

The background features a dark blue field with stylized gears in orange, red, and grey on the right side. At the bottom left, there is a decorative zigzag pattern of light blue lines.

Radical Collaboration on the Living with Machines project

Ruth Ahnert, Federico Nanni, Daniel Wilson
The Alan Turing Institute

The BL and The Alan Turing Institute



- The **British Library** has digitised millions of pages from its collections.
- The national institute for data science and artificial intelligence, **The Alan Turing Institute**, offers the expertise to harness this data to answer research questions at scale.

a data-driven history project/ a history-focused data science project.

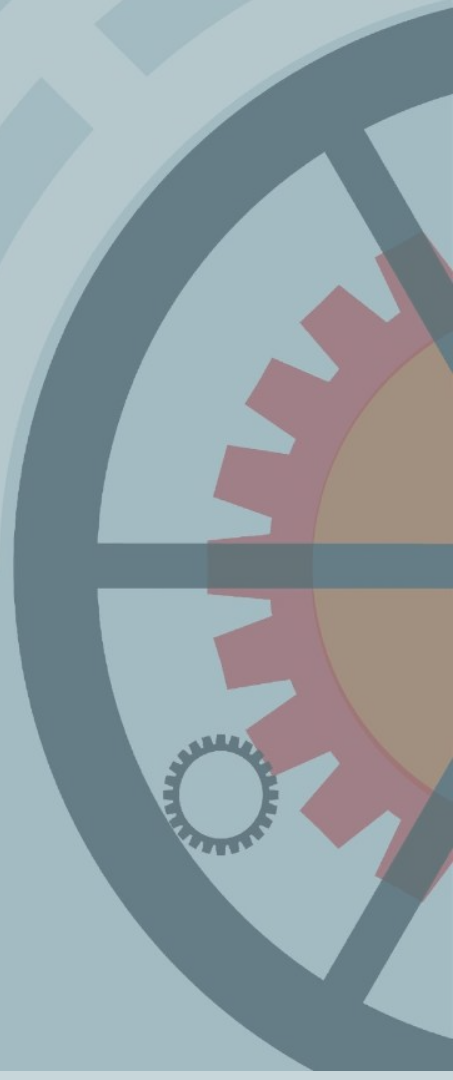
The untapped depth of digital archives means that there is still much to learn about the human, social, and cultural consequences of the coming of the machine age (c.1780-1920).



Objectives








































We aim to:

- provide computational means of marshalling at scale the UK's growing number of digitised historical texts, and documents;
- develop innovative computational models, tools, code, and infrastructure that will be transferable to other research projects;
- facilitate new historical findings about the effects of mechanisation on the lives of ordinary people;
- deliver the above through radical collaboration.



The Living with Machines Team

Principal and Co-Investigators Project team

 <p>Ruth Ahmert (QMUL)</p> <p>Read more</p>	 <p>David Beavan (Turing)</p> <p>Read more</p>	 <p>Emma Griffin (UEA)</p> <p>Read more</p>	 <p>Claire Austin Rights Assurance</p> <p>Read more</p>	 <p>Kasper Beelen Digital Humanities Senior Research Associate</p> <p>Read more</p>	 <p>Mariona Coll Ardanuy Computational Linguistics Senior Research Associate</p> <p>Read more</p>	 <p>André Piza Research Project Manager</p> <p>Read more</p>	 <p>Griffith Rees Research Data Scientist</p> <p>Read more</p>	 <p>Kalle Westerling Research Software Engineer</p> <p>Read more</p>	 <p>Daniel Wilson History Senior Research Associate</p> <p>Read more</p>	 <p>Rosie Wood Research Data Scientist</p> <p>Read more</p>
 <p>Timothy Hobson (Turing), Research Software Engineer</p> <p>Read more</p>	 <p>Jon Lawrence (Exeter)</p> <p>Read more</p>	 <p>Maja Maricevic (British Library)</p> <p>Read more</p>	 <p>Lótlé Demertzí Programme Coordinator</p> <p>Read more</p>	 <p>Luke Hare Research Data Scientist</p> <p>Read more</p>	 <p>Sherman Lo Research Data Scientist</p> <p>Read more</p>	 <p>Filipe Bento Research Associate</p> <p>Read more</p>	 <p>Giovanni Colavizza (Turing)</p> <p>Read more</p>	 <p>Joel Dearden (1979-2020) Research Software Engineer</p> <p>Read more</p>	 <p>Lucy Havens Research Software Engineer Intern</p> <p>Read more</p>	 <p>James Hetherington (Turing)</p> <p>Read more</p>
 <p>Barbara McGillivray (Turing / King's College London)</p> <p>Read more</p>	 <p>Mia Ridge (British Library)</p> <p>Read more</p>	 <p>Alan Wilson (Turing)</p> <p>Read more</p>	 <p>Katie McDonough Senior Research Associate</p> <p>Read more</p>	 <p>Federico Nanni Senior Research Data Scientist</p> <p>Read more</p>	 <p>Nilo Pedrazzini Corpus-Based Digital Humanities Research Assistant</p> <p>Read more</p>	 <p>Adam Farquhar (British Library)</p> <p>Read more</p>	 <p>Rosa Filgueira (EPCC), Data Architect</p> <p>Read more</p>	 <p>Sarah Gibson Research Software Engineer</p> <p>Read more</p>	 <p>Michael Jackson (EPCC), Software Architect</p> <p>Read more</p>	 <p>Joshua Rhodes Senior Research Associate</p> <p>Read more</p>
 <p>Guy Solomon Research Associate</p> <p>Read more</p>	 <p>Giorgia Tofo Data And Content Manager</p> <p>Read more</p>	 <p>Daniel Van Strien Digital Curator</p> <p>Read more</p>	 <p>Kasra Hosseini Research Data Scientist</p> <p>Read more</p>	 <p>Yann Ryan British Library Curator, Digital Newspapers</p> <p>Read more</p>	 <p>Olivia Vane Digital Humanities Research Software Engineer</p> <p>Read more</p>					

