Collectomics in plant biodiversity research - looking into the past to understand the present and shape the future

Solveig Franziska Bucher [1, 2, 3, 4*], Sebastian Gebauer [1, 2], Jonas Grieb [5], Matthias Körschens [2, 3, 6], Jochen Müller [1, 2], Christiane M. Ritz [3, 7, 8], Rajapreethi Rajendran [5], Claus Weiland [5], Karsten Wesche [3, 7, 8], Kristin Victor [1, 2], Christine Römermann [1, 2, 3, 4]

[1] Senckenberg Institute for Plant Form and Function Jena (SIP) Jena, Philosophenweg 12, Jena, Germany

[2] Institute of Biodiversity, Ecology and Evolution, Friedrich Schiller University Jena, Philosophenweg 16, Jena, Germany

[3] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Puschstrasse 4, Leipzig, Germany

[4] Michael Stifel Center Jena for Data-driven and Simulation Science, Friedrich Schiller University Jena, Jena, Germany

[5] Senckenberg - Leibniz Institution for Biodiversity and Earth System Research, Senckenberganlage 25, Frankfurt, Germany

[6] Institute of Computer Science, Friedrich Schiller University Jena, Ernst-Abbe-Platz 2, Jena, Germany

[7] Senckenberg Museum of Natural History Görlitz, Am Museum 1, Görlitz, Germany

[8] International Institute Zittau, Dresden University of Technology, Markt 23, Dresden, Germany

*Solveig-franziska.bucher@senckenberg.de

Abstract: Global biodiversity is changing at unprecedented rates during the Anthropocene. Whereas current biodiversity patterns can be observed directly, information from the recent past is far less easily retrieved yet urgently needed to understand present observations and predict future developments. For plants, herbaria offer such a unique glimpse into the past. Evaluation of plant specimens allows determining a wide range of attributes like species identity, morphological and phenological traits and even signs of biotic interactions. Specimen's labels convey data such as species identity (and identification history), date and locality of collection, as well as the surrounding biotic and abiotic environment. Current methodological developments in sensor technology and computer vision increasingly enable us to extract this information in a high throughput and automated way. Equally vast developments in data science allow to integrate data from other sources for much more comprehensive analyses than before. With millions of specimens already digitized and digitization schemes running in many institutions, we will be increasingly able to determine characteristics of species and link them via distribution records to large-scale climate change scenarios. This allows us to better predict species' threat levels, and to develop scenarios on the consequences of biodiversity change for ecosystem functioning. The present contribution reviews recent herbaria research and describes potential avenues with respect to Museomics and the Extended Specimen, and we propose Collectomics as a new framework to unravel, understand, and cope with the Anthropocene biodiversity change.

Keywords: Anthropocene Biodiversity loss; Computer vision: Herbaria; Extended Specimen; Natural history collection

1. Introduction

The term Anthropocene and its exact start are still under debate, but clearly humans have become a major force on planet Earth. This is defined by easily traceable golden spikes in geological records (Lewis and Maslin 2015, Zalasiewicz et al. 2015), e.g. the discovery of the so called "New World", and the 16th century subsequent spread of infectious diseases as well as exchange of biodiversity between continents (Lewis and Maslin 2015), or the radioactivity peak after launch of the first nuclear weapons, i.e., 1950 or 1964 (Lewis and Maslin 2015, Zalasiewicz et al. 2015). Whichever date will be finally agreed upon, humanity has become a major driving force of change in life on Earth, which has subsequently and irreversibly changed at unprecedented rates during the Anthropocene. Not only have human actions caused a shift in global nutrient cycles as well as climate conditions (IPCC 2007, Hoegh-Guldberg et al. 2018) but they have also triggered a massive decline in biodiversity (Johnson et al. 2017). It is now widely acknowledged that these changes started to affect the functioning of entire ecosystems and will do more so in the future. This may ultimately also threaten the well-being of humankind.

While few long-term datasets span decades or even centuries, natural history collections offer valuable archives to study the changes during the Anthropocene (Meineke et al. 2019b). Such information from the past is urgently needed to analyse and understand patterns and drivers of biodiversity change. Based on this, we can build better models that can predict the future consequences of global change and take measures to mitigate the impact of our actions. In this work, we want to outline novel approaches that we propose to summarise under the term Collectomics, which aims to ever more strongly combining scientific collections with informatics (Sigwart et al. 2025). This 'omics' approach does not only allow to retrieve and use data from natural history collections where we have digital specimens, but also to do this automatically and thus very quickly, opening the avenue for high-throughput data analyses for a vastly extended values chain. We also aim to highlight the impacts of Anthropocene changes on plant life as documented in herbaria. Based on this, we present our vision for modern ways of analysing these changes, and apply this framework to e.g., zoological collections as well.

2. Historical Herbarium collections – unlocking the treasure trove

Since the 16th century, scientists have collected herbarium specimens mainly of vascular plant species, but also of bryophytes and other cryptogams, and the value of these collections for modern science is widely acknowledged in various fields of biodiversity research (Johnson et al. 2011, Lavoie 2013, Meineke et al. 2018, Lang et al. 2019, Raxworthy and Smith 2021, Baldini et al. 2022, Davis 2023, Flannery 2023, Jones et al. 2024, Karbstein et al. 2024). These specimens are preserved and curated in herbaria, which do not only house the plant material itself, but for most specimens (at least from the 19th century onwards) also preserve records on where, when and by whom specimens were collected. We can use this data as information from the past, especially in the light of global change. Worldwide, herbarium collections currently house nearly 400 million specimens (Davis 2023, Thiers 2024), and they continue to grow. We are still just beginning to unlock this vast treasure trove, and further progress offers the possibility to shape the future of biodiversity research.

Herbarium specimens have been used mainly in taxonomy and floristics (e.g., species description and identification, biogeography, floristic accounts), but these fields had been increasingly declining as many herbaria faced cut of funding; entire collections were closed which also led to a lack of taxonomists and curators (Miller et al. 2020, Edwards et al. 2024). More recently, herbaria have, however, experienced a renaissance for more advanced scientific research (e.g., molecular systematics, macroecology and evolution, future predictions), as a wider circle of scientists has begun to recognize their value. Herbarium specimens can document shifts in geographic ranges of plant species, with many studies reporting shifts to higher elevation and latitude in the course of ongoing climate warming (Graham et al. 2004, Newbold 2010, Feeley 2012). A particularly easily monitored facet of Anthropocene biodiversity change is the arrival of neophytes. Since new plant species have always been of vast interest to collectors, their distribution has been particularly well recorded (Lavoie et al. 2007, Crawford and Hoagland 2009). The analysis of historic plant material from herbaria can furthermore help to assess whether a species is naturally rare or has declined due to anthropogenic activities (Albani Rocchetti et al. 2021, Vörös et al. 2025), one of the several applications for assessments in the context of IUCN Red Lists (Zizka et al. 2021, Zizka et al. 2022). It is also estimated that more than 50% of undescribed species are already harboured in our herbaria and simply need to be identified (Little et al. 2020).

