c^*)$ , що мінімізує значення функції похибки на множині  $X$ :

$$x^* = \operatorname{argmin} E(l, f, c) \quad \text{при} \quad (l, f, c) \in X, \quad (4.2).$$

За умови обмеження на гранично допустиму похибку  $\delta$ :

$$E(l^*, f^*, c^*) \leq \delta. \quad (4.3)$$

### 4.3 Створення датасету фотозображень насінини соняшнику

Створення колекції фотозображень насінини соняшнику відбувалося зі застосуванням стаціонарного приладу для фенотипування, який розроблений в Інституті олійних культур Національної академії аграрних наук України (Алієв Е. Б., Ведмедєва К.В., 2024). Дослідження відбувалося в рамках виконання науководослідної роботи 17.01.01.15.Ф «Формування колекцій олійних культур, виділення та створення джерел і донорів цінних господарських ознак» в Інституті олійних культур Національної академії аграрних наук України, яке направлено на наповнення баз даних фенотипових ознак стандартизованими цифровими даними. Прилад для фенотипування дозволяє фотографувати об'єкти

розміром до 3 см та обирати умови зйомки, а саме інтенсивність освітлення, контрастність фону та кольоровий режим. Прилад для фенотипування оснащений кольоровими світлодіодними лампами та ізолюючою трубкою для усунення зовнішніх перешкод освітлення. Інтенсивність та колір освітлення можна регулювати за допомогою пульта дистанційного керування.

Колекція фотозображень насінини соняшнику містить 3000 фотографій 250 ліній соняшнику. Колекційні зразки насіння бувають різної форми: трикутної, овальної, закругленої, близької до прямокутної. Для отримання тривимірних параметрів насінини використовуються фотографії, що містять дві проєкції кожної насінини. Забарвлення має як різний колір, так й різний малюнок. Колір змінюється від білого до чорного з різними відтінками, від жовтого до темно-коричневого, або бордового. За малюнком бувають однокольорові, смугасті з невеликою або великою кількістю смужок (Gorohivets N. A., Vedmedeva E. V., 2016).

#### **4.4 Метод цифрового фенотипування рослинних об'єктів на основі зображень**

Розв'язок задачі (4.1)-(4.3) залежить від морфологічних ознак об'єкта фенотипування. Фенотипування насіння зі суттєво різними кольоровими властивостями потребує оптимізації параметрів зйомки. Насіння може відрізнитися за кольором, візерунком або текстурою поверхні. Для кожного нового типу насіння необхідно вибрати набір параметрів зйомки, який дозволить мінімізувати похибку в оцінці обраних морфологічних ознак.

Набір параметрів  $(l^*, f^*, c^*)$  зйомки об'єктів дослідження, який дозволяє досягти найменшої похибки вимірювання, тобто мінімум функції  $E(l, f, c)$  (4.1), на всіх можливих наборах значень параметрів, будемо називати оптимальним.

Метод цифрового фенотипування рослинних об'єктів на основі зображень (МЦФЗ) дозволяє визначити оптимальний набір параметрів зйомки (4.2)  $x^* = (l^*, f^*, c^*)$ .

МЦФЗ складається з таких етапів.

1. Формування еталонної морфологічної моделі  $\Omega$  та критеріїв оцінки. На цьому етапі визначаються еталонна морфологічна модель  $\Omega$  з ключовими морфологічними ознаками насінини (геометрія, колір, текстура), які мають бути ідентифіковані. Відбирається множина насінин  $S, |S| = N$ , що відповідають еталонній морфологічній моделі  $\Omega$ .

2. Генерація сценаріїв модельного досліду. Здійснюється серія ітераційних запусків зйомки довільних  $a$  насінин з множини  $S$  при всіх наборах параметрів зйомки  $(l, f, c) \in X$ , де  $l$  є рівень інтенсивності світла. Фон  $f$ : вибір підкладки з різною відбивною здатністю та колірним спектром. Колірний режим  $c$ : перемикання між монохромним або спектральними каналами. Вибір параметра  $a$  визначається дослідником з огляду на морфологічні ознаки насінини, природна різноманітність насіння, для якого підбираються умови зйомки, трудомісткість операцій. Результатом є сформована множина зображень  $P = \{p_{t,j}, t = \overline{1, a}; j = \overline{1, |X|}\}$ .

3. Формування підмножини  $X'$ . На цьому етапі проводиться обробка зображень  $p_{t,j} \in P, t = \overline{1, a}; j = \overline{1, |X|}$  з метою формування підмножини  $X' \subset X$ , яка включає такі набори параметрів зйомки  $(l_j, f_j, c_j)$ , що дозволяють визначити контури насінини на загальному фоні. Характеристична функція множини  $X'$

$$\chi_{X'}(l_j, f_j, c_j) = \begin{cases} 1, \text{ якщо контури визначено на зображеннях } p_{t,j} \forall t = \overline{1, a} \\ 0, \text{ в протилежному випадку} \end{cases}.$$

Отже, множина наборів параметрів зйомки, що дозволяють визначити контури насінини на загальному фоні, є множина  $X' = \{(l_i, f_i, c_i)\}, i = \overline{1, |X'|}$ .

Набори параметрів зйомки  $X \setminus X'$ , які не дозволяють розпізнати насінину на зображенні, далі не застосовуються.

#### 4. Формування вибіркової сукупності для статистичного експерименту.

Визначається гранично допустима похибка  $\delta$  для оцінки морфологічної ознаки  $R$ .

Здійснюється серія ітераційних запусків зйомки для насінин множини  $S$  при всіх наборах параметрів зйомки  $(l_i, f_i, c_i) \in X', i = \overline{1, |X'|}$ . Формується множина зображень  $P' = \{p'_{g,i}\}, g = \overline{1, N}; i = \overline{1, |X'|}$ , де  $g$  – номер насінини множини  $S, g = \overline{1, N}; i$  – номер набору параметрів зйомки  $(l_i, f_i, c_i) \in X', i = \overline{1, |X'|}$ . Для кожного зображення  $p'_{g,i} \in P', g = \overline{1, N}; i = \overline{1, |X'|}$  визначається морфологічна ознака насінини  $R_{g,i}$ . Далі визначається похибка оцінки морфологічної ознаки  $R_{g,i}$

$$E_g(l_i, f_i, c_i) = E_{g,i} = \frac{|R_{g,i} - R_{ref}|}{R_{ref}}.$$

Знаходиться середня арифметична похибки оцінки морфологічної ознаки при наборі параметрів зйомки  $(l_i, f_i, c_i) \in X'$  для  $N$  зображень

$$E(l_i, f_i, c_i) = E_i = \frac{1}{N} \sum_{g=1}^N E_{g,i}.$$

Перевіряється виконання умови для наперед визначеній гранично допустимій похибки  $\delta$

$$E_i \leq \delta. \quad (4.4)$$

При не виконанні умови (4.4) набір параметрів зйомки  $(l_i, f_i, c_i)$  вилучається з множини  $X'$ , тобто  $X' = X' \setminus (l_i, f_i, c_i)$ .

По завершенню етапу 4 сформовано множину  $X'$  наборів параметрів зйомки, що забезпечують гранично допустиму похибку  $\varepsilon$  при визначенні морфологічної ознаки  $R$ .

5. Присвоєння статусу «оптимальний» набору параметрів зйомки. Зі всіх наборів параметрів зйомки  $(l_i, f_i, c_i) \in X'$ ,  $i = \overline{1, |X'|}$ , який отримано на попередньому етапі, обираються два такі, що дали найменші похибки  $E_i$  у визначенні ознаки  $R$  та не перевищили  $\varepsilon$ .

Далі визначається статистична значимість знайдених переваг певних наборів параметрів. За результатами обирається набір параметрів, який назвемо «оптимальний», якщо статистично значимою є перевага цього набору. У разі відсутності переваги обидва режими отримують статус «оптимальний».

6. Формалізація цифрового представлення та формування рекомендацій по режиму зйомки для насіння такого типу. Кінцевий етап, на якому фізичний об'єкт отримує опис у вигляді структурованого набору числових даних з фіксацією оптимальних параметрів зйомки.

#### **4.5 Визначення оптимального набору параметрів зйомки**

Розглянемо теоретичні підстави для присвоєння статусу «оптимальний» певному набору параметрів на 5 етапі, а також побудови інтервалу довіри.

По завершенню етапу 4 сформовано множину  $X'$  наборів параметрів зйомки, що забезпечують гранично допустиму похибку  $\varepsilon$  середньої арифметичної похибки оцінки морфологічної ознаки  $R$ . Серед отриманих наборів параметрів необхідно вибрати такий, що забезпечує мінімальне значення похибки оцінки морфологічної ознаки  $R$ , що відповідає цільовій функції (4.1).

Результати експерименту отримано при впливі випадкових факторів, тобто результуюча похибка ідентифікації є випадковою величиною. Для отримання об'єктивних висновків щодо ефективності застосування певних наборів

параметрів зйомки подальший аналіз проводиться з використанням апарату статистичної перевірки гіпотез.

Для комплексного оцінювання точності визначення морфологічної ознаки R застосовується дисперсійний аналіз (Wackerly et al., 2008). Задача полягає у декомпозиції загальної дисперсії похибки вимірювання на систематичні складові, що зумовлені керованими факторами (освітленість  $I$ , спектральні характеристики фону  $f$  та вибір колірної каналу освітлення  $c$ ), та випадкову складову. У межах дисперсійного аналізу під факторами розуміють контрольовані чинники, що детермінують кінцевий результат дослідження. Конкретні прояви цих чинників визначаються як рівні фактора, що формалізуються у номінальних або порядкових шкалах. Кількісний показник досліджуваної ознаки, що змінюється під впливом зазначених факторів, визначається як відгук. Критерієм вибору оптимальної стратегії зйомки є мінімізація сумарного впливу значущих факторів та мінімізації відгуку - похибки оцінки (4.1).

Методи дисперсійного аналізу класифікують на дві основні категорії залежно від структури дослідження та математичного апарату.

Метод Фішера (ANOVA-Analysis of Variations). Застосовується в межах однофакторних моделей. Його сутність полягає в декомпозиції загальної варіації (дисперсії) спостережень на дві складові: внутрішньогрупову (зумовлену випадковими чинниками) та міжгрупову (зумовлену безпосереднім впливом досліджуваного фактора).

Метод загальної лінійної моделі (GLM). Використовується для реалізації багатофакторного аналізу. Даний підхід базується на синтезі методів кореляційного та регресійного аналізу, що дозволяє оцінювати не лише окремих вплив факторів, а й їхню взаємодію.

Для формалізації впливу параметрів середовища на точність ідентифікації геометричних ознак об'єкта, процес зйомки розглядається як багатофакторний

експеримент, де вимірювана похибка  $E$  описується адитивною лінійною моделлю:

$$E_{ij\dots n} = \mu + \alpha_i + \beta_j + \dots + (\alpha\beta)_{ij} + \dots + e_{ij\dots n},$$

де  $\mu$  — загальне середнє значення похибки;  $\alpha_i, \beta_j$  — головні ефекти контрольованих факторів;  $(\alpha\beta)_{ij}$  — ефект взаємодії факторів, що характеризує нелінійну залежність точності від комбінації чинників дії;  $e_{ij\dots n}$  — вектор випадкових похибок.

Математична модель представлена у матричному виді, що дозволяє враховувати кількість факторів у більше ніж один:

$$E = B\beta + \varepsilon,$$

де  $E$  — вектор спостережуваних значень розмірністю  $N \times 1$ ;  $B$  — матриця експерименту розмірністю  $N \times u$ ;  $\beta$  — вектор невідомих параметрів ефектів, оцінювання яких проводиться методом найменших квадратів;  $\varepsilon$  — вектор залишків (неврахованих стохастичних впливів).

Оцінювання параметрів моделі проводиться методом найменших квадратів:  $\hat{\beta} = (B^T B)^{-1} B^T E$ .

Для врахування факторів, що мають три та більше якісних рівнів, застосовується метод бінарного кодування. Фактор з  $k$  рівнями декомпоновано на  $(k-1)$  незалежних векторів-стовпців матриці плану  $B$ , що забезпечує її повний стовпцевий ранг та запобігає виникненню мультиколінеарності при розрахунку інформаційної матриці  $B^T B$ .

Застосування цих методів базується на виконанні припущень Гаусса-Маркова, які гарантують, що отримані оцінки  $\beta$  будуть найкращими лінійними незміщеними оцінками (Wackerly et al., 2008).

Нульове математичне сподівання похибки  $E(\varepsilon_i) = 0$  свідчить про відсутність систематичного зміщення моделі, при якому випадкові відхилення ідентифікованих параметрів насінини взаємно компенсуються.

Гомоскедастичність  $\text{Var}(\varepsilon_i) = \sigma^2 = \text{const.}$  Дисперсія випадкової похибки залишається постійною для всіх рівнів факторів. Перевірка проводиться тестом Левена.

Відсутність автокореляції  $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$  для  $i \neq j$ . Похибки окремих вимірювань є незалежними. Перевірка проводиться за критерієм Дурбіна-Уотсона.

Фактори впливу є незалежними. Матриця плану експерименту  $B$  має повний стовпцевий ранг, що забезпечує лінійну незалежність її векторів-стовпців. Це гарантує, що матриця  $B^T B$  є невиродженою (позитивно визначеною), що дозволяє отримати однозначний розв'язок задачі оцінювання параметрів методом найменших квадратів

Нормальність розподілу похибок  $\varepsilon \sim N(0, \sigma^2)$ . Виконання цієї умови є критичним для застосування F-критерію Фішера в межах дисперсійного аналізу.

Для перевірки відповідності емпіричних даних теоретичному закону розподілу застосовується критерій Шапіро-Вілка (Wackerly et al., 2008).

В однофакторному дисперсійному аналізі для статистичної обробки даних, що не підпорядковуються закону нормального розподілу, доцільно використовувати критерій Краскела – Уолліса (1952 р.). Як ранговий метод, він дозволяє перевірити однорідність вибірок за умови, що їхня кількість перевищує дві. Критерій є розширенням методу Манна – Уїтні та базується на припущенні, що за відсутності систематичного впливу фактора всі вибірки походять із однієї генеральної сукупності аналіз (Wackerly et al., 2008).

У випадках двофакторного дисперсійного аналізу, коли порушуються базові припущення щодо застосування параметричного критерію, зокрема, умова нормальності розподілу або гомоскедастичності, доцільно використовувати непараметричний ранговий критерій Фрідмана (Friedman, M.

1937). Цей критерій не залежить від типу розподілу статистичних даних і базується на аналізі рангових послідовностей у межах сформованих блоків спостережень.

У випадках, коли структура експериментальних даних передбачає багатофакторний вплив при порушенні припущень нормальності, для статистичного висновку доцільно використовувати метод вирівняного рангового трансформу (ART) або критерій Шеіра — Рея — Гейра, які дозволяють поширити логіку непараметричного аналізу на складні моделі, забезпечуючи перевірку гіпотез щодо впливу окремих чинників та їхніх взаємодій як декомпозицію варіації рангових показників.

Незважаючи на те, що дисперсійний аналіз дозволяє ідентифікувати наявність статистично значущих факторів, він не забезпечує диференціацію конкретних рівнів факторів, що зумовлюють ці розбіжності. Для вирішення цієї задачі та локалізації значущих відмінностей доцільно застосовувати метод множинних порівнянь Шеффе, адаптований для аналізу пов'язаних вибірок.

З метою ідентифікації гомогенних груп серед досліджуваних стратегій застосовується критерій Тьюкі (HSD — Honestly Significant Difference). Цей метод дозволяє провести попарне порівняння всіх рівнів фактора при забезпеченні контролю групової ймовірності помилки першого роду.

Для визначення частки варіації похибки, що зумовлена безпосередньо впливом досліджуваного фактору, розраховується показник  $\eta^2$ , що відображає потужність впливу контрольованого чинника відносно стохастичного шуму.

Формула для розрахунків

$$\eta^2 = \frac{SS_{\text{fact}}}{SS_{\text{total}}},$$

де

$SS_{\text{fact}}$  (Sum of Squares for Factor): сума квадратів відхилень, зумовлена впливом фактора.

$SS_{\text{total}}$  (Total Sum of Squares): загальна сума квадратів відхилень вибірки.

Значення  $\eta^2$  змінюється в діапазоні від 0 до 1.

Малий ефект  $0.00 \leq \eta^2 < 0.06$ : фактор майже не впливає. Середній ефект  $0.06 \leq \eta^2 < 0.14$ ): фактор має помітний вплив, але стохастичні процеси домінують. Великий ефект ( $\eta^2 \geq 0.14$ ): фактор є визначальним.

У випадку зіставлення показників, для яких не підтверджується нормальне розподілення, та які отримані на одній і тій же групі в двох різних умовах, є доцільним застосувати Т-критерій Вілкоксона. Т-критерій Вілкоксона — це непараметричний статистичний метод, призначений для оцінювання розбіжностей між двома залежними (пов'язаними) вибірками. Підтвердження достовірності зміни в середньому по групі проводиться по інтенсивності індивідуальних зрушень. Доведено, що Т-критерій Вілкоксона забезпечує збереження статистичної потужності для малих вибірок. При малих вибірках і порушенні нормальності Т-критерій Вілкоксона є більш потужним, ніж t-критерій Стьюдента.

Також потрібно вибирати параметр  $N$ , тобто кількість насінини множини  $S$ ,  $g = \overline{1, N}$ . Параметр  $N$  визначає об'єм вибірки зображень  $P'$ , що забезпечує статистичну потужність запропонованих методів.

На кроці 6 фізичний об'єкт отримує опис у вигляді структурованого набору числових даних, які визначено на основі вибіркової сукупності. Визначимо довірчий інтервал для морфологічної ознаки, що визначається за кількісною шкалою. Це дозволяє встановити межі, у яких із заданою надійністю  $p$  знаходиться істинне значення. Звуження довірчого інтервалу при фіксованому рівні значущості є додатковим критерієм стабільності обраного набору параметрів зйомки.

З припущення про нормальність розподілу визначеного значення морфологічної характеристики  $R$ , з урахуванням того, що дисперсія генеральної сукупності є апіорі невідомою, побудова довірчого інтервалу для математичного сподівання  $M[R]$  визначається з використанням t-розподілу Стьюдента. При обсязі вибірки  $n$  та заданому рівні надійності  $\alpha$ , довірчий

інтервал вибірки визначається на основі критичного значення t-статистики для  $n-1$  ступенів свободи

$$\bar{x} - t_{\alpha, n-1} \frac{s}{\sqrt{n}} \leq R \leq \bar{x} + t_{\alpha, n-1} \frac{s}{\sqrt{n}}$$

де:

$\bar{x}$  — вибіркове середнє арифметичне,

$s$  — вибіркове середньоквадратичне відхилення,

$n$  — обсяг вибірки,

$t_{\alpha, n-1}$  — критичне значення t-критерію Стьюдента.

