



SANA LOGOS – Республикалық ғылыми журнал., Том 1, №3-2026

ӨОЖ: 58:004.8

<https://doi.org/10.5281/zenodo.23477056>

Өсімдіктер систематикасындағы жасанды интеллект цифрлық гербарийлер түрлерді анықтау және таксономиялық зерттеулердің жаңа мүмкіндіктері (әдеби шолу)

А.А. Құдайберген^{1*} ,

¹әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан

e-mail: aiauzhankudaibergen@gmail.com

Қабылданған күні: 29.07.2026 | Қабылданды: 23.09.2026 | Жарияланды: 30.09.2026

Аңдатпа. Бұл әдеби шолудың мақсаты – цифрлық гербарийлер мен жасанды интеллект әдістерінің өсімдіктер систематикасындағы орнын, түрлерді автоматтандырылған анықтау мүмкіндіктерін және таксономиялық зерттеулерге ықпалын кешенді бағалау. Әдебиеттер Google Scholar арқылы табуға болатын рецензияланған жарияланымдар арасынан тақырыптық сәйкестік бойынша іріктеліп, DOI және баспа беттерімен салыстырыла тексерілді. Негізгі іздеу кезеңі 2015–2026 жылдарды қамтыды, ал тұжырымдамалық маңызы бар ертерек еңбектер қосымша енгізілді. Нәтижелер цифрлық гербарийлердің сурет пен жапсырма метадеректерін біріктіретін ауқымды зерттеу инфрақұрылымына айналғанын көрсетті. Терең оқыту үлгілері ірі және сапалы таңбаланған деректер жиындарында түрлерді жоғары дәлдікпен ажырата алады; сонымен бірге өсімдік мүшелерін сегментациялау, морфологиялық белгілерді өлшеу, фенологиялық фазаларды тану, жапсырма мәтінін шығару және ықтимал қате анықтамаларды белгілеу мүмкіндіктері кеңеюде. Алайда дәлдік көрсеткіштері деректердің теңгерімсіздігіне, географиялық және коллекциялық ығысуға, сирек таксондардың аз ұсынылуына, таксономиялық атаулардың өзгеруіне және модельдің бұрын көрмеген түрлерді тану қабілетіне тәуелді. Сондықтан жасанды интеллект таксономисті алмастыратын құрал ретінде емес, сарапшы басқаратын, шешім қабылдауды қолдайтын жүйе ретінде қарастырылуы тиіс. Болашақ бағыттар мультимодальды деректерді, түсіндірілетін модельдерді, ашық классты тануды және халықаралық дерек стандарттарын біріктіруге негізделеді.

Түйін сөздер: жасанды интеллект, өсімдіктер систематикасы, цифрлық гербарий, компьютерлік көру, түрлерді анықтау, машиналық оқыту, интегративті таксономия.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В СИСТЕМАТИКЕ РАСТЕНИЙ: ЦИФРОВЫЕ ГЕРБАРИИ, ИДЕНТИФИКАЦИЯ ВИДОВ И НОВЫЕ ВОЗМОЖНОСТИ ТАКСОНОМИЧЕСКИХ ИССЛЕДОВАНИЙ (ЛИТЕРАТУРНЫЙ ОБЗОР)

Аннотация. Цель данного литературного обзора – комплексно оценить роль цифровых гербариев и методов искусственного интеллекта в систематике растений, возможности автоматизированной идентификации видов и влияние этих технологий на таксономические



исследования. Литература отбиралась по тематическому соответствию среди рецензируемых публикаций, доступных для поиска через Google Scholar, с последующей перекрёстной проверкой DOI и страниц издательств. Основной период поиска охватывал 2015–2026 годы, при этом дополнительно включались более ранние работы, имеющие концептуальное значение. Результаты показали, что цифровые гербарии превратились в масштабную исследовательскую инфраструктуру, объединяющую изображения образцов и метаданные с этикеток. Модели глубокого обучения способны с высокой точностью различать виды на крупных и качественно аннотированных наборах данных. Кроме того, расширяются возможности сегментации органов растений, измерения морфологических признаков, распознавания фенологических фаз, извлечения текста с этикеток и выявления потенциально ошибочных определений. Однако показатели точности зависят от несбалансированности данных, географических и коллекционных смещений, недостаточной представленности редких таксонов, изменений таксономической номенклатуры и способности моделей распознавать ранее не встречавшиеся виды. Поэтому искусственный интеллект следует рассматривать не как инструмент замены таксономиста, а как систему поддержки принятия решений под контролем эксперта. Перспективные направления дальнейших исследований связаны с интеграцией мультимодальных данных, разработкой интерпретируемых моделей, применением методов распознавания открытого класса и внедрением международных стандартов данных.

