Digital Art Technical Sources for the Netherlands: Integration and Improvement of Sources on Glass for a Sustainable Future – Art DATIS

Carlotta Capurro[1] ✉ 📵
Utrecht University, The Netherlands

Vera Provatorova ✉ 📵
University of Amsterdam, The Netherlands

Marieke Hendriksen ✉ 📵
Huygens Institute, Royal Netherlands Academy of Arts and Sciences (KNAW), Amsterdam, The Netherlands

Evangelos Kanoulas ✉ 📵
University of Amsterdam, The Netherlands

Sven Dupré ✉ 📵
Utrecht University, The Netherlands

—— **Abstract** ——

Art DATIS (Digital Art Technical sources for the Netherlands: Integration and improvement of sources on glass for a Sustainable future) is a five-year research project (2018-2023) within the Netherlands Organisation for Scientific Research's (NWO) Big Data / Digital Humanities program. The project is a collaboration between the Universities of Utrecht and Amsterdam, RKD Netherlands Institute for Art History, the Vrij Glas Foundation, and Picturae. The project investigates how to approach the automatic transcription and documentation of heterogeneous archival resources. The central object of the project is the archive of the Dutch glass artist Sybren Valkema (1916–96). Documents were digitised, and their content was made searchable through the processes of OCR and HTR. Through digitisation and the analysis of archival documents, the project aims to understand how traditional knowledge and practices of glassmaking were innovated during the twentieth century.

1 Introduction

Digital Art Technical sources for the Netherlands: Integration and improvement of sources on glass for a Sustainable future (Art DATIS) is a five-year research project launched in 2018 and funded by the Netherlands Organization for Scientific Research (NWO) within the Big Data / Digital Humanities program [8]. With this grant, NWO aimed at funding proposals that contribute to the development of research in the field of creative industries and Big Data while boosting the research in the field of Digital Humanities [18]. The Art DATIS project sees the collaboration of five partners: the Universities of Utrecht (UU) and Amsterdam

---

[1] Corresponding Author

(UvA), the Netherlands Institute for Art History (RKD), the Vrij Glas Foundation, and the Dutch heritage digitisation company Picturae. It aims to understand how traditional knowledge and practices of glassmaking were innovated during the twentieth century by analysing the work of one of the key actors in this process, the Dutch glass artist Sybren Valkema (1916–96).

Sybren Valkema was one of the representatives of the Studio Glass movement that, in the Netherlands, took the name of Vrij Glas. Spreading internationally from the US during the 1960s, Studio Glass promoted the transformation of traditional glassmaking practice, encouraging the use of glass as an artistic medium. The movement also had a significant impact on the circulation of knowledge about glass and glassmaking techniques. Believing in knowledge sharing and collaborative improvement, many artists adhering to the Studio Glass movement took up educative roles in art academies and universities, establishing courses to teach glass-blowing as a form of art [10] [9]. In line with the Studio Glass philosophy, Valkema combined glass-blowing with an intense career as a teacher, working at the Dutch Institute for Art and Crafts (Instituut voor Kunstnijverheidsonderwijs), later renamed Gerrit Rietveld Academy.

Central to Art DATIS research is Valkema's personal archive. Sources containing historical recipes and techniques for glass production, as well as images and object documentation, are currently digitised by museums and research institutions around the world. The Valkema's archive contains over 103.000 pages documenting the private and professional life of the artist, including teaching materials, letters, designs, sketches, descriptions of technical processes and many glass recipes, which have been tested and revised with handwritten notes [Figure 1]. Due to the importance of these resources for the study of glass art in the twentieth century, Metamorfoze, the Netherlands' national programme for the preservation of paper heritage, financed the conservation and digital preservation of the archive. Between 2013 and 2015, the archive was digitised thanks to the joint efforts of RKD and the Rakow Research Library of the Corning Museum of Glass [23].

Through digitising and enriching the documents in Valkema's archive, the Art DATIS project worked towards creating a comprehensive digital resource for the study of contemporary glassmaking, allowing scholars, artists and practitioners to investigate the use of glass as an artistic material through time. One of the main issues it had to face was implementing a robust process for the datafication of archival resources. The concept of datafication is not limited to its literal meaning of transforming something into data but includes the social, cultural and political implications that such a process has on our lives [17]. When dealing with archival resources, their transformation into data requires attentive consideration of the possible biases and the implications they might have on our understanding of the past [13].

