Living with Machines Delivery Plan version 1, 2019

Ruth Ahnert, David Beavan, Giovanni Colavizza, Adam Farquhar, Emma Griffin, James Hetherington, Jon Lawrence, Maja Maricevic, Barbara McGillivray, Amit Mulji, Andre Piza, Mia Ridge and Alan Wilson

Preamble

Living with Machines is a five-year collaborative project. It aims to generate new perspectives on the effects of the mechanisation of labour on the lives of ordinary people in Britain during the 'long nineteenth century' (c.1780-1918), by developing computational and historical techniques and research questions for working with historical sources.

This document contains the project's first 'Delivery Plan', a formal document that defines the project management processes used to execute and control the project. We have published it in response to interest from a range of sources about how we set up a project of this sort. This first version was drawn up at the start of the project, before the full team was in place. It was updated in March 2019, a few months after the project's official start date. This preamble was largely written in January 2020. On this project we regard the Delivery Plan as a living document that has been, and will continue to be, updated as the project progresses. Subsequent versions will also be deposited.

This project, funded by the UK Research and Innovation (UKRI) Strategic Priority Fund (SPF), is a multidisciplinary collaboration delivered by the Arts and Humanities Research Council (AHRC), with The Alan Turing Institute, the British Library, the University of Cambridge, the University of East Anglia, Exeter University, and Queen Mary University of London. Established in 2018, the SPF builds on Sir Paul Nurse's vision of a 'common fund', to support high quality multidisciplinary and interdisciplinary research programmes, which could have otherwise been missed through traditional funding channels. Living with Machines was one of the projects selected in the first phase of this funding, and therefore had few UK-models to look to for models of how to plan and deliver a project of this size and disciplinary breadth.

One of the aims of large-scale interdisciplinary work is the creation of new insights through the collaborative efforts of experts in different fields. Such a process requires an emphasis on iterative development of research questions, methods, infrastructure. Our challenge then was to build a set of structures, milestones and KPIs for the project that enabled space for experimentation, exploration, and iteration. What this first version of the Delivery Plan emphasises are our *broad* questions and proposed approaches, and the emphasis on the research process rather than overly-defined outcomes. These questions and methodological approaches have been honed over time, and research focuses have emerged that we did not foresee at the beginning of the project. By comparing this Delivery Plan to these subsequent versions, we hope people will be able to see both the evolution of our thinking

and processes, as well as how we adapted to unforeseen challenges, over the course of the project.

The main headings under which the Delivery Plan is structured come from a template that the AHRC used for The Creative Industries Clusters Programme. Although there is no single UKRI template, it is one that the Council has used successfully for recent major investments. The AHRC requires a Delivery Plan if the project is a major, high profile, complex or high value investment. The template contains six main sections: 1) Summary and Objectives, 2) Updated proposal, 3) Work Plan, 4) Finance Profile 5) Monitoring and Evaluation, and 6) Communications and Engagement.

Under Section 2 we introduce the structure of the five 'Labs' we began with on the project. In the project grant proposal we introduced the idea of Labs as a mechanism to gather interest around key intersections of data, methods and historical research questions. We are used to thinking of a laboratory as a room or building filled with scientific equipment, a place for doing tests and experiments, for teaching science, or producing chemicals and medicines. However, if you work in the humanities, you will have noticed in recent years a growing trend for using the word Lab to describe research groups of centres, specifically in the subdisciplines of media studies and digital humanities. These organisations share an emphasis on technology, experimentation, and collaboration across the usual disciplinary boundaries.[1] We also adopted the idea to avoid the problems of creating discrete work packages, that might create silos on the project, and prevent cross-pollination of ideas, and ultimately integration. Membership of the Labs are therefore not exclusive; most members of the team were expected to belong to at least two.

The decision to implement a process of iterative development has obvious implications for the Work Plan in section 3. This is most apparent at 3.2 Delivery Milestones, where we have focused on project stages rather than overly specific outcomes: this freedom is not only possible because of the nature of the SPF as a funding mechanism, but also necessary given the emphasis on iterative development at the centre of our project. It is in this area that most changes will be apparent in subsequent versions of the Delivery Plan.