Ключевые слова: искусственный интеллект, систематика растений, цифровой гербарий, компьютерное зрение, идентификация видов, машинное обучение, интегративная таксономия.

ARTIFICIAL INTELLIGENCE IN PLANT SYSTEMATICS: DIGITAL HERBARIA, SPECIES IDENTIFICATION, AND NEW OPPORTUNITIES FOR TAXONOMIC RESEARCH (A LITERATURE REVIEW)

Abstract. The aim of this literature review is to comprehensively assess the role of digital herbaria and artificial intelligence methods in plant systematics, the potential of automated species identification, and the impact of these technologies on taxonomic research. The literature was selected from peer-reviewed publications searchable through Google Scholar based on thematic relevance, with DOI records and publisher websites cross-checked for verification. The main search period covered 2015–2026, while earlier studies of conceptual significance were also included. The findings indicate that digital herbaria have evolved into large-scale research infrastructures integrating specimen images with label metadata. Deep learning models can distinguish species with high accuracy when trained on large, high-quality annotated datasets. In addition, capabilities for plant organ segmentation, morphological trait measurement, phenological stage recognition, label text extraction, and the flagging of potentially erroneous identifications are expanding. However, accuracy metrics depend on data imbalance, geographic and collection biases, the underrepresentation of rare taxa, changes in taxonomic nomenclature, and the ability of models to recognize previously unseen species. Therefore,



artificial intelligence should be regarded not as a replacement for taxonomists but as an expert-supervised decision-support system. Future research directions include the integration of multimodal data, interpretable models, open-set recognition, and international data standards.

Keywords: *artificial intelligence, plant systematics, digital herbarium, computer vision, species identification, machine learning, integrative taxonomy.*

Кіріспе

Өсімдіктер систематикасы өсімдік алуантүрлілігін сипаттау, атау, жіктеу және эволюциялық байланыстарын түсіндіру міндеттерін атқарады. Бұл міндеттердің эмпирикалық негізі – түрдің морфологиялық белгілерін, таралуын, жиналған орнын және таксономиялық анықтамаларын сақтайтын гербарий үлгілері. Гербарий парағы тек кептірілген өсімдік емес: ол қайта тексеруге болатын заттай куәлік, жапсырмадағы кеңістік-уақыттық дерек және кейінгі ревизиялардың тарихы. Сондықтан гербарий коллекцияларын цифрландыру систематикадағы деректердің қолжетімділігі мен салыстырмалылығын түбегейлі өзгертті (Soltis, 2017; Meineke et al., 2019).

Жаппай цифрландыру нәтижесінде миллиондаған жоғары ажыратымдылықтағы кескіндер GBIF, iDigBio және институционалдық виртуалды гербарийлер арқылы таратылады. Бұл қорлар зерттеушіге бір коллекциямен шектелмей, әр аймақ пен тарихи кезеңдегі үлгілерді қатар талдауға мүмкіндік береді. Бірақ цифрлық дерек көлемінің өсуі жаңа кедергі туғызды:

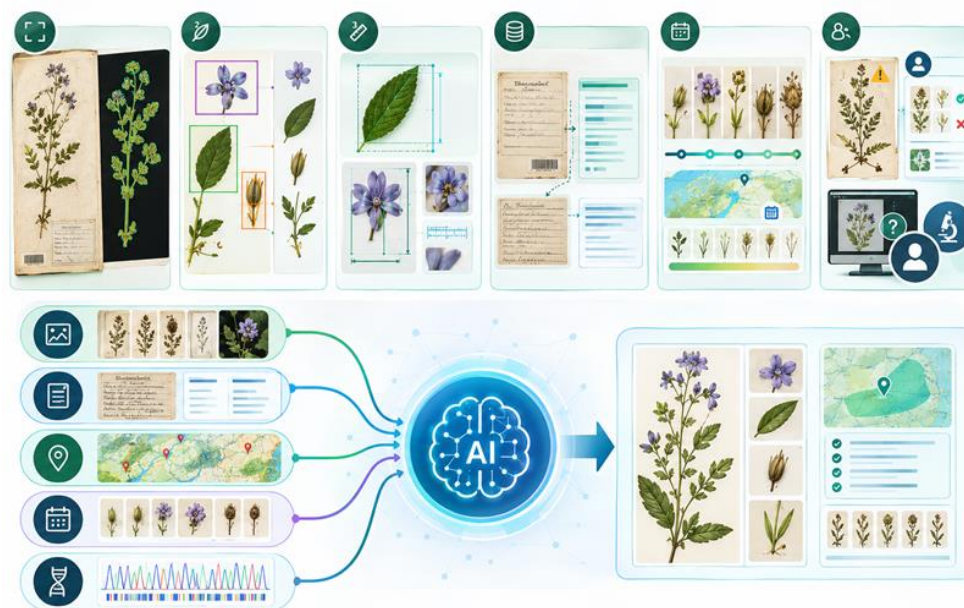
суреттерді қолмен анықтау, жапсырмаларды транскрипциялау, мүшелерді белгілеу және қате жазбаларды тексеру қарқынды дерек ағынына ілесе алмайды. Осы тұста компьютерлік көру, табиғи тілді өңдеу және машиналық оқыту жүйелері қайталанатын операцияларды жеделдететін құрал ретінде маңызға ие болды (Pearson et al., 2020; Hussein et al., 2022).