As a project developed in the Digital Humanities, Art DATIS has benefitted from two separate yet interlocking disciplinary approaches: information retrieval and digital art history. First, the project investigated the best approach to make the digital documents preserved in Valkema's archive accessible and usable as digital resources. Images were converted into machine-encoded text with the support of Text Recognition software (handwritten pages were processed using Trankribus, while typed documents were processed with the OCR engine Tesseract), and then the content was extended, enriched, and linked with open data. Second, this material was used to investigate Valkema's international network and shed light on the role of these relationships in his artistic production and the spread of Studio Glass in Europe.

This contribution aims to illustrate the developments within the ArtDATIS project and is structured as follows. The second section offers a historical introduction to the Studio Glass movement and the work of Sybren Valkema. The documents digitised and analysed

Figure 1 Example of the documents preserved in Sybren Valkema's archive. They include drawings, school notes, working notebooks, letters and booklets.

within the ArtDATIS framework are contextualised historically to illustrate better their relevance for the study of the international development of the Studio Glass movement. The third section describes the development of the project plan and the partners' role within the project's consortium, focussing on each member's role in the success of the project. The fourth and fifth sections describe, respectively, the issues faced and the solutions developed within the ArtDATIS project. Finally, the last section offers some conclusive remarks and lessons learned.

2 Art Historical Background: Sybren Valkema and the Vrij Glass Movement
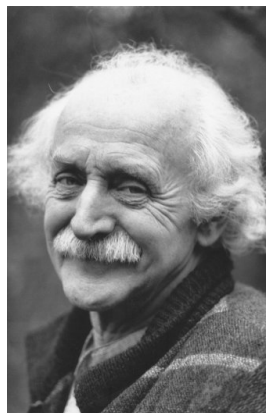
Figure 2 Sybren Valkema (Photo credit: Anna Carlgren, 1985).

From an art historical perspective, the Art DATIS project moved from the standpoint that the production of free glass saw significant technological innovations in the twentieth century. Due to the high temperatures and complex infrastructures that were necessary to melt glass, most of the early-twenty-century glass production happened within industrial facilities, where architects and designers were often invited to develop innovative and more sophisticated shapes capable of attracting the public's tastes [19]. With the intent of emancipating glass artists from industrial facilities in the early 1960s, Harvey Littleton (1922-2013) and Dominick Labino (1910-1987) prototyped an innovative studio furnace. It allowed artists to experiment freely with glass materials to produce aesthetically pleasant pieces [2].

In 1962, upon the invitation of the museum's director Otto Wittmann (1911-2001), Littleton organised two workshops on glass-blowing at the Toledo Museum of Glass. American artists and ceramists were invited to attend the workshops, experimenting with the studio furnace for melting and blowing glass. These events were considered a success for the knowledge they produced in setting up and running a studio production of glass, and are formally recognised as the inaugural events of the Studio Glass movement[16]. When, in 1964, the studio furnace was shown at the First World Craft Congress held in New York at Columbia University, Littleton's ideas encountered the support of a large number of artists and makers, determining the worldwide spread of the Studio Glass movement.

The Studio Glass movement encouraged artists to work independently in private studios, fostering a sense of artistic freedom and exploration. Within the glass-blowing studios, artists engaged in extensive research to push the boundaries of the glass medium, experimenting with new techniques and forms that emphasised individual expression and creativity. At the same time, the movement embraced education as a lever for transforming the art world, with universities playing a pivotal role in disseminating technical knowledge on glassmaking and cultivating a new generation of glass artists. In 1964, Littleton set up the first course on studio glass at the University of Wisconsin-Madison to promote glass-blowing among students. Between the second half of the 1960s and the early 1970s, the number of American universities offering courses in glass blowing grew considerably [16].

Sybren Valkema was among the participants at the 1964 World Craft Congress [Figure 2]. Despite his long and close collaboration as a design teacher with the Royal Leerdam Crystal, a Dutch factory producing glassware, Valkema had no direct experience in blowing glass [24]. In New York, he was fascinated by Littleton's demonstration and could experiment for the first time with the blowpipe. Once back in the Netherlands, he set up a glass-blowing laboratory on the premises of the Rietveld Academy in Amsterdam, building the first studio furnace in the Netherlands. In 1969, he inaugurated the first free glass curriculum, open to students of different disciplines [Figure 3]. During his life, Valkema experimented with glass and glass-blowing, studying historical sources and innovating them with newly developed materials and techniques, crucially contributing to the innovation the Studio Glass movement brought up.