При відхиленні від нормального розподілу визначеного значення морфологічної характеристики  $R$ , оцінювання довірчих інтервалів для математичного сподівання похибки ідентифікації здійснювалося за допомогою непараметричного методу Bootstrap (Efron В. 1979). Непараметричний метод Bootstrap базується на генерації 10 000 реплікацій вихідної вибірки шляхом рандомізованого вибору з поверненням. Побудований емпіричний розподіл дозволяє визначити межі довірчого інтервалу за перцентильним методом, що забезпечує високу репрезентативність оцінок в умовах апіорно невідомого закону розподілу експериментальних даних та наявності асиметрії відгуку. Після проведення 10 000 ітерацій бутстрепа всі отримані значення середніх упорядковуються у варіаційний ряд від мінімального до максимального. Для побудови 95% довірчого інтервалу застосовується перцентильний метод, який передбачає відсікання по 2,5% результатів з обох кінців розподілу. Діапазон, що залишився між цими двома точками, і є шуканим інтервалом, який з 95% імовірністю містить істинне значення досліджуваного параметра.

#### **4.6 Розробка програмного забезпечення вивчення морфологічних ознак насінини соняшнику за фотозображеннями**

Інтегровані набори даних (datasets) включають цифрове представлення морфологічних та фенотипових характеристик насінини, які визначено в цифровій моделі у розділі 2. Для формування наборів даних створено спеціалізоване програмне забезпечення. Програмне забезпечення складається з таких блоків: визначення метричних розмірів та форми, забарвлення насінини та фіксування смугастості за фотозображенням.

Програмні продукти для аналізу зображень реалізовано мовою програмування Python. Програму розрахунків розроблено в середовищі інтегрованої розробки PyCharm Community Edition 2022.2.4 із використанням базових та спеціалізованих модулів і бібліотек: os, cv2 (OpenCV), imutils, numpy, time, а також класу Pool модуля multiprocessing для паралельних обчислень.

Програмне забезпечення для розпізнавання метричних розмірів та форми насінини

На рисунку 4.6.1 представлено алгоритм розпізнавання метричних розмірів, форми, забарвлення насінини, який створено автором із застосуванням ШІ gemini ( промт в додатку А).

Програмне забезпечення розроблялося для аналізу фотозображень насіння соняшнику, що виконано в лабораторних умовах на стаціонарному приладі для фенотипування. Оригінальна фотографія містить зображення насінини, відцентроване по довжині вздовж діагоналі знімку. Це обумовлюється розташуванням предметного столику стаціонарного приладу, на якому розміщувалася насінинка для фотографування. Тому, на першому кроці зображення насінини повертається на кут в 315 градусів та обрізаються чорні ділянки зображення. Кут повернення постійний. Координати обрізки можна

змінювати у разі необхідності для кожного виду насінини. На другому етапі проводимо обробку зображення.

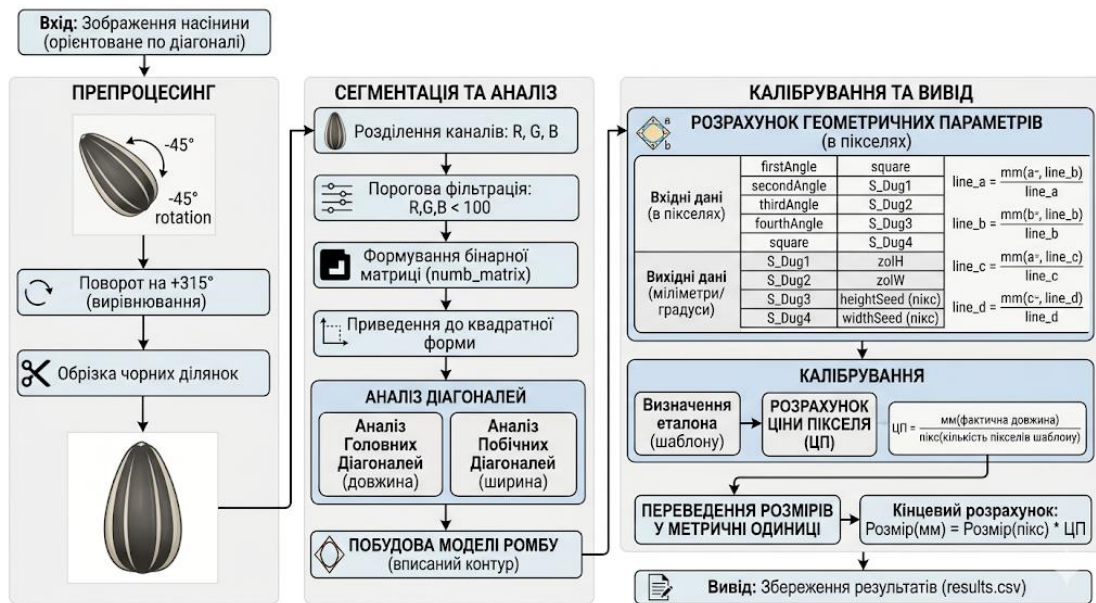


Рисунок 4.6.1 – Алгоритм розпізнавання метричних розмірів, форми, забарвлення насінини (створено автором із застосуванням ШІ gemini, промт в додатку А)

Для стабільного вимірювання параметрів насінини незалежно від дефектів її країв впроваджено модель вписаного ромба, орієнтованого за головною та допоміжною осями, що передбачає визначення лінійних розмірів (довжини та ширини) за діагоналями ромба; оцінку форми та видовженості за внутрішніми кутами та повною площею моделі; аналіз локальної асиметрії шляхом сегментації ромба на чотири симетричні трикутники з розрахунком їхніх площ. (рис.4.6.2).

Розрахунок відрізків довжини та ширини відбувається за формулами:

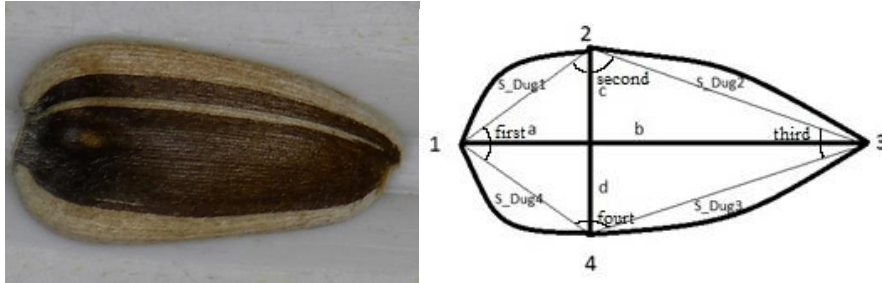
$$line\_a = y - y\_min\_r\_coordinate;$$

$$line\_b = y\_max\_r\_coordinate - y;$$

$$line\_c = x - x\_min\_c\_coordinate;$$

$$line\_d = x\_max\_r\_coordinate - x,$$

де  $line\_a$  та  $line\_b$  -складові довжини насінини,  $lineC$  та  $lineD$  – складові ширини насінини (рис. 4.6.2 б).



а) фото смугастої насінини б) модель для вимірювання метричних характеристик та форми насінини

Рисунок 4.6.2 – Фото насінини соняшнику та модель для розрахунків метричні характеристики насінини

Набір даних (datasets) формується як результати розрахунків метричних характеристик по кожному фотозображенню, які організовано в таблицю. Для кожного фотозображення створюється окремий запис з унікальним ключем у вигляді ім'я файлу відповідного фотозображення. Перший стовпчик – ім'я файлу фотозображення. Далі розташовано виміри кутів, площі трикутників та площі ділянок за межами трикутників, які позначені:  $firstAngle$  - перший кут;  $secondAngle$  - другий кут;  $thirdAngle$  - третій кут;  $fourthAngle$  - четвертий кут;  $square$  - площа насінини в пікселях;  $S\_Dug1$  – площа 1 ділянки за межами трикутника;  $S\_Dug2$  - площа 2 ділянки за межами трикутника;  $S\_Dug3$  - площа 3 ділянки за межами трикутника;  $S\_Dug4$  - площа 4 ділянки за межами трикутника;  $zolH$ - співвідношення половинок ширини;  $zolW$ - співвідношення половинок довжини;  $heightSeed$  - ширина насінини;  $widthSeed$  - довжина насінини.

*Алгоритм вимірювання геометричних параметрів насіння.*

Зображення насінини повертається на кут в 315 градусів та обрізаються чорні ділянки зображення.

Зображення розділяється на колірні канали: Red, Green, Blue. Застосовується порогова фільтрація. Створюється бітова маска, де пікселі, що відповідають умові ( $R < 100$ ,  $G < 100$ ,  $B < 100$ ), вважаються частиною об'єкта.

Шляхом попіксельного множення масок каналів отримується підсумкова бінарна матриця (`numb_matrix`), де 1 — об'єкт, 0 — фон.

Для спрощення математичних обчислень отримана матриця приводиться до квадратного вигляду для коректного обчислення діагоналей.

*Розрахунок габаритних розмірів (у пікселях).*

Аналіз головних діагоналей: у циклі по всіх діагоналях, паралельних головній, підраховує суму одиниць у кожній. Максимальна сума відповідає одному з лінійних розмірів (висоті).

Аналіз побічних діагоналей: аналогічний процес проводиться для діагоналей, паралельних побічній, для визначення другого габаритного розміру (ширині).

Калібрування та конвертація в метричні одиниці. Для переведення пікселів у міліметри використовується «шаблон» (об'єкт відомого розміру). Розрахунок ціни пікселя:

Розмір пікселя = Фактична довжина шаблону (мм)/(Кількість пікселів шаблону).

Отримані розміри насіння у пікселях множаться на фізичну величину одного пікселя.

Дані записуються у файл `results.csv` за структурою стовпці: назва файлу; висота; ширина. Формат роздільник — крапка з комою (;).

Часова складність алгоритму  $O(N^2)$ , де  $N$  — сторона квадратної матриці.

Для мінімізації похибки, яка залежить від розміру пікселя, необхідно:

Використовувати штатив для уникнення розмиття (`blur`).

Фіксувати відстань від камери до об'єкта.

Проводити повторне калібрування при зміні камери або фокусної відстані.

Обчислення відбуваються у циклі, для обробки усіх фотозображень, поміщених у теку. Фрагмент таблиці, що містить інформацію про метричні характеристики насінини, представлено на рис. 4.6.3.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1	Name	lineA	lineB	lineC	lineD	firstAngle	secondAn	thirdAngl	fourthAng	square	S_Dug1	S_Dug2	S_Dug3	S_Dug4	zolH	zolW	heightSe	widthSeed	
2	1-1-1b.jpg	259.0	361.0	153.0	129.0	57.048268	126.45989	63.103819	113.38801	137405	11328.961	15790.559	13313.609	9551.8693	1.8431372	2.3938223	282.0	620.0	
3	1-1-1p.jpg	261.0	387.0	245.0	143.0	71.906893	104.47424	62.607851	121.01101	207092	20697.483	30689.372	17912.572	12080.572	1.5836734	2.4827586	388.0	648.0	
4	1-1-2b.jpg	257.0	448.0	192.0	128.0	63.238541	120.03867	56.888658	119.83412	165960	11627.336	20268.663	13512.442	7751.5574	1.6666666	2.7431906	320.0	705.0	
5	1-1-2p.jpg	235.0	464.0	285.0	123.0	78.120076	97.948425	54.903261	129.02823	230752	20702.712	40876.845	17641.586	8934.8550	1.4315789	2.9744680	408.0	699.0	
6	1-1-3b.jpg	231.0	484.0	175.0	93.0	59.076274	122.97483	47.865936	130.08295	150682	11576.039	24254.557	12889.565	6151.8378	1.5314285	3.0952380	268.0	715.0	
7	1-1-3p.jpg	218.0	595.0	244.0	155.0	83.634216	109.48120	54.723337	112.16123	209335	7730.1207	21098.265	13402.586	4910.5275	1.6352459	3.7293577	399.0	813.0	
8	1-1-4b.jpg	182.0	508.0	179.0	82.0	68.777788	116.06559	44.023073	131.13353	136577	8417.5661	23495.184	10763.157	3856.0917	1.4581005	3.7912087	261.0	690.0	
9	1-1-4p.jpg	239.0	532.0	213.0	154.0	74.503619	116.47217	57.687081	111.33712	217227	13627.967	30335.058	21932.389	9853.0847	1.7230046	3.2259414	367.0	771.0	
10	1-1-5b.jpg	202.0	481.0	196.0	45.0	56.695146	113.69356	35.100680	154.51061	133922	12416.291	29565.525	6788.0033	2850.6791	1.2295918	3.3811881	241.0	683.0	
11	1-1-5p.jpg	259.0	469.0	189.0	169.0	69.244176	121.93196	63.751141	105.07272	204309	13898.286	25167.168	22503.975	12427.568	1.8941798	2.8108108	358.0	728.0	
12	1-1-6b.jpg	214.0	507.0	216.0	60.0	60.928596	111.65778	38.599837	148.81378	155736	13063.304	30949.043	8596.9565	3628.6956	1.2777777	3.3691588	276.0	721.0	
13	1-1-6p.jpg	195.0	548.0	223.0	140.0	84.508677	109.02463	54.263806	112.20287	219906	13712.795	38536.472	24193.301	8608.9301	1.6278026	3.8102564	363.0	743.0	

Рисунок 4.6.3 – Табличний результат роботи програми визначення метричних складових насіння соняшнику

Файли зображень насінини мають бути формату JPG, назву латиницею, або у вигляді цифр. Фотозображення розміщуються в одну теку з програмою. Файл результатів result.csv розміщується в цій же теці. Результати вимірювань фіксуються у форматі .csv, що є стандартним для обміну науковими даними. Файл result.csv є кросплатформним та сумісним із будь-яким сучасним ПЗ для аналізу структурованої інформації, що гарантує відтворюваність отриманих результатів дослідження.

Визначення забарвлення насіння з можливістю розпізнавати кількість смуг. Насіння соняшнику характеризуються складними текстурними ознаками, зокрема смугастістю з різним градієнтом яскравості — від висококонтрастних переходів до малопомітних змін тону. Оскільки типова вибірка зображень демонструє наявність 3–7 морфологічних смуг, для підвищення роздільної здатності методу та мінімізації похибки дискретизації було обґрунтовано використання восьми діапазонів аналізу ознак. Обґрунтування базується на теоремі Найквіста-Шеннона (Shannon, С.Е.,1949).

Для комп'ютерного зору, відповідно до теореми Найквіста-Шеннона, для того щоб безперервний просторовий текстурний сигнал зображення з обмеженим спектром міг бути точно відновлений за дискретними відліками без ефекту накладання частот (аліасингу), просторова частота дискретизації має щонайменше вдвічі перевищувати максимальну просторову частоту аналізованої текстури. Якщо розглядати смугастість насінини соняшнику як періодичний просторовий сигнал зміни яскравості за однією з координат, то вибір восьми інтервалів задовольняє умову для верхньої межі типової вибірки, дозволяючи ідентифікувати навіть слабо виражені стохастичні характеристики забарвлення. Такий дискретний підхід обрано ітеративно, що дозволяє мінімізувати обчислювальну складність алгоритмів комп'ютерного зору та знизити трудомісткість цифрового фенотипування.

В алгоритмі проводиться аналіз кольорних характеристик шляхом сканування масиву пікселів за шириною об'єкта. Процедура сегментації смуг ґрунтується на ідентифікації міжпіксельних градієнтів з найбільшою амплітудою. Сформовані таким чином вісім зон підлягають статистичній обробці, в результаті якої для кожної ділянки формується кольорна характеристика у вигляді середнього значення в системі RGB.

*Алгоритм сегментації та спектрального аналізу смугастості насіння.*

Визначення лінії сканування. На основі побудованого ромба обирається горизонтальний зріз, що проходить через центр насінини за її шириною.

Формується одновимірний масив, що містить значення R, G, B кожного пікселя вздовж цієї лінії.

Розраховується різниця значень пікселя  $i$  та пікселя  $i+1$ .

$$\Delta C = C_{i+1} - C_i,$$

де  $C$  — інтенсивність каналу.

Фіксація найбільших значень  $\Delta C$ , які визначають перехід від темної смуги до світлої. На основі виявлених  $\Delta C$  та загальної ширини об'єкта лінія сканування розподіляється на вісім рівних діапазонів. Вибір восьми діапазонів є надлишковим щодо кількості фізичних смуг (від 3 до 7), що дозволяє точно зафіксувати як самі смуги, так і перехідні зони між ними.

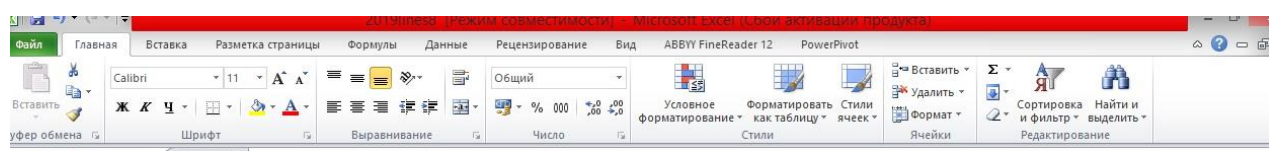
Для кожного з 8 діапазонів обчислюється середнє арифметичне значення для кожного колірної каналу Mean RGB.

Формування спектрального профілю. Результатом є вектор з 24 значень (8 діапазонів по 3 канали), який описує «цифровий паспорт» забарвлення насінини.

Оцінка контрастності. Порівняння середніх значень між сусідніми діапазонами дозволяє визначити рівень контрастності смуг (висококонтрастні або близькі за забарвленням).

Отримані середні значення додаються до підсумкової таблиці results.csv як додаткові стовпці для кожного фото.

Фрагмент набору даних (datasets), що містить інформацію про забарвлення смужок насінини соняшнику, представлено на рис. 4.6.4.