Жасанды интеллектінің өсімдік систематикасындағы қолданылуы тек «суреттен түр атауын табумен» шектелмейді. Қазіргі жүйелер гербарий парағындағы өсімдік пен фон элементтерін бөледі, жапырақ, гүл, жеміс сияқты мүшелерді табады, морфометриялық белгілерді өлшейді, жапсырма мәтінін құрылымдандырады, фенологиялық күйді бағалайды және коллекциядағы күмәнді анықтамаларды сарапшыға ұсынады (1-сурет). Соңғы еңбектер мұны бірнеше дерек түрін – кескін, мәтін, география, фенология және генетикалық ақпаратты – біріктіретін интегративті таксономиялық AI бағытына ұластырады (Guo et al., 2025; Takano et al., 2026).



SANA LOGOS

SANA LOGOS – Республикалық ғылыми журнал., Том 1, №3-2026



1-сурет. Жасанды интеллекттің өсімдік систематикасындағы қолданылуы: интегративті таксономиялық талдау

Мақаланың мақсаты – цифрлық гербарийлерге негізделген жасанды интеллект әдістерінің түрлерді анықтау мен таксономиялық зерттеулерге қосатын үлесін әдеби деректер негізінде жүйелеу, олардың сенімділік шектерін көрсету және өсімдіктер систематикасындағы болашақ зерттеу бағыттарын айқындау.

Материалдар мен әдістер

Жұмыс эмпирикалық эксперимент емес, құрылымдалған нарративтік әдеби шолу ретінде орындалды. РСС логикасы қолданылды: зерттеу нысаны – гербарий үлгілері мен өсімдік таксондары; негізгі ұғым – жасанды интеллект, машиналық оқыту және компьютерлік көру; контекст – цифрлық гербарийлер, түрлерді анықтау және таксономиялық зерттеу. Үш сұрақ қойылды: 1) цифрлық гербарий деректері AI үшін қандай инфрақұрылым қалыптастырады; 2) AI өсімдік

түрін анықтау және гербарийді өңдеу міндеттерінде қандай нәтижелер көрсетеді; 3) бұл әдістерді таксономиялық шешімге айналдыруға қандай шектеулер кедергі жасайды? Ақпарат іздеу стратегиясы Google Scholar-да табылатын ғылыми еңбектерді қамту мақсатымен ағылшын тіліндегі мына ұғымдар жеке және біріктіріліп пайдаланылды: “artificial intelligence”, “machine learning”, “deep learning”; “plant taxonomy”, “plant systematics”, “species identification”, “species delimitation”; “digital herbarium”, “digitized herbarium specimen”; “computer vision”; “label extraction”; “trait extraction”; “open-set recognition”. Нәтижелердің библиографиялық дұрыстығы журналдың ресми беті, DOI жазбасы, PubMed Central немесе ұйымдық репозиторий арқылы салыстырыла тексерілді. Тікелей тақырыптық еңбектердің әдебиеттер тізімдері бойынша кері



дәйексөздік іздеу де жүргізілді. Google Scholar библиографиялық іздеу құралы ретінде пайдаланылғанымен, басылымның сенімділігі оның Scholar-да көрінуімен ғана емес, рецензиялануы, баспасы, әдісінің ашықтығы және дерекке сәйкестігі арқылы бағаланды.