Through his artistic research, Valkema established long-lasting friendships with most of the artists involved in the Studio Glass movement. Located in several countries worldwide, from the US to Australia and Japan, passing from the United Kingdom, Germany, and the Czech Republic, Valkema entertained a close epistolary exchange with artists such as Harvey Littleton, Erwin Eisch (1927-2022), Sam Herman (1936-2020), Stanislav Libenský (1921-2002) and Jaroslava Brychtová (1924-2020). With them, he exchanged information on techniques, designs, and events from the glass-blowing community. At the same time, students from the various glass courses were invited to spend a period abroad practising in the workshop of a different academy, generating a dynamic exchange of knowledge and ideas. The documents

Figure 3 Sybren Valkema (in the background) observes his son Durk blowing glass, assisted by Anna Carlgren, 1982 (Archive Sybren Valkema).

preserved in the Valkema archive, among which many witness the correspondence the artist had with friends, colleagues, and students, are invaluable for studying this lively environment. Therefore, one of the main research lines developed within the Art DATIS project investigated how knowledge circulated within the Studio Glass movement by reconstructing the network of Valkema's contacts.

3   The Consortium

In its aims, Art DATIS aligned with Studio Glass' ambitions of making glass-related knowledge widely available, this time using the digital medium. This was mainly made possible by the specific composition of the consortium. It was formed by five partners: two from academia (Utrecht University and the University of Amsterdam) and three public (RKD) and private institutions (Vrij Glas Foundation and Picturae). Within the consortium, each member had specific competencies in art history (UU, RKD and Vrij Glass Foundation), data science (UvA), data management (UvA, RKD and Picturae) and digital humanities (UU).

The project's idea stemmed from the collaboration between Utrecht University and RKD within the framework of the ARTECHNE project [1]. ARTECHNE: Technique in the Arts 1500–1950 was a European Research Council-funded project hosted between 2015 and 2021 at Utrecht University and the University of Amsterdam. The project developed an online database of art technological sources containing very heterogeneous data: fully searchable texts in six languages from the period 1500 to 1900, images, records on historical people, geotags, and historical and current names of artistic materials and techniques [14]. Intending to provide for the long-term sustainability of the database, ARTECHNE partnered with

RKD, which added it to the list of digital resources offered within its search engine RKD Explore. This collaboration led RKD, which in the meantime had acquired Sybren Valkema's digital archive, to propose linking the historical sources in its collection with the documents in the archive and other existing big data in the field of art history and technical art history. This idea would have contributed to enriching Valkema's digital archive, falling within the scope of the call on big data and smart culture issued by the NWO. Thanks to Art DATIS, the ARTECHNE database was enriched with new entries on glass and glassmaking (i.e. historical recipes and descriptions of glassmaking procedures from historical sources).

Art DATIS, therefore, was the natural follow-up of the digitisation of Valkema's archive. In the process of digitisation of the archival resources, the work performed by the company Picturae, which specialised in collaborating with cultural heritage institutions to produce high-quality digital duplicates, was crucial. Due to its competence with the management of digital cultural data and its knowledge of the archive, Picturae was involved in the development of the Art DATIS project.

With the aim of forming a robust cohort of partners, RKD contacted the Vrij Glas Foundation, the legal owner of the Valkema archive, and involved them in the project. The Foundation is directed by Durk Valkema, Sybren's son, and his wife, Anna Carlgren. They both shared an artistic career working with glass, learning glassblowing at the Rietveld Academy under Sybren's supervision, and witnessing the evolution of the Studio Glass Movement in the Netherlands and internationally. As the gatekeeper of Valkema's heritage, the Foundation was enthusiastic about the prospect of making the archive widely available, emphasising the continuity between its digitisation and the Studio Glass support of open and free circulation of technical knowledge about glass. The intimate knowledge of the archival material and the first-hand experience of Valkema's glass-blowing practice made the presence of Vrij Glas in the consortium an invaluable resource.

