

Giovanni Colavizza

University of Amsterdam
BG 1, Turfdraagsterpad 9, 1012 XT Amsterdam

☎ +31 6292 26788

✉ g.colavizza@uva.nl – giovanni.colavizza@odoma.ch

🌐 www.gioannicolavizza.com

Research interests.....

Science and Technology Studies; Digital Arts and Humanities; Data Science; Blockchain.

Current positions

Academic.....

University of Amsterdam <i>Assistant Professor of Digital Humanities</i> Department of Media Studies and Institute for Logic Language and Computation.	Amsterdam <i>Since 2019</i>
---	---------------------------------------

The Alan Turing Institute <i>Visiting Fellow</i>	London <i>Since 2019</i>
--	------------------------------------

Centre for Science and Technology Studies (CWTS) <i>Permanent Visiting Fellow</i>	Leiden <i>Since 2017</i>
---	------------------------------------

Industry.....

Nifty-art.com <i>Co-Founder, www.nifty-art.com</i> Nifty-art.com is the first market intelligence and discovery platform for digital and blockchain art.	To be incorporated <i>Since 2020</i>
--	--

Odoma Sàrl <i>Co-Founder and CTO, www.odoma.ch</i> Blending AI technology and UI/UX design to bring new life and value from heritage collections and patrimonies. Our clients include Longines, the EPFL and the Olympic Studies Centre.	Lausanne <i>Since 2016</i>
---	--------------------------------------

Education

École Polytechnique Fédérale de Lausanne (EPFL) <i>PhD in Management of Technology</i>	Lausanne <i>2014 – 2018</i>
--	---------------------------------------

Universities of Milan, Padua, Udine and Venice <i>BSc in Computer Science; BA and MA in History</i>	Italy <i>2004 – 2014</i>
---	------------------------------------

Public profiles and spoken languages.....

Twitter: @giovanni1085 — LinkedIn: giovanni-colavizza-850b3342
ORCID: 0000-0002-9806-084X — Google Scholar: LQEgr6UAAAAJ

I am fluent in English, French and Italian.

Previous positions

The Alan Turing Institute

London

Senior Research Data Scientist

Jan 2018 – June 2019

I was part of the Research Engineering Group, Co-investigator on the "Living with Machines" project (9.2M GBP, 5 years) and research engineering lead for the "Data Science for Science" program.

École Polytechnique Fédérale de Lausanne (EPFL)

Lausanne

Doctoral Assistant

Jul 2014 – Dec 2017

Doctoral program in Management of Technology. Unit: Digital Humanities Laboratory.

University of Oxford

Oxford

Research Assistant (SNF doc.Mobility)

March – Aug 2017

Advisor: *Dr. Cristina Dondi*. I focused on the analysis of citation networks of humanities publications, to understand how scholars rely on shared research technologies, such as reference works and catalogues.

Centre for Science and Technology Studies (CWTS)

Leiden

Research Assistant (SNF doc.Mobility)

Sept 2016 – Feb 2017

Advisors: *Dr.s Ludo Waltman and Nees J. van Eck*. I worked on mapping the intellectual and social structures of humanities scholars, and on the large-scale text analysis of scientific publications.

École Polytechnique Fédérale de Lausanne (EPFL)

Lausanne

Operations Manager of the Venice Time Machine

Jun 2014 – Aug 2016

I headed a team of six to digitize and index archival and library holdings, and was responsible for IT.

École Polytechnique Fédérale de Lausanne (EPFL)

Lausanne

Student Research Assistant

Jan – Jun 2014

Advisor: *Dr. Frédéric Kaplan*. I worked on information extraction from historical texts, textual reuse and adapting methods from bioinformatics to work with very noisy data.

Leibniz Institute of European History (IEG)

Mainz

Research Assistant

Aug 2013 – Jan 2014

Advisor: *Dr. Michael Piotrowski*. I worked on controlled vocabularies for humanities research, focusing on their semantic integration relying on backbone ontologies.

Aranea Srl

Venice

Co-Founder and CTO

Jul 2013 – Feb 2014

Aranea developed book discovery systems relying on their contents. I was in charge for technology and research. I sold shares after delivering a first successful seed funding round in order to start my PhD.