Деректерді шығару және синтездеу

Әр еңбек бойынша авторы, жылы, дерек көзі, зерттеу мақсаты, дерек жиынтығының сипаты, модель немесе аналитикалық тәсіл, негізгі нәтиже, авторлар көрсеткен шектеулер және систематикаға қолданылуы тіркелді. Дәлдік, үлгі саны және таксон саны тек бастапқы мақалада немесе ресми толық мәтінде нақты көрсетілген жағдайда ғана пайдаланылды. Әртүрлі міндеттердің метрикалары бір-біріне тікелей теңестірілмеді: классификация дәлдігі, объектіні анықтау көрсеткіші және белгіні өлшеу сапасы бөлек түсіндірілді. Зерттеулер әдістемелік гетерогенді болғандықтан мета-талдау жасалмады; нәтижелер тақырыптық синтез арқылы цифрлық инфрақұрылым, автоматты идентификация, дерек шығару, таксономиялық интерпретация және тәуекелдер топтарына бөлінді. Бұл жұмыс толық жүйелі шолу немесе PRISMA есебі ретінде ұсынылмайды, себебі барлық дерекқорлар бойынша тәуелсіз қосарланған скрининг жүргізілген жоқ.

Нәтижелер

Цифрлық гербарийлер AI үшін дерек инфрақұрылымы ретінде

Гербарийлерді цифрландыру физикалық коллекцияны әлемдік деңгейде іздеуге болатын дерек жүйесіне айналдырды. Цифрлық жазба әдетте үлгі кескінін, таксон атауын, жинаушыны, күнді, географиялық

орынды, коллекция нөмірін және кейінгі анықтау белгілерін қамтиды. Осындай құрылым модельді оқытуға қажетті кескін-таңба жұптарын береді. Soltis (2017) цифрландырудың тарихи үлгілерді жаңа зерттеу сұрақтарына қайта пайдалануға жол ашқанын көрсетті; Meineke et al. (2019) табиғи-тарихи коллекцияларды антропоцендегі биоалуантүрлілік өзгерістерін түсіндіретін ұзақ мерзімді дерек ретінде бағалады.

AI үшін гербарий кескіндерінің маңызды артықшылығы – көпшілігінің мамандар берген таксономиялық атаумен байланысуы. Сонымен қатар бұл атаулар толық мінсіз «ақиқат белгі» емес: кейбір жазбалар ескірген номенклатураны сақтайды, кейбір үлгілер қате анықталған, ал бір таксон әр коллекцияда әртүрлі атпен берілуі мүмкін. Цифрландыру қарқыны да біркелкі емес; ірі мекемелер мен кең таралған түрлер басым, шағын гербарийлер, сирек түрлер және географиялық тұрғыдан аз зерттелген аймақтар жеткіліксіз көрінеді. Демек, цифрлық гербарий – дайын және бейтарап оқу жиынтығы емес, кураторлық тарихы бар дерек көзі.

Түрлерді автоматтандырылған анықтау

Carranza-Rojas et al. (2017) гербарий кескіндерін терең нейрондық желілермен жіктеудің жүзеге асатынын ірі масштабта көрсетті. Бұл еңбек кейінгі гербарийлік компьютерлік көру зерттеулеріне негіз болды. Little et al. (2020) ұйымдастырған Herbarium 2019 байқауында Melastomataceae тұқымдасының 46 469 үлгісі мен 683 түрі қолданылып, үздік төрт модель 88%-дан жоғары дәлдікке жетті. Мұндай нәтиже алгоритмдердің маман жұмысын жылдамдата



алатынын көрсеткенімен, деректер бір жаңа, гибридті, толық емес немесе таксономиялық топпен шектелді және анықталмаған материалмен кездеседі. түрлердің оқу мысалдары тең бөлінбеді.

Жапония флорасының 500 554 кескіні мен 2171 таксонында жасалған жүйе 96,4% дәлдік көрсетті және модель назар аударған кескін бөліктерін талдау арқылы қате анықталуы ықтимал үлгілерді белгілеуге мүмкіндік берді (Takano et al., 2022). Бұл нәтиже жоғары сапалы, аймақтық тұрғыдан үйлесімді коллекцияда AI өте тиімді болатынын дәлелдейді. Алайда көрсеткішті басқа құрлықтарға, басқа гербарий стандарттарына немесе оқу жиынтығында жоқ түрлерге тікелей көшіруге болмайды. Жабық классты сынақта барлық тест түрлері алдын ала белгілі болады; нақты гербарий жұмысында жүйе