Within the project, Utrecht University and the University of Amsterdam supervised the scientific research development, collaborating to make the Valkema archive digitally accessible and searchable. The research team was composed of three researchers and three supervisors. The project trained two PhDs, one in the field of art history and one in data science. They were based, respectively, at Utrecht University (History and Art History Department) and the University of Amsterdam (Informatics Institute). Additionally, a Postdoctoral researcher (Utrecht University, History and Art History Department) joined the project to facilitate the process of datafication of the archival resources. The research team closely collaborated during the whole development of the project, benefitting from the support of the other members of the consortium. Each partner offered its competencies on specific aspects of the work. The Vrij Glas Foundation, which inherited Valkema's artistic practice and is in charge of preserving his archive, offered support with its know-how on glassmaking; RKD offered the digital space to protect and enrich digital resources, granting their long-term sustainability, while the digital cultural heritage company Picturae provided advice on technical solutions and best practices in digital preservation.

4   Problem addressed

Art DATIS represents an outstanding case study for approaching the mass digitisation and automatic documentation of heterogeneous archival resources. The project investigated how to efficiently deal with the datafication of heterogeneous archival resources, connecting them with existing digital art technical sources on artistic glass production, such as object documentation, technical texts, images, and research data. Therefore, the project's ultimate

goal was to contribute to the creation of a comprehensive art historical resource about glass and glassmaking, allowing scholars, artists and practitioners to investigate the use of glass as an artistic material through time.

The diverse nature of the resources in the archive requires each a specific treatment to be transformed from simple digital images into meaningful data resources. The first challenge Art DATIS had to overcome was finding a way to describe and document each entry in the collection to process it adequately. This includes creating a set of metadata containing enough information to allow retrieving the resource in the appropriate context. As a matter of fact, when the archive was digitally remediated, resources were grouped into broad thematic categories, which offered those who approached the collection a general idea of their context but needed to be renamed and documented in detail.

The lack of prior information about the resources also represented an optimal opportunity to experiment with innovative archival practices and develop a system allowing resource documentation with minimal content interpretation. Scholars have closely scrutinised the practices of object documentation applied by archives and cultural heritage institutions, recognising that, at times, they have been subjected to cultural and political biases, which affect the production and dissemination of new knowledge [12] [3] [4]. For this reason, Art DATIS worked to implement a rich documentation system based on the very content of each document, aiming to reduce their interpretation.

To address these issues, the project had to investigate the technical aspects of datafication. The archive comprises various heterogeneous resources, including handwritten and typed documents, designs, and pictures. Art DATIS focussed on written resources, aiming at making them searchable. To make it possible, documents had to be categorised and analysed. It was first necessary to separate visual documentation from written resources; then, the latter had to be processed with the most appropriate tool to extract the textual information.

To make content searchable, we had to convert images into machine-encoded text, a process named Text Recognition [11]. As the following section will better describe, in practice, this required developing an algorithm capable of automatically identifying written sources and separating them from visual contents. Then, a second algorithm was designed to distinguish typed from handwritten documents. The former type of documents were then transcribed using an OCR tool, while the latter required HTR software. Once transcribed, the documents became searchable, and their content could eventually be used to generate automatic metadata. To facilitate this process, an algorithm allowing automatic entity recognition and linking was developed [22].

The joint collaboration of art historians and data scientists was crucial during the whole development of the project. It allowed for research in the field of digital humanities, developing a tool at the service of the community of scholars and artists. The archival resources, enriched with data generated by external resources and made available through the RKD servers, will allow researchers and practitioners to extensively investigate the transformation of glass production throughout time, browsing old and new recipes to discover the role of traditional practice in contemporary glass-blowing.

5   Key results

To address the issues at the core of the research project, it was first necessary to develop a strategy to sort out resources according to their type (i.e. drawings, pictures, typed texts or handwritten texts) and process each of them with the most appropriate tool. The focus of Art DATIS was primarily documentary resources. Therefore, we mostly worked to develop

Table 1 Evaluation of OCR's results.

type	examples	number of pages in the sample
true positive	typed pages; some mixed pages (where the majority of the text is typed but some handwritten notes or drawings are present)	27 (24 typed, 3 mixed)
false positive	handwritten pages with a very clear writing style (not a problem, can be OCRed); some mixed pages (where the amount of handwritten text is sufficient to make HTR necessary)	5 (1 handwritten, 3 mixed, 1 back)
true negative	handwritten pages, drawings, backsides, blank pages, some mixed pages (where only a small amount of typed text is present)	60 (14 handwritten, 35 backsides, 5 blank pages, 6 mixed)
false negative	typed pages missed by the classifier	8 (7 typed, 1 mixed)

optimal OCR and HTR solutions. To identify which documents are typed pages that can be processed with OCR algorithms, we developed a rule-based binary classifier using the following heuristic: if running the Tesseract OCR engine on a page detects Latin script, then the page is likely typed [15]. To validate this approach, we randomly sampled 100 pages from the archive and performed a manual evaluation. Out of the 100 pages, 35 were typed. The classifier was found to have 84% precision, i.e. it predicted 32 pages as typed, out of which 27 were actually typed, and 77% recall, i.e. out of all 35 typed pages, it successfully identified 27. Detailed results of the evaluation can be seen in Table 1.