We can also derive information beyond species identity, facilitated by rapidly developing techniques for studying specimens (Figure 1). Changes in species morphology and physiology due to changes in temperature, increases in nitrogen and heavy metal depositions and the rising CO_2 concentration of the atmosphere can be documented over time, using samples derived from the preserved plant tissues such as stomatal imprints or isotopic signals (Woodward 1987, Law and Salick 2005, Miller-Rushing et al. 2009, Bonal et al. 2011, Leger 2013, Lang et al. 2019, Heberling 2022). Genomic DNA can be extracted from the preserved plant material to study phylogenetic relationships, phylogeographic patterns, population and range dynamics, or the spread of plant diseases (Gugerli et al. 2005, Wandeler et al. 2007, Yoshida et al. 2013, Meineke et al. 2018, Gutaker et al. 2019, Burbano and Gutaker 2023). Modern spectroscopic methods allow for determination of nutrient concentrations in leaves without the necessity of destructive sampling (Kothari et al. 2023, Kühn et al. 2024, Kühn et al. 2025). Likewise, specific secondary compounds of plants can be analysed, and the evolution of metabolic pathways thus be monitored (Lankau et al. 2009, Mendes Resende et al. 2020, Barnes et al. 2023). Herbaria have proven instrumental in studies on various aspects of plant phenology, the recurring events in life history such as flowering or senescence, and collections thus allowed to extend the time span of phenological observations into the past (Primack et al. 2004, Bolmgren and Lönnberg 2005, Miller-Rushing et al. 2006, Everill et al. 2014, Ramirez-Parada et al. 2022, Ahlstrand et al. 2023, Lee et al. 2024, Iwanycki Ahlstrand et al. 2025). Some studies also use plant collections to assess the impact of insect herbivory of the past with admittedly varying degrees of success (Meineke et al. 2018, Meineke et al. 2019a, Kozlov et al. 2020, Ward 2024), or the interactions between flowering species and their pollinators themselves (Ziska et al. 2016, Streher et al. 2024). First studies also showed success in reviving already extinct plant species from propagules, which are sometimes included in herbarium specimens (Abeli et al. 2020). Not only has the range of biodiversity facets that can be studied with specimens increased, but also are the data almost exclusively digital, enabling a completely new level of integration.

3. Herbarium digitization

Soltis (2017) gives a nice overview on the use of digitization in herbaria and argues that just like fine wine, the value of these collections increases with age (Soltis 2017, Younis et al. 2020b, Karbstein et al. 2024). Herbarium collections are only recently used in Computer Science or in collaborations between informatics and biodiversity research, which opens new directions of herbarium research. Given the new methods available in computer vision, a huge field of potential and fruitful collaborations emerges (Heberling et al. 2019, Hedrick et al. 2020, Hardisty et al. 2022), especially with respect of analysing species and trait responses to Anthropocene biodiversity change. During the past decades, many herbaria conducted massive digitization efforts of their collections (Willis et al. 2017, Hedrick et al. 2020). Meanwhile, information on millions of herbarium specimens available through research platforms such as the Global Biodiversity Information Facility (GBIF, https://www.gbif.org/), the integrated Digitized Biological Collections (iDigBio, https://www.idigbio.org/) or the pan-European Research Infrastructure Distributed System of Scientific Collections (DiSSCo, https://www.dissco.eu) will greatly facilitate global access to specimens avoiding personal travels and shipment of loans (Davis 2023). Digitization efforts also lead to a decentralisation and a more democratic use of herbarium specimens, which can empower a more diverse research community (Drew et al. 2017). This also helps to address colonial legacies of many herbaria (Park et al. 2023) and allows for much more equal access. Dedicated software for curation of specimens is developed and matching protocols are pivotal for a joint effort in digitizing herbaria and to implement the FAIR principles of data collection and accessibility (data must be Findable, Accessible, Interoperable and Reusable; Borsch et al. 2020, Manzano and Julier 2021). Currently, even metadata such as sampling locality and date of collection are available for less than 30% of the specimens within herbaria, whereas for less than 10% are digital images available (Park et al. 2023). German herbaria hold around 23 million specimens, 87% of which are not yet digitized (Borsch et al. 2020). One of the biggest issues now is that digitized specimens are hosted on different data portals. Therefore, there is an urgent need for a global federated collection management system which facilitates the effective mobilization of already digitized specimens (Weaver and Smith 2023).

This digital information is very valuable, as information and specimens from very remote places and from long ago are now available and can be accessed remotely, thus from anywhere in the world. The biggest game-changer is, however, not only the improved accessibility and connectivity, but the increasingly comprehensive information that can be gained from specimens. Manual data assembly from specimens used to be very laborious, but information from herbarium collections can now be assessed using machine learning, such as deep learning, artificial neural networks, and optical character recognition (Albani Rocchetti et al. 2021, Goëau et al. 2022, Hussein et al. 2022). However, until now, comprehensive benchmark datasets of herbarium specimens with annotations for algorithm training are still scarce, and those published mostly focus on the species identification (Tan et al. 2019, de Lutio et al. 2022). This could serve as starting point for the development of new techniques, but would require an enormous effort to find suitable, representative specimens and to perform thorough data annotation of individual plant organs (Hussein et al. 2022). These and any subsequent efforts need strong collaborations between biologists and computer scientists to bring this research field forward.

4. Applications of artificial intelligence (AI) in herbaria

One of the most promising avenues are AI methods using herbarium collections are summarized in Figure 1. In the following, we will exemplify the potential of these new applications and address some of the still existing obstacles.

With respect to the metadata, the information provided on the label can be deciphered using handwriting recognition tools, and can be automatically transferred to the metadata of the collection (Weaver et al. 2023). Automatic species identification with AI tools (Carranza-Rojas et al. 2017b, Hussein et al. 2022, Shirai et al. 2022, Karbstein et al. 2024) helps, within certain limitations, to filter out specimens which are incorrectly identified or were not identified at all (de Lutio et al. 2022, Chulif et al. 2023). This is particularly valuable, since botanists who can identify plants are becoming increasingly rarer as described above (Miller et al. 2020, Edwards et al. 2024). Artificial intelligence also increasingly contributes to keep pace with the growing knowledge in taxonomy, as it can automatically compare existing accepted taxonomies (i.e., taxonomic synopses and revisions, national checklists), as well as valid nomenclatural names and their synonyms (reviewed in Hussein et al. (2022)). With AI-tools improving further, increasingly precise automated species identification, correct assignment of previously unresolved names and misinterpreted homo- and heterotypic synonyms, and pattern recognition from the non-digital literature (i.e., diagnoses, descriptions, handwritten annotations, taxonomic keys, etc.) together will open new avenues. Since the labels typically include information about the sampling locality and the date of collection, occurrence data on present and past distribution of the plants can be extracted and compared (Jones et al. 2024). Sometimes, further information such as on the habitat where the specimen was collected is indicated, providing information on species' ecology and e.g., potential shifts of its ecological niche. Furthermore, the name of the collector is given on the label opening the possibility to analyse the collection from a cultural and historical perspective. If the collector is not specified, handwriting recognition may assist in identifying the author and thus contribute to unveil the history of the collection.

