

Zweitveröffentlichung



Scheltjens, Werner

Genetische Kritik gedruckter historischer Datensammlungen am Beispiel der dänischen Sundzolltabellen (1906–1953) : ein Werkstattbericht

Datum der Zweitveröffentlichung: 23.10.2023

Verlagsversion (Version of Record), Beitrag in Sammelwerk

Persistenter Identifikator: urn:nbn:de:bvb:473-irb-91310

Erstveröffentlichung

Scheltjens, Werner: Genetische Kritik gedruckter historischer Datensammlungen am Beispiel der dänischen Sundzolltabellen (1906–1953) : ein Werkstattbericht. In: Digitale Methoden in der geschichtswissenschaftlichen Praxis : Fachliche Transformationen und ihre epistemologischen Konsequenzen: Konferenzbeiträge der Digital History 2023 ; Berlin, 23.-26.5.2023. Berlin: Zenodo, S. 1-11. DOI: 10.5281/zenodo.8322345

Rechtehinweis

Dieses Werk ist durch das Urheberrecht und/oder die Angabe einer Lizenz geschützt. Es steht Ihnen frei, dieses Werk auf jede Art und Weise zu nutzen, die durch die für Sie geltende Gesetzgebung zum Urheberrecht und/oder durch die Lizenz erlaubt ist. Für andere Verwendungszwecke müssen Sie die Erlaubnis des/der Rechteinhaber(s) einholen.

Für dieses Dokument gilt eine Creative-Commons-Lizenz.



Die Lizenzinformationen sind online verfügbar:

<https://creativecommons.org/licenses/by/4.0/legalcode>

Genetische Kritik gedruckter historischer Datensammlungen am Beispiel der dänischen Sundzolltabellen (1906–1953)

Ein Werkstattbericht

Werner Scheltjens

Otto-Friedrich-Universität Bamberg, Bamberg, Deutschland

ORCID: 0000-0002-5209-9052

Abstract: Am Beispiel der "dänischen Sundzolltabellen", eine renommierte Sammlung von Statistiken über Schifffahrt und Warenverkehr zwischen Nord- und Ostsee von 1497 bis 1783, zeigt dieser Beitrag wie digitale Tools und Methoden eine emische Perspektive auf die Genese historischer Datensammlungen ermöglichen. Die emische Perspektive versucht sich den Praktiken der Forscherinnen und Forschern wertschätzend und ‚auf Augenhöhe‘ anzunähern und verzichtet bewusst auf externe, a priori festgelegte Bewertungskriterien. Die materiellen Bedingungen für eine genetische Datenkritik der "dänischen Sundzolltabellen" sind einzigartig. Die uneingeschränkte digitale Verfügbarkeit der archivalischen Grundlage der "dänischen Sundzolltabellen" auf Sound Toll Registers Online und eine umfassende Historiographie machen es möglich, in die Quelle ‚einzutauchen‘ und die Genese der vielen Tabellen fast mikroskopisch zu betrachten und zu rekonstruieren. Der Beitrag zeigt, wie eine genetische Datenkritik über die Praktiken im Umgang mit historischen Daten informiert und so zu neuen Erkenntnissen über die Entwicklung und Anwendung von quantifizierenden Verfahren in der historischen Forschung beiträgt.

Keywords: Genetische Datenkritik, emische Perspektive, Sundzolltabellen, Sound Toll Registers Online

1. Einleitung

Gedruckte historische Datensammlungen sind in der sozial- und wirtschaftshistorischen Forschung unabdingbar. Manche Sammlungen, z.B. *European Historical Statistics* oder *British Historical Statistics*¹, enthalten Zeitreihen, die aus einer Vielzahl von Quellen zusammengestellt wurden und für die Beschäftigung mit ‚großen Themen‘ wie etwa Wachstum, Ungleichheit, Industrialisierung oder Globalisierung von zentraler Bedeutung sind. Andere Sammlungen sind vorrangig als Nachschlagewerk für historische Studien zu Handel, Gewerbe und Transport verdienstvoll, wie etwa das achtbändige *Handbuch der Historischen Metrologie*.² Hinzu kommt eine schwer durchschaubare Vielfalt an offiziellen und semi-offiziellen Statistiken, die seit dem 18. Jahrhundert veröffentlicht wurden. Obwohl seit Anfang des 21. Jahrhunderts wichtige historische Datenbeständen und Statistiken vermehrt digital erschlossen werden, verlieren die bereits vorhandenen gedruckten Sammlungen kaum ihre Relevanz für die Forschung.

Das gilt auch für die sogenannten *Sundzolltabellen*, eigentlich: *Tabeller over skibsfart og varetransport gennem Øresund* – in der Historiographie häufig abgekürzt als STT (*Eng.* Sound Toll

¹ Brian R. Mitchell, *European historical statistics 1750-1970*, New York 1975; Brian R. Mitchell, *British historical statistics*, Cambridge 1990.

² Harald Witthoft (Hrsg.), *Handbuch der historischen Metrologie* (8 vols.), St. Katharinen 1991–2007.

Tables). Die STT sind eine renommierte historische Datensammlung von etwa 3.000 Seiten mit Statistiken über Schifffahrt und Warenverkehr durch die Meerenge Øresund zwischen Nord- und Ostsee von 1497 bis 1783.³ Die STT erschienen in insgesamt sieben Bänden in der Zeit von 1906 bis 1953. Ziel der Initiatorin der STT, Nina Ellinger Bang (1866–1928), war es, den Zugang zu der gigantischen Archivsammlung der dänischen Sundzollregister (*Eng.* Sound Toll Registers; kurz: STR) zu verbessern, die etwa 60 Meter Regalfläche einnimmt.⁴ Trotz erheblicher Quellenkritik haben die STT eine enorme Bedeutung für die Erforschung des frühneuzeitlichen europäischen Warenverkehrs.⁵ Zwischen 2009 und 2020 wurde im Rahmen des Projekts *Sound Toll Registers Online* (kurz: STRO) eine elektronische Datenbank der Schiffspassagen durch den Øresund einschließlich der beförderten Ladungen und gezahlten Steuern erstellt (www.soundtoll.nl). Im Zuge dieses Projektes wurden auch die Scans der dänischen Sundzollregister als der gemeinsamen Grundlage von STT und STRO vollumfänglich online verfügbar gemacht.