Self-employed

Milan/Venice

Software Developer and System Administrator

before 2013

Various positions and independent activity as Java, SQL, PHP and JavaScript developer, Linux system administrator and IT network administrator.

Teaching and mentoring

Example courses.....

- **AI for Society** (6EC 3rd-year BSc in Artificial Intelligence, University of Amsterdam): https://github.com/Giovanni1085/UvA_AIforSociety_2021.
- **Coding the Humanities** (6EC 2nd-year BA in Media and Information, University of Amsterdam): https://github.com/Giovanni1085/UvA_CDH_2020.
- **Text Mining** (6EC 3rd-year BSc in Computer Science, Amsterdam University College): https://github.com/Giovanni1085/AUC_TMCI_2019.
- **Applied Data Analysis** (1-week summer school, University of Oxford): <https://github.com/mromanello/ADA-DHOxSS>.

Coursework as lecturer.....

2020-2021, University of Amsterdam

Faculties of Social Sciences and Humanities: **Introduction to Digital Humanities; Coding the Humanities; Digital Humanities Lab; Cultural and Social Data Analysis**. From 30 to 70 students per course, both bachelor and master levels.

Faculty of Science: **Text Mining; AI for Society**. 30 students per course, bachelor level.

Supervised 8 students for theses and internships.

2019-2020, University of Amsterdam

Faculties of Social Sciences and Humanities: **Archives, Information and Society; Coding the Humanities; Information Analytics and Digital Humanities; Media and Information; Media Theory**. From 30 to 70 students per course, both bachelor and master levels. Media and Information had, instead, 600+ bachelor students.

Faculty of Science: **Text Mining** (30 bachelor students).

Supervised 12 students for theses and internships.

2019 to 2021, University of Oxford

Applied Data Analysis: a 1-week course at the Oxford Digital Humanities Summer School.

2018-2019, The Alan Turing Institute

Research Software Engineering with Python for PhD students and researchers: alan-turing-institute.github.io/rsd-engineeringcourse.

Coursework as teaching assistant.....

2017-2018, École Polytechnique Fédérale de Lausanne (EPFL)

Head teaching assistant (TA) for **Applied Data Analysis** (350 master students: dlab.epfl.ch/teaching/fall2017/cs401) and **Foundations of Digital Humanities** (10 master students). **I was awarded among the best teaching assistants at the Computer Science School.**

2015-2016, École Polytechnique Fédérale de Lausanne (EPFL)

Teaching assistant (TA) for **Digital Humanities** (60 master students) and **Business Information Systems** (30 master students).

Supervised 12 students for individual projects.

2014-2015, École Polytechnique Fédérale de Lausanne (EPFL)

Teaching assistant (TA) for **Digital Humanities** (55 master students) and **Business Information Systems** (29 master students).

Supervised 7 students for individual projects.

2013-2014, École Polytechnique Fédérale de Lausanne (EPFL)

Teaching assistant (TA) for **Digital Humanities** (50 master students).
Supervised 3 students for individual projects.

Grants and research projects

Research projects.....

These projects are organised according to thematically-coherent lines of research, and have been or are funded either by a single or multiple grants.

Ongoing:

- **The impact of COVID-19 on research and researchers:** this project focuses on COVID-19 as a major crisis which impacting scientific research and researchers. My current focus is on the following three thematic areas:
 - How researchers are coping with COVID information overload, and the need to engage in societal debates on COVID measures.
 - How public knowledge infrastructure is behaving during the crisis (e.g., Wikipedia's reliability).
 - How (and if) societal debates online make use of scientific evidence (e.g., on Twitter).

An example of recent output: Roex, Karlijn and Giovanni Colavizza. 2020. *Doing Science in Times of COVID: A Survey*. <https://zenodo.org/record/4313278>. This dataset contains two rounds of surveys across Dutch-based researchers, on the impact of COVID-19 on their work and scientific information seeking/sharing behaviour.

Online presence: <https://covid19.humanities.uva.nl>.

- **Public understanding of science:** this project aims to understand how science is shared, debated and understood publicly, on online and social media, and what are the interdependencies of different media part of the online ecosystem in this respect. My current focus is on Wikipedia and Twitter.