On the other hand, the handwritten material needed to be processed with appropriate HTR algorithms. Art DATIS consortium used Transkribus, a platform enabling AI-powered text recognition. Transkribus relies on the principles of machine learning and employs artificial neural networks to assist researchers in training models that can read and transcribe their collections automatically. [6]. Performing HTR on the Valkema archive was particularly complex due to two main issues. First, the archive includes not only documents written by Valkema at any stage of his life, therefore presenting a variation in his handwriting, but also documents written by other people in several languages. Current HTR methods require extensive manual annotation efforts to create training data: every language and every handwriting have unique features and ideally need the use of a separate model. Transkribus suggests utilizing a minimum of 10,000 words per unique writer to effectively train a model [7]. Second, due to the fact that digital resources are undocumented, the lack of metadata about the language and author makes it impossible to automatically sort out the material for processing it with a specific model.

Using Transkribus, it was possible to train different language models and experiment with their efficiency in transcribing the collection. Due to the lack of metadata on the language used in the documents, we identified the more common languages used in the archive (Dutch, English, and German) and developed a model capable of producing automatic transcription with an acceptable margin of error for all the languages. Therefore, we conceived a series of experiments to assess the efficiency of a multilingual, multi-author model trained on a sample of documents from the Valkema archive. Its efficiency was tested on a sample of documents randomly selected in the archive and compared to the results produced by models trained on documents in a single language. Table 2 shows the monolingual and multilingual models we have trained and their performances. The multilingual model performed adequately in comparison with the monolingual ones, allowing for the automatic transcription of the whole archive [5].

Once all the images were converted into textual information using OCR and HTR, the archive became searchable. The second main challenge faced within Art DATIS was then designing a system to automatically identify meaningful entities embedded in the text and link them with existing open data. Identifying named entities, such as people and cities, is a useful tool for digital archives: it allows for the improvement of the metadata and

**Table 2** Overview of the HTR models trained during the experiment and their performances.

Language	Training details	Title & Transkribus ID	Performance during training (average CER)	Performance of the model calculated on our collection (average CER)
EN	base model: Transkribus English Handwriting M3 (ID 37646) training data: 87 pages (10% used for validation)	ArtDATIS_English_ latest (53449)	CER on train: 4.7% CER on validation: 2.1%	CER on test: 11.1% (STD 12.8%)
NL	base model: Transkribus Dutch Handwriting M2 (ID 45422) training data: 215 pages (10% used for validation)	ArtDATIS_Dutch_ latest (53462)	CER on train: 9.8% CER on validation: 6.1%	CER on test: 18.9% (STD 9.5%)
Multi	base model: Transkribus Dutch Handwriting M2 (ID 45422 training data: EN – 87 pages NL – 215 pages (10% in each language is used for validation)	ArtDATIS_Dutch_ English (53444)	CER on train: 6.6% CER on validation: 3.2%	CER on test: 20.1% all languages (STD 14.8%) 12.5% EN (STD 13.1%) 30.7% DE (STD 14.9%) 17.1% NL (STD 9.5%)

makes the documents easily searchable. Entity linking takes the process one step further by connecting unstructured text with a structured knowledge base, such as Wikipedia or Wikidata. Entity linking in digital archives comes with a number of challenges: (a) OCR and HTR errors confuse entity identification and linking algorithms, while (b) entities of the past are overshadowed by entities of the present due to popularity bias in entity identification and linking algorithms. We studied the latter problem [20] and proposed solutions for robustifying the entity linking algorithm [21].

The automatic identification of entities related to people allowed us to generate an original visualisation of the information contained in the archive: Valkema's social network [Figure 4]. Such a visualisation shed light on how innovation in glassmaking travelled across countries in the second half of the Twentieth century through the social connection of artists and the exchange of students.