Man könnte jetzt glauben, dass STRO die STT überflüssig gemacht hat, aber auch wenn die STT nach modernen Maßstäben ‚veraltet‘ und ‚überholt‘ sind, ist eine solche Behauptung aus erkenntnistheoretischer Sicht nicht gerechtfertigt. Als Beispiel für eine historische Datensammlung aus der Zeit vor der Erfindung des Computers bleiben die STT ein höchst relevantes Objekt für wissenschaftshistorische Untersuchungen. Gerade die Kombination von Archiv (Scans), Datenbank (STRO) und Retrodigitalisat der STT, flankiert durch eine umfassende Historiographie der dänischen Sundzollregister⁶, ist für die wissenschaftshistorische Betrachtung der STT eine einzigartige Ausgangslage. Unter Nutzung dieser Konstellation kann Neues über die Methoden der Quantifizierung und Datenmodellierung und über die Epistemologie gedruckter historischer Datensammlungen entdeckt werden. Dieser Werkstattbericht zeigt, wie eine wissenschaftshistorische Perspektive über Praktiken im Umgang mit historischen Daten informieren und so zu neuen Erkenntnissen über die Genese historischer Datensammlungen beitragen kann.

2. Datengenerese als Forschungsproblem

Nachschlagewerke und Datensammlungen wie die STT sind dank der massenhaften Retrodigitalisierung von ‚alten Drucken‘, leichter zugänglich als je zuvor. Neu erschlossene historische Datensammlungen kommen hingegen eher spärlich hinzu. Denn, ob digital oder gedruckt, die Erschließung historischer Datensammlungen ist in der Regel eine langfristige und kollaborative Unternehmung, die nach Vollendung kaum ersetzt, geschweige denn wiederholt werden kann. In dieser Hinsicht ist der Stellenwert von gedruckten und retrodigitalisierten historischen Datensammlungen mit der Bedeutung der sogenannten *legacy collections* in der Archäologie vergleichbar.⁷ Sie wurden weder nach den FAIR-

³ Nina Ellinger Bang / Knud Korst (Hrsg.), *Tabeller over skibsfart og Varetransport gennem Øresund / Tables de la navigation et du transport des marchandises passant par le Sund 1497–1783*. København 1906–1953.

⁴ Erik Gøbel, *The Sound Toll Registers Online project, 1497–1857*, in: *International Journal of Maritime History* 22 (2010), S. 305–324.

⁵ Pierre Jeannin, *Les comptes du Sund comme source pour la construction de l'activité économique en Europe (XVIe–XVIIIe siècle)*, in: *Revue historique* 121 (1964), S. 55–102; 307–340; Jan Willem Veluwenkamp / Werner Scheltjens / Siem Van der Woude, *Sound Toll Registers Online*, in: *TSEG-The Low Countries Journal of Social and Economic History* 18 (2021), S. 147–160. <https://doi.org/10.18352/tseg.1203>

⁶ Ole Degn (Hrsg.), *The Sound Toll at Elsinore. Politics, shipping and the collection of duties*, Copenhagen 2018.

⁷ Julia A. Kelly u.a., *A Critical Literature Review of Historic Scientific Analog Data: Uses, Successes, and Challenges*, in: *Data Science Journal* 21 (2022), S. 1–11. <https://doi.org/10.5334/dsj-2022-014>; Kathryn MacFarland / Arthur W. Vokes, *Dusting off the data: curating and rehabilitating archaeological legacy and orphaned collections*, in: *Advances in Archaeological Practice* 4 (2016), S. 161–175; Hilary Szu Yin Shiue u.a., *Assessing Legacy Collections for Scientific Data Rescue*, in: Katharina Toepe / Hui Yan / Samuel Kai Wah Chu (Hrsg.), *Diversity, Divergence, Dialogue: 16th International Conference, iConference 2021, Beijing, March 17–31 2021, Proceedings, Part II, Cham 2021, S. 308–318*. https://link.springer.com/chapter/10.1007/978-3-030-71305-8_25.

Prinzipien aufbereitet noch hinreichend dokumentiert, aber trotz aller Mängel ist es unmöglich und keinesfalls wünschenswert, auf sie zu verzichten.⁸

Allerdings verändern sich die Anforderungen im Umgang mit historischen Daten seit einigen Jahren grundlegend. Fragen der Nachprüfbarkeit und der Reproduzierbarkeit von Forschungsergebnissen stehen zunehmend im Fokus.⁹ Historiker:innen, die auf Retrodigitalisate von gedruckten historischen Datensammlungen mit mangelhafter Dokumentation angewiesen sind, werden mit Fragen konfrontiert, die in der herkömmlichen Quellenkritik nur eine geringe Rolle spielen. Ist die Herkunft der Daten in der gedruckten Sammlung nachprüfbar? Wurden alle Daten erfasst oder nur eine Auswahl getroffen? Ist *jede* Auswahl begründet? Wie wurden die Daten kategorisiert und formatiert, damit sie auf eine Druckseite passen? Welche Daten wurden quantifiziert? Wie wurde quantifiziert? Zu welchen Verlusten hat die Quantifizierung geführt? Sind die Ergebnisse der Quantifizierung reproduzierbar? Solche Fragen sind an die Genese der historischen Datensammlung und so auch an die Arbeit früherer Generationen von Forscher:innen zu stellen.

