

<https://doi.org/10.46341/PI2026002>

UDC 57.082 + 57.087.3

SHORT COMMUNICATION

## Short notice on choosing the herbarium digitization strategy

 Andriy Novikov

State Museum of Natural History of the NAS of Ukraine, Teatralna str. 18, 79008 Lviv, Ukraine; novikoffav@gmail.com

**Received:** 18.03.2026 | **Accepted:** 03.04.2026 | **Published:** 14.04.2026

---

### Abstract

Herbarium digitization is a strategic process that transforms physical collections into accessible, standardized, and interoperable digital format, thereby improving their preservation, accessibility, and scientific value. The paper emphasizes that successful digitization requires long-term planning, including clear objectives, workflows, resource allocation, quality control, and sustainable data management. It proposes organizing digitization through three complementary teams (i.e., Digitization, Archiving, and Implementation) and frames their roles through the DIPA<sup>2</sup> concept, which covers data, imaging, publishing, archiving, and application.

**Keywords:** herbarium digitization, digitization strategy, digital specimen, DIPA<sup>2</sup> concept, biodiversity data, collection management

---

**Funding:** None.

**Competing Interests:** The author has declared that no competing interests exist.

---

Herbarium digitization has become an essential direction in the modernization of natural history collections. Its primary aims are to improve access to collection- and collectors-based information, support the long-term preservation of collection knowledge, and increase the scientific, educational, and social value of herbaria. By converting physical specimens and their associated data into digital resources, institutions can reduce barriers to access, facilitate collection management, and enable broader participation in biodiversity research. The digital transformation of herbaria and the growing interoperability of collection data have facilitated the development of a so-called metaherbarium, defined by accessible, standardized, and interconnected data resources (Davis, 2023).

The prospects of herbarium digitization extend far beyond the creation of digital copies. Digitized collections can serve as a foundation for taxonomic, ecological, biogeographical, conservation, and historical studies, while also supporting education, public outreach, and international data exchange (Heberling et al., 2019; Paton et al., 2020, 2025). In this sense, digitization should be understood not as an end in itself, but as a strategic process, which transforms herbarium collections into active and interoperable research infrastructure (Davis, 2023). The long-term profit of digitization therefore depends not only on the efficiency of data capture, but also on how well digitized materials are preserved, published, integrated, and applied in future work.

Before initiating any herbarium digitization project, it is essential to define

not only technical parameters but also the overall strategy that will guide the entire digitization lifecycle. Digitization is not simply the mechanical conversion of physical specimens into digital format. It is a long-term commitment that requires combination of infrastructure, workflows, quality control, and sustainable data management (Nieva de la Hidalga, 2020; Borsch et al., 2020; Dillen et al., 2024). Therefore, strategic planning should address several interconnected aspects: the objectives of digitization, overall digitization scope, prioritization within the collection, data standards, resource allocation, and the intended long-term use of the derived digital materials.

In large or distributed collections, strategic planning becomes even more critical (e.g., Thiers et al., 2016; Dillen et al., 2024). A formal digitization plan should define clear timelines, resource requirements, responsibilities, and expected outcomes for each stage of the process. It should also specify how digitized assets will be stored, curated, and made accessible both for internal use and public. Without such planning, digitization efforts risk becoming fragmented, unsustainable, or underutilized.

A practical approach is to structure digitization activities around three complementary teams with distinct responsibilities. The *Digitization Team* focuses on capturing data and producing digital surrogates, ensuring standardized acquisition and metadata capture. The *Archiving Team* ensures long-term preservation, integrity, and security of digital assets, including backups and preservation formats. The *Implementation Team* works to maximize the value of digitized materials through data integration, publication, research, education, and secondary dissemination. Such division of labor helps to ensure both efficiency and long-term sustainability.

The roles and tasks of these teams can be effectively conceptualized using the *DIPA<sup>2</sup> concept* (Novikov, 2025; Novikov & Nachychko, 2025a), which distinguishes five functional layers of work with digitized collections:

- D – Data (database creation, as well as data mobilization, structuring, and management);
- I – Image (imaging capturing, structuring, and management);

- P – Publishing (making data publicly available);
- Ar – Archiving (long-term preservation and storage);
- Ap – Application (use of data in research, education, policy, etc.).