An example of recent output: Gligorić, Kristina et al. 2020. *Experts and authorities receive disproportionate attention on Twitter during the COVID-19 crisis*. <https://arxiv.org/abs/2008.08364>. This paper, currently under review, uses the large-scale Twitter COVID stream and shows how experts, such as scientists and healthcare professionals, received much higher than usual attention on Twitter during the peak of the first wave (March, April 2020), but then this attention receded.

- **Digital methods: Teaching and training:** this project focuses on designing and developing pedagogical offerings on digital methods and computational thinking, with a special focus on students and professionals in the arts, humanities and social sciences. Furthermore, it aims to do the same for arts, humanities and social sciences pedagogical offerings to science students and professionals.

Online presence: <https://lnu.se/en/research/searchresearch/forskningsprojekt/project-digital-methods-platform-for-arts-and-humanities>.

- **Cultural Analytics and AI for Cultural Heritage:** the goal of this project is to develop fit-for-purpose methods to analyse cultural data, in particular texts. The current focus is a collaboration with the National Library of the Netherlands and the Rijksmuseum, to develop methods for scaling-up the indexation of large textual collections by combining human expertise with modern deep learning techniques.

An example of recent output: Todorov, Konstantin and Giovanni Colavizza. *Transfer Learning for Historical Corpora: An Assessment on Post-OCR Correction and Named Entity Recognition*. Computational Humanities Research Conference. <http://ceur-ws.org/Vol-2723/long32.pdf>. This paper assesses transfer learning techniques, which allow to greatly reduce the need for linguistic resources on common tasks such as named entity recognition, when applied to historical textual corpora.

- **Digital Blockchain Art:** blockchain art is digital art registered and exchanged on blockchains. This art market is currently growing rapidly, attracting the interest of researchers given its novelty and data availability.

An example of recent output: Franceschet, Massimo, Giovanni Colavizza, et al. *Crypto art: A decentralized view*. Leonardo. <https://arxiv.org/abs/1906.03263>.

Online presence: <http://www.nifty-art.com> (start-up currently seeking angel funding).

Previously:

- **Living with Machines:** A large-scale UKRI-funded collaboration between The Alan Turing Institute and the British Library (9.2M GBP, 5 years, 2019-2024). This project ambitions to study the social and cultural impact of the industrial revolution in England, via big historical data. I have co-written the project's bid and led the development of the research infrastructure during the early stages of this project.

An example of output: <https://arxiv.org/abs/2005.11140>.

Online presence: <https://livingwithmachines.ac.uk>.

- **Venice Time Machine:** The goal of the Venice Time Machine was to create a multimedia information system to study the history of Venice. I led the team in charge for the digitization of archival and library holdings during the first 2.5 years of the project.

An example of output: <https://arxiv.org/abs/1804.10371>.

Online presence: <https://www.epfl.ch/research/domains/venice-time-machine>.

- **Linked Books:** This project has created an experimental citation index for the historiography on Venice, starting from the digitization of library holdings to the release of a Web application to search the network of citations they contain. I wrote the FNS grant which funded the project, and led it to completion as part of my PhD.

An example of output: <https://venicescholar.dhlab.epfl.ch>.

Online presence: <https://scholarindex.eu>.

Grants.....

2020

- NWO (Dutch equivalent of the SNF) Industry collaboration (KIEM). Role: PI.
"From books to knowledge graphs"
Amount: 40k EUR + 20k matched by the industry partner Brill.
- Global Digital Cultures research grant (University of Amsterdam). Role: Co-I.
"Global digital art: Perspectives on categories and economic values in the crypto art market"
Amount: 30k EUR.
- Wikicite research grant. Role: Co-I.
"From Wikipedia Citations to Wikicite"
Amount: 10k EUR.

- NWO (Dutch equivalent of the SNF) Corona fast-track data,. Role: PI.
“Doing science in times of covid-19: A survey”
Amount: 50k EUR + 50k more matched by the University of Amsterdam.
- NWO (Dutch equivalent of the SNF) Idea generator. Role: PI.
“Improving access to historical collections via transfer learning”
Amount: 50k EUR.
- Erasmus EU, *Digital Methods Platform for the Arts and Humanities (DiMPAH)*. Role: Co-I.
“Digital methods platform for the arts and humanities”
Amount: 390k EUR (direct: approx. 60k).
- Furthermore, I sponsored a China Scholarship Council (CSC) for a 4-year PhD student.