Artificial intelligence will also boost working with specimens themselves. Standardized scans provide a scale and a colour code to assist in analysing the size and colour of the specimen, or of more specific size-related traits such as leaf area or the diameter of inflorescence (Rehman et al. 2019). Although herbarium specimens exhibit notable changes in colour and appearance caused by drying, pressing and other conserving treatments, important information can be gained from them. Based on the colour of the leaves, potentially performance traits such as chlorophyll content could also be evaluated. This also makes it easier to analyse the phenological stage of the specimen, such as flowering, fruiting or senescence (Hardisty et al. 2022). Based on the colour, not only the visible spectrum and size of the flowers, but also potential pollination syndromes such as bird or bee pollination can be inferred (Streher et al. 2024). Pattern recognition allows to segment and analyse different plant organs to infer information (e.g., Younis et al. 2020b). Additionally, the herbivory status of the plants can be assessed to some extend or a potential infestation with pathogens such as gall wasps or fungal infections can be detected, and thus biotic interactions can be analysed. Moreover, X-ray techniques allow to determine leaf venation, which gives valuable insights into predicting the effects of climate change as it relates to water supply and demand (Schneider et al. 2017, Schneider et al. 2018).

In the past years, with the development of modern techniques, an increasing number of tools for automated specimen analysis have been published, yet most of them focus on specific aspects of plants such as species identification (Carranza-Rojas et al. 2017a, Carranza-Rojas et al. 2017b, Karbstein et al. 2024), morphological traits (Carranza-Rojas et al. 2017a, Zhu et al. 2017, Weaver and Smith 2023) or growth phenology (Goëau et al. 2022). There are applications available such as the LeafMachine2 – a program comprising a suite of modular machine learning and computer vision tools which automatically extract leaf traits from digitized herbarium specimens and texts from labels (Weaver and Smith 2023). However, we envision the extraction of a specimens' information as a whole and propose an automatic link of these results to other relevant databases as outlined below, which is described by the Collectomics approach (Sigwart et al. 2025).

5. Extending the Extended Specimen – a technical perspective

On account of the wealth of different specimen-related data sources, Webster (2017) introduced a new concept, describing the Extended Specimen as a combination of specimen and data types, which aims at representing the multidimensional phenotype and genotype of a physical specimen. Here, we propose to further analyse and combine these Extended Specimens.

The European Research Infrastructure Distributed System of Scientific Collections (DiSSCo) developed its foundational data model, the Digital Specimen, in close coordination with both the Extended Specimen concept and the FAIR Digital Objects (FDO) approach (Alliance for biodiversity knowledge 2017). These offer a technical implementation path for policies associated with FAIR principles for complex digital objects distributed over the internet (Islam et al. 2020, Hardisty et al. 2022). The goal is to bundle the disparate information about such an object into one cohesive package (Wittenburg et al. 2023). In line with this approach, a Digital Specimen as an FDO type contains or persistently links relevant information artefacts about the physical specimen including sequence data, images, descriptions of locations and habitats, geochemical measurements or taxonomic determinations (Hardisty et al. 2021). These resources can be available locally as files, or are linked and stored in domain-specific repositories such as the European Nucleotide Archive (Leinonen et al. 2010). In this manner, an object-centred and machine-interpretable digital twin of the physical specimen and its relationships is realized. Concepts of the Extended Specimen and “Holistic Sampling” provide the structure for modelling the specimen’s domain-specific content (Schindel and Cook 2018, Lendemer et al. 2020), while compliance with FDO specifications enables software agents or, in brief, machines to autonomously process the Digital Specimen and makes it therewith FAIR and AI-ready at the same time (Jacobsen et al. 2020). Building on this unified data model, the concept of “Machine learning as a Service” (MLaaS) was developed within the context of DiSSCo (Grieb et al. 2021). As shown in Figure 2, MLaaS facilitates the integration of Machine Learning-assisted services for data enrichment embedded by DiSSCo partners into the central Digital Specimen architecture (DSArch, Koureas et al. 2024).

As an extension of a previous study on extracting specimen data from herbarium sheets (Younis et al. 2020b), we developed an initial service for the implementation of the Collectomics concept using the Digital Specimen concept and DSArch as blueprints (Figure 2). Features extracted by the pipelines of the service, such as plant organs and morphological traits, are stored in addition to contextual data in a digital object model based on our previous studies (Grieb et al. 2021). Reusing an annotated dataset (Younis et al. 2020a), the study was further extended to perform plant organ segmentation to uniquely identify and extract additional information of each object in the digital herbarium specimen. In accordance with similar studies (Sapkota et al. 2024), we switched the model from Mask R-CNN to

YOLO11, because the latter that performed better for our use case (He et al. 2017, Jocher et al. 2024). YOLO11 is then applied and trained on the annotated dataset from the previous study. The segmented organs are further used for morphological analysis such as surface area calculation from digital images. To perform surface area calculation, the YOLO11 detection model is trained to detect scales from the herbarium images. Following scale detection, the tesseract OCR (Smith 2007) is applied on detected scales to calculate the pixel distance between the digits. The relative pixel distance is transformed into an absolute measurement in square centimetre (cm^2) using the detected scale. Then the transformed absolute measurement is applied on segmented organs and surface area is calculated for segmented regions of interest as an additional feature (Figure 3). To promote the reusability of the digitization output, we store the annotations together with the processed images in a container format compatible with the FDO approach called RO-Crate (Soiland-Reyes et al. 2024). This enables lightweight packaging of research outputs along with structured metadata. The RO-Crate comprising the complete dataset on which Figure 3 is based and is provided as supplementary material. It includes the processed images and annotations semantically mapped to the Flora Phenotype Ontology, a framework which allows cross-domain re-usability and enables the consecutive analysis of plant trait data (Hoehndorf et al. 2016).

6. Conclusions – towards a digital Collectomics framework

The renewed interest in herbaria is not without problems, as some methods require destructive sampling. As curators, we strongly advocate to handle specimens as sustainably as possible (Davis et al. 2024), and we also trust that technology will increasingly develop more and more non-destructive approaches such as NIRs technology or X-ray (Schneider et al. 2018, Kühn et al. 2024). We do, however, firmly believe in benefits of active research for collection maintenance and development, and thus new opportunities should be welcomed. Therefore, the present contribution is also an appeal to the scientific community to collect more specimens in current ecological and biological projects and to acknowledge the valuable contribution of natural history collections such as nicely summarized in Miller et al. (2020). Research from German herbaria could demonstrate, for example, that there is a current reduction of specimens being collected, especially when it comes to other species than rare ones or neophytes, which will obviously imply a bias in future studies (Renner and Rockinger 2016). Of course, broadening the perspective requires additional resources, such as those for storage and curation of the collections as well as digitization, but we are firmly convinced this is worth investing.

When handled wisely, preserved specimens in herbaria offer a unique opportunity to do both, study the past and predict the future of ecosystems and species distributions. We propose to link modern data retrieval techniques, AI-data science approaches and biodiversity research to analyse patterns in plant biodiversity and the changes of plant life throughout the Anthropocene using herbarium specimens. We strongly rely on the continued efforts to digitize and collect such specimens worldwide, but we are confident that more and more become available online, as this was already identified as a huge need by the scientific community (Borsch et al. 2020).