Moreover, the flexible nature of our Milestones in the first version has ramifications for the design of Section 5 on Monitoring and Evaluation, specifically the Key Performance Indicators (KPIs). KPIs are measurable values that demonstrate how effectively the project is achieving its key objectives. These KPIs are structured by the Impact Plan in Section 2 and seek to quantitatively measure the research process as well as the outputs so that we have meaningful ways of evaluating our progress throughout the project.

The appended terms of reference for the Project Management Board and the Advisory Board may also be useful documents for other projects seeking to set up monitoring frameworks, and are adapted from the terms of reference used for similar boards across The Alan Turing Institute.

---

[1] See Urszula Pawlicka, 'Data, Collaboration, Laboratory: Bringing Concepts from Science into Humanities Practice', *English Studies* 98 (2017), 526-41; and Darren Wershler, Lori Emerson and Jussi Parikka, *The Lab Book: Situated Practices in Media Studies* (University of Minnesota Press), chapter drafts and interviews available here: https://manifold.umn.edu/projects/the-lab-book.

The following document is shared with the spirit of openness and accountability that is central to our project charter, and we welcome feedback and engagement.

Please note that we have made minor edits to the Delivery Plan to improve clarity, and redacted personal or commercial information, including the aspirational metrics assigned to each KPI. Redactions are noted in square brackets.

The list of authors includes Project Management Board members who have since moved on from the project. We would also like to credit team member Daniel Wilson, whose job application influenced a section of the Plan.

Living with Machines Delivery Plan version 1: pre-launch, 2019

Table of Contents

Section 1 – Summary and objectives

Living with Machines is both a research project, and a bold proposal for a new research paradigm. In this ground-breaking partnership between The Alan Turing Institute, the British Library, and partner universities (the University of Cambridge, University of East Anglia, University of Exeter, and Queen Mary University of London), historians, data scientists, research software engineers, geographers, computational linguists, and curators have been brought together to examine the human impact of the industrial revolution.

It is widely recognised that Britain was the birthplace of the world's first industrial revolution. One consequence of this was an explosion in the creation and collection of documentary sources. Yet historians rarely have the training to work at scale with the huge collections that the Victorians created, so there is still much to learn about the human, social, and cultural consequences of this historical moment. Focussing on the long nineteenth century (c.1780-1918), the Living with Machines project aims to harness the combined power of massive digitised historical collections and computational analytical tools to examine the ways in which technology altered the very fabric of human existence on a hitherto unprecedented scale. The central theme - the mechanisation of work practices - speaks directly to present debates about how society can accommodate the revolutionary consequences of AI and robotics. To understand the fraught co-existence of human and machine, this project contends that we need research methods that combine technological innovation and human expertise.

The project will utilise the British Library's National Newspaper collection, and event-based records (such as census, electoral registration, birth / marriages / death records and trade directories) collected by Findmypast and other partner bodies. By developing intuitive computational interfaces, and adapting collaborative practices developed in the field of software development, we will enable close interaction between computational methods and historical inquiry.

Outreach and Engagement will be central to the project from the outset, and will take two forms: familiar outcomes such as television programmes and a regional exhibition; and working with individuals and communities to create common understandings of their shared histories. Participatory aspects will embody best practices in crowdsourcing and citizen history.

Project benefits:

- The UK's first large-scale synergy between data science, artificial intelligence research, and the arts and humanities, building capacity and catalysing new research areas.

- The development of new computational techniques to marshal the UK's rich historical collections (digitised and born-digital), to enable new research questions to be posed.
- Enriched and interlinked data holdings for the British Library, to add additional context and value to content.
- The development of generalisable tools, code, and infrastructure that can be adapted for and inspire future interdisciplinary research projects.
- New perspectives on the effects of the mechanisation of labour and associated changes on the lives of ordinary people during the long nineteenth century.
- The creation of computational models to represent how language and meanings change across time and geography.
- The mapping of occupational and demographic change and of internal migration during the nineteenth century.
- Research breakthroughs maintaining UK global leadership in Digital Humanities and driving large-scale international partnerships and opportunities.