Name	red_avg2	green_avg2	blue_avg2	red_avg3	green_avg3	blue_avg3	red_avg4	green_avg4	blue_avg4	red_avg5	green_avg5	blue_avg5	red_avg6	green_avg6	blue_avg6	red_avg7	green_avg7	blue_avg7	#
162-1b.jpg	63	52	35	56	45	29	41	32	14	34	26	11	26	22	17	146	146	146	
162-1p.jpg	85	78	58	68	61	42	56	49	31	45	37	26	25	22	19	66	59	51	
162-2b.jpg	27	24	19	46	43	36	81	76	54	74	67	47	92	88	77	46	41	37	
162-2p.jpg	44	40	33	33	30	27	33	31	29	49	50	46	88	83	75	65	61	47	
162-3b.jpg	60	49	33	79	72	63	111	107	101	122	118	94	68	59	41	65	53	30	
162-3p.jpg	32	24	18	54	48	44	21	17	13	64	61	55	147	148	147	63	82	94	
162-4b.jpg	42	38	25	90	83	59	72	66	42	60	54	31	36	32	25	63	59	54	
162-4p.jpg	32	28	26	58	54	53	76	70	69	66	57	48	38	32	26	146	147	146	
162-5b.jpg	68	60	40	47	36	11	21	16	12	25	20	17	56	56	54	69	69	69	
162-5p.jpg	33	28	20	25	21	20	63	59	57	149	149	149	83	96	110	68	83	99	
167-1b.jpg	19	16	13	40	38	39	25	21	19	30	27	22	70	69	70	83	82	88	
167-1p.jpg	18	14	12	39	36	32	58	53	47	70	67	58	89	87	80	151	152	152	
167-2b.jpg	74	68	46	54	50	39	45	38	22	13	10	7	22	19	18	149	150	150	
167-2p.jpg	81	73	36	63	55	19	16	13	12	22	19	17	60	59	54	75	75	74	
167-3b.jpg	77	71	47	60	53	25	42	35	10	18	13	10	23	22	21	148	149	149	
167-3p.jpg	100	93	63	86	78	41	66	57	22	19	16	11	52	49	40	149	150	150	
167-4b.jpg	29	25	20	38	37	31	73	69	53	35	27	19	71	70	67	149	149	149	
167-4p.jpg	12	9	8	12	9	8	44	41	36	13	10	8	56	51	48	94	89	86	
167-5b.jpg	27	22	18	48	46	35	76	73	60	25	19	13	52	46	46	19	15	15	
167-5p.jpg	58	43	12	21	16	11	16	13	10	25	19	15	71	66	55	155	156	156	

Рисунок 4.6.4 – Фрагмент набору даних (datasets), що містить інформацію про забарвлення смужок насінини соняшнику

#### 4.7 Результати експериментів застосування МЦФЗ

Для тестування спеціалізованого програмного продукту використовувалися шість кольорів фону та чотири кольори освітлення.

На першому етапі застосування МЦФЗ для подальшого опрацювання було обрано три типи насіння соняшнику: чорне, біле та смугасте. На другому етапі МЦФЗ було створено колекцію зображень об'єктів фенотипування. Для виконання наступних етапів МЦФЗ було розроблено спеціалізоване програмне забезпечення на мові програмування Python. Перший модуль програмного забезпечення обчислює площу зерна, ширину зерна (`heightSeed`) та його підрозділи (`lineC` та `lineD`), а також довжину зерна (`widthSeed`) та його підрозділи (`lineA` та `lineB`), максимальні та мінімальні пороги кольору в колірній моделі RGB (Kwiek, P., & Jakubowska, M. 2024). Відповідні змінні граничного кольору включають: `LimitRedMax`, `LimitRedMin`, `LimitGreenMax`, `LimitGreenMin`, `LimitBlueMax` та `LimitBlueMin`.

На третьому етапі МЦФЗ виділення об'єктів дослідження (насіння) від фону ґрунтується на аналізі векторних значень інтенсивності пікселів. Реалізація обчислювальних алгоритмів на піксельному рівні, попри підвищену ресурсомісткість, забезпечує необхідну прецизійність при ідентифікації морфологічних ознак. Застосування 8-бітного квантування рівнів інтенсивності гарантує трансформацію спектральної відбивної здатності біологічних об'єктів у детермінований масив цифрових значень, структурований за базовими спектральними складовими моделі RGB. Для об'єктивізації вибору параметрів зйомки впроваджено критерій максимізації цифрового контрасту. Оптимальним визначено режим, що забезпечує максимальну відносну різницю інтенсивностей між об'єктом фенотипування та фоном, що дозволяє мінімізувати похибку сегментації контурів і підвищити вірогідність цифрової моделі.

Зображення зерна були отримані за різних умов з використанням білого або червоного фону та чотирьох типів освітлення: білого, червоного, синього та

зеленого. Розрахунок розміру та площі не проводиться, якщо кольори зерна та фону занадто схожі, що може погіршити точність сегментації. Наприклад, зразок з позначкою 255-1 відповідає смугастому насінню з білими смугами, що створює труднощі для розрізнення за кольором. Репрезентативні зображення, що ілюструють вплив різних параметрів фотографії, представлені на рисунку 4.7.1.

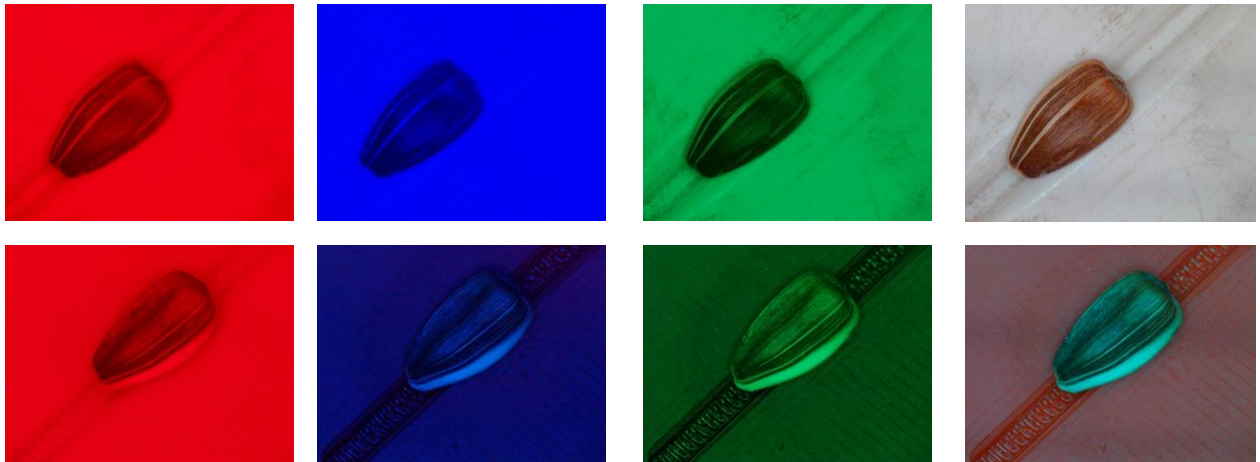


Рисунок 4.7.1. - Зображення смугастого насіння за різних умов зйомки: верхній ряд — на білому фоні, нижній ряд — на червоному фоні; зліва направо — фотографії у червоному, синьому, зеленому та білому світлі.

Представлені зображення, разом із фотографіями інших контрастних візерунків, були оброблені за допомогою розробленого програмного забезпечення. Результати експортовано у файл.

Назви полів таблиці такі:

Name - назва файлу обробки, в якому перші цифри - назва зразка, перша літера - колір фону, друга літера - колір освітлення;

lineA - перша частина довжини;

lineB - друга частина довжини;

lineC - перша частина ширини;

lineD - друга частина ширини;

square - площа насінини;

heightSeed - ширина насінини;  
widthSeed - довжина насінини;  
LimitRedMin - мінімальне значення червоного;  
LimitRedMax - максимальне значення червоного;  
LimitGreenMin - мінімальне значення зеленого;  
LimitGreenMax - максимальне значення зеленого;  
LimitBlueMin - мінімальне значення синього;  
LimitBlueMax - максимальне значення синього.

При відсутності значимого колірного контрасту між насінням та фоном визначення початкових точок ширини та довжини насіння є неможливим. У таких випадках значення `max_row_coordinates` та `max_col_coordinates` залишаються порожніми (тобто елементи масиву не повертаються).

Наприклад, насіння, розміщене на світло-блакитному фоні, показало такі максимальні значення RGB у вибраній області: R: 3, G: 3, B: 248. Мінімальні значення були: R: 0, G: 0, B: 104. Червоний та зелений канали не показали жодної помітної різниці між насінням та фоном. Автоматично обирається колірний канал з найбільшим цифровим контрастом. У цьому випадку синій (B) канал показав різницю понад 100 одиниць, що робить його найбільш інформативним для сегментації.

Під час початкових експериментів вибір оптимальних параметрів зйомки базується на результатах, представлених у табличній формі. Комбінації освітлення та фону, які повертають "Похибку", автоматично виключаються з подальшого розгляду.

На четвертому етапі МЦФЗ проводиться порівняння отриманих результатів із еталонними показниками для виявлення відхилень та похибок ідентифікації. Еталонними показниками будемо вважати такі, що отримані людиною.

Після збору та обробки зображень, що отримані при всіх можливих наборах параметрів зйомки, проводиться об'єднання окремих CSV-файлів в один набір даних, обчислюється відсоткова різниця між площами, визначеними оператором та програмно. Виділяються два найкращих результата.

На п'ятому етапі МЦФЗ виконується перевірка статистичної значимості переваги одного набору параметрів перед іншим. Далі відображається відповідне ім'я файлу (або двох файлів) фотографії, яке кодує конкретні використані параметри зйомки.

Для апробації та оцінювання ефективності розробленого методу МЦФЗ на першому етапі досліджень з колекції Інституту олійних культур НААН України було сформовано тестову вибірку смугастого насіння соняшнику (рис. 4.7.2).

Для експерименту було відібрано два фенотипові типи насінневого матеріалу: однорідне біле забарвлення (без виражених смуг та елементів малюнка на поверхні), смугасте забарвлення (біле насіння з вираженою поздовжньою смугастістю бежевого/коричневого відтінку). Вибір цих двох типів зумовлений необхідністю перевірки стійкості алгоритмів сегментації контуру та розпізнавання елементів забарвлення (етап 1 та етап 3 МЦФЗ).

Монохромне біле насіння є складним об'єктом для виділення геометричних меж при використанні світлих підкладок, тоді як смугастий перикарпій вимагає точного налаштування колірних каналів для коректного вимірювання площі та інтенсивності самих смуг.

На другому етапі МЦФЗ було створено колекцію зображень насінин кожного виду. Зображення були отримані за допомогою стаціонарного пристрою для фенотипування, розробленого в Інституті олійних культур НААН України (Алієв Е.Б., Ведмедева К.В., 2024).

На третьому етапі було проведено обробку фотозображень з метою виділення геометричних характеристик обраних насінин. На четвертому етапі МЦФЗ було створено колекцію по 30 зображень однієї множини насінин для тих наборів параметрів зйомки, які забезпечили виділення контуру насінини на тлі.

За кожним зображенням було визначено геометричні характеристики обраних насінин, площу насінини, а також визначено еталонне значення площі цієї насінини.

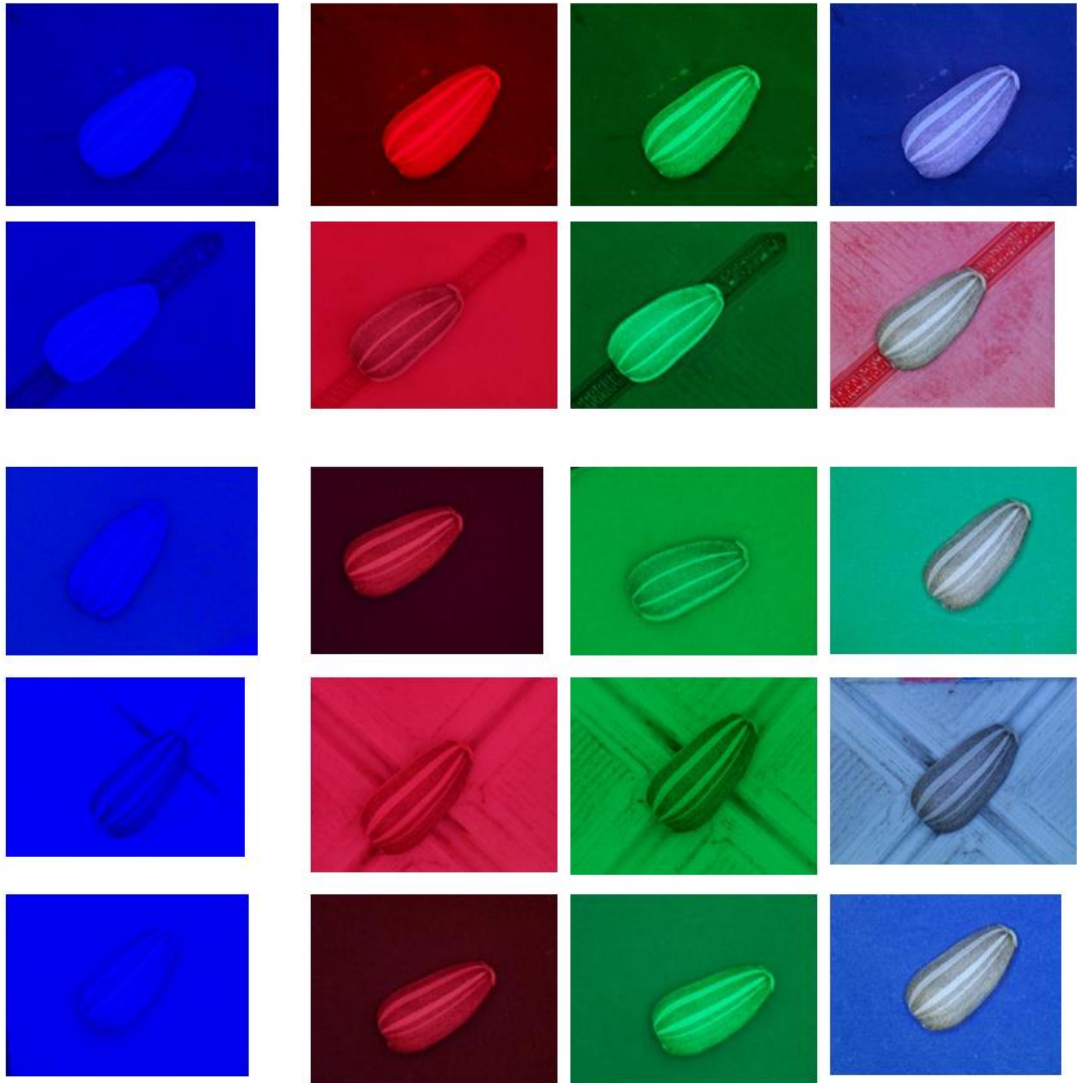


Рисунок 4.7.2 - Зображення смугастого насіння за різного спектрального освітлення: зверху вниз по стовпцях: синє (90 CRI), червоне (24 CRI), зелене (45 CRI), біле (80 CRI); по рядках на різному тлі: чорне, червоне, зелене, біле, синє

Визначено вибірккову середню відносної похибки оцінки морфологічної ознаки  $R_{g,i}$  за формулою (1), та перевірено виконання умови (4). Вибрано набори параметрів зйомки, що забезпечили гранично допустиму похибку  $\varepsilon=0,5\%$  при

визначенні площі насінини. На рисунку 4.7.3 представлено розрахунки для двох режимів зйомки при постійному білому освітленні, освітленості 300 lx, на червоному та чорному фоні (рис.4.7.3). Саме ці набори параметрів зйомки забезпечили найменші вибіркові середні похибки по визначенню площини насінини.

Реалізація п'ятого етапу МЦФЗ. За результатами третього та четвертого етапів було обрано два режими зйомки – постійне біле освітлення інтенсивності 300 lx, на червоному та чорному фоні. З трьох можливих параметрів впливу на якість оцінки площі насінини на фотозображенні мав вплив тільки один параметр «колір тла» при сталих параметрах кольору освітлення та освітленості. Це визначило вибір методів по оцінці статистичної значимості вибору кольору тла між чорним та червоним для фенотипування світлих насінин.

Розглянемо розподіл похибок як реалізацію двох випадкових величин  $Y_1^1$  (біла насінина на чорному тлі) та  $Y_2^1$  (біла насінина на червоному тлі),  $Y_1^2$  (біла з бежевими смужками насінина на чорному тлі) та  $Y_2^2$  (біла з бежевими смужками насінина на червоному тлі), що мають певні закони розподілу (D. Wackerly, W. Mendenhall, R. L. Scheaffer, 2008).

Було перевірено та відхилено гіпотезу про нормальність розподілу випадкових величин  $Y_1^1$ ,  $Y_2^1$ ,  $Y_1^2$ ,  $Y_2^2$  із застосуванням критерію  $\chi^2$  квадрат при рівні значущості  $\alpha = 0,1$ .

Далі було сформульовано основну гіпотезу  $H_0$  = «функції розподілу випадкових величин  $Y_1^i$  та  $Y_2^i$  рівні між собою», проти альтернативної  $H_1$  = «відсутня рівність функцій розподілу випадкових величин  $Y_1^i$  та  $Y_2^i$ » для  $i = 1,2$ . Було застосовано непараметричний критерій Т-критерій Вілкоксона (Wilcoxon T-test). Для білих насінин не біло знайдено статистично значимої різниці між двома режимами зйомки на рівні значущості  $\alpha = 0,1$ . Для білих з бежевими смужками насінин було доведено статистично значиму різницю між двома режимами на рівні значущості  $\alpha = 0,106$ . Знайдену тенденцію було

перевірено через збільшення виборки до 60 одиниць. Було доведено статистично значиму різницю між двома режимами зйомки для білих насінин з бежевими смужками на рівні значущості  $\alpha = 0,05$ .