Challenges

- Beginning from a standing start
 - No pilot projects
 - Almost no prior inter-team collaborations
 - Some team members with no collaborative experience
 - Data not yet acquired or ingested
 - Infrastructure not in place
- Different professional and disciplinary norms:
 - Expectations and definitions of success
 - Research processes
 - Publication and dissemination cultures.

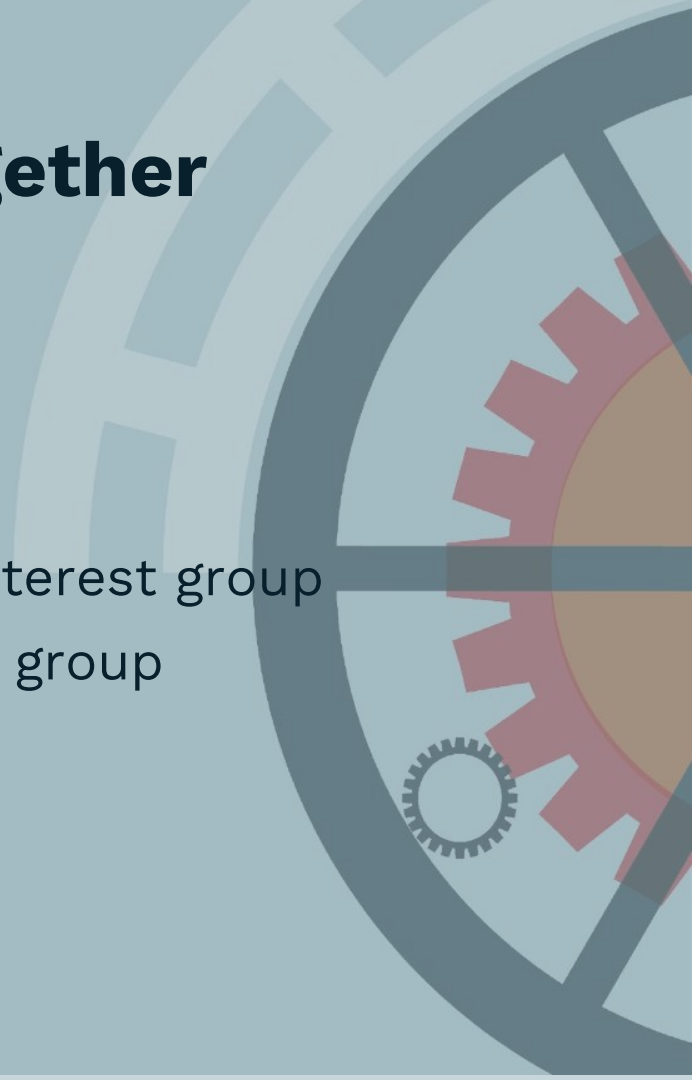


The Project Charter

- Increasing numbers of models for this practice: Scholars Lab, Colored Conventions Project, Stan Ruecker and Milena Radzikowska etc.
- Aim:
 - to create shared vision so that people knew what they were signing up for
 - lists a set of informal policies or values that we commit to abiding by, but also revisiting and revising where they are not working or require nuancing.
- Version 1: written by PI and Co-Is for hiring
- Version 2: revised once team on board
- Plan to revisit at intervals (Pandemic!)

