A presentation of ManuscriptAI, an upcoming project aiming to build an AI-tool to facilitate the integration, accessibility, and usability of heterogeneous cultural heritage data on medieval manuscripts

Data on our cultural heritage hold enormous potential for Europe's economic growth, for the construction of a more inclusive narrative of its past, and for the future collective identity of its citizens. The biggest obstacle to unlocking that potential is the current lack of integration and interoperability among the countless datasets describing the holdings of European heritage institutions. ManuscriptAI is an ERC Proof of Concept project starting in April 2024 that aims to help remove that obstacle for the data on Europe's medieval written heritage: manuscripts. Premodern handwritten books are a pivotal category of our heritage, yet they are currently underrepresented in large research infrastructures and their catalogue data locked in digital silos. ManuscriptAI will employ machine learning algorithms to construct a model capable of facilitating the automatic integration of distinct data sources describing medieval manuscripts, under a predefined set of machine-understandable vocabulary terms. The model will be made accessible through a human engagement interface and tested during a pilot in a real-world setting. Crucially, this project aims to avoid the creation of another data silo. Instead, it will function as a practical instrument for integrating existing datasets and information. This approach is designed to enhance data reusability, link disparate pieces of information, and ensure interoperability among datasets, all while considering the digital sustainability of the already existing datasets and information. This effort will establish the foundation for creating a unified global view of our cultural heritage holdings and descriptions, simultaneously addressing, and overcoming the issue of data redundancy caused by duplicated information. The project aims to fill two important desiderata: (1) a user-friendly AI-tool to allow heritage professionals to convert their metadata on manuscripts to Linked Open Data, and (2) a dedicated ontology for the description of medieval manuscripts to complete CIDOC-CRM extensions for the cultural heritage domain. The project, building on the ERC-2018-stg PASSIM, is a collaborative effort of Radboud University and the Netherlands Organisation for Applied Scientific Research (TNO), supported by domain experts and heritage professionals. In this paper, we will present the ManuscriptAI-project, its context, set-up, aims, and challenges.

GIULIA BIAGIONI   giulia.biagioni@tno.nl

Giulia Biagioni is a scientist at the Netherlands Organisation for Applied Scientific Research (TNO), within the Advanced Computing Engineering department. Her research focuses on utilizing Artificial Intelligence (AI), particularly semantic technologies, to facilitate interoperability among platforms, systems, and datasets. Additionally, she focuses on creating technologies that equip machines with the capability to comprehend and reason upon source of norms and regulations. This enables autonomous compliance checks within operations, aligning with international laws, policies, and regulations. Biagioni' s key interests lie in normative systems and digital culture, with a strong commitment to the FAIR data principles—aiming for data to be findable, accessible, interoperable, and reusable.

RECENT PROJECT COLLABORATIONS

2024-2027	ECOFeD: Next Generation Cloud Infrastructures and Services (EU IPCEI-CIS)
2024-2025	ManuscriptAI (EU Horizon ERC PoC)
2024	Climate and Air Quality (ERP 2024)
2023-present	ITEA OpenLog (Status: Final Project Proposal)
2023-present	Norm Engineering Program (Dutch Ministry of the Interior and Kingdom Relations)
2023	TDI500 (Duurzaam Team NL)
2022-2023	FEDeRATED (EU and CEF Project)
2022-2023	MAGPIE: Smart Green Ports (EU Horizon 2020)
2022-2023	Interconnect (EU Horizon 2020)

COLLABORATIONS IN STANDARDIZATION ACTIVITIES

ETSI EN 303 760 – STF 641. Guidelines for IoT Semantic Interoperability
ETSI TS 103 410 – V.1.1.3. Part 1: Energy Domain

SELECTED ARTICLES AND CONFERENCES

- Biagioni G., "The Handwriting of Diplomacy" in The Mediaeval Journal,vol.10:2 (2020), Brepols.
- Daniele L., Biagioni G., Bouter C., Large Scale Interoperability Solutions in InterConnect Using SAREF4ENER: Experiences and Lessons Learned, Presentation in the session for Industry at the 18th International Conference on Semantic Systems co-located with 18th International Conference on Semantic Systems (SEMANTiCS 2022).
- Stathis G., Biagioni G., Trantas A., van den Herik, J., Custers B., (2023),A Visual Analysis of Hazardous Events in Contract Risk Management, In the Proceedings of the 12thInternational Conference on Data Science, Technology and Applications(DATA 2023).
- Stathis G., Biagioni G., Trantas A., Graaf K.A., Custers B., (2023), The Value of Proactive Data for Intelligent Contracts, In the Proceedings ofthe 7th International Conference on Smart Trends in Systems, Security& Sustainability (WorldS4 2023).
- Stathis G., Trantas A., Biagioni G., Graaf K.A., Custers B., (2023), Designing an Intelligent Contract with Communications and Risk Data, Spinger Nature [Status: Under Publication]
- Breteler, J., van Gessel, T., Biagioni, G. & van Doesburg, R. (2023) The FLINT Ontology: An Actor-Based Model of Legal Relations. *Knowledge Graphs: Semantics, Machine Learning, and Languages. Proceedings of the 19th International Conference on Semantic Systems, 20-22 September 2023, Leipzig, Germany.* Studies on the Semantic Web volume 56, IOS Press.
- Van Gessel, T., Biagioni, G., Breteler, J., Tolios, I. & Boertjes, E. (2023) A Toolset for Normative Interpretations in FLINT. *Proceedings of the Posters and Demo Track of the 19th International Conference on Semantic Systems co-located with 19th International Conference on Semantic Systems (SEMANTiCS 2023).*
- Breteler, J., van Gessel, T., Biagioni, G. & van Doesburg, R. (2023) The FLINT Ontology: An Actor-Based Model of Legal Relations. Knowledge Graphs: Semantics, Machine Learning, and Languages. Proceedings of the 19th International Conference on Semantic Systems, 20-22 September 2023, Leipzig, Germany. Studies on the Semantic Web volume 56, IOS Press.
-