чорний фон											червоний фон										
Name	square	LimitRed Min	LimitRed Max	LimitGree nMin	LimitGree nMax	LimitBlue Min	LimitBlue Max	Image_ar ea	red_min devicion	max_min divicion	Name	square	LimitRed Min	LimitRed Max	LimitGree nMin	LimitGree nMax	LimitBlue Min	LimitBlue Max	Image_ar ea	red_min devicion	max_min divicion
gray/1.jpg	332261.0	49	177	71	196	92	196	333144	6	10	1.jpg	363503.0	103	198	2	194	33	208	364214	103	103
gray/2.jpg	330974.5	51	177	74	196	87	196	331540	8	8	2.jpg	390921.0	78	181	30	199	47	208	391574	78	78
gray/3.jpg	352023.0	23	205	39	219	58	215	352302	23	58	3.jpg	356559.5	110	189	41	192	67	200	357444	110	110
gray/4.jpg	360486.5	48	189	66	200	76	200	360007	48	76	4.jpg	358091.0	109	208	18	188	44	193	358847	109	109
gray/5.jpg	300378.5	51	195	69	206	91	205	300137	51	91	5.jpg	304908.0	117	178	30	187	58	194	305664	117	117
gray/6.jpg	353428.5	50	188	73	207	85	207	353926	50	85	6.jpg	386611.0	99	177	49	189	60	202	387370	99	99
gray/7.jpg	320857.5	40	199	53	205	75	201	321246	40	75	7.jpg	355242.5	114	185	38	194	65	198	356135	114	114
gray/8.jpg	331736.5	58	193	76	205	87	200	332126	58	87	8.jpg	350287.5	115	174	61	193	68	198	351076	115	115
gray/9.jpg	361332.0	46	184	68	198	79	196	361690	46	79	9.jpg	364223.0	125	194	41	193	50	198	365077	125	125
gray/10.jpg	321023.5	48	185	68	204	78	198	320834	48	78	10.jpg	357528.5	116	170	54	193	69	203	358294	116	116
gray/11.jpg	310666.0	34	199	53	213	72	205	311286	34	72	11.jpg	319383.5	105	175	20	195	42	202	320152	105	105
gray/12.jpg	326454.0	59	197	73	212	88	209	327102	59	88	12.jpg	360383.0	116	177	34	189	54	194	361289	116	116
gray/13.jpg	325838.0	59	189	78	206	85	201	325694	59	85	13.jpg	339762.5	112	173	47	188	77	192	340485	112	112
gray/14.jpg	346374.0	53	193	68	205	83	199	347029	53	83	14.jpg	344632.5	121	179	20	188	56	195	345475	121	121
gray/15.jpg	210258.5	67	204	87	223	95	216	210878	67	95	15.jpg	376208.5	110	176	60	187	70	192	377117	110	110
gray/16.jpg	321694.5	58	197	71	210	54	203	321407	58	71	16.jpg	340198.0	111	174	72	190	64	200	341060	111	111
gray/17.jpg	304545.0	59	191	80	203	76	197	305193	59	80	17.jpg	322371.0	107	170	20	188	35	189	323283	107	107
gray/18.jpg	333394.0	49	183	65	199	86	193	334151	49	86	18.jpg	354398.0	113	173	61	189	78	190	355294	113	113
gray/19.jpg	369479.0	39	193	58	208	70	208	370226	39	70	19.jpg	360427.0	112	177	30	184	59	193	361309	112	112
gray/20.jpg	338954.5	32	183	53	198	71	196	339569	32	71	20.jpg	356831.0	103	165	47	184	64	188	357760	103	103
gray/21.jpg	319839.5	48	195	67	205	79	201	320437	48	79	21.jpg	63419.5	129	190	66	191	103	200	63527	129	129
gray/22.jpg	338241.0	47	201	64	212	86	206	338919	47	86	22.jpg	379599.0	100	188	11	191	36	200	380364	100	100
gray/23.jpg	320570.0	61	190	81	207	86	208	321245	61	86	23.jpg	320349.5	124	196	22	185	48	188	320861	124	124
gray/24.jpg	355563.5	61	195	79	204	83	201	356117	61	83	24.jpg	388604.0	113	183	70	197	82	202	389307	113	113
gray/25.jpg	354042.5	60	185	78	197	62	191	354843	60	78	25.jpg	329732.5	86	173	55	179	55	179	240411	86	86
gray/26.jpg	359618.0	49	195	67	206	82	207	359977	49	82	26.jpg	353668.5	90	168	38	180	38	184	354416	90	90
gray/27.jpg	320235.0	45	194	56	211	61	211	320994	45	61	27.jpg	368828.5	118	179	53	195	72	204	369644	118	118
gray/28.jpg	339398.0	60	181	80	196	85	192	339820	60	85	28.jpg	353989.5	101	164	19	182	42	186	354737	101	101
gray/29.jpg	345561.5	65	194	83	209	88	205	346216	65	88	29.jpg	378875.0	112	177	68	192	79	197	379280	112	112
gray/30.jpg	332835.5	56	200	72	206	80	202	335279	56	80	30.jpg	362869.5	115	177	90	187	91	190	363474	115	115
Друга лінія - мілка біль жовтувата																					
gray/31.jpg	205958.0	68	201	88	224	79	226	206704	68	88	31.jpg	231501.0	87	180	12	201	43	208	232280	87	87
gray/32.jpg	259914.0	52	199	58	204	42	198	260508	52	58	32.jpg	269279.0	66	185	40	207	44	220	270040	66	66
gray/33.jpg	248816.5	67	196	85	217	87	219	249524	67	87	33.jpg	251346.5	90	176	16	199	48	211	251473	90	90
gray/34.jpg	238139.0	69	210	91	233	90	229	238649	69	91	34.jpg	258491.5	85	191	9	210	38	227	259311	85	85
gray/35.jpg	219676.0	74	205	88	230	86	231	220801	74	88	35.jpg	254541.5	96	197	3	209	28	216	255148	96	96
gray/36.jpg	243335.0	70	191	86	201	85	196	243721	70	86	36.jpg	252352.0	95	192	37	191	57	193	253116	95	95
gray/37.jpg	212381.5	59	194	68	217	72	210	212827	59	72	37.jpg	278315.5	71	180	6	202	34	208	278858	71	71
gray/38.jpg	223227.0	71	204	94	227	82	225	223660	71	94	38.jpg	250404.0	98	188	23	202	58	199	251134	98	98
gray/39.jpg	259400.0	52	196	72	213	74	203	259542	52	74	39.jpg	28370.5	93	186	19	204	44	215	283729	93	93
gray/40.jpg	248679.5	49	189	64	206	64	205	248229	49	64	40.jpg	265352.5	81	212	6	211	38	226	266173	81	81
gray/41.jpg	273923.0	54	198	73	211	58	206	274044	54	73	41.jpg	281949.0	94	186	11	188	40	190	282277	94	94
gray/42.jpg	194862.5	56	205	64	225	46	222	195395	56	64	42.jpg	228680.5	87	206	6	203	40	220	229232	87	87
gray/43.jpg	246860.0	66	191	80	201	57	193	247280	66	80	43.jpg	287836.5	87	202	13	212	45	228	288495	87	87
gray/44.jpg	263419.0	47	194	69	221	79	217	263863	47	79	44.jpg	273233.0	86	202	12	218	31	235	273431	86	86
gray/45.jpg	197122.0	58	187	81	207	86	194	197776	58	86	45.jpg	208421.5	93	190	18	205	35	209	209079	93	93
gray/46.jpg	205287.0	62	206	82	234	79	235	205955	62	82	46.jpg	243457.0	98	191	21	214	59	230	244211	98	98
gray/47.jpg	229754.5	69	180	85	187	75	177	230351	69	85	47.jpg	243514.0	86	182	17	179	55	179	244070	86	86
gray/48.jpg	211942.0	72	201	91	223	87	217	212551	72	91	48.jpg	249613.0	90	190	11	195	51	200	249774	90	90
gray/49.jpg	226493.5	52	183	63	201	56	191	227161	52	63	49.jpg	250923.0	81	182	24	202	55	215	251652	81	81
gray/50.jpg	222744.0	64	192	79	212	57	207	223439	64	79	50.jpg	254562.5	81	190	13	202	43	215	255140	81	81
gray/51.jpg	245790.5	54	194	61	212	64	206	246403	54	64	51.jpg	260758.0	82	182	22	205	39	207	261495	82	82
gray/52.jpg	242194.5	52	194	64	212	64	206	242909	52	64	52.jpg	243583.0	83	189	11	184	30	186	244391	83	83
gray/53.jpg	244901.0	54	194	64	212	64	206	245621	54	64	53.jpg	236520.5	85	196	9	202	32	205	237314	85	85
gray/54.jpg	245998.0	50	194	61	212	64	206	246715	50	64	54.jpg	228589.5	106	212	16	206	39	201	229090	106	106
gray/55.jpg	246303.0	54	194	61	212	64	206	247079	54	64	55.jpg	227858.5	95	179	23	197	39	202	228484	95	95
gray/56.jpg	243923.0	56	194	64	212	64	206	244684	56	64	56.jpg	230987.0	92	177	73	198	66	210	231629	92	92
gray/57.jpg	246579.5	54	194	61	212	64	206	247357	54	64	57.jpg	257733.0	88	183	44	202	55	217	258309	88	88
gray/58.jpg	243540.0	55	194	61	212	64	206	244264	55	64	58.jpg	233364.0	102	187	48	192	72	198	233928	102	102
gray/59.jpg	247094.0	54	194	62	212	64	206	247879	54	64	59.jpg	243275.5	85	184	17	207	60	226	243937	85	85
gray/60.jpg	248236.0	51	194	61	212	64	206	248965	51	64	60.jpg	222216.5	89	189	48	210	59	216	222797	89	89

Рисунок 4.7.3 - Результати розрахунків площі насінини при двох режимах зйомки – постійне біле освітлення інтенсивності 300 lx, на червоному та чорному фоні

На шостому етапі було надано рекомендації по проведенню процедур визначення геометричних розмірів білих насини з оптимальними режимами зйомки «постійне біле освітлення, освітленість 300 lx, на червоному або чорному фоні», а для білих з бежевими смужками: «постійне біле освітлення, освітленість 300 lx, на чорному фоні».

## 4.8 Висновки до розділу 4

У четвертому розділі удосконалено цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets), яка забезпечує семантично узгоджене формування наборів даних відповідно до онтологічної моделі HELIANTUS та їх подальшу інтеграцію до бази знань. Розроблено метод цифрового фенотипування рослинних об'єктів на основі зображень та виконано програмно-алгоритмічну реалізацію розробленого методу, проведено комплексні експериментальні дослідження, а також здійснено валідацію та оцінку ефективності створеного інструментарію при обробці реальних біологічних об'єктів.

За результатами розділу сформульовано такі висновки:

1. Вперше розроблено метод цифрового фенотипування рослинних об'єктів на основі зображень, який реалізує оптимізацію параметрів реєстрації даних. Метод забезпечує автоматизоване виділення та високоточне оцінювання морфологічних, структурних і колірних характеристик насіння соняшнику за кількісними дескрипторами.

2. Розроблено та протестовано спеціалізоване програмне забезпечення, яке дозволяє в автоматичному режимі здійснювати прецизійну сегментацію об'єктів і генерувати структуровані цифрові паспорти для кожної одиниці аналізованого насіння. Метрики охоплюють геометричні характеристики, показники форми (ексцентриситет, округлість) та спектральні показники забарвлення, забезпечуючи високу роздільну здатність ідентифікації навіть для низькоконтрастних та візуально схожих селекційних зразків.

4. Проведено комплексну апробацію та валідацію цифрової технології на репрезентативній вибірці насіння соняшнику з різними типами текстурних візерунків перикарпію (монохромне біле, сіре та смугасте забарвлення). Результати експериментальних досліджень підтвердили високу ефективність методу для мінімізації мінливості цифрових даних. Встановлено, що

сформований робочий процес має високу обчислювальну масштабованість і повністю готовий до безпосередньої трансляції отриманих датасетів у простір фактів (ABox) онтології HELIANTUS для подальшого дедуктивного аналізу у системі «генотип — середовищний стрес».

5. Практична цінність розробленого інструментарію підтверджується успішним впровадженням результатів досліджень у селекційно-генетичну практику Інституту олійних культур НААН України. Створена інформаційна технологія, цифрові моделі та структуровані бази даних забезпечують наскрізну синтаксичну стандартизацію науково-дослідних робіт, гарантуючи несуперечливий перехід від суб'єктивних візуальних оцінок до високоточних методів автоматизованого системного аналізу вихідного селекційного матеріалу.

Основні наукові результати розділу опубліковано в працях автора (Ведмедєв, 2024; Ведмедєв та Терещенко, 2026; Ведмедєва та Ведмедєв, 2026; Махова та ін., 2024; Aliiev et al., 2024).

## ВИСНОВКИ

1. Проведено системний аналіз сучасних процесів цифрової трансформації аграрної галузі та визначено вимоги до формування цифрового середовища вирощування і селекції соняшнику в умовах Agriculture 4.0. Встановлено, що ефективна цифрова трансформація аграрних процесів потребує переходу від фрагментарного накопичення інформації до інтегрованого використання різномірних цифрових даних, які характеризують генотипові, фенотипові, агротехнологічні, кліматичні та екологічні параметри. Обґрунтовано необхідність застосування інтелектуальних систем підтримки прийняття рішень, що базуються на поєднанні методів математичного моделювання, онтологічного представлення знань та аналізу цифрових даних. Визначено ключову науково-практичну суперечність, яка полягає у невідповідності між зростанням обсягів цифрових агротехнологічних даних та недостатнім рівнем їх структурованості, стандартизації й семантичної сумісності для автоматизованого використання. Це обґрунтувало необхідність розроблення онтологічної моделі соняшнику та цифрової технології формування семантично сумісних наборів даних як інформаційної основи інтелектуальної системи підтримки прийняття рішень у цифровій селекції.

2. Удосконалено цифрову модель соняшнику на основі моделювання взаємодії «генотип — середовище», що забезпечує інтеграцію різномірних даних про генотипові, фенотипові, агротехнологічні та екологічні характеристики рослини. Запропонована чотирирівнева архітектура Digital Phenotyping System забезпечує послідовне об'єднання процесів збору, зберігання, обробки та аналізу цифрових даних, отриманих у результаті лабораторного і польового фенотипування, дистанційного моніторингу та оцінювання параметрів навколишнього середовища. Розроблена структура цифрової моделі створює інформаційну основу для формування цифрового двійника рослини соняшнику та застосування методів математичного і комп'ютерного моделювання для дослідження закономірностей її розвитку та адаптивності в різних умовах вирощування.

3. Удосконалено онтологічну модель агросистеми вирощування соняшнику HELIANTHUS, яка забезпечує інтеграцію різномірних даних про генотипові характеристики, фенотипові ознаки, селекційні процеси, агротехнологічні параметри, кліматичні показники та екологічні фактори у єдиному семантичному інформаційному просторі. Запропонована модель забезпечує формалізацію знань предметної області на рівнях TBox та ABox, що дозволяє виконувати логічне виведення, встановлювати причинно-наслідкові зв'язки у системі «генотип — середовище» та формувати інформаційну основу для функціонування інтелектуальних систем підтримки прийняття рішень у цифровій селекції соняшнику.

4. Розроблено метод цифрового фенотипування рослинних об'єктів на основі зображень, який забезпечує автоматизоване виділення та кількісну оцінку морфологічних, структурних і колірних характеристик об'єктів за цифровими зображеннями. Запропонований метод обробки зображень підвищує достовірність отримання фенотипових дескрипторів, що формують стандартизовані цифрові ознаки рослини та забезпечують їх подальшу інтеграцію в онтологічну модель, базу знань і інтелектуальну систему підтримки прийняття рішень у цифровій селекції соняшнику.

5. Удосконалено цифрову технологію фенотипування із формуванням інтегрованих наборів даних (datasets) для цифрової селекції соняшнику, яка забезпечує автоматизований збір, стандартизацію, семантичне узгодження та збереження фенотипових даних відповідно до онтологічної моделі HELIANTUS. Запропонована технологія забезпечує трансформацію результатів цифрового фенотипування у структуровані інформаційні ресурси, придатні для наповнення бази знань та подальшого використання в інтелектуальних системах підтримки прийняття рішень, методах інтелектуального аналізу даних і предиктивного моделювання.

6. Розроблено концептуальні та технологічні засади створення інтелектуального інформаційного середовища цифрової селекції соняшнику, що

базується на інтеграції цифрової моделі рослини, онтологічного представлення знань, методів комп'ютерного зору та технології формування семантично сумісних наборів даних, створюючи інформаційну основу для функціонування інтелектуальних систем підтримки прийняття рішень при вирощуванні соняшнику.

Отримані результати роблять внесок у вирішення визначеної ключової науково-практичної суперечності, яка полягає у невідповідності між зростанням обсягів цифрових агротехнологічних даних та недостатнім рівнем їх структурованості.

## СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Abebe, A., Kim, Y., Kim, J., Kim, S., & Baek, J. (2023). Image-Based High-Throughput Phenotyping in Horticultural Crops. *PLANTS-BASEL*, 12(10). doi:10.3390/plants12102061.
2. Adebola T., Zhang Y., Kim J., Navlakha S. (2025) GrowSplat: Dynamic 3D Gaussian splatting for plant morphology alignment. arXiv preprint. arXiv:2505.10923. URL: <https://arxiv.org/abs/2505.10923>.
3. AgEval: A Benchmark for Zero-Shot and Few-Shot Plant Stress Phenotyping with Multimodal LLMs. arXiv preprint. 2023. URL: <https://arxiv.org/abs/2306.05431>.
4. AgroFIMS: Agricultural Field Information Management System [Electronic resource]. Alliance of Bioversity & CIAT, 2024. Access mode: <https://agrofims.org>.
5. Aliiev E., Vedmedeva K., Machova T., Vedmedev S. (2024) Study of the distribution of phenotypic characteristics of sunflower seeds in a head of different genotypes. *Bulgarian Journal of Crop Science*, 62, No 4, 2024 p.73-89. <https://doi.org/10.61308/JIXX8922>.
6. Aliyev E.B. (2019) Development of a device for automatic phenotyping of sunflower seed material. *Machinery & Energetics. Journal of Rural Production Research*. 2019. Vol. 10, No. 1. P. 11–17. ISSN 2663-1334. (Kyiv, Ukraine).
7. Alkhamash, E. (2022). Graphical Transformation of OWL Ontologies to Event-B Formal Models. *Cmc-computers materials & continua*, 70(2), 3733-3750. doi:10.32604/cmc.2022.015987.
8. Andres F., Wiechers D., Langensiepen M., Kage H. Field Robot Platform for Phenotyping Maize. *Computers and Electronics in Agriculture*. 2021. Vol. 182. Article ID: 105991. DOI: <https://doi.org/10.1016/j.compag.2021.105991>.
9. An Intelligent System Based on Ontologies for Determining the Similarity of User Preferences / O. Oborska, M. Teliatynskiy, D. Dosyn, V. Lytvyn, S. Kostenko. *Proceedings of the 7th International Conference on Computational Linguistics and Intelligent Systems (COLINS-2023)* (Kharkiv, Ukraine, April 20–21, 2023). CEUR

Workshop Proceedings. 2023. Vol. 3403. P. 283–292. URL: <https://ceur-ws.org/Vol-3403/paper23.pdf>.