Ірі деректердегі өкілдік оқыту да маңызды бағыт болып отыр. Walker et al. (2022) бес гербарийден алынған 2,25 миллионнан астам кескін мен 64 500 түрді пайдаланып, кескіндердің жалпыға тасымалданатын сандық сипаттамаларын үйренді. Үштік желі арқылы алынған кеңістік бірнеше кейінгі міндетке жақсырақ ауысқан. Мұндай морфологиялық кеңістік ұқсас үлгілерді топтауға, атипиялық жазбаларды табуға және сарапшыға салыстыру үшін жақын үлгілерді ұсынуға мүмкіндік береді.

Қолдану бағыты	Негізгі дерек	AI нәтижесі	Таксономиялық мәні
Түрді жіктеу	Гербарий парағының кескіні	Таксонға ықтималдық береді	Алдын ала анықтау және ревизияны басымдықтау
Сегментация	Өсімдік пен парақ элементтері	Өсімдікті фоннан бөледі	Шу әсерін азайтып, өлшеуді тұрақтандырады
Мүше тану	Жапырақ, гүл, жеміс, сабақ	Мүшелерді табады және санайды	Диагностикалық белгілер мен фенологияны шығарады
Мәтін шығару	Жапсырма және аннотация	OCR және NER арқылы өрістерді бөледі	Локалитет, күн, жинаушы және анықтамаларды құрылымдайды
Морфологиялық кеңістік	Көп үлгінің үйренілген белгілері	Ұқсастықтар мен аномалияларды көрсетеді	Түр шекарасын тексеруге гипотеза ұсынады

1-кесте. Цифрлық гербарийлердегі AI міндеттері және олардың систематикадағы қызметі

Морфологиялық белгілер мен LeafMachine және оның кейінгі нұсқалары фенологиялық ақпаратты автоматты шығару жапырақтарды табу, өлшеу және Таксономия үшін тек соңғы түр атауы морфологиялық белгілерді көп мөлшерде емес, оны негіздейтін белгілер маңызды. шығару жұмысын автоматтандырады (Weaver et al., 2020). GinJinn сияқты объектіні анықтау құбырлары Younis et al. (2020) жапырақ пен гербарий кескіндеріндегі құрылымдарды сабақтың жақсы танылғанын, ал ұсақ немесе автоматты табуға арналған (Ott et al., 2020). қабаттасқан гүл мүшелерінің қиынырақ



болғанын көрсетті. Бұл айырмашылық таксономиялық белгілердің визуалды қолжетімділігі модель нәтижесіне тікелей әсер ететінін білдіреді.

Фенологиялық кезеңдерді анықтау гүлдеу және жеміс беру уақытын тарихи қатарлар бойынша зерттеуге мүмкіндік береді. Pearson et al. (2020) гербарий кескіндерінен фенологиялық дерек алуға арналған модульдік машиналық оқыту жұмыс үдерісін сипаттады. Мұнда модель адам еңбегін түгел жоймайды: сарапшы оқу дерегін белгілейді, модель көп кескінді өңдейді, ал қателері бақылау іріктемесінде бағаланады. Осы схема таксономиялық зерттеуге де сәйкес келеді, себебі морфологиялық белгіні өлшеу мен оның таксономиялық мағынасын түсіндіру екі бөлек кезең болып қала береді.

Жапсырма мәтіні мен коллекциялық метадеректерді өңдеу

Гербарий парағындағы жапсырма түр атауымен қатар үлгінің қай жерде, қашан және кім жинағанын көрсетеді. OCR мәтінді таниды, ал атаулы мәндерді тану алгоритмдері сол мәтіннен ел, елді мекен, координата, жинаушы, күн және коллекция нөмірі секілді өрістерді бөледі. Соңғы жүйелер объектіні анықтау, OCR және табиғи тілді өңдеуді бір құбырға біріктіріп, қолжазба, типографиялық әркелкілік және көптілді жапсырмалар жағдайында жұмыс істеуге ұмтылады. Мұндай автоматтандыру үлгілерді картаға түсіруді, синонимдерді салыстыруды және тарихи анықтамаларды іздеуді жеделдетеді, бірақ нәтижелерді Darwin Core сияқты стандарттарға сәйкестендіріп, адамның сапа бақылауынан өткізу қажет.