2018-19

- UK Government Communications HQ and The Alan Turing Institute. Role: Co-I.
“Scaleable representation learning on graphs”
Amount: 80k GBP.
- National Library of the Netherlands, *Researcher in residence*.
“Assessing the impact of OCR quality on downstream tasks”
Amount: 50k EUR (buyout).
- Europeana events grant, *Masterclasses for digital cultural heritage*.
Amount: 5k EUR.
- Arts and Humanities Research Council (AHRC) and UK Research Innovation (UKRI) grant (9.2M GBP, 5 years), *Living with Machines*. Role: Co-I in charge for research infrastructure.
Amount: 9.2M GBP in total.

Previously

- SNF doc.Mobility fellowship, *Understanding Citations in the Humanities*.
Amount: 60k CHF.
- SNF regular grant, *Linked Books: An Algorithmic History of the History of Venice*. I wrote the grant, which was signed by Prof. Frédéric Kaplan.
Amount: 290k CHF.

Selected publications

Highlights to give a sense of my recent work. *My name is in bold, co-authors under my direct or joint supervision are underlined. Each reference includes a direct link to an Open Access version of the paper.*

- **Colavizza, Giovanni**, Thomas Franssen and Thed van Leeuwen. 2019. *An empirical investigation of the Tribes and their Territories: Are research specialisms rural and urban?* Journal of Informetrics. <https://arxiv.org/abs/1806.09899>.
This paper considers Becher and Trowler's theory of academic tribes and territories, according to which 'hard' sciences are more focused on few questions and thus have higher intellectual and social cohesion than 'soft' sciences focusing on many, seemingly disconnected questions. We operationalise this theory using network science methods, and find that empirical evidence partially supports it.
What this paper shows: capability to conduct theory-driven empirical research on a key STS topic as leading author.
- Franceschet, Massimo and **Giovanni Colavizza**. 2020. *Quantifying the higher-order influence of scientific publications*. Scientometrics. <https://arxiv.org/abs/2006.03561>.
This paper proposes a method to quantify the citation impact of high-order citations, i.e., those beyond direct citation relations. The proposed method is based on linear algebra and analysed formally. We find that, using it, approximately 40% of citation impact accrues via higher-order citations, which are not accounted for in commonly-used citation-based indicators.
What this paper shows: capability to design a n STS measurement for a well-posed problem, and study its theoretical as well as empirical properties.
- Piccardi, Tiziano, Miriam Redi, **Giovanni Colavizza** and Robert West. 2020. *Quantifying Engagement with Citations on Wikipedia*. World Wide Web Conference. <https://arxiv.org/abs/2001.08614>.
In this paper, we study the user interaction with external references in Wikipedia, and show how users seek information in Wikipedia. If an article is low-quality, external references are more frequently clicked on, suggesting that when a Wikipedia article is of high-quality, most users find the information they need in it.
What this paper shows: capability to co-supervise research leveraging a large and sensitive dataset, in collaboration with a societal stakeholder (Wikimedia Foundation), in order to deliver insights to improve a public and open knowledge infrastructure.
- Sigh, Harshdeep, Robert West and **Giovanni Colavizza**. 2020. *Wikipedia Citations: A comprehensive dataset of citations with identifiers extracted from English Wikipedia*. Quantitative Science Studies. <https://arxiv.org/abs/2007.07022>.
This is a recent dataset paper, contributing the largest, to date, dataset of citations extracted from Wikipedia. This dataset includes over 23M citations to scientific and non-scientific sources cited from Wikipedia, including code for reproducibility.
What this paper shows: capability to contribute open STS research datasets and code, as senior author.
- **Colavizza, Giovanni**. 2020. *COVID-19 research in Wikipedia*. Quantitative Science Studies. https://www.mitpressjournals.org/doi/pdf/10.1162/qss_a_00080.
This paper studies the reliability and representativity of Wikipedia's sources on COVID-19, and finds that Wikipedia's editor community has managed to curate the integration of high-quality information in Wikipedia, despite the unprecedented number of new contributions and time pressures.