10. APSIM Initiative. (n.d.). *APSIM (Agricultural Production Systems Simulator)*. Retrieved June 23, 2026, from <https://www.apsim.info/>

11. Araus, J., & Cairns, J. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends in plant science*, 19(1), 52-61. doi:10.1016/j.tplants.2013.09.008.

12. Arend, D., Junker, A., Scholz, U., Schüler, D., Wylie, J., & Lange, M. (2016). Daniel Arend, Astrid Junker, Uwe Scholz, Danuta Schüler, Juliane Wylie, Matthias Lange, *PGP repository: a plant phenomics and genomics data publication infrastructure*, Database, Volume 2016, 2016, baw033, <https://doi.org/10.1093/database/baw033>.

13. Arvidsson, S., Pérez-Rodríguez, P., & Mueller-Roeber, B. (2011). A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects. *New Phytologist*, 191(3), 895-907.

14. Atkinson, N.J., Lilley, C.J., Urwin, P.E. Genome-Wide Association Mapping of Time-Dependent Growth Responses to Moderate Drought Stress in *Arabidopsis*. *Plant Cell and Environment*. 2013. Vol. 36, Issue 2, pp. 262–276. DOI: <https://doi.org/10.1111/j.1365-3040.2012.02598.x>

15. Auernhammer, H. (2001). Precision farming - the environmental challenge. *Computers and electronics in agriculture*, 30(1-3), 31-43. doi:10.1016/S0168-1699(00)00153-8.

16. Awais, M., Anwar-ul-Haq, M., Mubeen, M., Akhtar, J., & Wajid, S. (2020). Adaptation strategies for sunflower production under climate change for arid environment by using cropgro and apsim-sunflower models. *Pakistan journal of agricultural sciences*, 57(6), 1529-1538. doi:10.21162/PAKJAS/20.515.

17. Bakurova A., Vedmedeva K., Vedmedev S., Tereschenko E.(2023) Ontological model of *Helianthus* cultivation in Ukrainian conditions. *CEUR Workshop*

*Proceedings*. 3396, 130–140. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85160858256&partnerID=40&md5=01df4a1fe6a8617729393e0965e4a00f>.

18. Bakurova A., Vedmedeva K., Vedmedev S., Tereschenko E., Shyrokorad D. (2023) Development of the System for the Digital Model of the Helianthus Phenotype. *Proceedings of the VI International Scientific Congress Society of Ambient Intelligence 2023 (ISC SAI 2023)*. 20–25 November 2023. 20–25. DOI: 10.46489/ISCSAI-23-30.

19. Balachandran, S., Hurry, V. M., Kelley, S. E., Osmond, C. B., Robinson, S. A., Rohozinski, J., et al. (1997). Concepts of plant biotic stress. Some insights into the stress physiology of virus infected plants, from the perspective of photosynthesis. *Physiologia Plantarum*, 100(2), 203-213.

20. Balliau, T., Duruflé, H., Blanchet, N., Blein-Nicolas, M., Langlade, N., & Zivy, M. (2021). Proteomic data from leaves of twenty-four sunflower genotypes under water deficit. *Ocl-oilseeds and fats crops and lipids*, 28. doi:10.1051/ocl/2020074.

21. Bancic, J., Greenspoon, P., Gaynor, R., & Gorjanc, G. (2025). Plant breeding simulations with AlphaSimR. *Crop science*, 65(1). doi:10.1002/csc2.21312.

22. Bao, T., Melenka, G. W., Ljubotina, M. K., Carey, J. P., & Cahill, J. F. (2018). A new method for the rapid characterization of root growth and distribution using digital image correlation. *New Phytologist*, 218(2), 835-846. doi:10.1111/nph.15009.

23. Bashir, M. 1375 (2025) 'FAIR and Beyond: Evolving Principles for Modern Data Ecosystems' Designing the conceptual landscape for a xair validation infrastructure, *DCLXVI 2024*. 2025-01-01, 170-189.

24. Bauriegel, E., Giebel, A., & Herppich, W. B. (2011). Hyperspectral and chlorophyll fluorescence imaging to analyse the impact of Fusarium culmorum on the photosynthetic integrity of infected wheat ears. *Sensors*, 11(4), 3765-3779.

25. Berger, B., Parent, B., & Tester, M. (2010). High-throughput shoot imaging to study drought responses. *Journal of experimental botany*, 61(13), 3519-3528.

26. Bin Nasir, F., & Li, J. (2024). Understanding machine learning predictions of wastewater treatment plant sludge with explainable artificial intelligence. *Water environment research*, 96(10). doi:10.1002/wer.11136.
27. Brewer, M. T., Lang, L., Fujimura, K., Dujmovic, N., Gray, S., & van der Knaap, E. (2006). Development of a controlled vocabulary and software application to analyze fruit shape variation in tomato and other plant species. *Plant physiology*, 141(1), 15-25.
28. Buttigieg P.L. et al. (2016) The Environment Ontology in 2016: bridging domains with increased scope, semantic density, and interoperation. *J Biomed Semantics*. 7:57.
29. Castillo-Barrera, F., Durán-Limón, H., Médina-Ramírez, C., & Rodríguez-Rocha, B. (2013). A method for building ontology-based electronic document management systems for quality standards—the case study of the ISO/TS 16949:2002 automotive standard. *Applied intelligence*, 38(1), 99-113. doi:10.1007/s10489-012-0360-1.
30. Chern, C.-G., Fan, M.-J., Yu, S.-M., Hour, A.-L., Lu, P.-C., Lin, Y.-C., et al. (2007). A rice phenomics study—phenotype scoring and seed propagation of a T-DNA insertion-induced rice mutant population. *Plant molecular biology*, 65(4), 427-438.
31. Chorny A., Dosyn D. G. Development of a Unified Output Format for Text Parsers in the Ontology Construction System from Text Documents. *Information Systems and Networks*. 2025. Vol. 17. P. 170–188. <https://doi.org/10.23939/sisn2025.17.170>.
32. Choudhury, S., Samal, A., & Awada, T. (2019). Leveraging Image Analysis for High-Throughput Plant Phenotyping. *Frontiers in plant science*, 10. doi:10.3389/fpls.2019.00508.
33. Clark, R. T., MacCurdy, R. B., Jung, J. K., Shaff, J. E., McCouch, S. R., Aneshansley, D. J., et al. (2011). Three-dimensional root phenotyping with a novel

imaging and software platform. *Plant physiology*, 156(2), 455-465. doi:10.1104/pp.110.169102.

34. Clarke, J., Cooper, L., Poelchau, M., Berardini, T., Elser, J., Farmer, A., et al. (2023) Data sharing and ontology use among agricultural genetics, genomics, and breeding databases and resources of the Agbiodata Consortium. *Database-the journal of biological databases and curation*. doi:10.1093/database/baad076.

35. Connor, D., & Sadras, V. (1992). Physiology of yield expression in sunflower. *Field crops research*, 30(3-4), 333-389. doi:10.1016/0378-4290(92)90006-U.

36. Cook, C., Stroe, O., Cochrane, G., Birney, E., & Apweiler, R. (2020). The European Bioinformatics Institute in 2020: building a global infrastructure of interconnected data resources for the life sciences. *Nucleic acids research*, 48(D1), D17-D23. doi:10.1093/nar/gkz1033.

37. Cooper, M., Messina, C. D., Podlich, D., Totir, L. R., Baumgarten, A., Hausmann, N. J., et al. (2014). Predicting the future of plant breeding: complementing empirical evaluation with genetic prediction. *Crop & Pasture Science*, 65(4), 311-336. doi:10.1071/cp14007.

38. Crop Ontology is licensed under a (Creative Commons Attribution 4.0 International License) <https://croponontology.org/>

39. Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O., Jarquín, D., de los Campos, G., et al. (2017). Genomic Selection in Plant Breeding: Methods, Models, and Perspectives. *Trends in plant science*, 22(11), 961-975. doi:10.1016/j.tplants.2017.08.011.

40. Danzi, D., De Paola, D., Petrozza, A., Summerer, S., Cellini, F., Pignone, D., et al. (2022). The Use of Near-Infrared Imaging (NIR) as a Fast Non-Destructive Screening Tool to Identify Drought-Tolerant Wheat Genotypes. *Agriculture-basel*, 12(4). doi:10.3390/agriculture12040537.

41. de Reffye, P., Blaise, F., & Houllier, F. (1998) 'Modelling plant growth and architecture: Some recent advances and applications to agronomy and forestry' *second*

*international symposium on models for plant growth, environmental control and farm management in protected cultivation*. 1998-01-01. 105-116.

42. Debbagh, M., Sun, S., & Lefsrud, M. (2025). Predictive modeling, pattern recognition, and spatiotemporal representations of plant growth in simulated and controlled environments: A comprehensive review. *Plant phenomics*, 7(3). doi:10.1016/j.plaphe.2025.100089.

43. Demydov, O., Zamlila, N., Novytska, N., Kyrylenko, V., & Miliar, B. (2024). Assessment of the stability of common winter wheat breeding lines in multi-environment tests. *Scientific Horizons*, 27(7). <https://doi.org/10.48077/scihor7.2024.62>

44. Di Tocco, J., Lo Presti, D., Massaroni, C., Cinti, S., Cimini, S., De Gara, L., et al. (2023). Plant-Wear: A Multi-Sensor Plant Wearable Platform for Growth and Microclimate Monitoring. *Sensors*, 23(1), 549.

45. Dias, P. M. B., Brunel-Muguet, S., Dürr, C., Huguet, T., Demilly, D., Wagner, M.-H., et al. (2011). QTL analysis of seed germination and pre-emergence growth at extreme temperatures in *Medicago truncatula*. *Theoretical and Applied Genetics*, 122(2), 429-444.

46. Ukraine Agriculture Resilience Initiative (AGRI-Ukraine): Digital Solutions Report. USAID, 2024. 35 p <https://agritechobservatory.fao.org/digital-technologies-in-the-grain-sector-of-ukraine>.

47. Dordea, A. C., Bray, M. A., Allen, K., Logan, D. J., Fei, F., Malhotra, R., et al. (2016). An open-source computational tool to automatically quantify immunolabeled retinal ganglion cells. *Experimental Eye Research*, 147, 50-56. doi:10.1016/j.exer.2016.04.012.

48. DSSAT Foundation. (n.d.). *DSSAT: Decision Support System for Agrotechnology Transfer*. Retrieved June 23, 2026, from <https://dssat.net/>

49. Duan, L. F., Yang, W. N., Huang, C. L., & Liu, Q. (2011). A novel machine-vision-based facility for the automatic evaluation of yield-related traits in rice. *Plant Methods*, 7. doi:10.1186/1746-4811-7-44.

50. Dumschott, K., Dörpholz, H., Laporte, M., Brillhaus, D., Schrader, A., Usadel, B., et al. (2023). Ontologies for increasing the FAIRness of plant research data. *Frontiers in plant science*, 14. doi:10.3389/fpls.2023.1279694
51. Efron B. (1979) Bootstrap methods: another look at the jackknife. *The Annals of Statistics*. 7(1), 1–26.
52. Falconer, D. (1981). Introduction to quantitative genetics. Copublished in the United States with John Wiley Sons, inc, 605.
53. Fan, J., Zhang, Y., Wen, W., Gu, S., Lu, X., & Guo, X. (2021). The future of Internet of Things in agriculture: Plant high-throughput phenotypic platform. *Journal of cleaner production*, 280. doi:10.1016/j.jclepro.2020.123651.
54. Fang, W., Feng, H., Yang, W., & Liu, Q. (2016). A fast 3D Reconstruction for Wheat Plant Architecture Studies in Phenotyping. *Journal of Agricultural Science and Technology*, 18(2), 95-101.
55. Faux, A., Gorjanc, G., Gaynor, R., Battagin, M., Edwards, S., Wilson, D., et al. (2016). AlphaSim: Software for Breeding Program Simulation. *Plant genome*, 9(3). doi:10.3835/plantgenome2016.02.0013.
56. Fernández-López, M., & Gómez-Pérez, A. (2002). Overview and analysis of methodologies for building ontologies. *Knowledge engineering review*, 17(2), 129-156. doi:10.1017/S0269888902000462.
57. Fiorani, F., & Schurr, U. (2013). Future scenarios for plant phenotyping. *Annu. Rev. Plant Biol*, 64(1), 267-291.
58. Fisher, R. A. (1919). XV.—The correlation between relatives on the supposition of Mendelian inheritance. *Earth and Environmental Science Transactions of the Royal Society of Edinburgh*, 52(2), 399-433.
59. Flavel, R. J., Guppy, C. N., Tighe, M., Watt, M., McNeill, A., & Young, I. M. (2012). Non-destructive quantification of cereal roots in soil using high-resolution X-ray tomography. *Journal of Experimental Botany*, 63(7), 2503-2511.

60. Flügel, S., Glauer, M., Neuhaus, F., & Hastings, J. (2025). When one logic is not enough: Integrating first-order annotations in OWL ontologies. *Semantic web*, 16(2). doi:10.3233/SW-243440.
61. Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675–701. <https://doi.org/10.1080/01621459.1937.10503522>
62. Gaggion, N., Boccardo, N. A., Bonazzola, R., Legascue, M. F., Mammarella, M. F., Rodriguez, F. S., et al. (2025). ChronoRoot 2.0: An open AI-powered platform for 2D temporal plant phenotyping. arXiv preprint arXiv:2504.14736.
63. Gao Y., Qi X. Neural 3D Gaussian Splatting for Plant Structure Reconstruction. *ACM Transactions on Graphics*. 2024. Vol. 43, No. 4. P. 1–14.
64. García-Coronado, H., Ojeda-Contreras, A., Berumen-Varela, G., Robles-Parra, J., Handa, A., & Tiznado-Hernández, M. (2025). Engineering Crops for Enhanced Drought Stress Tolerance: A Strategy for Sustainable Agriculture. *Agronomy-basel*, 15(8). doi:10.3390/agronomy15081912.
65. Gauch, H. (2013). A Simple Protocol for AMMI Analysis of Yield Trials. *Crop science*, 53(5), 1860-1869. doi:10.2135/cropsci2013.04.0241.
66. Giuliani, A., & Baron, H. (2025). The CAP (Common Agricultural Policy): A Short History of Crises and Major Transformations of European Agriculture. *Forum for social economics*, 54(1), 68-94. doi:10.1080/07360932.2023.2259618.
67. Godin, C., & Sinoquet, H. (2005). Functional–structural plant modelling. *New phytologist*, 166(3), 705-708.
68. Golzarian, M. R., Frick, R. A., Rajendran, K., Berger, B., Roy, S., Tester, M., et al. (2011). Accurate inference of shoot biomass from high-throughput images of cereal plants. *Plant methods*, 7(1), 1-11.
69. Gorohivets N. A., Vedmedeva E. V. (2016) Inheritance of epidermis pigmentation in sunflower achenes. *Cytology and Genetics*, 50(2), 116-120. <https://doi.org/10.3103/S0095452716020031>

70. Gosseau, F., Blanchet, N., Varès, D., Burger, P., Campergue, D., Colombet, C., et al. (2019) Heliaphen, an Outdoor High-Throughput Phenotyping Platform for Genetic Studies and Crop Modeling. *Frontiers in plant science*, 9. doi:10.3389/fpls.2018.01908.
71. Grobler, E., & Celano, G. (2025). Photogrammetric and LiDAR Scanning with iPhone 13 Pro: Accuracy, Precision and Field Application on Hazelnut Trees. *Sensors*, 25(18). doi:10.3390/s25185629.
72. Gruber, T. (1991) 'The role of common ontology in achieving sharable, reusable knowledge bases' principles of knowledge representation and reasoning. 1991-01-01. 601-602.
73. Guo, Y., Ma, Y. T., Zhan, Z. G., Li, B. G., Dingkuhn, M., Luquet, D., et al. (2006). Parameter optimization and field validation of the functional-structural model GREENLAB for maize. *Annals of Botany*, 97(2), 217-230. doi:10.1093/aob/mcj033.
74. Hartmann, A., Czauderna, T., Hoffmann, R., Stein, N., & Schreiber, F. (2011). HTPPheno: An image analysis pipeline for high-throughput plant phenotyping. *Bmc Bioinformatics*, 12. doi:10.1186/1471-2105-12-148.
75. Hatfield, J., Gitelson, A., Schepers, J., & Walthall, C. (2008). Application of spectral remote sensing for agronomic decisions. *Agronomy journal*, 100(3), S117-S131. doi:10.2134/agronj2006.0370c.
76. Heitjan, D. F. (1999). Causal inference in a clinical trial: A comparative example. *Controlled Clinical Trials*, 20(4), 309-318. doi:10.1016/s0197-2456(99)00012-4.
77. Hennig, H., Rees, P., Blasi, T., Kamensky, L., Hung, J., Dao, D., et al. (2017). An open-source solution for advanced imaging flow cytometry data analysis using machine learning. *Methods*, 112, 201-210. doi:10.1016/j.ymeth.2016.08.018.
78. Hnatiienko, V. H., Hnatiienko, H. M., Zozulya, O. L., Snytyuk, V. Ye., & Schwartau, V. V. (2025). Site-specific sunflower yield forecasting based on spatial analysis and machine learning. *Доповіді Національної академії наук України*, (4), 17—26. <https://doi.org/10.15407/dopovidi2025.04.017>

79. Hnatiienko, V., & Hnatiienko, H. (2024). Integration of machine learning and deep learning methods for sunflower yield prediction. *Management of Development of Complex Systems*, (59), 225–234. <https://doi.org/10.32347/2412-9933.2024.59.225-234>
80. Holzworth, D. et al. (2024) APSIM – Agricultural Production Systems Simulator [Electronic resource]. APSIM Initiative, Access mode: <https://www.apsim.info>;
81. Horizon Europe: Digital Twins for Crops Initiative [Electronic resource]. European Commission, 2023. Access mode: <https://research-and-innovation.ec.europa.eu>].
82. Hoyos-Villegas, V., Houx, J. H., Singh, S. K., & Fritschi, F. B. (2014). Ground-based digital imaging as a tool to assess soybean growth and yield. *Crop Science*, 54(4), 1756-1768.
83. Hu, B.-G., De Reffye, P., Zhao, X., Yan, H.-P., & Kang, M. Z. (2003) 'Greenlab: A new methodology towards plant functional-structural model--structural part' Plant growth modelling and applications. *TsingHua University Press and Springer*, pp. 21-35.
84. Hudzenko, V., Buniak, N., Tsentylo, L., Demydov, O., Fedorenko, I., Fedorenko, M., et al. (2022). Evaluation of grain yield performance and its stability in various spring barley accessions under condition of different agroclimatic zones of Ukraine. *Biosystems diversity*, 30(4), 406-422. doi:10.15421/012240.
85. Ikeda, M., Hirose, Y., Takashi, T., Shibata, Y., Yamamura, T., Komura, T., et al. (2010). Analysis of rice panicle traits and detection of QTLs using an image analyzing method. *Breeding Science*, 60(1), 55-64.
86. ImageBreed: Image-Based Phenotyping Tools [Electronic resource]. – 2023. – Access mode: <https://imagebreed.org>].
87. Intelligent Monitoring of Stress Induced by Water Deficiency in Plants using Deep Learning. arXiv preprint. 2022. URL: <https://arxiv.org/abs/2205.09364>