Section 2 – Updated proposal and impact plan

The Living with Machines project is multi-levelled. At the most abstract level it is concerned with proving the viability of a new research paradigm. That paradigm is defined by radical collaboration that seeks to close the gap between computational sciences and the arts and humanities by creating a space of shared understanding, practices, and practices of publication and communication. We want to create both a data-driven approach to our cultural past, and a human-focused approach to data science. In other words, providing a harmonious, scholarly way of living with machines.

That co-existence, however, is historically fraught. In 1933 the sociologist William Ogburn published *Living with Machines*, in which he advanced his theory of 'cultural lag': the contention that inventions outstrip societies' ability to understand and evaluate them.[2] We seek to find new ways to ensure that humanistic study keeps pace with the changes introduced by the digital era - especially the rapid digitisation and creation of born-digital archives. We will do this by turning our collaborative attention to a period in national history when the machine transformed the very fabric of human existence.

For many generations, historians have been seeking to make sense of the industrial revolution from a wide kaleidoscope of perspectives - quantitative, cultural, social, visual. However, since the 1960s, attention has shifted towards the lived experience of industrialisation. How did the fabric of everyday life change when people drifted

---

[2] Here we credit team member Daniel Wilson, who allowed Ruth to incorporate material from his cover letter applying for a post on the team in part of this section.

from the countryside to the cities, and traded a life in agriculture for one that was increasingly driven by machines? What was the impact of mechanisation on the lives of ordinary men, women, and children? Recent responses to these questions have become stratified within a 'quantitative' tradition, that seeks to count and measure social change, and a 'cultural' one, which looks at everything else. We contend that our understanding of that moment could be transformed by situating the qualitative insights alongside a quantitative approach. If mass digitisation could be harnessed, there is a genuine opportunity to ask a whole new set of research questions, which could in turn be answered on a more empirically grounded and representative basis.

Focussing on the long nineteenth century (c.1780-1918), the Living with Machines project will utilise the British Library's National Newspaper collection, and event-based records (census, electoral registration, births/marriages/deaths, trade directories) collected by contributing partners including Findmypast. This is the applied level of the project, which is defined by the collaborative development of computational tools and code to answer historical questions about the impact of the mechanisation of work and industrial infrastructure on where people lived, how often they moved, the nature of their work, the makeup of the working population, and how people conceived of these changes.

The lessons we hope to learn from this examination of the historical sources will not only contribute to our historical understanding of the first and second industrial revolutions, but also to the ongoing debates about how society can accommodate the revolutionary consequences of artificial intelligence (AI) and robotics in what has become known as the fourth industrial revolution.

What ties together the multiple layers of this project is the rigorous re-evaluation of the impact of technology on society. This project seeks to create a new scholarly space at the intersection of data science and humanistic inquiry to examine the role of technology, and its social and political impacts.

Collaboration

We propose an uncompromisingly collaborative research philosophy that will be iterative, self-reflexive, and designed to evolve. Our central datasets and research questions will drive the development of infrastructure, computational methods, and tools; the outcomes of these methods will iteratively nuance research questions; research questions will help to hone and improve code; and the cyclical process will help our research team develop best practice for collaboration and exchange, seamlessly joining qualitative and quantitative approaches. Moreover, our engagement with a wider audience will be planned into the evolving research programme: by inviting family and local historians to engage with the project in-progress, they will help direct our research questions, our methodological approach, and our research outputs.