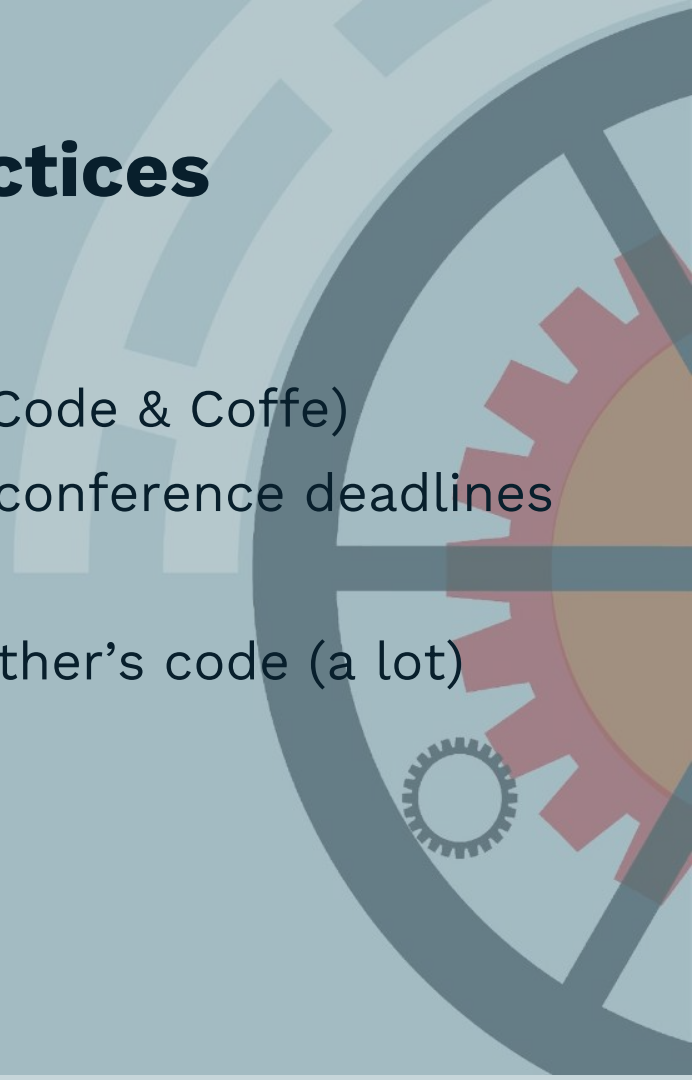
How to brainstorm ideas together

- HypGen: hypothesis generation group
- IdeasLab
- NLP reading group
- Computer vision for digital heritage interest group
- Humanities & data science discussion group



How to embed best RSE practices

- Offering git-flow overviews
- Being available for informal support (Code & Coffe)
- Having milestones independent from conference deadlines
- Having regular stand-up meetings
- Coding together and reviewing each other's code (a lot)



How to recognise all contributions

Conceptualization

Mariona Coll Ardanuy^{1,5}
Daniel CS Wilson^{1,5}

Methodology

Mariona Coll Ardanuy
Federico Nanni¹
Kasra Hosseini¹

Implementation

Federico Nanni
Kasra Hosseini
Mariona Coll Ardanuy
Kaspar Beelen^{1,5}

Reproducibility

Kasra Hosseini
Federico Nanni

Interpretation

Kaspar Beelen
Mariona Coll Ardanuy
Katherine McDonough^{1,5}
Daniel CS Wilson
Ruth Ahnert⁵
Jon Lawrence⁴
Giorgia Tolfo²

Historical Analysis

Daniel CS Wilson
Katherine McDonough
Kaspar Beelen
Jon Lawrence

Data Curation

Kaspar Beelen
Mariona Coll Ardanuy
Federico Nanni
Giorgia Tolfo

Annotation

Giorgia Tolfo
Ruth Ahnert
Kaspar Beelen
Mariona Coll Ardanuy
Jon Lawrence
Katherine McDonough
Federico Nanni
Daniel CS Wilson

Writing and Editing

Mariona Coll Ardanuy
Federico Nanni
Ruth Ahnert
Kaspar Beelen
Kasra Hosseini
Jon Lawrence
Katherine McDonough
Barbara McGillivray^{1,3}
Daniel CS Wilson

Supervision

Barbara McGillivray
Ruth Ahnert

Project Management

Barbara McGillivray
Ruth Ahnert
Mariona Coll Ardanuy