What this paper shows: capability to focus, as sole author, on a timely STS question. Furthermore, this contribution leverages data from public and private sources, and uses a mix of quantitative methods (citation analysis, text mining, regression modelling).

- **Colavizza, Giovanni**, Iain Hrynaszkiewicz, Isla Staden, Kirstie Whitaker and Barbara McGillivray. 2020. *The citation advantage of linking publications to research data*. PLOS ONE. <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0230416>. This paper considers whether research articles providing access to code and data for reproducibility are better cited than papers which do not. We find that, on average, papers providing access to code and data receive 25% more citations. This paper has been highly-cited and read already during its publication year.

What this paper shows: capability to contribute research on pressing societal issues (open science) as leading author, and applied statistics (logistic regression modelling).

Full list of publications

My name is in bold, co-authors under my direct supervision are underlined. Only authors, year, title and venue are given. Each publication is retrievable via a Web search.

In preparation or under review.....

The following manuscripts are currently in preparation (just the title):

- *Wikipedia and Open Access*
- *Wikipedia's Sources: A Comprehensive Assessment of their Political Bias and Factual Reliability*
- *Scientists in times of COVID: An Investigation of Information Seeking and Sharing Behaviours*
- *Science in times of COVID: A Large-Scale Study of Debates on Twitter*
- *Citation Context Embeddings and their Applications for Mapping Science*
- *Open Computational Challenges in Digital Art Markets*
- *Accelerated Indexation of Cultural Heritage Corpora via Active Learning*

The following manuscripts are currently under review (just the title):

- *Experts and authorities receive disproportionate attention on Twitter during the COVID-19 crisis*
- *Data Engineering for Data Analytics: A Classification of the Issues, and Case Studies*

Science and Technology Studies.....

2021

- Piccardi, Tiziano, Miriam Redi, **Giovanni Colavizza** and Robert West. *The Value of Wikipedia as Web Gateway*. WWW Conference.
- **Colavizza, Giovanni**, Rodrigo Costas, Vincent Traag, Nees J. van Eck, Thed van Leeuwen and Ludo Waltman. *A scientometric overview of COVID-19*. PLOS ONE.

2020

- **Colavizza, Giovanni**. *COVID-19 research in Wikipedia*. Quantitative Science Studies.
- Piccardi, Tiziano, Miriam Redi, **Giovanni Colavizza** and Robert West. *Quantifying Engagement with Citations on Wikipedia*. WWW Conference.
- Sigh, Harshdeep, Robert West and **Giovanni Colavizza**. *Wikipedia Citations: A comprehensive dataset of citations with identifiers extracted from English Wikipedia*. Quantitative Science Studies.
- Waltman, Ludo, Kevin W. Boyack, **Giovanni Colavizza** and Nees J. van Eck. *A principled methodology for comparing relatedness measures for clustering publications*. Quantitative Science Studies.
- **Colavizza, Giovanni**, Iain Hrynaszkiewicz, Isla Staden, Kirstie Whitaker and Barbara

McGillivray. *The citation advantage of linking publications to research data*. PLOS ONE.

- Spinaci, Gianmarco, **Giovanni Colavizza** and Silvio Peroni. *Preliminary Results on Mapping Digital Humanities Research*. AIUCD.
- Franceschet, Massimo and **Giovanni Colavizza**. *Quantifying the higher-order influence of scientific publications*. Scientometrics.
- Daquino, Marilena, et al. *The OpenCitations data model*. International Semantic Web Conference.

2019

- **Colavizza, Giovanni** and Matteo Romanello. *Citation Mining of Humanities Journals: the Progress to Date and the Challenges Ahead*. Journal of European Periodical Studies.
- **Colavizza, Giovanni**, Thomas Franssen and Thed van Leeuwen. *An empirical investigation of the Tribes and their Territories: Are research specialisms rural and urban?* Journal of Informetrics.
- **Colavizza, Giovanni** et al. *Linked Books: un indice citazionale per la storia di Venezia*. DigItalia.

2018

- **Colavizza, Giovanni**. *A diachronic study of historiography*. Scientometrics.
- Boyack, Kevin W., Nees Jan van Eck, **Giovanni Colavizza** and Ludo Waltman. *Characterizing in-text citations in scientific articles: A large-scale analysis*. Journal of Informetrics.
- Rodrigues Alves, Danny, **Giovanni Colavizza** and Frédéric Kaplan. *Deep Reference Mining from Scholarly Literature in the Arts and Humanities*. Frontiers in Research Metrics and Analytics.