In the new era, we can take digital Extended Specimens even one step further and envision the automatic link between the information extracted from the specimens such as plant name, sampling locality, etc. with further information such as climate data, the species' distributional range, genetic data or threat levels. This will significantly expand the concept of Museomics, which currently mainly involves extracting DNA data from specimens (Raxworthy and Smith 2021). We thus echo pleas to think in a larger framework, recently being termed Collectomics by Sigwart et al. (2025). In that view, Collectomics is not only more precise than Museomics, as it specifically addresses collections and does not invoke other core museum activities such as for example exhibitions. It is also not restricted to biological specimens only. Collectomics also embraces the ground-breaking idea of the Extended Specimen and puts this in context of a full-scale values chain ranging from specimens to derived data ultimately inform fields as diverse as large-scale modelling, or studies on cultural history (Sigwart et al. 2025, and examples therein such as (Tobias et al. 2022)).

Herbaria already play a pioneering role not only because imaging techniques are well and taxonomical backbones are relatively stable. Concepts such as trait-based analysis (Díaz et al. 2016), phenological studies (Parmesan and Yohe 2003) or non-invasive measurement of nutrient contents (Kühn et al. 2024) are particularly advanced in plant biodiversity research, and the pivotal role of plants for ecosystem primary production also makes plants crucial for Earth System Modelling (Thavhana et al. 2024). In this way, we expect that herbarium research will make big step forward, enabling us to process large amounts of data in a fast, yet meaningful manner, thereby opening the treasure trove of scientific plant collections to the scientific community. Lessons learnt will ultimately benefit other natural historical collections, and potentially even collections of cultural or technical artefacts.

6. Acknowledgements

We acknowledge support from the Leibniz-Gesellschaft for developing our Collectomics approach. We would also like to thank the many universities and scientific institutions that see and saw the relevance of collections and provided budget for maintaining and developing them in the future. Comments of two anonymous reviewers greatly helped us to sharpen our ideas.

7. Data availability statement

The dataset consisting of 100 annotated, digitized herbarium specimens, which is referenced in section 4, is available as an RO-Crate under a Creative Commons license and openly published under https://doi.org/10.12761/w2c1-x551.

8. Authors contributions:

SFB, SG, JG, MK, JM, CMR, RR, CW, KW, KV, CR- conceptualization

JG, JM, RR, CW, KV- Data curation

JG, RR, CW- Formal analysis

SFB, JG, RR, CW- Visualisation

SFB, SG, JG, MK, JM, CMR, RR, CW, KW, KV, CR- Writing original draft

SFB, SG, JG, MK, JM, CMR, RR, CW, KW, KV, CR- Writing review & editing

9. References

Abeli, T., S. Dalrymple, S. Godefroid, A. Mondoni, J. V. Müller, G. Rossi, and S. Orsenigo. 2020. Ex situ collections and their potential for the restoration of extinct plants. Conservation biology **34**:303-313.

Ahlstrand, N. I., H. M. Elvery, and R. B. Primack. 2023. Grass flowering times determined using herbarium specimens for modeling grass pollen under a warming climate. Science of The Total Environment **885**:163824.

Albani Rocchetti, G., C. G. Armstrong, T. Abeli, S. Orsenigo, C. Jasper, S. Joly, A. Bruneau, M. Zytaruk, and J. C. Vamosi. 2021. Reversing extinction trends: new uses of (old) herbarium specimens to accelerate conservation action on threatened species. New Phytologist **230**:433-450.

Alliance for biodiversity knowledge. 2017. Converging Digital and Extended Specimens: Towards a global specification for data integration.

Baldini, R. M., G. Cristofolini, and C. Aedo. 2022. The extant herbaria from the Sixteenth Century: a synopsis. Webbia **77**:23-33.

Barnes, M., J. Sulé-Suso, J. Millett, and P. Roach. 2023. Fourier transform infrared spectroscopy as a non-destructive method for analysing herbarium specimens. Biol Lett **19**:20220546.

Bolmgren, K., and K. Lönnberg. 2005. Herbarium data reveal an association between fleshy fruit type and earlier flowering time. International Journal of Plant Sciences **166**:663-670.

Bonal, D., S. Ponton, D. Le Thiec, B. Richard, N. Ningre, B. Hérault, J. Ogée, S. Gonzalez, M. Pignal, and D. Sabatier. 2011. Leaf functional response to increasing atmospheric CO2 concentrations over the last century in two northern Amazonian tree species: a historical δ13C and δ18O approach using herbarium samples. Plant, Cell & Environment **34**:1332-1344.

Borsch, T., A.-D. Stevens, E. Häffner, A. Güntsch, W. G. Berendsohn, M. Appelhans, C. Barilaro, B. Beszteri, F. Blattner, and O. Bossdorf. 2020. A complete digitization of German herbaria is possible, sensible and should be started now. Research Ideas and Outcomes **6**:e50675.

Burbano, H. A., and R. M. Gutaker. 2023. Ancient DNA genomics and the renaissance of herbaria. Science **382**:59-63.

Carranza-Rojas, J., H. Goeau, P. Bonnet, E. Mata-Montero, and A. Joly. 2017a. Going deeper in the automated identification of Herbarium specimens. BMC evolutionary biology **17**:1-14.

Carranza-Rojas, J., A. Joly, P. Bonnet, H. Goëau, and E. Mata-Montero. 2017b. Automated herbarium specimen identification using deep learning. Page e20302 *in* TDWG: Biodiversity Information Standards.

Chulif, S., S. H. Lee, Y. L. Chang, and K. C. Chai. 2023. A machine learning approach for cross-domain plant identification using herbarium specimens. Neural Computing and Applications **35**:5963-5985.

Crawford, P. H., and B. W. Hoagland. 2009. Can herbarium records be used to map alien species invasion and native species expansion over the past 100 years? Journal of biogeography **36**:651-661.

Davis, C. C. 2023. The herbarium of the future. Trends in ecology & evolution **38**:412-423.

Davis, C. C., E. Sessa, A. Paton, A. Antonelli, and J. K. Teisher. 2024. Guidelines for the effective and ethical sampling of herbaria. Nature Ecology & Evolution:1-8.

de Lutio, R., J. Y. Park, K. A. Watson, S. D'Aronco, J. D. Wegner, J. J. Wieringa, M. Tulig, R. L. Pyle, T. J. Gallaher, and G. Brown. 2022. The herbarium 2021 half–earth challenge dataset and machine learning competition. Frontiers in plant science **12**:787127.

Díaz, S., J. Kattge, J. H. Cornelissen, I. J. Wright, S. Lavorel, S. Dray, B. Reu, M. Kleyer, C. Wirth, and I. C. Prentice. 2016. The global spectrum of plant form and function. Nature **529**:167-171.

Drew, J. A., C. S. Moreau, and M. L. Stiassny. 2017. Digitization of museum collections holds the potential to enhance researcher diversity. Nature Ecology & Evolution **1**:1789-1790.

Edwards, E. J., B. D. Mishler, and C. D. Davis. 2024. University herbaria are uniquely important. Trends in plant science.

