Session Proposal - Baltic Connections Conference
University of Jyväskylä, 16-17 June 2022

Session organisers
Werner Scheltjens, University of Bamberg
Jeremy Land, University of Helsinki

Session title
Digital methods for the advancement of maritime history

Session description
Maritime history is a relatively small, but vibrant subdiscipline of the historical sciences with a long tradition of exploring and applying novel research methods to a wide variety of sources. Since the breakthrough of Web 2.0 at the turn of the twenty-first century, several international projects have been carried out, which used the web as a means to edit and publish historical source materials in novel ways. Moreover, the community of maritime historians has made considerable efforts to make sources of maritime history easier to access and explore on the web.

The proposed session brings together historians and digital humanists, who apply cutting-edge digital methods to their research and thus contribute to the advancement of the discipline of maritime history. The papers in the session reflect both the wide variety of maritime-historical sources as well as the different kinds of digital approaches that are implemented to deal with these sources. For example, the session includes papers that deal with customs registers (Veluwenkamp; Marzagalli & Plumejeaud-Perreau), but also with ship logbooks, crew lists, account books and business correspondence (Delis, Fafalios & Doerr; Oakley) or sources about ships' insurance (Iodice). The range of digital approaches included in this session comprises the development of relational databases as well as the elaboration of domain ontologies, the use of web technology to facilitate volunteer projects or the construction of a web interface for data exploration and analysis.

The goal of our session is to critically appraise digital research methods applied to maritime-historical resources, and to address potential synergies between various approaches. Thus, the session aims to make a contribution to the discipline of maritime history as a field of inherently interdisciplinary research, pointing out the strengths and weaknesses of selected digital research methods and providing directions for future research. At the same time, the results of the session also provide insights about the use of digital methods that may be useful to other subdisciplines of the historical social sciences as well.

Papers

1. Working with Volunteers. Sound Toll Registers Online 2013-2020
Jan Willem Veluwenkamp, University of Groningen

In 2020 the University of Groningen and Tresoar, the Frisian Historical and Literary Centre in Leeuwarden, The Netherlands, have completed Sound Toll Registers Online (STRO), the online database providing instant and free access to the complete Sound Toll Registers (STR) at www.soundtoll.nl. All together STRO contains the data of 2.15 million entries for 1.8 million Sound passages from the period 1497-1857. Breed, the sheltered workplace for the city of Nijmegen, entered data from 2009 to 2013, using digitized microfilms of the original

source. In 2013, volunteers took over. They entered the data of the remaining 651 thousand STR entries and completed this work in 2020. This contribution will show how we organized and managed the work of the STRO volunteers, how we taught them the old Danish handwriting, the ever changing source format and the data entry, and how we kept them involved and motivated.

2. Visualizing 18th century maritime trade in the Portic experience: How to match uncertain historical maritime data and some data standardization "requirements" in Digital Humanities
Silvia Marzagalli, Université Côte d'Azur, Centre de la Méditerranée moderne et contemporaine / Christine Plumejeaud-Perreau, UMR 7301 Migrinter, Poitiers

PORTIC (PORTs & Information and Communication Sciences and Technology. Querying and visualizing eighteenth-century shipping and trade dynamics in the digital era, https://anr.portic.fr/en/) is an ongoing interdisciplinary project financed by the French National Research Agency (2019-2023). Portic aims at producing new knowledge on shipping and maritime trade in eighteenth-century France through a collaboration between digital humanities experts, historians and economists, and at providing user-friendly and accurate visualization tools to explore collected data online. The paper presents the current achievements, the considerable benefit of interdisciplinary collaboration for data curation and exploratory visualizations, but also the relentless renegotiation of disciplinary boundaries to make collaboration fruitful, notably concerning the formalisation of uncertain data.