Түр шекаралары және интегративті таксономия

Классификатор белгілі атаулардың бірін таңдайды, ал түрді делимитациялау деректердегі шекаралардың өзін тексереді. Сондықтан автоматты идентификация мен жаңа түрді сипаттау бір міндет емес. Бақылаусыз оқыту, кластерлеу және өлшемділікті азайту морфологиялық немесе генетикалық кеңістіктегі табиғи топтарды көрсетуге көмектеседі, бірақ кластер өздігінен биологиялық түр болып саналмайды. Species delimitation 4.0 тұжырымдамасы AI көмегімен морфология, геномика, экология және географияны бірге талдап, субъективтілікті азайтуды ұсынады; түпкі шешім бірыңғай түр тұжырымдамасы мен таксономиялық сараптамаға сүйенуі керек.

Takano et al. (2026) ұсынған интегративті таксономиялық AI тәсілінде жеке кескінге бір атау беру емес, көптеген үлгіні ортақ үйренілген морфологиялық кеңістікте салыстыру басым. Бұл кеңістік түрішілік өзгергіштікті, ықтимал гибридтерді, шеткі үлгілерді, қате анықтамаларды және қазіргі түр шекарасына сәйкес келмейтін топтарды көрсетуі мүмкін. AI мұнда жаңа таксономиялық гипотеза түзеді; гипотеза типтік материал, диагноз, номенклатура, география, молекулалық филогения және сарапшының морфологиялық бағалауымен тексеріледі.

Талқылау

Ғылыми және практикалық артықшылықтар

Әдебиеттер AI-дың ең күшті әсері ауқым мен басымдықта екенін көрсетеді. Біріншіден, модель миллиондаған кескін арасынан белгілі



бір таксонға, мүшеге немесе фенологиялық күйге сәйкес жазбаларды жылдам іріктейді. Екіншіден, күмәнді немесе атипиялық үлгілерді белгілеп, сарапшы уақытын ең күрделі жағдайларға бағыттайды. Үшіншіден, бірдей алгоритмдік ереже көптеген коллекцияға қолданылғанда өлшеу қайталанымдылығы жоғарылайды. Төртіншіден, цифрлық кескіндер физикалық үлгіні қайта-қайта ұстау қажеттігін азайтып, сирек және типтік материалды қашықтан салыстыруға жағдай жасайды.

Жасанды интеллект әсіресе «адам–алгоритм» бірлескен үдерісінде тиімді. Модель алғашқы болжамды және сенімділік мәнін береді; таксономист диагностикалық белгілерді, атаулардың қолданылуын және типтік үлгілерді тексереді; расталған түзетулер қайтадан оқу дерегіне қосылады. Осындай белсенді оқыту циклі сирек таксондарға арналған таңбаланған деректі біртіндеп жақсартады және модель қатесін жасырмай, кураторлық шешімге айналдырады.

Сенімділік шектері және әдіснамалық тәуекелдер

Бірнеше кең таралған түрдің мыңдаған кескіні бар, ал сирек түрлердің бірен-саран үлгісі ғана кездеседі. Жалпы дәлдік жоғары болғанымен, сирек немесе қорғауды қажет ететін таксондағы қате үлесі жасырын қалуы мүмкін. Сондықтан macro-F1, әр таксон бойынша сезімталдық, top-k дәлдігі және калибрленген сенімділік секілді көрсеткіштер қажет. Модель өсімдік морфологиясымен қатар гербарий қағазының түсін, масштаб сызғышын, штрихкодты, мекеме мөрін немесе кескін түсіру стилін үйренуі мүмкін. Бір

коллекцияда жақсы нәтиже көрсеткен жүйе басқа гербарийде әлсірейді.

Көп модель тест кескінін оқу кезінде көрген түрлердің біріне мәжбүрлеп жатқызады. Нақты таксономиялық жұмыста анықталмаған, жаңа, гибриді немесе оқу жиынтығынан тыс таксондар бар. Ашық классты тану жүйесі мұндай үлгіні «белгісіз» деп тоқтатып, сарапшыға жіберуі тиіс. Жылу картасы модель назар аударған аймақты көрсете алады, бірақ ол ботаникалық диагностикалық белгіні пайдаланғанын толық дәлелдемейді. Сондықтан түсіндірме морфологиялық негізбен және қарсы мысалдармен тексерілуі керек (Guo et al., 2025).