Everill, P. H., R. B. Primack, E. R. Ellwood, and E. K. Melaas. 2014. Determining past leaf-out times of New England's deciduous forests from herbarium specimens. American Journal of Botany **101**:1293-1300.

Feeley, K. J. 2012. Distributional migrations, expansions, and contractions of tropical plant species as revealed in dated herbarium records. Global Change Biology **18**:1335-1341.

Flannery, M. C. 2023. In the herbarium: The hidden world of collecting and preserving plants. Yale University Press.

Goëau, H., T. Lorieul, P. Heuret, A. Joly, and P. Bonnet. 2022. Can Artificial Intelligence Help in the Study of Vegetative Growth Patterns from Herbarium Collections? An Evaluation of the Tropical Flora of the French Guiana Forest. Plants **11**:530.

Graham, C. H., S. Ferrier, F. Huettman, C. Moritz, and A. T. Peterson. 2004. New developments in museum-based informatics and applications in biodiversity analysis. Trends in ecology & evolution **19**:497-503.

Grieb, J., C. Weiland, A. Hardisty, W. Addink, S. Islam, S. Younis, and M. Schmidt. 2021. Machine Learning as a Service for DiSSCo's Digital Specimen Architecture. Biodiversity Information Science and Standards **5**:e75634.

Gugerli, F., L. Parducci, and R. J. Petit. 2005. Ancient plant DNA: review and prospects. New Phytologist **166**:409-418.

Gutaker, R. M., C. L. Weiß, D. Ellis, N. L. Anglin, S. Knapp, J. Luis Fernández-Alonso, S. Prat, and H. A. Burbano. 2019. The origins and adaptation of European potatoes reconstructed from historical genomes. Nature Ecology & Evolution **3**:1093-1101.

Hardisty, A., W. Addink, F. Glöckler, A. Güntsch, S. Islam, and C. Weiland. 2021. A choice of persistent identifier schemes for the Distributed System of Scientific Collections (DiSSCo). Research Ideas and Outcomes **7**:e67379.

Hardisty, A. R., E. R. Ellwood, G. Nelson, B. Zimkus, J. Buschbom, W. Addink, R. K. Rabeler, J. Bates, A. Bentley, and J. A. Fortes. 2022. Digital extended specimens: Enabling an extensible network of biodiversity data records as integrated digital objects on the internet. Bioscience **72**:978-987.

He, K., G. Gkioxari, P. Dollár, and R. Girshick. 2017. Mask r-cnn. Pages 2961-2969 *in* Proceedings of the IEEE international conference on computer vision.

Heberling, J. M. 2022. Herbaria as big data sources of plant traits. International Journal of Plant Sciences **183**:87-118.

Heberling, J. M., L. A. Prather, and S. J. Tonsor. 2019. The changing uses of herbarium data in an era of global change: an overview using automated content analysis. Bioscience **69**:812-822.

Hedrick, B. P., J. M. Heberling, E. K. Meineke, K. G. Turner, C. J. Grassa, D. S. Park, J. Kennedy, J. A. Clarke, J. A. Cook, and D. C. Blackburn. 2020. Digitization and the future of natural history collections. Bioscience **70**:243-251.

Hoegh-Guldberg, O., D. Jacob, M. Taylor, M. Bindi, S. Brown, I. Camilloni, A. Diedhiou, R. Djalante, K. L. Ebi, F. Engelbrecht, J. Guiot, Y. Hijioka, S. Mehrotra, A. Payne, S. I. Seneviratne, A. Thomas, R. Warren, and G. Zho. 2018. Impacts of 1.5ºC Global Warming on Natural and Human Systems. *.in* V. Masson-Delmotte, P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P. R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J. B. R. Matthews, Y. Chen, X. Zhou, M. I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield, editors. Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.

Hoehndorf, R., M. Alshahrani, G. V. Gkoutos, G. Gosline, Q. Groom, T. Hamann, J. Kattge, S. M. De Oliveira, M. Schmidt, S. Sierra, E. Smets, R. A. Vos, and C. Weiland. 2016. The flora phenotype ontology (FLOPO): tool for integrating morphological traits and phenotypes of vascular plants. Journal of biomedical semantics **7**:1-11.

Hussein, B. R., O. A. Malik, W.-H. Ong, and J. W. F. Slik. 2022. Applications of computer vision and machine learning techniques for digitized herbarium specimens: A systematic literature review. Ecological Informatics **69**:101641.

IPCC. 2007. Climate change impacts, adaptation and vulnerability. Contribution of working group II to the fourth assessment report of the Intergovernmental Panel on Climate Change. Summary for policymakers-Brussels-April.

Islam, S., A. Hardisty, W. Addink, C. Weiland, and F. Glöckler. 2020. Incorporating RDA outputs in the design of a European research infrastructure for natural science collections. Data Science Journal **19**:50-50.

Iwanycki Ahlstrand, N., R. B. Primack, M. W. Austin, Z. A. Panchen, C. Römermann, and A. J. Miller-Rushing. 2025. The promise of digital herbarium specimens in large-scale phenology research. New Phytologist **accepted**.

Jacobsen, A., R. de Miranda Azevedo, N. Juty, D. Batista, S. Coles, R. Cornet, M. Courtot, M. Crosas, M. Dumontier, and C. T. Evelo. 2020. FAIR principles: interpretations and implementation considerations. Pages 10-29. MIT Press One Rogers Street, Cambridge, MA 02142-1209, USA journals-info ….

Jocher, G., J. Qiu, and A. Chaurasia. 2024. Ultralytics YOLO

Johnson, C. N., A. Balmford, B. W. Brook, J. C. Buettel, M. Galetti, L. Guangchun, and J. M. Wilmshurst. 2017. Biodiversity losses and conservation responses in the Anthropocene. Science **356**:270-275.

Johnson, K. G., S. J. Brooks, P. B. Fenberg, A. G. Glover, K. E. James, A. M. Lister, E. Michel, M. Spencer, J. A. Todd, and E. Valsami-Jones. 2011. Climate change and biosphere response: unlocking the collections vault. Bioscience **61**:147-153.

Jones, C. B., K. Stock, and S. E. Perkins. 2024. AI-based discovery of habitats from museum collections. Trends in ecology & evolution **39**:323-327.

Karbstein, K., L. Kösters, L. Hodač, M. Hofmann, E. Hörandl, S. Tomasello, N. D. Wagner, B. C. Emerson, D. C. Albach, and S. Scheu. 2024. Species delimitation 4.0: integrative taxonomy meets artificial intelligence. Trends in ecology & evolution **39**:771-784.

Kothari, S., R. Beauchamp-Rioux, E. Laliberté, and J. Cavender-Bares. 2023. Reflectance spectroscopy allows rapid, accurate and non-destructive estimates of functional traits from pressed leaves. Methods in Ecology and Evolution **14**:385-401.

Koureas, D., L. Livermore, W. Addink, E. Alonso, J. Alonso, A. Casino, F. Dusoulier, V. Ferreira, J. Grieb, and Q. Groom. 2024. DiSSCo Transition Abridged Grant Proposal. Research Ideas and Outcomes **10**.