2017

- **Colavizza, Giovanni**. *The structural role of the core literature in history*. Scientometrics.
- **Colavizza, Giovanni**. *The core literature of the historians of Venice*. Frontiers in Digital Humanities.
- **Colavizza, Giovanni**, Kevin W. Boyack, Nees Jan van Eck and Ludo Waltman. *The Closer the Better: Similarity of Publication Pairs at Different Co-Citation Levels*. Journal of the Association for Information Science and Technology.
- Franceschet, Massimo and **Giovanni Colavizza**. *TimeRank: A dynamic approach to rate scholars using citations*. Journal of Informetrics.

2016

- **Colavizza, Giovanni** and Massimo Franceschet. *Clustering Citation Histories in the Physical Review*. Journal of Informetrics.

Digital/Computational Arts and Humanities.....

2020

- Todorov, Konstantin and **Giovanni Colavizza**. *Transfer Learning for Historical Corpora: An Assessment on Post-OCR Correction and Named Entity Recognition*. Computational Humanities Research Conference.
- Todorov, Konstantin and **Giovanni Colavizza**. *Transfer Learning for Named Entity Recognition in Historical Corpora*. CLEF Initiative (Conference and Labs of the Evaluation Forum).
- Lazzari, Gianrocco, **Giovanni Colavizza** et al. *A digital reconstruction of the 1630-1631 large plague outbreak in Venice*. Scientific Reports.
- McGillivray, Barbara et al. *The challenges and prospects of the intersection of humanities and data science: A white paper from The Alan Turing Institute*.
- van Strien, Daniel, Kaspar Beelen, Mariona Ardanuy, Ksra Hosseini, Barbara McGillivray and **Giovanni Colavizza**. *Assessing the Impact of OCR Quality on Downstream NLP Tasks*. ARTIDIGH ICAART.
- Franceschet, Massimo, **Giovanni Colavizza**, et al. *Crypto art: A decentralized view*. Leonardo.

2019

- **Colavizza, Giovanni**. *Are We Breaking the Social Contract?* Journal of Cultural Analytics.
- **Colavizza, Giovanni**, Maud Ehrmann, Fabio Bortoluzzi and Frédéric Kaplan. *Index-driven Digitization and Indexation of Historical Archives*. Frontiers in Digital Humanities.
- Filgueira Rosa, et al. *defoe: A Spark-based Toolbox for Analysing Digital Historical Textual Data*. E-Science Conference.
- Madge, Jim, **Giovanni Colavizza**, James Hetherington, Weisi Guo and Alan Wilson. *Assessing Simulations of Imperial Dynamics and Conflict in the Ancient World*. Cliodynamics.
- **Colavizza, Giovanni**, Riccardo Cella and Anna Bellavitis. *Apprenticeship in Early Modern Venice*, In Prak, Maarten and Patrick Wallis, *Apprenticeship in Early Modern Europe*. Cambridge University Press.

2018

- Albertin, Fauzia, Eleonora Balliana, Giorgia Pizzol, **Giovanni Colavizza**, Elisabetta Zendri and Dorit Raines. *Printing materials and technologies in the 15th-17th century book production: an undervalued research field*. Microchemical Journal.
- **Colavizza, Giovanni** and Matteo Romanello. *Annotated References in the Historiography on Venice: 19th-21st centuries*. Journal of Open Humanities Data.
- **Colavizza, Giovanni**. *Understanding the history of the humanities from a bibliometric perspective: Expansion, conjunctures and traditions in the last decades of Venetian historiography (1950-2013)*. History of the Humanities.

2017

- **Colavizza, Giovanni**, Matteo Romanello and Frédéric Kaplan. *The References of References: A Method to Enrich Humanities Library Catalogs with Citation Data*. International Journal on Digital Libraries.
- **Colavizza, Giovanni**. *A View on Venetian Apprenticeship through the Garzoni Database*. In Garzoni Apprendistato, Lavoro e Società a Venezia e in Europa, XVI-XVIII secolo.