Гербарий қорлары жинау экспедицияларының, қолжетімді жолдардың, саяси шекаралардың және ғылыми басымдықтардың ізін сақтайды. AI осы үлгіні қайталайды: деректе аз аймақ пен аз таксон модельде де әлсіз болады. Бұған қоса, қабылданған атаулар мен синонимдер уақыт өте өзгереді. Оқу жиынтығын тұрақты таксономиялық анықтамалықпен нұсқалау, бастапқы verbatim identification өрісін сақтау және модель шығарған атаудың қай номенклатуралық нұсқаға сәйкес екенін тіркеу қажет.

Зерттеу инфрақұрылымына қойылатын талаптар

Сенімді жүйе құру үшін кескіннің тұрақты идентификаторы, коллекция коды, таксон атауының авторы, анықтаушы мен анықтау күні, географиялық дерек, лицензия және таксономиялық нұсқа сақталуы тиіс. Оқу, валидация және тест жиынтықтарын кездейсоқ кескін деңгейінде ғана бөлу бір үлгінің



көшірмелері немесе бір жинау оқиғасының сериясы арқылы дерек ағып кетуіне әкелуі мүмкін. Сондықтан бөлу үлгі идентификаторы, мекеме, география немесе уақыт бойынша жүргізіліп, сыртқы коллекцияда сынақ жасалғаны дұрыс. Код, модель салмақтары, дерек іріктеу ережелері және қате матрицалары ашық берілгенде ғана нәтиже қайталанып, басқа флораға бағаланады.

Қазақстан және Орталық Азия гербарийлері үшін басым қадам – бірден күрделі универсал модель жасау емес, сапалы цифрлық корпус қалыптастыру. Алдымен типтік және флористикалық маңызды үлгілерді жоғары сапада түсіру, жапсырмаларды екі немесе үш тілде құрылымдау, синонимдерді біріздендіру және сарапшы растаған шағын эталон жиынтықтарын құру қажет. Одан кейін аймақтық модельді сыртқы гербарий материалымен сынап, сирек түрлер бойынша қателерді бөлек көрсету керек. Бұл тәсіл жергілікті флораны халықаралық дерек желілеріне қосып, сонымен бірге жергілікті таксономиялық білімнің жоғалуына жол бермейді.

Бұл шолу барлық библиографиялық дерекқорды толық қамтыған жүйелі шолу емес. Негізгі назар ағылшын тіліндегі, онлайн тексеруге болатын рецензияланған әдебиетке аударылды; сондықтан жергілікті тілдердегі және индекстелуі әлсіз еңбектер толық көрінбеуі мүмкін. Өртүрлі зерттеу деректері, таксон саны және метрикалары біркелкі болмағандықтан, нәтижелер сандық түрде біріктірілмеді. Кейбір жаңа еңбектер 2025–2026 жылдары жарияланғандықтан, олардың

тәуелсіз қайталануы мен ұзақ мерзімді әсері әлі толық бағаланбаған.

Қорытынды

Цифрлық гербарийлер өсімдіктер систематикасына жоғары көлемді, тексерілетін және тарихи терең дерек берді; жасанды интеллект сол деректі іздеу, жіктеу және өлшеу мүмкіндігін кеңейтті. Әдебиеттер терең оқыту жүйелерінің түрлерді анықтау, мүшелерді табу, белгілерді шығару, фенологияны бағалау және жапсырма мәтінін құрылымдау міндеттерінде нақты нәтижелер бергенін көрсетеді. Ең перспективалы өзгеріс – жеке кескінге атау беруден көптеген үлгіні ортақ морфологиялық және мультимодальды кеңістікте салыстыруға көшу.

Сонымен бірге AI нәтижесі таксономиялық акт емес. Жоғары дәлдік деректер жиынтығының шегінде ғана мағыналы; жаңа түрді тану, түр шекарасын өзгерту немесе номенклатуралық шешім қабылдау үшін типтік материал, диагноз, әдебиет, генетикалық және экологиялық деректер, сондай-ақ маманның жауапкершілігі қажет. Сондықтан болашақтағы тиімді модель – түсіндірілетін, белгісіз таксонды қабылдайтын, бірнеше дерек модальдігін біріктіретін және адам бақылауын жүйенің міндетті бөлігі ретінде сақтайтын модель. Осындай жағдайда AI өсімдіктер систематикасын алмастырмайды, оның дәлелдік ауқымын және зерттеу жылдамдығын арттырады.