Kozlov, M. V., I. V. Sokolova, V. Zverev, A. A. Egorov, M. Y. Goncharov, and E. L. Zvereva. 2020. Biases in estimation of insect herbivory from herbarium specimens. Scientific reports **10**:12298.

Kühn, P., T. Proß, C. Römermann, K. Wesche, and H. Bruelheide. 2024. Using near-infrared spectroscopy to predict nitrogen and phosphorus concentrations of herbarium specimens under different storage conditions. Plant methods **20**:19.

Kühn, P., R. Umazekabiri, C. Römermann, H. Bruelheide, and K. Wesche. 2025. Nitrogen content of herbarium specimens from arable fields and mesic meadows reflect the intensifying agricultural management during the 20th century. Journal of Ecology.

Lang, P. L., F. M. Willems, J. Scheepens, H. A. Burbano, and O. Bossdorf. 2019. Using herbaria to study global environmental change. New Phytologist **221**:110-122.

Lankau, R. A., V. Nuzzo, G. Spyreas, and A. S. Davis. 2009. Evolutionary limits ameliorate the negative impact of an invasive plant. Proceedings of the National Academy of Sciences **106**:15362-15367.

Lavoie, C. 2013. Biological collections in an ever changing world: Herbaria as tools for biogeographical and environmental studies. Perspectives in plant ecology, evolution and systematics **15**:68-76.

Lavoie, C., Y. Jodoin, and A. G. De Merlis. 2007. How did common ragweed (Ambrosia artemisiifolia L.) spread in Quebec? A historical analysis using herbarium records. Journal of biogeography **34**:1751-1761.

Law, W., and J. Salick. 2005. Human-induced dwarfing of Himalayan snow lotus, Saussurea laniceps (Asteraceae). Proceedings of the National Academy of Sciences **102**:10218-10220.

Lee, B. R., E. F. Alecrim, T. K. Miller, J. R. Forrest, J. M. Heberling, R. B. Primack, and R. D. Sargent. 2024. Phenological mismatch between trees and wildflowers: Reconciling divergent findings in two recent analyses. Journal of Ecology **112**:1184-1199.

Leger, E. A. 2013. Annual plants change in size over a century of observations. Global Change Biology **19**:2229-2239.

Leinonen, R., R. Akhtar, E. Birney, L. Bower, A. Cerdeno-Tárraga, Y. Cheng, I. Cleland, N. Faruque, N. Goodgame, and R. Gibson. 2010. The European nucleotide archive. Nucleic acids research **39**:D28-D31.

Lendemer, J., B. Thiers, A. K. Monfils, J. Zaspel, E. R. Ellwood, A. Bentley, K. LeVan, J. Bates, D. Jennings, and D. Contreras. 2020. The extended specimen network: A strategy to enhance US biodiversity collections, promote research and education. Bioscience **70**:23-30.

Lewis, S. L., and M. A. Maslin. 2015. Defining the anthropocene. Nature **519**:171-180.

Little, D. P., M. Tulig, K. C. Tan, Y. Liu, S. Belongie, C. Kaeser-Chen, F. A. Michelangeli, K. Panesar, R. V. Guha, and B. A. Ambrose. 2020. An algorithm competition for automatic species identification from herbarium specimens. Applications in plant sciences **8**:e11365.

Manzano, S., and A. C. Julier. 2021. How FAIR are plant sciences in the twenty-first century? The pressing need for reproducibility in plant ecology and evolution. Proceedings of the Royal Society B **288**:20202597.

Meineke, E. K., A. T. Classen, N. J. Sanders, and T. Jonathan Davies. 2019a. Herbarium specimens reveal increasing herbivory over the past century. Journal of Ecology **107**:105-117.

Meineke, E. K., T. J. Davies, B. H. Daru, and C. C. Davis. 2019b. Biological collections for understanding biodiversity in the Anthropocene. Page 20170386. The Royal Society.

Meineke, E. K., C. C. Davis, and T. J. Davies. 2018. The unrealized potential of herbaria for global change biology. Ecological Monographs **88**:505-525.

Mendes Resende, J. V., N. M. D. de Sá, M. T. L. de Oliveira, R. C. Lopes, R. Garrett, and R. Moreira Borges. 2020. Chemical profiling of herbarium samples of solanum (Solanaceae) using mass spectrometry. Phytochemistry Letters **36**:99-105.

Miller-Rushing, A. J., R. B. Primack, D. Primack, and S. Mukunda. 2006. Photographs and herbarium specimens as tools to document phenological changes in response to global warming. American Journal of Botany **93**:1667-1674.

Miller-Rushing, A. J., R. B. Primack, P. H. Templer, S. Rathbone, and S. Mukunda. 2009. Long-term relationships among atmospheric CO2, stomata, and intrinsic water use efficiency in individual trees. American Journal of Botany **96**:1779-1786.

Miller, S. E., L. N. Barrow, S. M. Ehlman, J. A. Goodheart, S. E. Greiman, H. L. Lutz, T. M. Misiewicz, S. M. Smith, M. Tan, and C. J. Thawley. 2020. Building natural history collections for the twenty-first century and beyond. Bioscience **70**:674-687.

Newbold, T. 2010. Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. Progress in physical geography **34**:3-22.

Park, D. S., X. Feng, S. Akiyama, M. Ardiyani, N. Avendaño, Z. Barina, B. Bärtschi, M. Belgrano, J. Betancur, and R. Bijmoer. 2023. The colonial legacy of herbaria. Nature Human Behaviour:1-10.

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature **421**:37-42.

Primack, D., C. Imbres, R. B. Primack, A. J. Miller-Rushing, and P. Del Tredici. 2004. Herbarium specimens demonstrate earlier flowering times in response to warming in Boston. American Journal of Botany **91**:1260-1264.

Ramirez-Parada, T. H., I. W. Park, and S. J. Mazer. 2022. Herbarium specimens provide reliable estimates of phenological responses to climate at unparalleled taxonomic and spatiotemporal scales. Ecography **2022**:e06173.

Raxworthy, C. J., and B. T. Smith. 2021. Mining museums for historical DNA: advances and challenges in museomics. Trends in ecology & evolution **36**:1049-1060.

Rehman, A., S. Naz, and M. I. Razzak. 2019. Writer identification using machine learning approaches: a comprehensive review. Multimedia Tools and Applications **78**:10889-10931.

Renner, S. S., and A. Rockinger. 2016. Is plant collecting in Germany coming to an end? Willdenowia **46**:93-97.

Sapkota, R., D. Ahmed, and M. Karkee. 2024. Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments. Artificial Intelligence in Agriculture **13**:84-99.

Schindel, D. E., and J. A. Cook. 2018. The next generation of natural history collections. PLoS biology **16**:e2006125.

Schneider, J. V., J. Habersetzer, R. Rabenstein, J. Wesenberg, K. Wesche, and G. Zizka. 2017. Water supply and demand remain coordinated during breakdown of the global scaling relationship between leaf size and major vein density. New Phytologist **214**:473-486.

Schneider, J. V., R. Rabenstein, J. Wesenberg, K. Wesche, G. Zizka, and J. Habersetzer. 2018. Improved non-destructive 2D and 3D X-ray imaging of leaf venation. Plant methods **14**:1-15.