INVITED SPEAKER

Trade Facilitation Innovation Days 2023 hosted by UNCTAD (under the World Economic Forum's Centre for Regions, Trade and Geopolitics program) and broadcasted internationally, 19 – 20 September 2023.

SHARI BOODTS
shari.boodts@ru.nl
https://www.ru.nl/en/people/boodts-s

My research centers on the use of computational methods to study the textual transmission and knowledge dissemination in the Middle Ages. My recent research projects have focused on textual corpora from early Christian literature, with a particular focus on patristic sermons. I aim at a productive synergy of history, textual criticism, transmission and reception studies, and Digital Manuscript Studies. Additionally, I am interested in questions surrounding heritage policy, sustainable digital tools and data, and Open Access.

ACADEMIC APPOINTMENTS

2018-present	Radboud University, Faculty of Arts, Nijmegen (NL), Assistant professor
2013-2018	KU Leuven, Faculty of Arts, Leuven (BE), Postdoctoral fellow
2016	University of Ghent, Faculty of Arts and Philosophy, Ghent (BE), Guest professor
2015	Universität Salzburg, CSEL, Vienna (AU), Visiting fellow (Oct.-Nov.)
2013-2014	IRHT and ÉPHÉ, Paris (FR), Visiting fellow (Oct.-Jan.)
2009-2013	PhD candidate, Literary Studies, KU Leuven, Faculty of Arts, Leuven (BE)

FELLOWSHIPS/GRANTS/AWARDS/MEMBERSHIPS (since 2017)

2024-2025	ERC Proof of Concept "ManuscriptAI", European Research Council (€ 150.000)
2023	Radboud Team Science Award, ERC-StG "PASSIM" team (€ 10.000)
2021-present	Dutch Young Academy, Member
2021-present	Associate Editor, *Zeitschrift für Antikes Christentum* (ISSN 1612-961X)
2020-present	Radboud Young Academy, Founding member
2020	Early career award, Dutch Royal Academy of Sciences (KNAW) (€15.000)
2019-2024	ERC Starting Grant "PASSIM", European Research Council (€ 1.500.000)
2018-present	Editorial Board Member, *Augustiniana* (ISSN 0004-8003)
2018-2022	VENI grant "ALANUS", NWO (€ 250.000)
2017	Pontifical Medal of the Academia Pontificia Latinitatis

KEY PUBLICATIONS

1. Boodts, Schmidt, Macchioro, Denis, Rempt, Komen, Hermsen, *PASSIM Research Tool, beta-version*. Radboud Institute for Culture and History, 2022. Last accessed 2 February 2024 (https://passim.rich.ru.nl/);
2. Denis, Boodts, 'The pseudo-Augustinian S. App. 121 and its medieval textual connections. A testcase for the development of network visualizations in the PASSIM research tool', *Journal of Historical Network Research* 9 (2023), 1-35 (https://doi.org/10.25517/jhnr.v9i1.222);
3. Boodts, 'Navigating the Vast Tradition of St. Augustine's Sermons: Old Instruments and New Approaches', *Augustiniana*, 69 (2019), 83-115;
4. Dupont, Boodts, Partoens, Leemans (eds), *Preaching in the Patristic Era: Sermons, Preachers, and Audiences in the Latin West* (A New History of the Sermon, 6), Leiden: Brill, 2018;
5. *Sancti A. Augustini Sermones in epistolas apostolicas II (ss. 157-183)*, ed. Boodts, cuius seriei 11 serm. ed. Dolbeau, Partoens, Torfs, Weidmann (CCSL 41Bb), Turnhout: Brepols, 2016;
6. Boodts, Meijns, Leemans (eds), *Shaping Authority. How did a person become an authority in Antiquity, the Middle Ages and the Renaissance?* (LECTIO, 3), Turnhout: Brepols, 2016;
7. Boodts, 'Augustine's *sermo* 180 on Iac. 5, 12: *Ante omnia nolite iurare*. A new critical edition with introduction', *Recherches augustiniennes et patristiques*, 37 (2013), 1-50;

ORGANISATION OF PARTICIPATION IN SCIENTIFIC MEETINGS

Between 2011 and 2023, I co-organised nine international conferences. Since 2018, I delivered 25 papers at international conferences. I was an invited speaker at seven international conferences.