88. Iqbal, J., Xu, R., Halloran, H., & Li, C. (2020). Development of a Multi-Purpose Autonomous Differential Drive Mobile Robot for Plant Phenotyping and Soil Sensing. *Electronics*, 9(9). doi:10.3390/electronics9091550.
89. Iyer-Pascuzzi, A. S., Symonova, O., Mileyko, Y., Hao, Y. L., Belcher, H., Harer, J., et al. (2010). Imaging and Analysis Platform for Automatic Phenotyping and Trait Ranking of Plant Root Systems. *Plant Physiology*, 152(3), 1148-1157. doi:10.1104/pp.109.150748.
90. Jacobsen, A., Azevedo, R., Juty, N., Batista, D., Coles, S., Cornet, R., et al. (2020). FAIR Principles: Interpretations and Implementation Considerations. *Data intelligence*, 2(1-2), 10-29. doi:10.1162/dint\_r\_00024.
91. Jansen, M., Gilmer, F., Biskup, B., Nagel, K. A., Rascher, U., Fischbach, A., et al. (2009). Simultaneous phenotyping of leaf growth and chlorophyll fluorescence via GROWSCREEN FLUORO allows detection of stress tolerance in *Arabidopsis thaliana* and other rosette plants. *Functional Plant Biology*, 36(10-11), 902-914. doi:10.1071/fp09095.
92. Jones, H. G., & Vaughan, R. A. (2010). Remote sensing of vegetation: principles, techniques, and applications. *OUP Oxford*. 363.
93. Joosen, R. V. L., Arends, D., Willems, L. A. J., Ligterink, W., Jansen, R. C., & Hilhorst, H. W. M. (2012). Visualizing the genetic landscape of *Arabidopsis* seed performance. *Plant physiology*, 158(2), 570-589.
94. Kayad, A., Sozzi, M., Paraforos, D. S., Jr, F. A. R., Cohen, Y., Fountas, S., et al. (2022). How many gigabytes per hectare are available in the digital agriculture era? A digitization footprint estimation. *Computers and Electronics in Agriculture*, 198. doi:10.1016/j.compag.2022.107080.
95. Kumar, P., Huang, C., Cai, J., & Miklavcic, S. J. (2014). Root phenotyping by root tip detection and classification through statistical learning. *Plant and soil*, 380(1), 193-209.

96. Kundu, S., Saini, D., Meena, R., Bahuguna, R., & Jagadish, S. (2024). High-throughput phenotyping and AI technologies for deciphering crop resilience to heat stress. *Plant physiology reports*, 29(4), 699-715. doi:10.1007/s40502-024-00821-4.
97. Li Q., Wang Z., Chen L., Xu Y. Integrating 3D Gaussian Fields with NeRF for High-Fidelity Plant Reconstruction. *Sensors*. 2025. Vol. 25, No. 3. Article ID: 987. DOI: <https://doi.org/10.3390/s25030987>.
98. Li Z., Guo R., Li M., Chen Y., Li G. (2020) A review of computer vision technologies for plant phenotyping. *Computers and electronics in agriculture*, 176: 105672. <https://doi.org/10.1016/j.compag.2020.105672>.
99. Li, J., Qi, X., Nabaei, S., Liu, M., Chen, D., Sun, Q., et al. (2025). A survey on 3D reconstruction techniques in plant phenotyping: From classical methods to Neural Radiance Fields (NeRF), 3D Gaussian Splatting (3DGS), and beyond. *Plant phenomics*, 7(4). doi:10.1016/j.plaphe.2025.100137.
100. Li, L., Zhang, Q., & Huang, D. F. (2014). A Review of Imaging Techniques for Plant Phenotyping. *Sensors*, 14(11), 20078-20111. doi:10.3390/s141120078.
101. Li, Y. W., Wu, X. Y., Xu, W. Z., Sun, Y. D., Wang, Y., Li, G. J., et al. (2021). High-Throughput physiology-based stress response phenotyping: Advantages, applications and prospective in horticultural plants. *Horticultural Plant Journal*, 7(3), 181-187. doi:10.1016/j.hpj.2020.09.004.
102. Liu S., Wang H., Yan J., Chen Y. 3D Plant Phenotyping with Voxel-Based Reconstruction Techniques. *Plant Methods*. 2020. Vol. 16. Article ID: 116. DOI: <https://doi.org/10.1186/s13007-020-00652-1>.
103. Liu, F., Wang, S., & Zhao, L. (2025a). Research Progress and Prospect of Intelligent High-Throughput Crop Phenotyping Platform. *Journal of crop health*, 77(5). doi:10.1007/s10343-025-01228-3.
104. Lucido, A., Basallo, O., Marin-Sanguino, A., Eleiwa, A., Martinez, E., Vilaprinyo, E., et al. (2025). Multiscale Mathematical Modeling in Systems Biology: A Framework to Boost Plant Synthetic Biology. *Plants-basel*, 14(3). doi:10.3390/plants14030470.

105. Lyons, E. (2021) 'BreedBase: Breeding Database Management System' 2021-12-23T19:04:41Z. <https://breedbase.org>: Cornell University.
106. Macedo, N., & Cunha, A. (2016). Least-change bidirectional model transformation with QVT-R and ATL. *Software and Systems Modeling*, 15(3), 783-810. doi:10.1007/s10270-014-0437-x.
107. Manianga A., Bose C., Sreeraj S. (2024) Sustainable applications of phytochemicals and nutritive components derived from selected underutilized seeds: a review. *Acta scientiarum polonorum-technologia alimentaria*. 23(1), 87-122. <https://doi.org/10.17306/J.AFS.001204>
108. Martin-Brualla R., Radwan N., Sajjadi M. S. M., Barron J. T., Dosovitskiy A., Duckworth D. NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. P. 7210–7219.
109. Menda, N., Ellerbrock, B., Simoes, C., Karaikal, S., Nyaga, C., Flores-Gonzalez, M., et al. (2025). Post-composing ontology terms for efficient phenotyping in plant breeding. *Database-the journal of biological databases and curation*, 2025. doi:10.1093/database/baaf020.
110. Mildenhall B., Srinivasan P. P., Tancik M., Barron J. T., Ramamoorthi R., Ng R. NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *Communications of the ACM*. 2021. Vol. 65, No. 1. P. 99–106.
111. Monforte, A. J., Diaz, A., Caño-Delgado, A., & Van Der Knaap, E. (2013). The genetic basis of fruit morphology in horticultural crops: lessons from tomato and melon. *Journal of experimental botany*, 65(16), 4625-4637.
112. Morales, N., Ogbonna, A., Ellerbrock, B., Bauchet, G., Tantikanjana, T., Teclé, I., et al. (2022). Breedbase: a digital ecosystem for modern plant breeding. *G3-genes genomes genetics*, 12(7). doi:10.1093/g3journal/jkac078.
113. Mulla, D. (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems engineering*, 114(4), 358-371. doi:10.1016/j.biosystemseng.2012.08.009.

114. Murphy, K., Ludwig, E., Gutierrez, J., & Gehan, M. (2024). Deep Learning in Image-Based Plant Phenotyping. *Annual review of plant biology*, 75, 771-795. doi:10.1146/annurev-arplant-070523-042828.
115. Neumann, K., Klukas, C., Friedel, S., Rischbeck, P., Chen, D., Entzian, A., et al. (2015). Dissecting spatiotemporal biomass accumulation in barley under different water regimes using high-throughput image analysis. *Plant cell and environment*, 38(10), 1980-1996. doi:10.1111/pce.12516.
116. Neveu, P., Tireau, A., Hilgert, N., Nègre, V., Mineau-Cesari, J., Bricchet, N., et al. (2019). Dealing with multi-source and multi-scale information in plant phenomics: the ontology-driven Phenotyping Hybrid Information System. *New phytologist*, 221(1), 588-601. doi:10.1111/nph.15385.
117. Papoutsoglou, E. A. (2021) Paving the way for FAIR data in plant phenotyping. Plant Breeding Data Competence Centre Research output: Thesis › internal PhD, *Wageningen University and Research EPS*. DOI: 10.5281/zenodo.5572771.
118. Paproki, A., Sirault, X., Berry, S., Furbank, R., & Fripp, J. (2012). A novel mesh processing based technique for 3D plant analysis. *Bmc Plant Biology*, 12. doi:10.1186/1471-2229-12-63.
119. PhenoApps. (n.d.). Field Book. Retrieved June 23, 2026, from <https://play.google.com/store/apps/details?id=com.fieldbook.tracker>
120. Phenotype And Trait Ontology. 2024 OBO Technical WG <https://obofoundry.org/ontology/pato.html>.
121. Pichura, V., & Potravka, L. (2025). Impact of war on natural and climatic transformation of territories in the irrigation zone of Ukraine. *Discover applied sciences*, 7(7). doi:10.1007/s42452-025-07404-4.
122. Pichura, V., Potravka, L., Domaratskiy, Y., & Petrovas, S. (2023). Spatiotemporal patterns and vegetation forecasting of sunflower hybrids in soil and climatic conditions of the Ukrainian Steppe zone. *Ukrainian Black Sea Region Agrarian Science*, 27(3), 31-45.

123. Pichura, V., Potravka, L., Stoiko, N., & Dudych, H. (2025). Scenarios for the Functioning of the Kakhovka Reservoir Territory. *Journal of landscape ecology*, 18(3), 118-154. doi:10.2478/jlecol-2025-0023.
124. Pierce, F., & Nowak, P. 67 (1999) 'Aspects of precision agriculture' *advances in agronomy*, VOL 67. 1999-01-01. pp. 1-85.
125. Pierz, L., Heslinga, D., Buell, C., & Haus, M. (2023). An image-based technique for automated root disease severity assessment using PlantCV. *Applications in plant sciences*, 11(1). doi:10.1002/aps3.11507.
126. Pilarczyk, R., Bąkowska, M., Sablik, P., Tomza-Marciniak, A., & Pilarczyk, B. (2024). Agriculture 4.0 – Innovations for sustainable development. *Acta Scientiarum Polonorum Zootechnica*, 23(2), 3–12.
127. Pinto, H., Staab, S., & Tempich, C. 110 (2004) 'DILIGENT: Towards a fine-grained methodology for Distributed, loosely-controlled and evolving Engineering of ontologies' *ECAI 2004: 16th european conference on artificial intelligence, proceedings*. 2004-01-01. pp. 393-397.
128. Plant Trait Ontology. 2024 OBO Technical WG <https://obofoundry.org/ontology/to.html>.
129. PlantCV PlantCV: Open-Source Image Analysis for Plant Phenotyping [Electronic resource]. 2024. Access mode: <https://plantcv.danforthcenter.org>  
ImageBreed ImageBreed: Image-Based Phenotyping Tools [Electronic resource]. 2023. Access mode: <https://imagebreed.org>.
130. Prusinkiewicz, P. (2004). Modeling plant growth and development. *Current opinion in plant biology*, 7(1), 79-83.
131. Ramos-Quiroga, J., Bosch, R., & Casas, M. (2009). Diagnostic assessment and multilevel phenotyping in the IMpACT study group. *European neuropsychopharmacology*, 19, S184-S185. doi:10.1016/S0924-977X(09)70216-0/.
132. Rao, N. K. S., & Laxman, R. H. (2013). Phenotyping horticultural crops for abiotic stress tolerance. *Climate-Resilient Horticulture: Adaptation and Mitigation Strategies* (pp. 147-157). Springer.

133. Richards, R. A., Rebetzke, G. J., Watt, M., Condon, A. G. T., Spielmeier, W., & Dolferus, R. (2010). Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. *Functional Plant Biology*, 37(2), 85-97.
134. Rife, T., & Poland, J. (2026). *Fieldbook-Flexport: Create a database, as easily as a spreadsheet* [Computer software]. GitHub. <https://github.com/fieldbook>.
135. Rincent, R., Charpentier, J., Faivre-Rampant, P., Paux, E., Le Gouis, J., Bastien, C., et al. (2018). Phenomic Selection Is a Low-Cost and High-Throughput Method Based on Indirect Predictions: Proof of Concept on Wheat and Poplar. *G3-genes genomes genetics*, 8(12), 3961-3972. doi:10.1534/g3.118.200760.
136. Saritha, S., & Thangaraja, G. (2023). Prediction of crop disease using Rank Regressive Learning and Proaftn Fuzzy Classification models. *Soft computing*. doi:10.1007/s00500-023-08357-9.
137. Schneiter, A., & Miller, J. (1981). Description of sunflower growth-stages. *Crop science*, 21(6), 901-903. doi:10.2135/cropsci1981.0011183X002100060024x.
138. Selby, P., Abbeloos, R., Adam-Blondon, A., Agosto-Perez, F., Alaux, M., Alic, I., et al. (2025). BrAPI v2: real-world applications for data integration and collaboration in the breeding and genetics community. *Database-the journal of biological databases and curation*, 2025. doi:10.1093/database/baaf048.
139. Sen, C., Summers, J., & Mocko, G. (2013). A Formal Representation of Function Structure Graphs for Physics-Based Reasoning. *Journal of computing and information science in engineering*, 13(2). doi:10.1115/1.4023167.
140. Shanahan, J. F., Schepers, J. S., Francis, D. D., Varvel, G. E., Wilhelm, W. W., Tringe, J. M., et al. (2001). Use of remote sensing imagery to estimate corn grain yield. *Agronomy journal*, 93(3), 583-589.
141. Shannon, C.E. (1949) Communication in the Presence of Noise. Proceedings of the IRE, 37, 10-21. <http://dx.doi.org/10.1109/JRPROC.1949.232969>

142. Shi, J., Xie, S., Li, W., Wang, X., Wang, J., Chen, Y., et al. (2025). RPT: An integrated root phenotyping toolbox for segmenting and quantifying root system architecture. *Plant biotechnology journal*, 23(6), 2095-2109. doi:10.1111/pbi.70040.
143. Sievänen, R., Godin, C., DeJong, T. M., & Nikinmaa, E. (2014). Functional–structural plant models: a growing paradigm for plant studies. *Annals of botany*, 114(4), 599-603.
144. Singh, R., Prasad, P., & Reddy, K. 129 (2015) 'Climate Change: Implications for Stakeholders in Genetic Resources and Seed Sector' *advances in agronomy*, 129. 2015-01-01. 117-180.
145. Sinha, D., Dasmandal, T., Yeasin, M., Mishra, D., Rai, A., & Archak, S. (2025). GB5mCPred: Cross-species 5mc Site Predictor Based on Bootstrap-based Stochastic Gradient Boosting Method for Poaceae. *Current bioinformatics*, 20(2), 139-148. doi:10.2174/0115748936285544231221113226.
146. SpringerLink. High-Throughput Plant Phenotyping: Methods and Protocols. 2022. URL: <https://link.springer.com/book/10.1007/978-1-0716-2537-8>].
147. Stoma, S., Frohlich, M., Gerber, S., & Klipp, E. (2011). STSE: Spatio-Temporal Simulation Environment Dedicated to Biology. *Bmc Bioinformatics*, 12. doi:10.1186/1471-2105-12-126.
148. International Maize and Wheat Improvement Center (CIMMYT). (2022). *Strengthening wheat breeding in Ukraine through digital tools* (18 p.). CIMMYT.
149. Tardieu, F., Cabrera-Bosquet, L., Pridmore, T., & Bennett, M. (2017). Plant Phenomics, From Sensors to Knowledge. *Current biology*, 27(15), R770-R783. doi:10.1016/j.cub.2017.05.055.
150. Tardieu, F., Simonneau, T., & Muller, B. (2018). The Physiological Basis of Drought Tolerance in Crop Plants: A Scenario-Dependent Probabilistic Approach. *Annual Review of Plant Biology*, Vol 69, 69, 733-759. doi:10.1146/annurev-arplant-042817-040218.
151. Tartir, S., Arpinar, I. B., Moore, M., Sheth, A. P., & Aleman-Meza, B. (2005). OntoQA: Metric-Based Ontology Quality Analysis. *Proceedings of the IEEE*

*ICDM 2005 Workshop on Knowledge Acquisition from Distributed, Heterogeneous, and Dynamic Sources*, 769–774. <https://doi.org/10.1109/ICDMW.2005.105>.

152. Integrated Breeding Platform. (2019). *The Breeding Management System (BMS)*. Retrieved February 26, 2026, from <https://www.integratedbreeding.net/>.

153. Trusova N., Vasylieva O., Kolokolchikova I., Konovalenko A., Herasymenko I. Marketing support of corporate social responsibility of agri-food enterprises. *Scientific Horizons*. 2022. 25(7). 101–114.

154. United States Agency for International Development. (2024). *Ukraine Agriculture Resilience Initiative (AGRI-Ukraine): Digital solutions report (35 p.)*. USAID.

155. UPOV TG/81/7 Sunflower Guidelines. Sunflower (*Helianthus annuus* L.) : Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability. Geneva : International Union for the Protection of New Varieties of Plants (UPOV), (2011), 40.

156. Vasylyeva, O. O., Sokolov, A. V., Lisnichenko, M. O., Halan, O. Y., & Butenko, T. V. (2023). Socio-Economic Dimension of Human Potential of World Countries in the Digital Space of the Labor Market. *Qubahan Academic Journal*, 3(4), 106–129.

157. Vasylyeva, O., Sokolov, A., Morozov, D., Horoshkova, L., Yuriechko, V. (2025). Modeling the Sustainable Development of Agrarian Sphere. In: Solovieva, V., Hushko, S. (eds) *Sustainable Development in Economics, Technology and Environmental Engineering. ISC SAI 2023. Sustainable Economy and Ecotechnology*. Springer, Cham. [https://doi.org/10.1007/978-3-031-91953-4\\_43](https://doi.org/10.1007/978-3-031-91953-4_43).