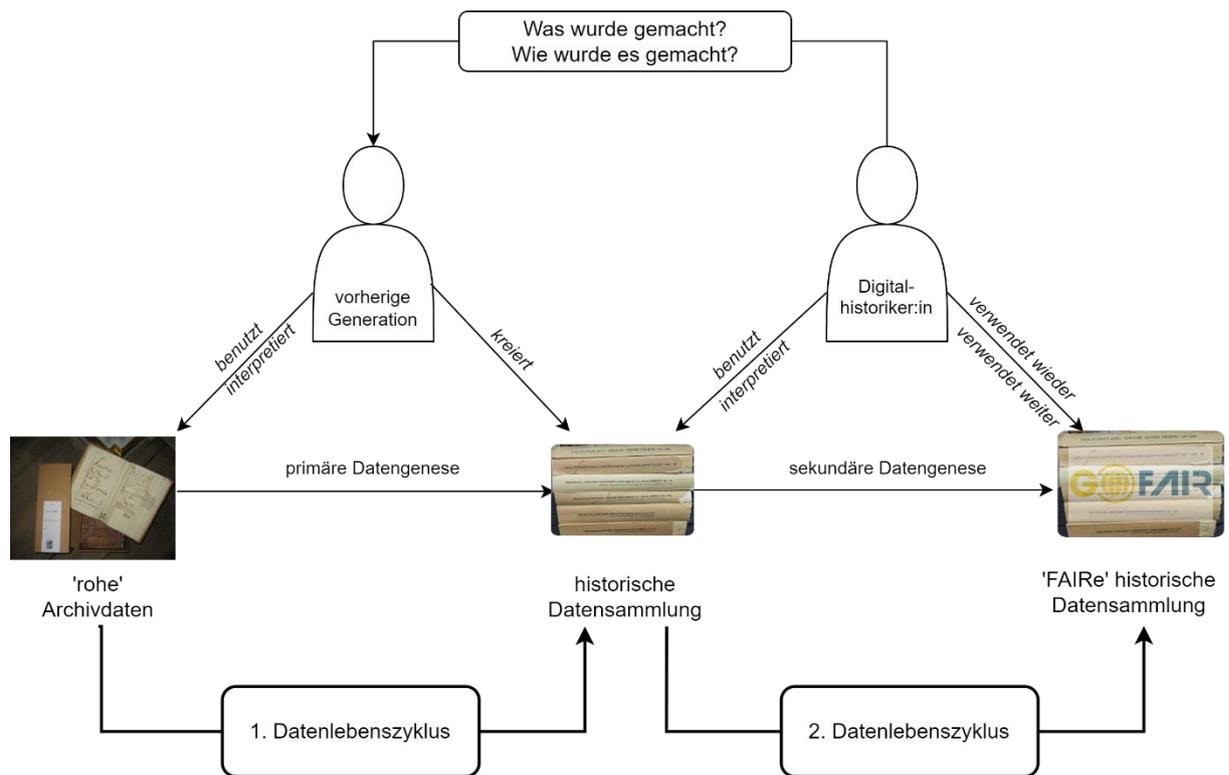


Abb. 1: Datengenese als Forschungsproblem.

Ausgehend von einer gedruckten historischen Datensammlung, die zunächst aus primären Archivdaten entstanden ist, und nun wieder- oder weiterverwendet werden soll, lässt sich die Datengenese an zwei Stellen verorten (siehe Abbildung 1). Die primäre Datengenese umfasst die Transformation von den ‚rohen‘ Archivdaten zur gedruckten historischen Datensammlung. Die sekundäre Datengenese umfasst die Transformation von der gedruckten zur FAIRen, reproduzierbaren und nachhaltig verfügbaren Datensammlung. Der Fokus der Datengenese liegt also auf den Transformationsprozessen, die zum Entstehen neuer Daten führen. Weil bei jedem Entstehen neuer Daten prinzipiell auch ein neuer Datenlebenszyklus anfängt, könnte man Datengenese auch beschreiben

⁸ Werner Scheltjens, Upcycling historical data collections. A paradigm for digital history?, in: Journal of Documentation, Vol. ahead-of-print (2023). <https://doi.org/10.1108/JD-12-2022-0271>.

⁹ Christof Schöch u.a., Replication and Computational Literary Studies, in: Digital Humanities Conference 2020 (DH2020), Ottawa. <https://doi.org/10.5281/zenodo.3893428>; Charlotte Schubert, Digital Humanities auf dem Weg zu einer Wissenschaftsmethodik: Transparenz und Fehlerkultur, in: Digital Classics Online 7 (2021), S. 39–53. <https://doi.org/10.11588/dco.2021.7.82371>.

als ein Transformationsprozess, der zwischen Datenlebenszyklen stattfindet und diese miteinander verbindet.

Die Herausforderung besteht nun darin, dass gerade bei historischen Datensammlungen die Prozesse, die zur Transformation von ‚rohen‘ Archivdaten in eine gedruckte Quelle beigetragen haben, häufig unsichtbar, nicht oder nur geringfügig dokumentiert, und in den meisten Fällen sogar nicht (mehr) nachvollziehbar sind. Für die Wieder- und Weiterverwendung von historischen Datensammlungen ist die Unsichtbarkeit der Datengenese von zentraler Bedeutung. Weil die Transformationsprozesse zwischen vergangenen Datenlebenszyklen unsichtbar und (häufig) nicht dokumentiert sind, verlangt die Erforschung der Datengenese einen hermeneutischen Rekonstruktionsakt, der die Prozesse, die zwischen den Datenlebenszyklen stattfinden, quasi sichtbar macht und dokumentiert. Dieser Rekonstruktionsakt ist immer durch Unsicherheit und Unvollständigkeit geprägt. Dennoch liefert er einen wesentlichen Beitrag zu epistemologisch fundiertem Wissen über historische Datensammlungen. Anlehnend an die genetische Kritik (*Eng. genetic criticism*) in den Literaturwissenschaften¹⁰, nennen wir die erkenntnistheoretische Methode des Rekonstruierens der Datengenese, d.h. der Prozesse zwischen Datenlebenszyklen, genetische Datenkritik.

3. Genetische Datenkritik

Die genetische Datenkritik (GDK) könnte als datenorientierter Counterpart des *genetic criticism* (GC) verstanden werden. Beide Formen der Kritik beschäftigen sich mit den kognitiven Prozessen, die sich bei der Genese von literarischen Texten (GC) und Datensammlungen (GDK) entfalten, und benutzen dafür eine möglichst breite Palette an verfügbaren Quellen. Für die Rekonstruktion des Schreibprozesses von Autoren wie etwa James Joyce orientiert sich die genetische Kritik an Notizen, Skizzen, Tagebucheinträgen, Marginalien und verschiedenen handschriftlichen Versionen eines Textes. Für die Rekonstruktion des Transformationsprozesses von ‚rohen‘ Archivdaten zur gedruckten historischen Datensammlung stehen solche Unterlagen meist nicht zur Verfügung. Das liegt auch daran, dass im Gegensatz zur genetischen Kritik, in welcher die Kreativität eine zentrale Rolle spielt, für GDK eher die Konstruktion im Vordergrund steht. Vielleicht noch mehr als die genetische Kritik muss die GDK sich *mittelbar* Einsicht in die Methoden der Konstruktion von historischen Datensammlungen verschaffen. Dank der ersten Massendigitalisierung und dem Durchbruch digitaler Forschungsmethoden in den Geisteswissenschaften stehen dafür mehr als je zuvor die notwendigen Materialien zur Verfügung.