Shirai, M., A. Takano, T. Kurosawa, M. Inoue, S. Tagane, T. Tanimoto, T. Koganeyama, H. Sato, T. Terasawa, T. Horie, I. Mandai, and T. Akihiro. 2022. Development of a system for the automated identification of herbarium specimens with high accuracy. Scientific reports **12**:8066.

Sigwart, J. D., M. Schleuning, A. Brandt, M. Pfenninger, H. Saeedi, T. Borsch, E. Häffner, R. Lücking, A. Güntsch, and H. Trischler. 2025. Collectomics–towards a new framework to integrate museum collections to address global challenges. Natural History Collections and Museomics **2**:1-20.

Smith, R. 2007. An overview of the Tesseract OCR engine. Pages 629-633 *in* Ninth international conference on document analysis and recognition (ICDAR 2007). IEEE.

Soiland-Reyes, S., L. Castro, R. Ravinder, C. Weiland, J. Grieb, A. Rogers, C. Blanchi, and H. Van de Sompel. 2024. BioHackEU23 report: Enabling FAIR Digital Objects with RO-Crate, Signposting and Bioschemas.

Soltis, P. S. 2017. Digitization of herbaria enables novel research. American Journal of Botany **104**:1281-1284.

Streher, N. S., T. Budinsky, K. Halabi, I. Mayrose, and T.-L. Ashman. 2024. The effect of polyploidy and mating system on floral size and the pollination niche in Brassicaceae. International Journal of Plant Sciences **185**:89-99.

Tan, K. C., Y. Liu, B. Ambrose, M. Tulig, and S. Belongie. 2019. The herbarium challenge 2019 dataset. arXiv preprint arXiv:1906.05372.

Thavhana, M. P., T. Hickler, M. Urban, K. Heckel, and M. Forrest. 2024. Dynamics and Drivers of Net Primary Production (NPP) in Southern Africa Based on Estimates from Earth Observation and Process-Based Dynamic Vegetation Modelling. Pages 759-786 Sustainability of Southern African Ecosystems under Global Change: Science for Management and Policy Interventions. Springer.

Thiers, B. 2024. the Index Herbariorum

Tobias, J. A., C. Sheard, A. L. Pigot, A. J. Devenish, J. Yang, F. Sayol, M. H. Neate-Clegg, N. Alioravainen, T. L. Weeks, and R. A. Barber. 2022. AVONET: morphological, ecological and geographical data for all birds. Ecology letters **25**:581-597.

Vörös, W. A., I. Hensen, S. Träger, C. Rosche, J. Müller, C. Römermann, A. Baasch, and M. Lehnert. 2025. From herbarium specimens to contemporary surveys: Tracing the extinction dynamics and performance of endangered plant species in central Germany. Flora **323**:152664.

Wandeler, P., P. E. Hoeck, and L. F. Keller. 2007. Back to the future: museum specimens in population genetics. Trends in ecology & evolution **22**:634-642.

Ward, N. 2024. The Potential of Herbarium Specimens in Capturing Historical Herbivory: A test of the consequences of global insect decline. Stanford University.

Weaver, W. N., B. R. Ruhfel, K. J. Lough, and S. A. Smith. 2023. Herbarium specimen label transcription reimagined with large language models: Capabilities, productivity, and risks. American Journal of Botany **110**:e16256.

Weaver, W. N., and S. A. Smith. 2023. From leaves to labels: Building modular machine learning networks for rapid herbarium specimen analysis with LeafMachine2. Applications in plant sciences:e11548.

Webster, M. S. 2017. The extended specimen: emerging frontiers in collections-based ornithological research. CRC Press.

Willis, C. G., E. R. Ellwood, R. B. Primack, C. C. Davis, K. D. Pearson, A. S. Gallinat, J. M. Yost, G. Nelson, S. J. Mazer, and N. L. Rossington. 2017. Old plants, new tricks: Phenological research using herbarium specimens. Trends in ecology & evolution **32**:531-546.

Wittenburg, P., U. Schwardmann, C. Blanchi, and C. Weiland. 2023. Fdos to enable cross-silo work.*in* Proceedings of the Conference on Research Data Infrastructure.

Woodward, F. I. 1987. Stomatal numbers are sensitive to increases in CO2 from pre-industrial levels. Nature **327**:617-618.

Yoshida, K., V. J. Schuenemann, L. M. Cano, M. Pais, B. Mishra, R. Sharma, C. Lanz, F. N. Martin, S. Kamoun, and J. Krause. 2013. The rise and fall of the Phytophthora infestans lineage that triggered the Irish potato famine. elife **2**:e00731.

Younis, S., M. Schmidt, and S. Dressler. 2020a. Plant organ detections and annotations on digitized herbarium scans *in* PANGAEA, editor.

Younis, S., M. Schmidt, C. Weiland, S. Dressler, B. Seeger, and T. Hickler. 2020b. Detection and annotation of plant organs from digitised herbarium scans using deep learning. Biodiversity Data Journal **8**.

Zalasiewicz, J., C. N. Waters, M. Williams, A. D. Barnosky, A. Cearreta, P. Crutzen, E. Ellis, M. A. Ellis, I. J. Fairchild, and J. Grinevald. 2015. When did the Anthropocene begin? A mid-twentieth century boundary level is stratigraphically optimal. Quaternary international **383**:196-203.

Zhu, Y., T. Durand, E. Chenin, M. Pignal, P. Gallinari, and R. Vignes-Lebbe. 2017. Using a deep convolutional neural network for extracting morphological traits from herbarium images. Biodiversity Information Science and Standards **1**:e20400.

Ziska, L. H., J. S. Pettis, J. Edwards, J. E. Hancock, M. B. Tomecek, A. Clark, J. S. Dukes, I. Loladze, and H. W. Polley. 2016. Rising atmospheric CO2 is reducing the protein concentration of a floral pollen source essential for North American bees. Proceedings of the Royal Society B: Biological Sciences **283**:20160414.

Zizka, A., T. Andermann, and D. Silvestro. 2022. IUCNN–Deep learning approaches to approximate species' extinction risk. Diversity and Distributions **28**:227-241.

Zizka, A., O. Rydén, D. Edler, J. Klein, A. Perrigo, D. Silvestro, S. C. Jagers, S. I. Lindberg, and A. Antonelli. 2021. Bio-Dem, a tool to explore the relationship between biodiversity data availability and socio-political conditions in time and space. Journal of biogeography **48**:2715-2726.