3. Digital Seafaring: Digitising, Curating and Exploring Archival Sources of Maritime History
Apostolos Delis, Pavlos Fafalios, Martin Doerr, ICS-FORTH

The European funded project SeaLiT (ERC) has developed a set of innovative tools for supporting maritime historians in digitising, curating and exploring archival sources of maritime history. The tools are the result of the interdisciplinary work between maritime historians of the Institute of Mediterranean Studies of FORTH and researchers and data engineers of the Centre for Cultural Informatics of the Institute of Computer Science of FORTH. In this paper, we present the methodology we followed and the challenges we faced for supporting the historians of SeaLiT in digitizing and exploring their unique historical sources. These information sources range from ship logbooks, crew lists, payrolls and student registers, to civil registers, business records, account books and consulate reports, originating from different authorities in countries of the Mediterranean and the Black Sea, covering the period 1850s-1920s. We describe, in particular, the process of building a rich semantic network of integrated maritime data, which includes i) developing tools that allow historians to digitize and curate their archival sources collaboratively, ii) modeling and integrating the data using established documentation models (in particular, the ISO standard CIDOC-CRM / ISO 21127:2014) for supporting interoperability and long-term validity, and iii) offering data exploration and visualization tools that enable historians take advantage of the digitization and integration of the original information sources.

4. Triangulating on trade. Digital methods and the improvement of sources in maritime history
Eric Oakley, Kennesaw State University

The most common sources for maritime history are items such as ship logs, voyage narratives, account books, and arrival/departure records, but all of these sources contain temporal and spatial limitations. Throughout most of history, the physical nature of archival

materials limited our ability to place them in routine conversation. In contrast, modern digital methods (unimaginable during the age of sail!) enable historians to improve upon the information value of individual sources by "triangulating" between items. This presentation will use sources in Pacific World history to demonstrate how a robust database can be compiled, with particular interest upon how the interconnectedness of sources enables us to draw larger conclusions about maritime history.

5. Perils of the Sea, the AveTransRisk Database: Unforeseen Expenses and Transaction Costs in Early Modern European Shipping (1500-1800)

Antonio Iodice, University of Exeter (UK) / University of Genoa (Italy)

AveTransRisk (Average-Transaction Costs and Risk Management during the First Globalization, Sixteenth-Eighteenth Centuries) is an ongoing ERC project coordinated by Maria Fusaro. The project started in July 2017 and will end in December 2022. The resulting online database *AveTransRisk* is one of the project's principal outputs. It contains details on thousands of specific shipping voyages that were involved in a maritime average, a legal instrument, which underpins maritime trade by redistributing damages' costs across all interested parties, still applied nowadays. This contribution will show how we designed the database, the methodological challenges derived by working on different archives and different periods, and the possible use of these resources for digital interdisciplinary research in the historical social sciences.

6. Creating Datasets from Disparate Digital Archives: 18th Century Colonial American Merchant Networks

Jeremy Land, University of Helsinki / Werner Scheltjens, University of Bamberg

Colonial American history has earned, in general, a large amount of interest by social scientists of all kinds. The success of the United States in both the economic and political spheres has sparked many archives to make more and more of their resources available digitally, via scanning and organizing documents into freely or cheaply available online collections. Creating a dataset from disparate sources causes pragmatic concerns about what sorts of data to pull from the manuscripts themselves. Generating dataCreating a dataset that is *available* to everyone is easy. Creating a dataset that is **useful** to everyone is difficult. This paper explores the questions and methodologies that one should consider in creating datasets for the early modern period. The first section explores the problems and presents questions that arise when one attempts to create digitally available datasets from sources that transcend a single archive collection. In it the paper breaks down the "datafication" process into its most basic components, down the column and row headers in Excel files. While keeping an eye on the pragmatic concerns, the section also asks very broad questions about the sources themselves. Questions such as: what can be quantified? What information is useful to the overall project? What other types of data can be useful? Who can or might use this data (i.e. sociologists, economists, etc.), and for what purposes and/or research questions for which this data can be utilized? The second section subjects these methodologies and questions to a pilot study, wherein merchant bills of lading from 18th century colonial America are collected, analyzed, and placed into a working dataset from various digital archive collections across the United States. This dataset is built to include as much information as possible that can be used for as many purposes as possible, making sure to keep the data linked, digitally, to the original archival location to ensure total access by the researcher to confirm or dig further into the source's data. Unfortunately, each archive organizes their documents into collections in very different ways and pulling data from these sources presents difficulties that change with every single archive. The third

section addresses these issues, and thus contributes to the goal to create a framework by which datasets can be created that provides the widest possible use for scholars from any discipline.