6   Conclusion and looking ahead

One of the most remarkable results of the Art DATIS project is the creation of a simple procedure for the datafication of mass digitised archives, which can be easily imitated by cultural heritage institutions. Most of the archival collections are, in fact, constituted of heterogeneous material that needs to be processed with ad hoc instruments in order to become searchable. At the same time, the large number of documents that have to be processed and the amount of funding necessary for the operation make it very difficult for cultural institutions to process their entire collections. In this sense, the solutions developed within Art DATIS might represent a valuable reference for similar collections. Concretely, the multilingual HTR model trained using Transkribus will be freely available for reuse on
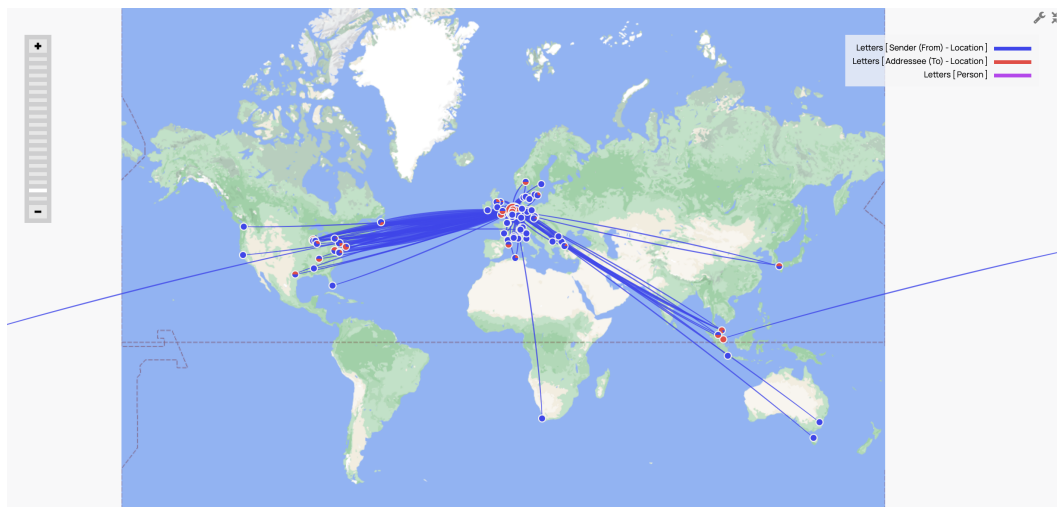
Figure 4 Visualisation of the geographical provenance of the correspondence in the Valkema's archive. Graph generated with the software Nodegoat.

the tool platform. Any institution or researcher dealing with handwritten documents will be able to use the model to transcribe their collections, saving time and resources in training a new ad-hoc model. In the same way, the entity recognition and entity linking algorithms developed to process the Valkema collection will be made available, a.o. as part of the toolset offered by the RKD online platform.

To conclude, the most important lesson that the project can offer to similar ones addressing the datafication of archival resources is to carefully estimate the time and resources necessary to reach their goals. The heterogeneous nature of archival resources makes them singular research objects, which poses specific and time-consuming issues when digitised and transformed into data. These aspects need to be taken into consideration in the project design phase in order not to jeopardise the success and financial viability of the project. In this sense, the close collaboration of experts from the humanities and technical disciplines is necessary already in the earliest phases of data creation in order to successfully complete digital humanities projects. Funding agencies should create financial opportunities for this, and as this project shows, investment in this phase of data creation can also lead to more widely useful results and tools.

References

1   ARTECHNE. ARTECHNE - Technique in the Arts, 1500-1950 | Technique in the Arts, 1500-1950: Concepts, Practices, Expertise. URL: `https://artechne.wp.hum.uu.nl`.

2   Joan Falconer Byrd and Harvey K. Littleton. *Harvey K. Littleton - a life in glass: founder of America's studio glass movement.* Skira Rizzoli, New York, 2011.

3   Fiona Cameron and Helena Robinson. Digital Knowledgescapes: Cultural, Theoretical, Practical, and Usage Issues Facing Museum Collection Databases in a Digital Epoch. In Fiona Cameron and Sarah Kenderdine, editors, *Theorizing Digital Cultural Heritage*, pages 165–191. The MIT Press, March 2007. `doi:10.7551/mitpress/9780262033534.003.0010`.

4   Carlotta Capurro and Gertjan Plets. Europeana, EDM, and the Europeanisation of cultural heritage institutions. *Digital Culture and Society*, 6(2):163–189, 2020. `doi:DOI10.14361/dcs-2020-0209`.