Am Beispiel der STT soll im Folgenden gezeigt werden, wie digitale Tools und Methoden eine *emische* Perspektive auf die Genese historischer Datensammlungen ermöglichen, d.h. eine Perspektive, die versucht, sich den Praktiken der Forscherinnen und Forscher wertschätzend und ‚auf Augenhöhe‘ anzunähern und dabei bewusst auf externe, *a priori* festgelegte Bewertungskriterien verzichtet.¹¹ Das einzigartige ‚genetische Dossier‘ der STT, das durch die gleichzeitige und uneingeschränkte Verfügbarkeit von Archiv (Scans), Datenbank (STRO) und Retrodigitalisat (STT) gekennzeichnet ist, macht es möglich, in die Quelle einzutauchen und die Genese der STT fast mikroskopisch zu betrachten und zu rekonstruieren.

4. Die Semantik der Kupferexporte

Die für eine GDK der STT in Frage kommende Datenmenge ist sehr groß. Das genetische Dossier umfasst etwa 120 Jahre Historiographie, 3.000 Seiten Statistiken in den STT, etwa 1,8 Millionen Durchfahrten und 6 Millionen Warenregistrierungen in STRO und 227.000 Scans der Originalregister. Eine umfassende GDK kann im Rahmen dieses Beitrags natürlich nicht realisiert werden. Jedoch wurde die Registrierung von Kupferexporten in der Zeit von 1497 bis 1569 als repräsentatives Beispiel

¹⁰ Dirk Van Hulle, *Genetic Criticism. Tracing Creativity in Literature*, Oxford 2022. <https://doi.org/10.1093/oso/9780192846792.001.0001>.

¹¹ Carlo Ginzburg, *Our Words, and Theirs. A Reflection on the Historian's Craft, Today*, in: Cromohs (*Cyber Review of Modern Historiography*) 18 (2013), S. 97–114.

ausgewählt, das zumindest erste Einsichten in die Chancen und Herausforderungen der GDK und erste Aussagen über die Methoden der Quantifizierung und Epistemologie der STT erlaubt.

4.1. Hintergrund

Die Kupferexporte aus dem Ostseeraum stehen in einem engen Zusammenhang mit dem Vordringen der oberdeutschen Kaufmannsfamilie Fugger in den Ostseeraum. 1494 begründeten die Fugger zusammen mit der Krakauer Kaufmannsfamilie Thurzo ein Unternehmen für den „Gemeinen Ungarischen Handel“.¹² Dieses Unternehmen war der Kupfergewinnung im ungarischen Neusohl gewidmet. Bis 1546 hielten die Fugger an ihrem Monopol an der Neusohler Kupfergewinnung fest.¹³ Von 1548 bis 1560 pachtete Melchior Manlich, der ebenfalls zum Augsburger Kaufmannsstand gehörte, die Neusohler Kupfergruben. Später übernahmen weitere Augsburger Firmen die nunmehr in königlicher Regie organisierte Pacht.¹⁴ Für den Kupfervertrieb spielte Danzig eine herausragende Rolle. Von dort aus wurde das Kupfer durch den Øresund nach Antwerpen und Amsterdam exportiert. Um einen reibungslosen Transport und steuerliche Begünstigungen zu erzielen, traten die Augsburger Firmen zudem in enge finanzielle Beziehungen mit dem dänischen König.¹⁵

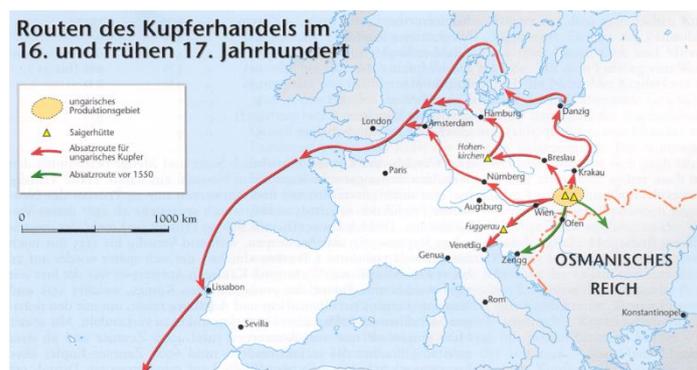


Abb. 2: Routen des Kupferhandels im 16. und frühen 17. Jahrhundert. Quelle: <https://www.hdbg.eu/karten/karten/detail/id/51>.

4.2. Kupfertransporte in den STT

In der Zeit von 1497 bis 1569 sind Kupferexporte in den STT sowohl in Band 1, Tabellen über die Schifffahrt 1497–1660, als auch in Band 2A, Tabellen über die Warentransporten 1497–1660, auffindbar. In Band 1 wurden im Prinzip nur die Durchfahrten durch den Øresund quantifiziert. Dennoch erhalten die frühesten tabellarischen Übersichten auch Angaben über die transportierten Mengen an Rheinischem Wein und Kupfer. Band 2A hingegen erhält eine Quantifizierung der Warentransporte erst ab 1557. Grund dafür ist die allmähliche Herausbildung einer festen Struktur der Schiffs- und Warenregistrierungen zwischen 1497 und 1569, die anschließend bis 1632 beibehalten wurde.¹⁶ Weil erst seit 1557 mit der systematischen Erfassung von Waren in den Registern experimentiert wurde, bildet dieses Jahr den Startpunkt für Band 2A. Im Einklang mit dem selbsterklärten Ziel, die Tabellen „[...] aussi simples et uniformes que possible [...]“¹⁷ zu gestalten, vermitteln die STT hingegen auch für die frühesten Register ein hohes Maß an Homogenität. Wichtig ist zudem, dass die Warentransporte für jedes zehnte Jahr (1565 bis 1635, 1646, 1655) im Detail aufbereitet wurden, während für die restlichen

¹² Mark Häberlein, Kaufleute als kulturelle Vermittler im Ostseeraum (Spätmittelalter und Frühe Neuzeit), in: Agnese Dubova / Ineta Balode / Konrad Schröder (Hrsg.), Sprach- und Kulturkontakte im Ostseeraum, Bamberg 2022, S. 21–51, hier S. 37. DOI: 10.20378/irb-56351

¹³ Häberlein, Kaufleute, S. 39.