Depending on institutional priorities, available resources, and strategic goals, different combinations of these layers can be implemented. The digitization pathway chosen, e.g., *Object-to-Data* (OD), *Object-to-Image* (OI), *Object-to-Image-to-Data* (OID), or *Object-to-Data-to-Image* (ODI), strongly influences the downstream workflows and resource requirements (Nelson et al., 2015; Novikov & Nachychko, 2025b). These four digitization models define how physical specimens are transformed into digital assets and what type of information is prioritized at each stage.

As summarized in [Table 1](#), the *DIPA<sup>2</sup>* framework yields 32 distinct scenarios depending on the chosen digitization model and the subsequent actions taken with the digitized materials. These scenarios illustrate the full spectrum of possible strategies, ranging from simple data capture without long-term storage or dissemination, to fully integrated workflows involving publication, archiving, and advanced application.

A key principle illustrated by these scenarios is that intended use should guide the digitization strategy. If long-term archiving is not planned, producing archival-quality images may be unnecessary. If public dissemination is not anticipated, strict adherence to international data standards may be deprioritized at least at the initial stages of the digitization process. However, digitization projects often evolve over time, and what may appear unnecessary at the outset may later become critical. Therefore, institutions are encouraged to adopt a future-oriented approach and design workflows that enable scalability, interoperability, and reuse. In practical terms, this means aiming for more comprehensive scenarios such as *IDPArAp* or *DIPArAp*, which combine data mobilization, imaging, publication, archiving, and their active or prospective application.

Ultimately, a well-designed digitization strategy enhances not only access and preservation but also scientific collaboration,

data integration, and the long-term impact of collections. Planning digitization with flexibility and future use in mind ensures that digital assets remain valuable and sustainable over time.

## References

Borsch, T., Stevens, A.-D., Häffner, E., Güntsch, A., Berendsohn, W. G., Appelhans, M., Barilaro, C., Beszteri, B., Blattner, F., Bossdorf, O., Dalitz, H., Dressler, S., Duque-Thüs, R., Esser, H.-J., Franzke, A., Goetze, D., Grein, M., Grünert, U., Hellwig, F., ... Jürgens, N. (2020). A complete digitization of German herbaria is possible, sensible and should be started now. *Research Ideas and Outcomes*, 6, Article e50675. <https://doi.org/10.3897/rio.6.e50675>

Davis, C.C. (2023). The herbarium of the future. *Trends in Ecology & Evolution*, 38(5), 412–423. <https://doi.org/10.1016/j.tree.2022.11.015>

Dillen, M., Abraham, L., Bogaerts, A., De Smedt, S., Engledow, H., Leliaert, F., Trekels, M., Dessein, S., & Groom, Q. (2024). The Meise Botanic Garden Herbarium data management plan. *Research Ideas and Outcomes*, 10, Article e124288. <https://doi.org/10.3897/rio.10.e124288>

Heberling, J.M., Prather, L.A., & Tonsor, S.J. (2019). The changing uses of herbarium data in an era of global change: an overview using automated content analysis. *BioScience*, 69(10), 812–822. <https://doi.org/10.1093/biosci/biz094>

Nelson, G., Sweeney, P., Wallace, L.E., Rabeler, R.K., Allard, D., Brown, H., Carter, J.R., Denslow, M.W., Ellwood, E.R., Germain-Aubrey, C.C., Gilbert, E., Gillespie, E., Goertzen, L.R., Legler, B., Marchant, D.B., Marsico, T.D., Morris, A.B., Murrell, Z., Nazaire, M. ... Mast, A.R. (2015). Digitization workflows for flat sheets and packets of plants, algae, and fungi. *Applications in Plant Sciences*, 3, Article 1500065. <https://doi.org/10.3732/apps.1500065>

Nieva de la Hidalga, A., Rosin, P., Sun, X., Bogaerts, A., De Meeter, N., De Smedt, S., Strack van Schijndel, M., Van Wambeke, P., & Groom, Q. (2020). Designing an herbarium digitisation workflow with built-in image quality management. *Biodiversity Data Journal*, 8, Article e47051. <https://doi.org/10.3897/bdj.8.e47051>

Novikov, A. (2025, June 27–29). Justification of the application of the DIP(a) concept in the digitization of natural history collections. In *Proceedings of the Scientific-Practical Conference “Current problems of studying the entomofauna of the western region of Ukraine”* (pp. 29–30). State Museum of Natural History of the NAS of Ukraine, Lviv.