5   Carlotta Capurro, Vera Provatorova, and Evangelos Kanoulas. Experimenting with training a neural network in transkribus to recognise text in a multilingual and multi-authored manuscript collection. *Heritage*, 6(1212):7482–7494, December 2023. `doi:10.3390/heritage6120392`.

6   Sebastian Colutto, Philip Kahle, Hackl Guenter, and Guenter Muehlberger. Transkribus. A Platform for Automated Text Recognition and Searching of Historical Documents. In *2019 15th International Conference on eScience (eScience)*, pages 463–466, San Diego, CA, USA, September 2019. IEEE. `doi:10.1109/eScience.2019.00060`.

7   Read Coop. Training models. data preparation, 2023. URL: `https://help.transkribus.org/data-preparation`.

8   Art DATIS. Art DATIS project. URL: `https://artdatis.nl`.

9   Sven Dupré. Art Educational Reforms and the Studio Glass Movement: Artists' Ambivalent Self-Representation as Actors of Lost Knowledge. In Johan Östling, David Larsson Heidenblad, and Anna Nilsson Hammar, editors, *Knowledge Actors: Revisiting Agency in the History of Knowledge*. Nordic Academic Press, Lund, 2023.

10   Susanne Kaatherine Frantz. *Artists and Glass. A History of International Studio Glass*. PhD thesis, The University of Arizona, 1987.

11   Maya R. Gupta, Nathaniel P. Jacobson, and Eric K. Garcia. Ocr binarization and image pre-processing for searching historical documents. *Pattern Recognit.*, 40:389–397, 2007. URL: `https://api.semanticscholar.org/CorpusID:15012496`, `doi:10.1016/J.PATCOG.2006.04.043`.

12   Stuart Hall. Constituting an archive. *Third Text*, 15(54):89–92, March 2001. `doi:10.1080/09528820108576903`.

13   Karin Hansson, Anna Näslund Dahlgren, and Teresa Cerratto Pargman. Datafication and cultural heritage: Critical perspectives on exhibition and collection practices. *Information I& Culture*, 57(1):1–5, March 2022. `doi:10.7560/IC57101`.

14   Marieke M. A. Hendriksen. Building the ARTECHNE Database: How to Develop a Multi-Purpose Database for an Interdisciplinary Project. *IDEAH*, May 2020. `doi:10.21428/f1f23564.e6779c76`.

15   Anthony Kay. Tesseract: an Open-Source Optical Character Recognition Engine | Linux Journal. *LINUX Journal*, July 2007. URL: `https://www.linuxjournal.com/article/9676`.

16   M.D. Lynn. *American Studio Glass, 1960-1990*. Hudson Hills Press, 2004. URL: `https://books.google.nl/books?id=b4KhNlxnyOEC`.

17   Sue Newell and Marco Marabelli. Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of "datification". *Journal of Strategic Information Systems*, 24(2644093), March 2015. `doi:10.2139/ssrn.2644093`.

18   NWO. Smart culture - nwo, February 2022. URL: `https://www.nwo.nl/en/researchprogrammes/creative-industry/smart-culture`.

19   Ada Polak. *Modern Glass*. Faber I& Faber, New York, thomas yoseloff edition, 1962.

20   Vera Provatorova, Samarth Bhargav, Svitlana Vakulenko, and Evangelos Kanoulas. Robustness evaluation of entity disambiguation using prior probes: the case of entity overshadowing. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 10501–10510, 2021. `doi:10.18653/V1/2021.EMNLP-MAIN.820`.

21   Vera Provatorova, Simone Tedeschi, Svitlana Vakulenko, Roberto Navigli, and Evangelos Kanoulas. Focusing on context is nice: Improving overshadowed entity disambiguation, 2022. `arXiv:2210.06164`, `doi:10.48550/arXiv.2210.06164`.

22   Vera Provatorova, Svitlana Vakulenko, Evangelos Kanoulas, and Johannes M van Hulst. Named entity recognition and linking on historical newspapers: Uva.ilps i& rel at. In *CEUR-WS*, volume 2696, Thessaloniki, 2020. URL: `https://ceur-ws.org/Vol-2696/paper_209.pdf`.

23   RKD. Archief Sybren Valkema. URL: `https://rkd.nl/nl/projecten-en-publicaties/projecten/265-archief-sybren-valkema`.

24   Sybren Valkema and Klaas Laansma. *Sybren Valkema*. De Prom, Baarn, 1994.