¹⁴ Attila Tózsá-Rigó, Infrastruktur und Logistik der Bergbauproduktion in Neusohl in der zweiten Hälfte des 16. Jahrhunderts, in: Michael Farrenkopf / Stefan Siemer (Hrsg.), Materielle Kulturen des Bergbaus / Material Cultures of Mining: Zugänge, Aspekte und Beispiele / Approaches, Aspects and Examples, Berlin 2022, S. 395–408 hier S. 401–402.

¹⁵ Häberlein, Kaufleute, S. 37.

¹⁶ Bang / Korst, Tabeller, 1, S. IV–VI.

¹⁷ Bang / Korst, Tabeller, 2A, S. IV.

Jahre nur die wichtigsten Waren erfasst sind.¹⁸ Unter diesen Bedingungen ist Kupfer im Zeitraum von 1497 bis 1569 meistens ausschließlich in der Form von Anmerkungen in den STT auffindbar. Nur in den Tabellen über die Warentransporte im Jahr 1565 werden die Kupferexporte aufgeschlüsselt und tabellarisch dargestellt (siehe Tabelle 1).

Jahr	Angaben über Kupfer in den Sundzolltabellen, 1497-1569
1497, 1503, 1528, 1536	k.A.
1537	"(...) af Kobber [fortoldedes] 464 Læst (for Fuggerne)"
1538	"Intet Kobber"
1539	"(...) af Kobber [fortoldedes] 312 ½ Læst 309 for Fuggerne"
1540	"(...) af Kobber [fortoldedes] 609 ¾ Læst 608 ¾ for Fuggerne"
1541	"(...) af Kobber [fortoldedes] 608 ¼ Læst 604 ¼ for Fuggerne"
1542, 1543	k.A.
1544	"(...) af Kobber [fortoldedes] 884 Læster for Fuggerne"
1545	"(...) af Kobber [fortoldedes] 486 ½ Læst 468 for Fuggerne"
1546	"(...) af Kobber [fortoldedes] 456 ½ Læst for Fuggerne"
1547	"(...) af Kobber [fortoldedes] 102 Læster (53 for Fuggerne)"
1548	k.A.
1554–1557	"(...) af Kobber 36 ½ Læst. 1554 · 22 L., 1555 · 476 L., 1556 · 524 L."
1558–1562	k.A.
1563	"(...) af Kobber [fortoldedes] 79 Læster"
1564	"(...) af Kobber [fortoldedes] 119 ½ Læster"
1565	Vol. 1: (...) af Kobber [fortoldedes] 97 Læst" Vol. 2A: "Foruden de under Afsnittet: Kobbertold opgivne 97 L. (som Hieronymus Schillinck af Danzig har udsendt paa Ulrich Lyncken af Augsburgs vegne)" findes spredt i Regnskabet 25 L. Kobber." Weitere Angaben befinden sich in Vol. 2A, Tabel 2 und Tabel 4a.
1566	"(...) af Kobber [fortoldedes] 133 Læster"
1567	"(...) af Kobber [fortoldedes] 282 Læster"
1568, 1569	k.A.

Tabelle 1: Kupfer in den STT, 1497–1569. Für die Angaben über die Jahre 1537–1541; 1544–1547; 1554–1557; und 1563–1567, siehe Vol. 1, Text vor Anmerkung A1 zur Tabelle „Skibe passeret Øresund i Aaret“ (Bang & Korst 1906: 6–16, 20–36, 40, 44, 48, 52). Für die Angaben über das Jahr 1565, siehe Vol. 2A, Tabel 2: Varer ført vestpaa fordelt efter Skibenes Hjemsted; J₁: Kobber, Læst; Anmerkung zu J₁: 1–6 und Tabel 4a: Varer ført vestpaa fordelt efter Skibenes Afgangshavn; Kobber, Læst (Bang & Korst 1922: 18, 21).

4.3. Kupfertransporte in STRO

Was ist nun die Genese dieser Daten – woher kommen sie und wie wurden sie verarbeitet? Um eine Antwort auf diese Fragen zu finden, wurde die Datenbank *Sound Toll Registers Online* (www.soundtoll.nl) konsultiert, die Einblick gibt in alle Schiffs- und Warenregistrierungen in den Registern der *Øresundregnskaberne*, also in die Buchführung des Øresundes. Eine möglichst umfassende Suche wurde gewährleistet, indem die Inhalte von STRO für den Zeitraum von 1497 bis 1569 vollständig als CSV-Dateien heruntergeladen wurden. Zusätzliche, nur auf der Webseite verfügbare Informationen über die Überschriften der erhaltenen Register und Zoll-Listen wurden gesondert heruntergeladen. Nach Beseitigung von *parsing errors* in den CSV-Dateien, wurde mit *regular expressions* nach Angaben über Kupfer in den Spalten ‚soort‘ (Ware), ‚opmerking_bron‘ (Anmerkungen in der Quelle), ‚naam‘ (Name des Zolles), ‚opmerkingen‘ (Anmerkungen zu den Zöllen)

¹⁸ Bang / Korst, Tabeller, 2A, S. IV.

und ‚section‘ (Listenüberschrift) gesucht.¹⁹ Das Ergebnis umfasst eine Datei mit Registrierungen, in der Kupfer erwähnt wird, und eine Liste von Bilddateien, auf denen die Registrierungen enthalten sind. Mithilfe dieser Dateien können wir nun erste Beobachtungen über die Datengenese der STT und die dafür eingesetzten Methoden formulieren.