158. Veal, A. (2021). Climate change 2021: the physical science basis, 6th report. *World leisure journal*, 63(4), 443-444. doi:10.1080/16078055.2021.2008646.

159. Vedmedeva K., Machova T. (2022) Study of the possibilities of using sunflower lines with different colours of seeds to create poultry feed, *Helia* 45(76), 23-40. <https://doi.org/10.1515/helia-2021-0016>.

160. Vedmedyev, S. (2024) Rozrobka prohramnoho zabezpechennya dlya zbyrannya fenotypichnykh danykh nasinnya sonyashnyku. *Information Technology: Computer Science, Software Engineering and Cyber Security*. 1, 3–10. <https://ir.nmu.org.ua/entities/publication/f7df8236-a9f8-4ca9-bae7-dc4b314a3915>.
161. Velasco, L., Perez-Vich, B., & Fernandez-Martinez, J. M. (2016). Research on resistance to sunflower broomrape: an integrated vision. *Ocl-Oilseeds and Fats Crops and Lipids*, 23(2). doi:10.1051/ocl/2016002.
162. Villalobos, F.J., Ritchie, J.T. DSSAT Sunflower Model Documentation [Electronic resource]. University of Florida, 2023. Access mode: <https://dssat.net>
163. Wackerly, D. D., Mendenhall, W., III, & Scheaffer, R. L. (2008). *Mathematical statistics with applications* (7th ed.). Brooks/Cole, Cengage Learning 215.
164. Wagner, J., Nair, S., Skakun, S., Duncan, E., Li, F., Oliinyk, O., et al. (2025). Monitoring cropland cultivation, abandonment, fallowing and recultivation dynamics with regard to conflict intensity in war-affected Ukraine. *Science of remote sensing*, 12. doi:10.1016/j.srs.2025.100326.
165. Walter, A., Silk, W. K., & Schurr, U. (2009). Environmental effects on spatial and temporal patterns of leaf and root growth. *Annual review of plant biology*, 60(1), 279-304.
166. Walter, A., Studer, B., & Kölliker, R. (2012). Advanced phenotyping offers opportunities for improved breeding of forage and turf species. *Annals of botany*, 110(6), 1271-1279.
167. Wang, J., et al. (2022) Transcriptome Profiles Reveal Response Mechanisms and Key Role of PsNAC1 in *Pinus sylvestris* var. *mongolica* to Drought Stress. *BMC Plant Biology*. 22, 72. DOI: <https://doi.org/10.1186/s12870-022-03475-7>
168. Wang, J., Li, Y., Wang, S., Shi, L., & Duan, B. (2024). Global impact analysis and implementation strategy recommendations of the FAIR principles. *Chinese science bulletin-chinese*, 69(9), 1183-1191. doi:10.1360/TB-2023-0585.

169. Wang, R., Qu, H., & Su, W. (2025). From sensors to insights: Technological trends in image-based high-throughput plant phenotyping. *Smart agricultural technology*, 12. doi:10.1016/j.atech.2025.101257.
170. Wang, Y. S., Pyorala, J., Liang, X. L., Lehtomaki, M., Kukko, A., Yu, X. W., et al. (2019). In situ biomass estimation at tree and plot levels: What did data record and what did algorithms derive from terrestrial and aerial point clouds in boreal forest. *Remote Sensing of Environment*, 232. doi:10.1016/j.rse.2019.111309.
171. Watt, M., Fiorani, F., Usadel, B., Rascher, U., Muller, O., & Schurr, U. 71 (2020) 'Phenotyping: New Windows into the Plant for Breeders' Annual review of plant biology, 71, 2020-01-01. 689-712.
172. White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Research*, 133, 101-112.
173. Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3(1), 160018. doi:10.1038/sdata.2016.18.
174. Xie C., Yang C. (2020) A review on plant high-throughput phenotyping traits using UAV-based sensors. *Computers and electronics in agriculture*. 178: 105731. <https://doi.org/10.1016/j.compag.2020.105731>.
175. Yagaliyeva, B., Ivashchuk, O., & Goncharov, D. (2025). Modeling Approaches for Digital Plant Phenotyping Under Dynamic Conditions of Natural, Climatic and Anthropogenic Factors. *Algorithms*, 18(11). doi:10.3390/a18110720.
176. Yan, W., & Kang, M. S. (2002). GGE biplot analysis: A graphical tool for breeders, geneticists, and agronomists. *CRC press*.
177. Yazdanbakhsh, N., & Fisahn, J. (2009). High throughput phenotyping of root growth dynamics, lateral root formation, root architecture and root hair development enabled by PlaRoM. *Functional Plant Biology*, 36(10-11), 938-946. doi:10.1071/fp09167.

178. Yen, G. S., Fujimoto, B. S., Schneider, T., Kreutz, J. E., & Chiu, D. T. (2019). Statistical Analysis of Nonuniform Volume Distributions for Droplet Based Digital PCR Assays. *Journal of the American Chemical Society*, 141(4), 1515-1525. doi:10.1021/jacs.8b09073.
179. Yoosefzadeh-Najafabadi, M. (2025). Merging traditional practices and modern technology through computational plant breeding. *Plant physiology*, 199(1). doi:10.1093/plphys/kiaf355.
180. Zatserkovnyi, V., Vorokh, V., Hloba, O., Mironchuk, T., & Plichko, L. (2025). Utilizing gis, gps, remote sensing, and ai in the study of soil characteristics. *Visnyk of taras shevchenko national university of kyiv-geology*, (3), 98-107. doi:10.17721/1728-2713.110.11.
181. Zhai, Z., Martínez, J. F., Beltran, V., & Martínez, N. L. (2020). Decision support systems for agriculture 4.0: Survey and challenges. *Computers and Electronics in Agriculture*, 170, Article
182. Zhang, P., Xu, L. H., Li, D. W., & Gu, X. M. 9443 (2014) 'Improved stereo matching applied to digitization of greenhouse plants' 6th International Conference on Graphic and Image Processing (ICGIP). Oct 24-26. *Beijing, peoples r china*. Available at: <Go to ISI>://WOS:000354613300077.
183. Zhang, S. W., Huang, W. Z., & Wang, Z. (2020). Plant species identification based on modified local discriminant projection. *Neural Computing & Applications*, 32(21), 16329-16336. doi:10.1007/s00521-018-3746-0.
184. Zhang, Y., & Zhang, N. (2018). Imaging technologies for plant high-throughput phenotyping: a review. *Frontiers of agricultural science and engineering*, 5(4), 406-419. doi:10.15302/J-FASE-2018242.
185. Zhang, Z., et al. (2020) Molecular Mechanisms of Drought Resistance Using Genome-Wide Association Mapping in Maize (*Zea mays L.*). *Frontiers in Plant Science*. 11, 1109. DOI: <https://doi.org/10.3389/fpls.2020.01109>.

186. Zhu H., Huang L., Li Y. (2024) High-resolution 3D reconstruction of plants using NeRF-based models in controlled environments. *Journal of Plant Research and Imaging*. 29, 2. 145–159.

187. Ziamtsov I., Navlakha S. GrowSplat: Data-driven modeling of plant architecture using voxelized imaging. *Bioinformatics*. 2020. Vol. 36, Suppl\_1. P. i125–i132. DOI: <https://doi.org/10.1093/bioinformatics/btaa432>.

188. Аверченко, В. І., & Самойленко, Н. М. (2018). *Ґрунтознавство: Навчальний посібник*. Мачулін. <https://repository.kpi.kharkov.ua/handle/KhPI-Press/43475>.

189. Алієв, Е. Б., Ведмедєва, К. В. (2024). Кількісне фенотипування генотипів соняшнику: *монографія*. Дніпро: ЛІРА. 204.

190. Міністерство аграрної політики та продовольства України; Український інститут експертизи сортів рослин. (2018). *Атлас морфологічних ознак сортів рослин соняшнику однорічного Helianthus annuus L. (наочне доповнення до Методики проведення кваліфікаційної експертизи на ВОС соняшнику однорічного)*. Київ: Український інститут експертизи сортів рослин.

191. Бакурова А.В., Ведмедєв С.Р., Терещенко Е.В. (2023) Предметна онтологія «Селекція соняшнику». *Економічна кібернетика: інструменти, моделі і методи інтелектуальної підтримки прийняття рішень: збірник наукових праць за матеріалами Всеукраїнської інтернет-конференції*, м. Дніпро, 1-2 березня 2023 р. Дніпро : УДУНТ, 46-51.

192. Білка, Д., Шаптала, С., & Миронова, Н. (2025). Розробка прототипу цифрового двійника робототехнічного пристрою для відтворення руху в просторі. *Електротехніка та електроенергетика*, 4, 48–58. <https://doi.org/10.15588/1607-6761-2024-4-5>.

193. Васильєва, О., & Душин, Б. (2026). Цифровізація як чинник структурної трансформації та сталого розвитку малого та середнього бізнесу України. *Економіка та суспільство*, (86).

194.Ведмедєв С. Р. (2023) Розробка системи фенотипування насіння соняшнику для селекції. *Збірник тез Міжнародної наукової інтернет-конференції «Олійні культури: сьогодення та перспективи»*. (21 березня 2023 р.). Запоріжжя. ІОК НААН, 13-14.

195.Ведмедєв С. Терещенко Е. (2026) Фенотипування на основі зображень: ідентифікація оптимальних параметрів отримання зображень для аналізу морфології насіння. *Information Technology: Computer Science, Software Engineering and Cyber Security*2, 45-54.

196.Ведмедєв С., Р. (2024). Розробка програмного забезпечення для збирання фенотипічних даних насіння соняшнику. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 1, 3–10.

197.Ведмедєв С.Р., Терещенко Е.В. (2022) Створення цифрової моделі соняшника. *Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій: Тези доповідей XI Міжнародної науково-практичної конференції* (12-14 грудня 2022 р., м.Запоріжжя). [Електронний ресурс] Електрон. дані. Запоріжжя: НУ «Запорізька політехніка», 136-137.

198.Ведмедєв С.Р., Терещенко Е.В. (2025) Цифрова модель рослини соняшника для фенотипування в задачах селекції. *Інформаційні технології та суспільство*. 3 (18). 32-39. DOI <https://doi.org/10.32689/maup.it.2025.3.4> .

199.Ведмедєв, С. (2024). Розробка програмного забезпечення для збирання фенотипічних даних насіння соняшнику. *Information Technology: Computer Science, Software Engineering and Cyber Security*, 4, 53–60, <https://doi.org/10.32782/IT/2024-4-7>.

200.Ведмедєва К.В., Ведмедєв СР. (2026) Можливості використання фенотипування в селекції рослин та біотехнології. *Актуальні питання біотехнології, екології та природокористування* [Електронний ресурс]: *матеріали Міжнар. наук. конф., 16–17 квітня 2026 р./* Держ. біотехнол. ун-т. Харків: ДБТУ, Режим доступу: <http://btu.kharkov.ua/nauka/konferentsiyi/>.

201. Гнатієнко, В. Г., Гнатієнко, Г. М., Зозуля, О. Л., & Снитюк, В. Є. (2024). Метод прогнозування врожайності сільськогосподарських культур з використанням мультифакторного аналізу та нейронних мереж. *Науковий вісник Ужгородського університету. Серія «Математика і інформатика»*, 44(1), 93–105. [https://doi.org/10.24144/2616-7700.2024.44\(1\).93-105](https://doi.org/10.24144/2616-7700.2024.44(1).93-105).
202. Гнатієнко, В., & Гнатієнко, Г. (2024). Integration of machine learning and deep learning methods for sunflower yield prediction. *Управління розвитком складних систем*, (59), 225-234.
203. Господаренко, Г. М., Черно, О. Д., Мартинюк, А. Т., & Бойко, В. П. (2021). Винесення основних елементів живлення з ґрунту культурами польової сівозміни за різного удобрення. *AgroChemistry and Soil Science*, 91, 31-40.
204. Грачов О. 2025. Plant monitoring system: ШІ-система для розумного моніторингу рослин. *Інформаційні технології та суспільство*. 1 (16). 59–64. <https://doi.org/10.32689/maup.it.2025.1.7>.
205. Григорук, П., & Співак, В. (2026). Аналіз сучасного стану аграрного сектора економіки України в контексті забезпечення економічної безпеки. *Modeling the development of the economic systems*, (1), 40-54.
206. ДСТУ 2240-93. Насіння сільськогосподарських культур. Сортові та посівні якості. Технічні умови. Київ : Держстандарт України, 1994.
207. ДСТУ 2949-94. Насіння сільськогосподарських культур. Терміни та визначення. Київ : Держстандарт України, 1995.
208. ДСТУ 4138-2002. Насіння сільськогосподарських культур. Методи визначення якості. Київ : Держспоживстандарт України, 2003.
209. ДСТУ 4362:2004. Якість ґрунту. Показники родючості ґрунтів. Статус: діє. Код УКНД: 13.080.01. Примітка: чинний з 01-01-06 згідно наказу Держспоживстандарту №273 —. К.: Держспоживстандарт України, 2006. — 19 с.

210. ДСТУ 6006:2008. Насіння овочевих, баштанних культур та кормових коренеплодів. Пакування, маркування, транспортування та зберігання. Технічні умови. Київ : Держспоживстандарт України, 2009.

211. ДСТУ 7160:2020. Насіння овочевих, баштанних, кормових і пряно-ароматичних культур. Сортові та посівні якості. Технічні умови. Київ : ДП «УкрНДНЦ», 2020.

212. Зозуля, О. Л., Швартау, О. Л., & Михальська, Л. М. (2023). Сучасні методи цифрового моніторингу в рослинництві: монографія. Київ: ІФРГ НАНУ 254.

213. Іванова І. Є., Герасько Т. В., Тимошук Т. М., Покопцева Л. А., Нежнова Н. Г. (2026) Основи наукових досліджень в агрономії: навчальний посібник . Запоріжжя: ВПЦ «Люкс», 241.

214. Кендзьора, Н. З. (2020). Динаміка температури атмосферного повітря і режим опадів як фактори змін феноритміки рослин в період 2010–2019 років. *Екологія, природокористування та охорона навколишнього середовища: прикладні аспекти*, 51-54.

215. Колісніченко, В. (2025). Системний підхід до цифрової трансформації аграрного виробництва: визначення та структурно-функціональна класифікація. *Економіка та суспільство*, (78).

216. Литвин В.В., Пасічник В.В., Яцишин Ю.В. (2009) Інтелектуальні системи, Видавництво "Новий світ - 2000", Львів, 309.

217. Махова Т. В., Ведмедєв С. Р., Поляков О. І. (2024) Створення баз даних фенотипових ознак та добір ліній соняшнику особливого призначення. *Науково-технічний бюлетень Інституту олійних культур НААН. Запоріжжя*, 35. 51-62  
[http://www.bulletin.imk.zp.ua/pdf/2023/35/Makhova\\_35.pdf](http://www.bulletin.imk.zp.ua/pdf/2023/35/Makhova_35.pdf)  
<https://doi.org/10.36710/ІОС-2023-35-05>.

218. Методика проведення експертизи сортів рослин групи олійних та прядивних на відмінність, однорідність і стабільність (ВОС): Соняшник

(*Helianthus annuus* L.). (2003) Київ: Український інститут експертизи сортів рослин, 169.

219.Міністерство захисту довкілля та природних ресурсів України. (2023, 11 липня). *Про державну реєстрацію пестицидів і агрохімікатів: Наказ № 494*. <https://mepr.gov.ua/>.

220.Осієвський С.В., Третяк В.Ф. (2022) модель інформаційно–аналітичної підтримки процесів розробки знання–орієнтованих інформаційних систем. у колективній монографії: *Сучасний стан проведення наукових досліджень у IT-технологіях, галузях електроніки, інженерії, нанотехнологіях та транспортній сфері*. С. ред. Валеренко Г.І. Вінниця: Європейська наукова платформа, 23-43

221.Примак, І. Д., Купчик, В. І., Лозінський, М. В., Войтовик, М. В., Панченко, О. Б., Косолап, М. П., Коваленко, В. П., Федорук, Ю. В., Левандовська, С. М., & Панченко, І. А. (2017). *Агрономічне ґрунтознавство*. ТОВ «Нілан-ЛТД». 580.

222.Про державну систему біобезпеки при створенні, випробуванні, транспортуванні та використанні генетично модифікованих організмів: Закон України від 31 трав. 2007 р. № 1103-V. URL: <https://zakon.rada.gov.ua/laws/show/1103-16> (дата звернення: 16.06.2026).

223.Про карантин рослин : Закон України від 30 черв. 1993 р. № 3348-XII. URL: <https://zakon.rada.gov.ua/laws/show/3348-12> (дата звернення: 16.06.2026).

224.Про насіння і садивний матеріал : Закон України від 26 груд. 2002 р. № 411-IV. URL: <https://zakon.rada.gov.ua/laws/show/411-15> (дата звернення: 16.06.2026).

225. Про охорону прав на сорти рослин : Закон України від 21 квіт. 1993 р. № 3116-XII. URL: <https://zakon.rada.gov.ua/laws/show/3116-12> (дата звернення: 16.06.2026).

226.Прозатвердження Методичних вимог у сфері насінництва щодо збереження сортових та посівних якостей насіння зернових культур : наказ

Міністерства економіки, довкілля та сільського господарства України № 813. (2025) URL: <https://zakon.rada.gov.ua> (дата звернення: 16.06.2026).

227.Пустовойт, В. (1975). Подсолнечник. М.: Колос, 6-13.

228.Рябчун, В., Кузьмишина, Н., Богуславський, Р., Бондаренко, В., Музафарова, В., Холод, С., et al. (2014). Шляхи збагачення Національного генбанку рослин України. Генетичні ресурси рослин, (14), 5-21.

229.Семенова І. Г. (2017) Синоптичні та кліматичні умови формування посух в Україні: монографія, Одеський державний екологічний університет. 236.

230.Теплий, Ю., та Досин, Д. (2026). Розробка онтологічного методу для масштабованої генерації моделей частково спостережуваного марковського процесу прийняття рішень (POMDP). *Технологічний аудит та виробничі резерви*, 3 (2(89)), 13–20.

231.Чуйко, Д., Кириченко, В., & Білик, В. (2025). Agrobiological Evaluation of Sunflower Hybrids in Eastern Ukraine. *Селекція і насінництво*, (127), 56-67.