Әдебиеттер тізімі

1. Carranza-Rojas, J., Goëau, H., Bonnet, P., Mata-Montero, E., & Joly, A. (2017). Going deeper in the automated identification of herbarium specimens. *BMC Evolutionary Biology*, 17, 181. <https://doi.org/10.1186/s12862-017-1014-z>
2. Guo, Y.-Y., Cai, H., Bramley, G. L. C., Atkins, H. J., Li, B., & Theodossiades, S. (2025). A review of artificial intelligence in herbarium specimen image analysis. *Artificial Intelligence Review*, 58, 402. <https://doi.org/10.1007/s10462-025-11408-2>
3. Hussein, B. R., Malik, O. A., Ong, W.-H., & Slik, J. W. F. (2022). Applications of computer vision and machine learning techniques for digitized herbarium specimens: A systematic literature review. *Ecological Informatics*, 69, 101641. <https://doi.org/10.1016/j.ecoinf.2022.101641>
4. Little, D. P., Tulig, M., Tan, K. C., Liu, Y., Belongie, S., Kaeser-Chen, C., Michelangeli, F. A., Panesar, K., Guha, R. V., & Ambrose, B. A. (2020). An algorithm competition for automatic species identification from herbarium specimens. *Applications in Plant Sciences*, 8(6), e11365. <https://doi.org/10.1002/aps3.11365>
5. Meineke, E. K., Davies, T. J., Daru, B. H., & Davis, C. C. (2019). Biological collections for understanding biodiversity in the Anthropocene. *Philosophical Transactions of the Royal Society B*, 374, 20170386. <https://doi.org/10.1098/rstb.2017.0386>
6. Ott, T., Palm, C., Vogt, R., & Oberprieler, C. (2020). GinJinn: An object-detection pipeline for automated feature extraction from herbarium specimens. *Applications in Plant Sciences*, 8(6), e11351. <https://doi.org/10.1002/aps3.11351>
7. Pearson, K. D., Nelson, G., Aronson, M. F. J., Bonnet, P., Brenskelle, L., Davis, C. C., Denny, E. G., Ellwood, E. R., Goëau, H., Heberling, J. M., Joly, A., Lorieul, T., Mazer, S. J., Meineke, E. K., Stucky, B. J., Sweeney, P., & White, A. E. (2020). Machine learning using digitized herbarium specimens to advance phenological research. *BioScience*, 70(7), 610–620. <https://doi.org/10.1093/biosci/biaa044>
8. Soltis, P. S. (2017). Digitization of herbaria enables novel research. *American Journal of Botany*, 104(9), 1281–1284. <https://doi.org/10.3732/ajb.1700281>
9. Takano, A., Moon, H., Shirai, M., Tagane, S., Vasques, D. T., Lee, C.-K., Khan, A., Seo, J., Han, S., Park, C., Lee, J., Akihiro, T., & Kim, S. (2026). Herbarium specimens in the age of artificial intelligence: From herbarium image identification to Integrative Taxonomic AI. *Journal of Plant Research*. <https://doi.org/10.1007/s10265-026-01743-y>
10. Takano, A., et al. (2022). Development of a system for the automated identification of herbarium specimens with high accuracy. *Scientific Reports*, 12, 8066. <https://doi.org/10.1038/s41598-022-11450-y>
11. Walker, B. E., Tucker, A., & Nicolson, N. (2022). Harnessing large-scale herbarium image datasets through representation learning. *Frontiers in Plant Science*, 12, 806407. <https://doi.org/10.3389/fpls.2021.806407>
12. Wäldchen, J., & Mäder, P. (2018). Plant species identification using computer vision techniques: A systematic literature review. *Archives of Computational Methods in Engineering*, 25, 507–543. <https://doi.org/10.1007/s11831-016-9206-z>
13. Weaver, W. N., Ng, J., & Laport, R. G. (2020). LeafMachine: Using machine learning to automate leaf trait extraction from digitized herbarium specimens. *Applications in Plant Sciences*, 8(6), e11367. <https://doi.org/10.1002/aps3.11367>
14. Younis, S., Weiland, C., Hoehndorf, R., Dressler, S., Hickler, T., Seeger, B., & Schmidt, M. (2018). Taxon and trait recognition from digitized herbarium specimens using deep convolutional neural networks. *Botany Letters*, 165(3–4), 377–383. <https://doi.org/10.1080/23818107.2018.1446357>
15. Younis, S., Schmidt, M., Weiland, C., Dressler, S., Seeger, B., & Hickler, T. (2020). Detection and annotation of plant organs from digitised herbarium scans using deep learning. *Biodiversity Data Journal*, 8, e57090. <https://doi.org/10.3897/BDJ.8.e57090>