4.4. Erste Beobachtungen

1. Der fiskalische Charakter der dänischen Sundzollregister, der unter anderem in parallel geführten Listen zu verschiedenen Zöllen zum Ausdruck kommt, ist während der Quantifizierung der Daten über Schifffahrt und Warenverkehr durch den Øresund weitgehend verloren gegangen. Als die Beamten im Øresund im Jahr 1557 anfangen, auch die transportierten Güter systematisch zu erfassen, führte dies zu einer Vielzahl von Teil-Listen und Mehrfachregistrierungen einzelner Schiffer und Ladungen. Die Aufteilung der Register in verschiedene Listen wird zwar in der Einleitung zu Band 1 diskutiert, aber die eigentliche tabellarische Darstellung berücksichtigt sie nur geringfügig (wenn überhaupt). In den STT wurden die Teil-Listen ohne Ausnahme in (fast) einheitliche Tabellen überführt. Dabei sind die in Listenüber- und -unterschriften enthaltenen semantischen Informationen verloren gegangen.

Für die Jahre 1537, 1539, 1540, 1541, 1544, 1545, 1546 und 1547 hingegen finden wir in den STT explizite Angaben zum Anteil der Augsburger Kaufmannsfamilie Fugger an den Kupfertransporten durch den Sund (Dk. *for Fuggerne*) (siehe Tabelle 1). Diese Angaben stammen aus entsprechende Notizen in den STR. Nach dem Ausscheiden der Fugger übernahmen andere Augsburger Kaufmannsfamilien die Kontrolle über die Neusohler Kupfergewinnung. Angaben zu diesen Kaufmannsfamilien in den STR wurden aber nicht mehr systematisch in die STT übernommen (siehe Tabelle 1). So finden wir für die nicht erhaltenen Register über die Jahre 1554, 1555 und 1556 zwar die Gesamtmengen des verschifften Kupfers in einer Anmerkung zur Tabelle über das Jahr 1557²⁰, aber über den Empfänger dieser Kupferladungen, den Augsburger Kaufmann Melchior Manlich, erfahren wir in den STT nichts (siehe Tabelle 1). In den detaillierten Angaben über das Jahr 1565 hingegen wurde präzisiert, dass außer den 97 L[ast], die unter der Rubrik Kupferzoll angegeben sind (und die Hieronymus Schillinck im Namen von Ulrich Lyncken von Danzig nach Augsburg geschickt hat), verstreut in den Rechnungen weitere 25 L[ast] Kupfer verzeichnet wurden (siehe Tabelle 1).²¹ Ausnahmsweise wird in dieser Anmerkung die Bezeichnung *Kaabertold* (Kobbertold; *Ger.* Kupferzoll) verwendet, die regelmäßig als Listenüberschrift in den Registern auftaucht.

2. Die Registrierung von Schiffen in verschiedenen parallel geführten Listen hatte Mehrfachregistrierungen zur Folge, die während der Quantifizierung stillschweigend aufgelöst wurden. Die Auflösung umfasste sowohl die Vereinheitlichung von Ortsnamen, Produktbezeichnungen, Maßen und Gewichten als auch das Zusammenführen von zerstreuten Angaben über einzelne Schiffer. Anhand der Kupferexporte des Jahres 1565 (siehe Tabelle 2) kann dieses Vorgehen zumindest teilweise rekonstruiert werden.

¹⁹ Die niederländischen Spaltenüberschriften in den CSV-Dateien wurden beibehalten. Bislang wurden folgende *regular expressions* verwendet, um möglichst viele variante Schreibweisen vom dänischen Wort ‚Kobber‘ (Kupfer) zu finden: `([C|c|K|k][o|a]+[b|p]+er)$` für die Ware; `([C|c|K|k][o|a]+[b|p]+er?)` in den Anmerkungen und Listenüberschriften; und `([C|c|K|k][o|a]+[b|p]+er)` für den Namen des Zolles.

²⁰ Bang / Korst, Tabeller, 1, S. 20.

²¹ Bang / Korst, Tabeller, 2A, S. 18, 21.

Kat.	Hjemsted	Læst
	<i>Nederlandene:</i>	
1	Holland	59,5
	<i>Amsterdam</i>	29
	<i>Hoorn</i>	5
3	Friesland	57
	<i>Harlingen</i>	3
	<i>Stavoren</i>	35,5
4	<i>Andre Provinser</i>	1,5
		118
	<i>Nordvest-Tyskland:</i>	
6	Bremen	4
27	<i>Tillsammen</i>	122

Tabelle 2: Kupferexporte 1565 nach Heimatshafen (Dk. Hjemsted) der Transporteure, in Last – Ausschnitt aus Tabel 2: Varer ført vestpaa fordelt efter Skibenes Hjemsted; J₁: Kobber, Læst. Quelle: Bang & Korst 1922, S. 18.

Die Angaben in der Tabelle 2 entsprechen lediglich einer Spalte in der tabellarischen Aufbereitung der Warentransporte durch den Øresund im Jahr 1565 – ein winzig kleiner Teil der STT. Die Grundlage für diese Angaben bilden insgesamt drei Zoll-Listen. Die erste Liste erfasst den Kupferzoll (*Dk.* Kaaber Told). Diese Liste umfasst nicht mehr als drei Folio-Seiten. Die anderen Listen sind viel umfangreicher und erfassen einerseits niederländische Schiffe und deren Waren (*Dk.* Skib og Gods, *Nederlander*), und andererseits Angaben über die Berechnungsgrundlage und Bezahlung des Feuergeldes (*Dk.* Fyrpenge). Schiffer, die zu Gunsten von Ulrich Lyncken Kupfer durch den Øresund transportierten, wurden in der Kupferzoll-Liste registriert. Dies betrifft 30 Schiffer, die insgesamt 97 Last Kupfer geladen hatten. In der Regel bezahlten diese Schiffer auch das Feuergeld und sind entsprechend auch in dieser Zoll-Liste auffindbar. Auch andere niederländische Schiffer transportierten Kupfer, ohne an den Geschäften des Ulrich Lyncken beteiligt zu sein. Diese Schiffer wurden in der „Skib og Gods“-Liste und in der „Fyrpenge“-Liste erfasst. Dies betrifft 11 Schiffer, die insgesamt 16,5 Last Kupfer transportierten. Schließlich gibt es noch drei Schiffer, die nur in der „Fyrpenge“-Liste auffindbar waren und weitere 8,5 Last Kupfer transportierten. Auffällig ist, dass die in den STT gemachten Angaben über die verschifften Mengen – 97 Last in der Kupferzoll-Liste und noch 25 Last verstreut in den Registern – mithilfe von Abfragen in STRO und einer weiterführenden interaktiven Suche auf der Webseite vollständig rekonstruiert werden konnten. Das weist darauf hin, dass Nina Ellinger Bang und ihre Mitarbeiter:innen ein System entwickelt hatten, dass das Zusammenfügen von Mehrfachregistrierungen auf der Grundlage von Personennamen ermöglichte. Wie dieses System ausgesehen haben könnte, bedarf der weiteren Erforschung.