232.Шаптала, С., & Миронова, Н. (2023). Впровадження технології цифрових двійників для робототехніки. *Управління розвитком складних систем*, 53, 45–51. <https://doi.org/10.32347/2412-9933.2023.53.45-51>

## ДОДАТОК А Інструментальні засоби розробки

Інструментальні засоби розробки. Для формалізації англomовних специфікацій графічного матеріалу, а також для текстуально-стилістичного аналізу та редагування розділів дисертаційного дослідження відповідно до вимог академічного стилю МОН України, було залучено інструментарій генеративного штучного інтелекту на базі розширених мовних моделей (зокрема, LLM Gemini від Google)»

Для генерації рисунку 4.1.1

A high-resolution scientific architecture diagram for a PhD dissertation in System Analysis, presented as a 3-step pipeline with clear Ukrainian text labels.

The image layout is split horizontally into three distinct, interconnected panels:

1. LEFT PANEL (Module 1: Computer Vision & Image Processing - Python, OpenCV):

- Starts with an input thumbnail image showing a close-up pile of black and striped sunflower seeds.
- An arrow leads into a block diagram flow: "Колірна конвертація (RGB -> HSV)" -> down arrow -> "Контурний аналіз & Бінарна сегментація" -> down arrow -> "ВИХІД: Автоматичне вилучення морфологічних дескрипторів".

2. MIDDLE PANEL (Module 2: Structured Repositories - Excel Datasets):

- A large, crisp screenshot of an open Microsoft Excel spreadsheet acting as a data repository.
- Columns are cleanly labeled with scientific features like: "Sample\_ID", "Area\_mm2", "Circularity", "Mean\_Hue\_HSV", "Defect\_Score", "Genotype".
- Rows are filled with numerical metrics and strings (e.g., Line\_A1, 398.00, 0.9955, 18.17365).

### 3. RIGHT PANEL (Module 3: Ontology Mapping Subsystem - Protégé 5.5, HELIANTUS):

- A clean UI screenshot of the Protégé 5.5 ontology editor.
- The Class Hierarchy panel is visible on top, showing an expanded tree: "SeedPhenotype" -> "MorphologicalDescriptor", "SpectralDescriptor".
- Below it, the "Datatype Properties" panel is highlighted, listing properties like: "hasAreaValue", "hasHueValue", "belongsToGenotype".

#### INTEGRATION AND CONNECTIONS:

- From the Left Panel (Output) to the Middle Panel, there is a prominent red arrow showing data flow.
- Between the Middle Panel (Excel columns) and the Right Panel (Datatype Properties), there are multiple parallel, dashed red arrows pointing horizontally to represent "БЕЗПОСЕРЕДНЄ КООРДИНАТНЕ ВІДОБРАЖЕННЯ (Mapping)".
- On the far right, a final output block reads: "ВИХІД: Семантично сумісні масиви даних, готові до завантаження у простір фактів (ABox)".

STYLE: Clean corporate and academic style, minimal light background, precise engineering layout, sharp readable fonts, professional vector look, no realistic human hands or artistic clutter.

Для генерації рисунку 4.6.1

A highly detailed, professional scientific flowchart for a PhD dissertation in System Analysis, illustrating an image processing pipeline for plant seed analysis. The diagram uses a clean, modern academic design with a light gray background and clear Ukrainian text labels. The layout flows sequentially across three main vertical panels from left to right.

### 1. INPUT & LEFT PANEL (ПРЕПРОЦЕСИНГ - Preprocessing):

- Top input block: "Вхід: Зображення насінини (орієнтоване по діагоналі)".
- Below it, a thumbnail shows a single sunflower seed tilted diagonally with the label "-45° rotation".
- An arrow leads to a process block: "Поворот на +315° (вирівнювання)" with a rotation icon.
- Next block: "Обрізка чорних ділянок" with a scissors icon.
- Bottom thumbnail shows the sunflower seed now perfectly aligned vertically.

### 2. MIDDLE PANEL (СЕГМЕНТАЦІЯ ТА АНАЛІЗ - Segmentation and Analysis):

- A continuous vertical pipeline connected by arrows:
  - \* Block 1: "Розділення каналів: R, G, B" with a small seed icon.
  - \* Block 2: "Порогова фільтрація: R,G,B < 100" with a settings/sliders icon.
  - \* Block 3: "Формування бінарної матриці (numb\_matrix)" with a binary grid icon.
  - \* Block 4: "Приведення до квадратної форми" with an axis/grid icon.
- Then, a split block titled "АНАЛІЗ ДІАГОНАЛЕЙ" containing two parallel sub-blocks: "Аналіз Головних Діагоналей (довжина)" and "Аналіз Побічних Діагоналей (ширина)".
- Bottom block of this panel: "ПОБУДОВА МОДЕЛІ РОМБУ (вписаний контур)" with a diamond/rhombus icon.

### 3. RIGHT PANEL (КАЛІБРУВАННЯ ТА ВИВІД - Calibration and Output):

- Top large section: "РОЗРАХУНОК ГЕОМЕТРИЧНИХ ПАРАМЕТРІВ (в пікселях)" with a rhombus geometry icon on the top left.
  - \* Inside, a structured table contains two sections on the left: "Вхідні дані (в пікселях)" lists [firstAngle, secondAngle, thirdAngle, fourthAngle, square] and [S\_Dug1, S\_Dug2, S\_Dug3, S\_Dug4]. "Вихідні дані (міліметри/градуси)" lists

[S\_Dug1, S\_Dug2, S\_Dug3, S\_Dug4] and [zolH, zolW, heightSeed (пікс), widthSeed (пікс)].

\* To the right of the table, mathematical equations are clearly rendered:

$$\text{line\_a} = \text{mm}(\text{a-}, \text{line\_b}) / \text{line\_a}$$

$$\text{line\_b} = \text{mm}(\text{b-}, \text{line\_b}) / \text{line\_b}$$

$$\text{line\_c} = \text{mm}(\text{a-}, \text{line\_c}) / \text{line\_c}$$

$$\text{line\_d} = \text{mm}(\text{c-}, \text{line\_d}) / \text{line\_d}$$

- Middle section: "КАЛІБРУВАННЯ" flowchart:

\* "Визначення еталона (шаблону)" -> arrow -> "РОЗРАХУНОК ЦІНИ ПІКСЕЛЯ (ЦП)".

\* Next to it, the formula box:  $\text{ЦП} = \text{mm}(\text{фактична довжина}) / \text{пікс}(\text{кількість пікселів шаблону})$ .

- Lower section: A horizontal merger: "ПЕРЕВЕДЕННЯ РОЗМІРІВ У МЕТРИЧНІ ОДИНИЦІ" and a box with the formula: "Кінцевий розрахунок:  $\text{Розмір}(\text{мм}) = \text{Розмір}(\text{пікс}) * \text{ЦП}$ ".

#### 4. OUTPUT & BOTTOM:

- A final horizontal block at the bottom right: "Вивід: Збереження результатів (results.csv)" with a document/file icon.

STYLE REQUIREMENTS: Linear, technical, precise layout, flat design, professional color scheme using shades of light blue, gray, and dark text. All text must be in clean, crisp, highly legible Cyrillic/Ukrainian sans-serif font. No artistic elements, no human hands, no 3D reflections.

## ДОДАТОК Б Акти впровадження



НААН  
 ІНСТИТУТ ОЛІЙНИХ КУЛЬТУР  
 НАЦІОНАЛЬНОЇ АКАДЕМІЇ АГРАРНИХ НАУК УКРАЇНИ  
 ІОК НААН

вул. Інститутська, 1, сел. Сонячне, Запорізький район, Запорізька обл., 70417, Україна  
 Тел./факс: (061) 223 99 50, E-mail: iocnaas@gmail.com, ЄДРПОУ 01296051

20.12.2023 № 448

На № \_\_\_\_\_ від \_\_\_\_\_

## А К Т

## впровадження результатів науково-дослідної роботи

Даним актом стверджується, що результати наукового проекту ДР ДБ05022 (Державний реєстраційний номер: 0121U113264) «Розвиток методів дослідження складних соціально-економічних систем на основі інтелектуальних технологій», кер. проф. Бакурова А.В., які отримані у НУ «Запорізька політехніка» впроваджені у лабораторії генетики та генетичних ресурсів Інституту олійних культур Національної академії аграрних наук України.

Розроблену систему цифровізації візуальної інформації для дослідження фенотипу соняшнику застосовано у процесах створення розробки приладу фенотипування насіння соняшнику та створення бази даних морфологічних ознак ліній соняшнику. Результати виконання цих робіт будуть у подальшому використовуватися у проведенні селекційних доборів, ідентифікації зразків, та створенні нового наукового продукту у вигляді ліній, гібридів, та методів оцінки.

Заступник директора  
 з наукової роботи, к.б.н. с.н.сп.

Катерина ВЕДМЕДСВА



впровадження результатів НДР

28 листопада 2025 р.

м. Запоріжжя

Складено комісією у складі:

Голови Анна БАКУРОВА  
(Ім'я ПРІЗВИЩЕ)

члени комісії Дмитро ШИРОКОРАД, Еліна ТЕРЕЩЕНКО, Вадим РИБІН  
(Ім'я ПРІЗВИЩЕ)

Комісія провела роботу по визначенню фактичного впровадження практично-науково-дослідної роботи **05014 Прийняття ефективних рішень на основі інтелектуальних технологій та відкритих даних**. Науковий керівник доц. Рябенко А.С.

назва та № роботи виконаної у НУ "Запорізька політехніка" згідно (назва тематичного плану) \_\_\_\_\_  
тематичного плану по кафедральних НДР на 2024-2027 н.р.

та установила, що результати вказаної роботи впроваджені на кафедрі «Системний аналіз та обчислювальна математика» НУ «Запорізька політехніка»

Назва організації ( підприємства), структурного підрозділу

Вид та об'єм впровадження.

Розроблено *програмний продукт* для визначення фенотипових характеристик насінника соняшника.

Результати виконання цієї роботи впроваджені в якості прикладу на лекціях і під час виконання лабораторних робіт з курсів «Теорія прийняття рішень», «Аналіз даних», «Основи системного аналізу» та «Моделювання соціально-економічних систем», а також як змістовна задача для курсового та дипломного проєктування за спеціальністю F4 «Системний аналіз та наука про дані» для першого (бакалаврського) та другого (магістерського) рівнів навчання вищої освіти.

Досягнуті технічні результати, їх рівень: підвищення якості підготовки фахівців.

Складено в 2 прим.: 1-й прим. В НДЧ НУ "Запорізька політехніка"

2-й прим. Виконавцю


Голова комісії



Анна БАКУРОВА

(Ім'я ПРІЗВИЩЕ)

Члени комісії



Дмитро ШИРОКОРАД

(Ім'я ПРІЗВИЩЕ)



Еліна ТЕРЕЩЕНКО

(Ім'я ПРІЗВИЩЕ)



Вадим РИБІН

(Ім'я ПРІЗВИЩЕ)



НААН  
ІНСТИТУТ ОЛІЙНИХ КУЛЬТУР  
НАЦІОНАЛЬНОЇ АКАДЕМІЇ АГРАРНИХ НАУК УКРАЇНИ  
ІОК НААН

вул. Інститутська, 1, сел. Сонячне, Запорізький район, Запорізька обл., 70417, Україна  
E-mail: iocnaas@gmail.com, <http://imk.zp.ua/>, ЄДРПОУ 01296051

18.06.2026 № 162

**Акт впровадження**

Результати дисертаційної роботи Ведмедева С. Р. «Підтримка прийняття рішень при вирощуванні соняшнику на основі цифрової технології фенотипування» у вигляді методу цифрового фенотипування об'єктів агросистем на основі аналізу цифрових зображень та онтологічну модель вирощування соняшника запроваджено для фенотипування насіння соняшнику у лабораторії генетики, генетичних ресурсів та інтродукції Інституту олійних культур НААН і використано для виконання фундаментальних завдань Національної академії аграрних наук України: ПНД 17 «Генетичні ресурси рослин» 17.01.01.15.Ф Формування колекцій олійних культур, виділення та створення джерел і донорів цінних господарських ознак, ДР№ 0120U105673; ПНД 25 Формування, використання і збереження генетичних ресурсів рослин (Генетичні ресурси рослин). 25.01.01.07.Ф Формування колекцій олійних культур, виділення та створення джерел і донорів цінних господарських ознак №ДР 0126U001156В наукових дослідженнях

В.о. директора



Петро БАЛАБАЙ

## ДОДАТОК В Фрагменти коду

### Лістинг В.1 — Модуль обробки зображень

```
import os
import numpy as np
import csv
from PIL import Image
import cv2
import math

IMAGES_DIRECTORY = '9-11'

def main():
    results = []

    for filename in os.listdir(IMAGES_DIRECTORY):
        file_path = os.path.abspath(os.path.join(IMAGES_DIRECTORY, filename))

        if not os.path.isfile(file_path):
            continue

        result = process_image(file_path)

        if result is not None:
            results.append(result)

    headers = [
        'Name', 'lineA', 'lineB', 'lineC', 'lineD',
        'firstAngle', 'secondAngle', 'thirdAngle', 'fourthAngle',
        'square', 'S_Dug1', 'S_Dug2', 'S_Dug3', 'S_Dug4',
        'zolH', 'zolW', 'widthSeed', 'heightSeed'
    ]

    with open('results.csv', 'w', newline='', encoding='utf-8') as f:
        writer = csv.writer(f, delimiter=';')
        writer.writerow(headers)
        writer.writerows(results)

if __name__ == "__main__":
    main()
```

### Лістинг В.2 — Обробка одного зображення та виділення ознак

```
def process_image(path):
    image = Image.open(path)

    image = image.rotate(325, resample=Image.Resampling.BILINEAR)
    image = image.crop((250, 258, 1437, 974))

    image_np = np.array(image)
    image_cv = cv2.cvtColor(image_np, cv2.COLOR_RGB2BGR)

    r, g, b = cv2.split(image_cv)

    binary_mask = (b > 37).astype(np.uint8)
```

```

area = np.sum(binary_mask)

if area < 40000:
    return None

max_row, max_col = find_coordinates(binary_mask)

if not max_row or not max_col:
    return None

(x1, y1), (x2, y2) = max_row
(x3, y3), (x4, y4) = max_col

lineA = abs(y1 - y2)
lineB = abs(y2 - y1)
lineC = abs(x3 - x4)
lineD = abs(x4 - x3)

width = lineA + lineB
height = lineC + lineD

return [os.path.basename(path), lineA, lineB, lineC, lineD,
        0, 0, 0, 0,
        area, 0, 0, 0, 0,
        0, 0, width, height]

```

### Лістинг В.3 — Пошук координат структури об'єкта

```

def find_sequences(lst, min_length=120):
    sequences = []
    start = -1

    for i, val in enumerate(lst):
        if val == 1:
            if start == -1:
                start = i
            else:
                if start != -1 and i - start >= min_length:
                    sequences.append((start, i - 1))
                start = -1

    if start != -1 and len(lst) - start >= min_length:
        sequences.append((start, len(lst) - 1))

    return sequences

def find_coordinates(matrix):
    max_row = None
    max_col = None
    max_row_len = 0
    max_col_len = 0

    rows, cols = len(matrix), len(matrix[0])

    for i in range(rows):
        for start, end in find_sequences(matrix[i]):

```

```

        if end - start > max_row_len:
            max_row_len = end - start
            max_row = ((i, start), (i, end))

    for j in range(cols):
        col = [matrix[i][j] for i in range(rows)]
        for start, end in find_sequences(col):
            if end - start > max_col_len:
                max_col_len = end - start
                max_col = ((start, j), (end, j))

    return max_row, max_col

```

Лістинг В.4 — Аналіз зображення за діагоналлю та формування спектральних ознак

```

import os
import csv
import numpy as np
import pandas as pd
from PIL import Image

IMAGES_DIRECTORY = '514'
STATISTICS_DIRECTORY = 'statistics514'
IMAGE_SIZE = 1200
LINES_COUNT = 8

def main():
    results = []

    for filename in os.listdir(IMAGES_DIRECTORY):
        file_path = os.path.abspath(os.path.join(IMAGES_DIRECTORY, filename))

        if os.path.isfile(file_path):
            results.append(process_image(file_path))

    headers = get_headers()

    with open('514.csv', 'w', newline='') as f:
        writer = csv.writer(f, delimiter=';')
        writer.writerow(headers)
        writer.writerows(results)

def get_headers():
    headers = ['Name']

    for i in range(1, LINES_COUNT - 1):
        headers += [f'red_avg{i+1}', f'green_avg{i+1}', f'blue_avg{i+1}']

    return headers

def process_image(path):
    image = Image.open(path).resize((IMAGE_SIZE, IMAGE_SIZE))
    pix = image.load()

```

```

filename = os.path.basename(path)

return [filename] + process_image_pix(filename, pix)

def process_image_pix(filename, pix):

    if not os.path.exists(STATISTICS_DIRECTORY):
        os.makedirs(STATISTICS_DIRECTORY)

    csv_filename = os.path.join(
        STATISTICS_DIRECTORY,
        os.path.splitext(filename)[0] + '.csv'
    )

    with open(csv_filename, 'w', newline='') as f:
        writer = csv.writer(f, delimiter=',')

        writer.writerow(['x', 'y', 'r', 'g', 'b', 'r_razn', 'g_razn', 'b_razn'])

        for i in range(1, IMAGE_SIZE):
            r, g, b = pix[i, i]

            diff = np.abs(
                np.array(pix[i, i]) - np.array(pix[i - 1, i - 1])
            )

            r_razn, g_razn, b_razn = diff

            writer.writerow([i, i, r, g, b, r_razn, g_razn, b_razn])

    statistics = pd.read_csv(csv_filename, delimiter=',')
    statistics = statistics.sort_values(by="r_razn", ascending=False)

    newdf = statistics.nlargest(LINES_COUNT, "r_razn")

    return [filename] + process_statistics(newdf, statistics)

def process_statistics(newdf, statistics):

    max_values = []
    for _ in range(0, LINES_COUNT - 1):
        max_val = max(x for x in newdf['x'] if x not in max_values)
        max_values.append(max_val)

    max_values.reverse()

    filtered = [statistics[statistics.x < max_values[0]]]

    for i in range(1, len(max_values)):
        filtered.append(
            statistics[
                (statistics['x'] < max_values[i]) &
                (statistics['x'] >= max_values[i - 1])
            ]
        )

```

```
filtered.append(statistics[statistics.x >= max_values[-1]])

results = []

for i in range(1, LINES_COUNT - 1):
    red_avg = int(filtered[i]['r'].mean())
    green_avg = int(filtered[i]['g'].mean())
    blue_avg = int(filtered[i]['b'].mean())

    results += [red_avg, green_avg, blue_avg]

return results

if __name__ == "__main__":
    main()
```