Table 1. DIPA<sup>2</sup> framework distinct scenarios.

Digitization model	Manipulations with digitized materials				
	None	Publishing	Archiving	Application	Publishing & Archiving
OD (Object-to-Data)	DN	DP	DAR	DAP	DPAp
OI (Object-to-Image)	IN	IP	IAR	IAP	IPAp
OID (Object-to-Image-to-Data)	IDN	IDP	IDAR	IDAP	IDPAp
ODI (Object-to-Data-to-Image)	DIN	DIP	DIAR	DIAP	DIPAap

- Novikov, A., & Nachychko, V. (2025a). *Digitization of herbarium collections: educational-methodological guide*. Helvetica, Odesa. (In Ukrainian)
- Novikov, A., & Nachychko, V. (2025b). The digitisation workflow of the herbarium of the State Museum of Natural History of the NAS of Ukraine (LWS). *Biodiversity Data Journal*, 13, Article e148861. <https://doi.org/10.3897/BDJ.13.e148861>
- Paton, A.J., Ameka, G.K., Antonelli, A., Asase, A., Barrett, R.L., Bogaerts, A., Cardoso, D., Carine, M., Culham, A., Dalimunthe, S.H., Davies, N., De Smedt, S., Demissew, S., Forzza, R.C., Groom, Q., Haston, E.M., Kartonegoro, A., Kersey, P., Larridon, I., ... Young, A. (2025). Life after herbarium digitisation: physical and digital collections, curation and use. *Plants, People, Planet, Early View*, 1–13. <https://doi.org/10.1002/ppp3.70078>
- Paton, A.J., Antonelli, A., Carine, M.A., Forzza, R.C., Davies, N.M., Demissew, S., Dröge, G., Fulcher, T., Grall, A., Holstein, N., Jones, M., Liu, U., Miller, J.T., Moat, J., Nicolson, N., Ryan, M.J., Sharrock, S., Smith, D.W., Thiers, B.M., Victor, J.E., Wilkinson, T., & Dickie, J.B. (2020). Plant and fungal collections: current status, future perspectives. *Plants, People, Planet*, 2, 499–514. <https://doi.org/10.1002/ppp3.10141>
- Thiers, B.M., Tulig, M.C., & Watson, K.A. (2016). Digitization of The New York Botanical Garden Herbarium. *Brittonia*, 68, 324–333. <https://doi.org/10.1007/s12228-016-9423-7>

## Коротке зауваження щодо вибору стратегії оцифрування гербарію

Андрій Новіков

Державний природознавчий музей НАН України, вул. Театральна, 18, Львів, 79008, Україна; novikoffav@gmail.com

Оцифрування гербаріїв – це стратегічний процес, який перетворює фізичні колекції в операбельний, стандартизований та сумісний цифровий формат, тим самим покращуючи їх збереження, доступність та наукову цінність. У статті наголошується, що успішне оцифрування вимагає довгострокового планування, включаючи чіткі цілі, робочі процеси, розподіл ресурсів, контроль якості та стале управління даними. У ній пропонується організувати оцифрування через три взаємодоповнюючі команди (а саме, оцифрування, архівування та впровадження) та визначає їхні ролі за допомогою концепції DIPA<sup>2</sup>, яка охоплює дані, візуалізацію, публікацію, архівування та застосування.

**Ключові слова:** оцифрування гербарію, стратегія оцифрування, цифровий зразок, концепція DIPA<sup>2</sup>, дані про біорізноманіття, управління колекцією