Auch die Verteilung der Kupfermengen nach Heimatort (siehe Tabelle 2) ist problemlos nachvollziehbar und zeigt, dass stillschweigend eine Vereinheitlichung der Ortsnamen vorgenommen wurde. Vermutlich haben die Mitarbeiter:innen von Nina Ellinger Bang ebenfalls stillschweigend eine *Art nominal record linkage* eingesetzt, um aus den verstreuten Mehrfachregistrierungen eine einheitliche Liste zu erstellen. Wahrscheinlich ist, dass die Listenüberschriften bei der Identifikation von Mehrfachregistrierungen eine große Rolle spielten. Möglicherweise setzten die Bearbeiter:innen der STR für die Datenerfassung *edge-notched cards* ein, die es seit den 1890er-Jahren ermöglichten, Lochkarten in einer ‚papieren‘ Datenbank zu sortieren und auszuwählen.

5. Diskussion

Die genetische Datenkritik (GDK) ist aus zweierlei Hinsicht nicht einfach umzusetzen. Es fehlt an Orientierung, sowohl im Sinne eines Leitfadens für Untersuchungen zur genetischen Datenkritik als auch in Bezug auf eine Methodik für die systematische Erfassung von Arbeitsschritten, die auf das Offenlegen der Datengeneese zielen. Während für einen Leitfaden grundsätzlich die notwendigen Best

Practice Beispiele fehlen, könnten *paradata* zumindest für die systematische Erfassung der datengenetischen Prozesse, die zu jeder historischen Datensammlung gehören, Auskunft geben. *Paradata* ist ein bekanntes Konzept in der Informationswissenschaft.²² Es dokumentiert „wie Daten erfasst oder erstellt wurden, unter welchen Bedingungen, mit welchen Instrumenten, wer was wann getan hat, welche Qualifikationen und Erfahrungen die Personen hatten und wer beispielsweise Entscheidungen unter welchen Voraussetzungen getroffen hat“²³. Im Zusammenhang mit der Frage der Reproduzierbarkeit von Forschungsergebnissen erregt der Begriff seit kurzem auch Aufmerksamkeit in den Digitalen Geisteswissenschaften.

Im Kontext der GDK haben *paradata* eine doppelte Bedeutung. Einerseits erfassen sie Informationen über die von Forschern in der Vergangenheit angewandten Methoden, d. h. Informationen über den Umgang früherer Generationen von Historiker:innen mit ‚rohen‘ Daten, mit denen sie gearbeitet haben. Andererseits erfassen *paradata* auch Informationen über die von digital arbeitenden Historiker:innen in der Gegenwart angewandten Methoden der Datenverarbeitung und -auswertung. Die Besonderheit der GDK gedruckter historischer Datensammlungen besteht nun darin, dass die *paradata* über den Umgang früherer Generationen mit ‚rohen‘ Daten auf einen Rekonstruktionsakt beruhen, der diesen Umgang erst offenlegt und nachvollziehbar macht. Anders gesagt: Ohne GDK sind genaue Aussagen über historische Methoden der Quantifizierung und Datenmodellierung kaum möglich. Die GDK erlaubt zumindest einen mittelbaren Einblick in die Methoden, die für die Erschließung von historischen Datensammlungen in der vordigitalen Zeit eingesetzt wurden. Während die Notwendigkeit und Nützlichkeit von *paradata* unbestritten ist und bereits verschiedene technische Lösungen für deren systematische Erfassung vorgeschlagen wurden²⁴, wurde die Frage, wie man das Verhältnis zwischen Datenlebenszyklen, Datengenese und Datenmodellierung in der historischen Forschung konfigurieren könnte, bislang noch nicht hinreichend beantwortet.

6. Ausblick

Dieser ‚Werkstattbericht‘ sollte zeigen, wie eine genetische Datenkritik über Praktiken im Umgang mit historischen Daten informieren und so zu neuen Erkenntnissen über die Entwicklung und Anwendung von quantifizierenden Verfahren in der historischen Forschung beitragen kann. Zum einen ging der Beitrag auf Entscheidungen über das Nicht-Erfassen von bestimmten Daten ein und erläuterte die Auswirkungen auf die Semantik der Zeitreihen in den STT. Zum anderen warfen die ersten Erkenntnisse der genetischen Datenkritik der STT neue Fragen auf über die Charakteristika und Methoden der datenorientierten historischen Forschungspraxis in der Zeit vor der Entwicklung des Computers.

Vor uns liegt nun eine große Aufgabe. Für die Weiterbearbeitung der Sundzolltabellen als repräsentative historische Datensammlung aus der ersten Hälfte des 20. Jahrhunderts wäre es erstens unabdingbar, das Datenmodell für die Register aus der Zeit bis 1632 neu zu denken und eine nachhaltige Lösung für den Umgang mit Mehrfachregistrierungen in verschiedenen Zoll-Listen zu finden. Für die perspektivische Automatisierung von Teilen der GDK wäre dieser Schritt sehr wichtig. Daran anknüpfend wäre, zweitens, zu überlegen, ob eine Volltexterschließung der Quelle notwendig sein wird, um Einsicht zu gewinnen in die Inhalte der vielen Scans, die nicht in STRO eingeflossen sind, weil sie keine Schiffs- oder Warenregistrierungen in Listenform enthalten. Auf diesen Scans befindet sich eine Menge noch nicht gesichteten Materials. Und schließlich sollte eine Methode entwickelt werden, welche die aus der Wechselwirkung von STT, STRO und Scans ersichtlichen Prozesse der Datengenese

²² Anna Bentkowska-Kafel / Hugh Denard / Drew Baker, *Paradata and transparency in virtual heritage*, Ashgate 2012; Isto Huvila, *Improving the usefulness of research data with better paradata*, in: *Open Information Science* 6 (2022), S. 28–48. <https://doi.org/10.1515/opis-2022-0129>

²³ Huvila, *Improving*, S. 31.