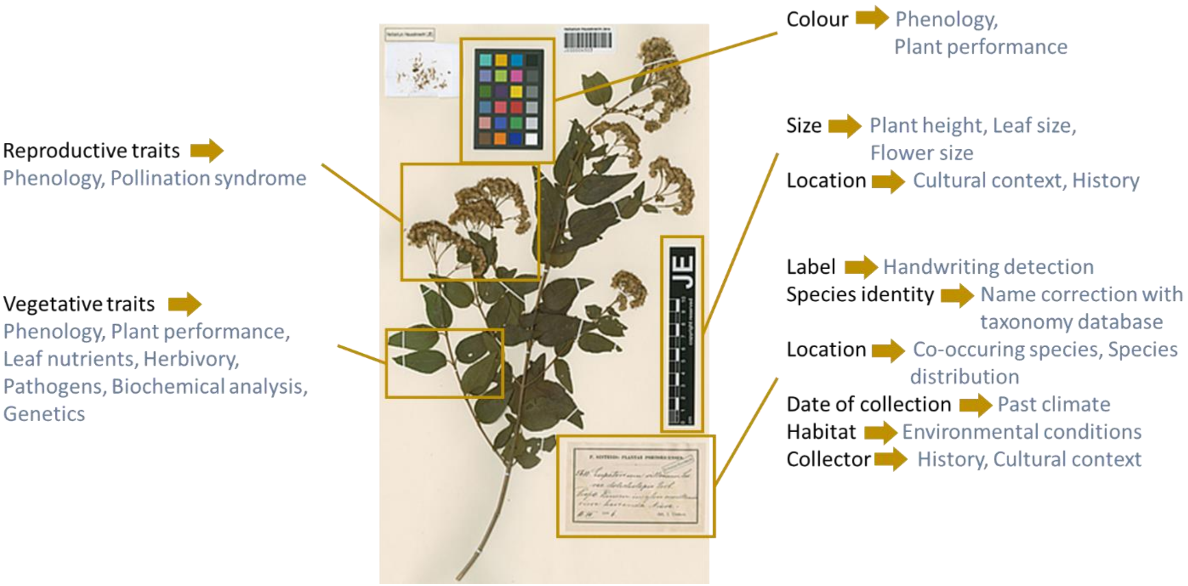

Figure 1: Schematic display of an herbarium specimen and its use as an Extended Specimen.

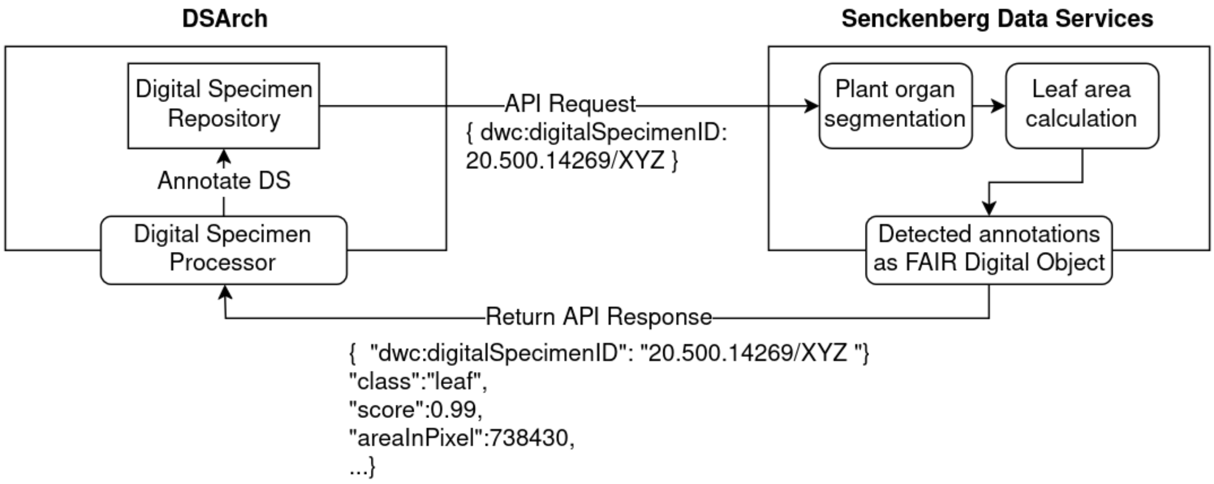

Figure 2: Blueprint for the integration of an annotation service into DiSSCo's core Digital Specimen architecture (DSArch). Identifiability, traceability and persistence of the processed research element is realized by a specific identifier, the digital SpecimenID, which is proposed as an extension for the Darwin Core vocabulary.

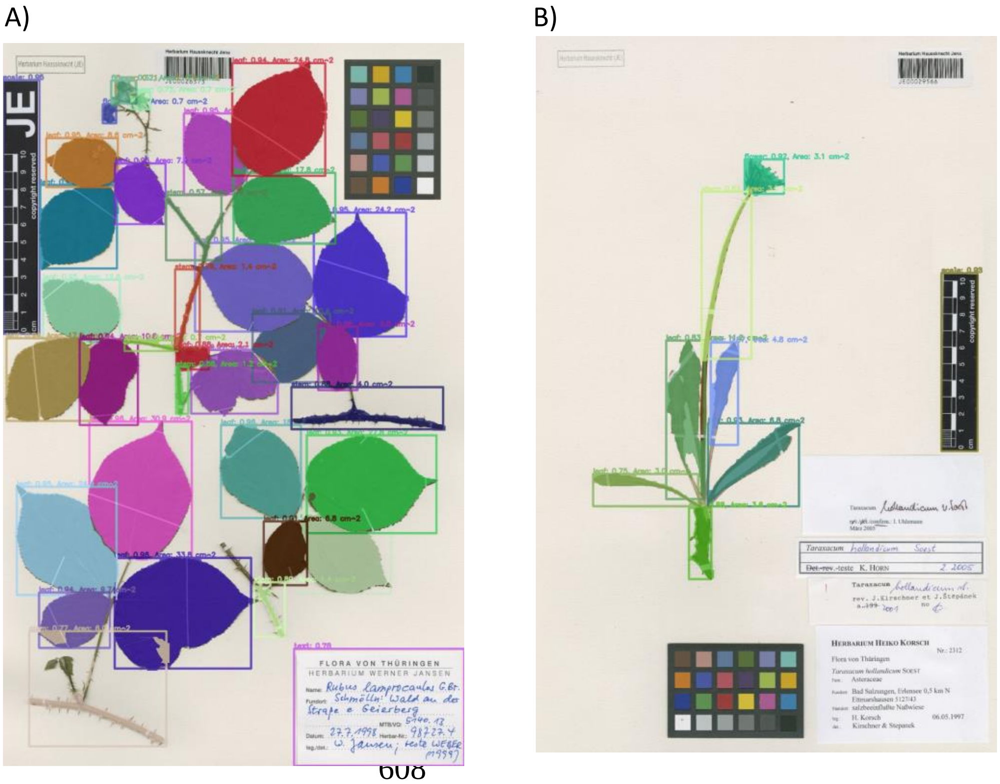

Figure 3: Detection of regions of interest (ROI) corresponding to categorized plant organs (leaf, flower, fruit, seed, stem, root) and subsequent instance segmentation of identified organs of two plant genera (*Rubus*, *Taraxacum*). Based on the segmentation in addition to scale bar detection, the visible surface area in cm² of the plant organs is computed and included in the resulting digital object. Shown are predicted annotations for two example specimens (A: *Rubus pruinosus* Arrh.; B: *Taraxacum subalpinum* Hudziok) from a compiled example dataset of 100 specimens from the Herbarium Senckenbergianum – Herbarium Haussknecht (JE).