2016

- Ehrmann, Maud, **Giovanni Colavizza**, Orlin Topalov, Riccardo Cella, Davide Drago, Andrea Erbooso, Francesca Zugno, Anna Bellavitis, Valentina Sapienza, Frédéric Kaplan. *From Documents to Structured Data: First Milestones of the Garzoni Project*. DHCommons.
- Ehrmann, Maud, **Giovanni Colavizza**, Yannick Rochat and Frédéric Kaplan. *Diachronic Evaluation of NER Systems on Old Newspapers*. In Proceedings of the 13th Conference on Natural Language Processing (KONVENS 2016).
- **Colavizza, Giovanni**. *Epidemics in Venice: On the Small or Large Nature of the Pre-Modern World*. In Bozic, Bojan, Gavin Mendel-Gleason, Christophe Debruyne and Declan O'Sullivan. IFIP Advances in Information and Communication Technology.

Selected conferences and talks

2020

- Doing Science in Times of Crisis, webinar series at CWTS (organizer).
- Computational Humanities Research (committee).
- Collect & Connect: Archives and Collections in a Digital Age Conference (committee).
- Digital traces, navigational paths, and intellectual mobilities workshop, EPFL (invited).
- World Wide Web Conference (committee).
- International AAAI Conference on Web and Social Media (ICWSM).

- Vossius Center talks (invited).
- International Semantic Web Conference (ISWC).

2019

- Social Data Seminar, The Alan Turing Institute (invited).
- Futures Thinking Conference, Oxford.
- International Conference on Scientometrics & Informetrics (ISSI), Rome (committee).
- Digital Humanities Conference, Utrecht (committee).
- Computational Archival Science Workshop, London.
- Software Citations Workshop, London (organizer).

2018

- WikiCite, San Francisco.
- Effective Altruism Global, London.
- Science and Technology Indicators (STI), Leiden (committee).
- OpenCitations Workshop, Bologna (committee).
- Infiniteconf, London (Invited).
- Digital Humanities Conference, Mexico City (committee).
- Computational Methods in the Humanities, Lausanne (committee).

2017

- International Conference on Scientometrics & Informetrics (ISSI), Wuhan (committee).
- Making of Humanities Conference, Oxford.
- Research Evaluation in the Social Sciences and Humanities, Antwerp (committee).
- Bruno Kessler Foundation, Trent (invited).

2016

- Leibniz Institute of European History, Mainz (invited).
- Making of Humanities Conference, Baltimore.
- Centre for Science and Technologies Studies, Leiden (invited).
- Effective Altruism x Oxford.
- International Workshop on Bibliometric-enhanced Information Retrieval, Padua.
- Society for the History of Authorship, Reading and Publishing (SHARP), Paris.
- Digital Humanities Conference, Krakow.
- Computational History and Data-Driven Humanities, Dublin.
- European Conference on Information Retrieval (ECIR), Padua.
- Renaissance Society Conference, Boston.

2015 and before

- Italian Conference on Computational Linguistics (Click-IT), Trent.
- *De l'argile au nouage : une archéologie des catalogues*, Geneva (invited).
- Renaissance Society Conference, Berlin.
- Histoinformatics, Barcelona.
- Sixteenth Century Society Conference (SCSC), New Orleans.
- Consortium of European Research Libraries Annual Meeting, Oslo (invited).
- Leipzig e-Humanities Seminars, Leipzig.

Academic service

Societies.....

Society for the History of the Humanities, International Society for Scientometrics and Informetrics, European Association for the Digital Humanities, Cultural Evolution Society, Oxford Bibliographical Society, Computational Humanities Research.

Conference service.....

Digital Humanities Conference, Corpus-based Research in the Humanities, Italian Natural Language Processing Conference, World Wide Web Conference, Computational Humanities Research Conference, Science and Technology Indicators Conference, International Conference on Scientometrics & Informetrics, and more.

Journal reviewer.....

Scientometrics, Research Evaluation, Journal of Informetrics, Journal of the Association for Information Science and Technology, Quantitative Science Studies, International Journal on Digital Libraries, Frontiers in Digital Humanities, Frontiers in Research Metrics & Analytics, PLOS ONE, PeerJ Computer Science, and more.