²⁴ Isto Huvila / Olle Sköld / Lisa Börjesson, *Documenting information making in archaeological field reports*, in: *Journal of Documentation* 77 (2021), S. 1107–1127. <https://doi.org/10.1108/JD-11-2020-0188>; L. Börjesson / O. Sköld / I. Huvila, *Paradata in Documentation Standards and Recommendations for Digital Archaeological Visualisations*, in: *Digital Culture & Society* 6 (2020), S. 191–220. <https://doi.org/10.14361/dcs-2020-0210>

systematisch erfasst. Hier könnte man sinnvoll anknüpfen bei Überlegungen von Fachkolleg:innen über *paradata* und aus einer historischen Perspektive zu deren Weiterentwicklung beitragen. Dies sind keine geringen Aufgaben, aber sie haben das Potenzial, aus den Sundzolltabellen mit ihrer vielfältigen Geschichte und ihrem herausragenden genetischen Dossier ein Paradebeispiel für die Erforschung der Geschichte der Quantifizierung und Datenmodellierung in der vordigitalen Zeit zu machen.

7. Literaturliste

- Bang, Nina Ellinger / Korst, Knud (Hrsg.), *Tabeller over skibsfart og Varetransport gennem Øresund / Tables de la navigation et du transport des marchandises passant par le Sund 1497–1783*. København 1906–1953.
- Bentkowska-Kafel, Anna / Denard, Hugh / Baker, Drew, *Paradata and transparency in virtual heritage*, Ashgate 2012.
- Börjesson, Lisa / Sköld, Olle / Huvila, Isto, *Paradata in Documentation Standards and Recommendations for Digital Archaeological Visualisations*, in: *Digital Culture & Society* 6, 2020, S. 191–220. <https://doi.org/10.14361/dcs-2020-0210>.
- Degn, Ole (Hrsg.), *The Sound Toll at Elsinore. Politics, shipping and the collection of duties*, Copenhagen 2018.
- Ginzburg, Carlo, *Our Words, and Theirs. A Reflection on the Historian's Craft, Today*, in: *Cromohs (Cyber Review of Modern Historiography)* 18 (2013), S. 97–114.
- Gøbel, Erik, *The Sound Toll Registers Online project, 1497–1857*, in: *International Journal of Maritime History* 22 (2010), S. 305–324.
- Häberlein, Mark, *Kaufleute als kulturelle Vermittler im Ostseeraum (Spätmittelalter und Frühe Neuzeit)*, in: *Agnese Dubova / Ineta Balode / Konrad Schröder (Hrsg.), Sprach- und Kulturkontakte im Ostseeraum*, Bamberg 2022, S. 21–51 dort S. 37. DOI: 10.20378/irb-56351.
- Huvila, Isto / Sköld, Olle / Börjesson, Lisa, *Documenting information making in archaeological field reports*, in: *Journal of Documentation* 77, 2021, S. 1107–1127. <https://doi.org/10.1108/JD-11-2020-0188>.
- Huvila, Isto, *Improving the usefulness of research data with better paradata*, in: *Open Information Science* 6, 2022, S. 28–48. <https://doi.org/10.1515/opis-2022-0129>.
- Jeannin, Pierre, *Les comptes du Sund comme source pour la construction de l'activité économique en Europe (XVIe–XVIIIe siècle)*, in: *Revue historique* 121 (1964), S. 55–102; 307–340.
- Kelly, Julia A. u.a., *A Critical Literature Review of Historic Scientific Analog Data: Uses, Successes, and Challenges*, in: *Data Science Journal* 21 (2022), S. 1–11. <https://doi.org/10.5334/dsj-2022-014>.
- MacFarland, Kathryn / Vokes, Arthur W., *Dusting off the data: curating and rehabilitating archaeological legacy and orphaned collections*, in: *Advances in Archaeological Practice* 4 (2016), S. 161–175. 10.7183/2326-3768.4.2.161.
- Mitchell, Brian R., *British historical statistics*, Cambridge 1990.
- Mitchell, Brian R., *European historical statistics 1750–1970*, New York 1975.
- Scheltjens, Werner, *Upcycling historical data collections. A paradigm for digital history?*, in: *Journal of Documentation*, Vol. ahead-of-print (2023). <https://doi.org/10.1108/JD-12-2022-0271>.
- Schöch, Cristoff u.a., *Replication and Computational Literary Studies*, in: *Digital Humanities Conference 2020 (DH2020)*, Ottawa 2020. <https://doi.org/10.5281/zenodo.3893428>.

- Schubert, Charlotte, Digital Humanities auf dem Weg zu einer Wissenschaftsmethodik: Transparenz und Fehlerkultur, in: Digital Classics Online 7 (2021), S. 39–53. <https://doi.org/10.11588/dco.2021.7.82371>.
- Shiue, Hilary Szu Yin u.a., Assessing Legacy Collections for Scientific Data Rescue, in: Katharina Toepe / Hui Yan / Samuel Kai Wah Chu (Hrsg.), Diversity, Divergence, Dialogue: 16th International Conference, iConference 2021, Beijing, China, March 17–31 2021, Proceedings, Part II, Cham 2021, S. 308–318. https://link.springer.com/chapter/10.1007/978-3-030-71305-8_25.
- Tózsza-Rigó, Attila, Infrastruktur und Logistik der Bergbauproduktion in Neusohl in der zweiten Hälfte des 16. Jahrhunderts, in: Michael Farrenkopf / Stefan Siemer (Hrsg.), Materielle Kulturen des Bergbaus / Material Cultures of Mining: Zugänge, Aspekte und Beispiele / Approaches, Aspects and Examples, Berlin / Boston 2022, S. 395–408 dort S. 401–402.
- Van Hulle, Dirk, Genetic Criticism. Tracing Creativity in Literature, Oxford 2022. <https://doi.org/10.1093/oso/9780192846792.001.0001>
- Veluwenkamp, Jan Willem / Scheltjens, Werner / Van der Woude, Siem, Sound Toll Registers Online, in: TSEG-The Low Countries Journal of Social and Economic History 18 (2021), S. 147–160. <https://doi.org/10.18352/tseg.1203>.
- Withhöft, Harald (Hrsg.), Handbuch der historischen Metrologie (8 vols.), St. Katharinen 1991–2007.