This iterative and collaborative ethos will be central at each stage of the process, and structured by forming 'laboratories' around the key methodological challenges posed by our research aims and historical questions. These are:

Sources Lab

This lab is focused on the opportunities provided by the digitised sources we will collate and create in the project. It will engage with issues of representativeness, genre balance and bias in the source base, and develop new methods for exposing, assessing and compensating for historical biases and absent voices, and for ensuring that the specificity of different source types is not lost in the building of larger corpora and data-sets. It will also be responsible for assessing the uncertainties and simplifications introduced through digitisation, optical character recognition, data extraction and processing, algorithmic analysis and visualisation, and for developing the techniques necessary to improve name linkage and disambiguation when combining a wide range of source types. We propose to respond to these technical and methodological challenges by integrating historical, curatorial and data science perspectives and expertise, informed by automated methods including topic modelling and outlier detection.

Key questions:

- How do we identify and adjust for the biases, assumptions and skews in large inter-linked datasets?
- How can we best ensure that datasets can be interrogated in ways that allow scholars to focus on specific spatial and temporal frames, or specific genres of source (e.g. plebeian radical newspapers, cheap daily papers, elite county papers etc.)?
- What cutting-edge digital methods can we develop to 'clean' data and improve readability and understanding?
- How can we improve cross-census and cross-source name disambiguation rates and develop robust means of estimating the probability of name matches between sources?
- On what geographic scale can we hope to develop micro-studies of key demographic indicators e.g. residential stability/mobility rates over time using census enumerators' schedules, trade directories etc.?
- What sources can be effectively marshalled or created to provide the underlying evidence required for the project?

Language Lab

By combining the power of computational linguistics, and sentiment and network analysis, the lab will model the complex attitudes of people towards machines as expressed in various historical texts, extending our understanding of the transition to a heavily mechanised workplace and society; the connections between

mechanisation and new forms of social conflict; and perceptions of industrial accidents, risk, speed, and safety.

Key questions:

- How did the concept of technology arise in the texts of newspapers of the time under investigation?
- How did social commentators understand the emergence of new technologies and write about the relationship between humans and machines?
- How can we use quantitative methods to make sense of the cultural response to mechanisation?
- How does text genre affect the way the relationship between humans and machines was talked about?
- Can we map the mutation of a deep-rooted historic sense of a North/South divide into a new divide rooted in industry?
- Who articulated negative narratives and fears about mechanisation and how did such accounts change over time?
- How central was the idea of the rise of machines to concepts of 'modernity' and 'the modern' and how did this change over time?
- What kind of linguistic context and non-linguistic information contributes to defining the changing meaning of words in our corpus?

Space and Time Lab

One widely recognised feature of the modernisation of Britain during the nineteenth century was the expansion of the state, and its concomitant desire to collect ever more information about its citizens. The inaugural census of 1801 took a head count of Britain's population, and subsequent censuses, held at ten-yearly intervals, gathered increasingly intricate information about where individuals lived, where they had been born, the people they lived with, and the jobs they performed. Along with the registration of births, marriages, and deaths, this provides us with billions of data points – an unparalleled trove of information about the lives of people during the world's first industrial revolution.

The difficulty is that historians rarely possess the skills or training to handle data on such a vast scale. Through micro-studies on specific locations, or short time periods, historians have demonstrated the vast potential of this data. The Cambridge Population Group's (Campop) important work on the census maps change over time within specific localities (GIS polygons). This lab will focus on linking people over time - improving name recognition techniques in order to combine census data and event-based records from Findmypast and other bodies to refine our understanding of urban demographic change, analysing the linked data using a combination of GIS technologies and time series analysis.

Key questions:

- What can we discover about the makeup of households and their variation in industrial areas?
- How did migration change in the industrial age (in terms of age of individuals, distance travelled, areas to which people moved, and from which they originated)?
- Did the rise of capital intensive machine-based factory production tend to increase or reduce residential stability in its classic locales such as East Lancashire, West Yorkshire and the Black Country compared with both urban areas that were less machine-based?
- How did opportunities for social and geographical mobility in Victorian Britain vary between industrial and non-industrial areas?
- Can we use census data to understand changing patterns of social mobility and economic activity amongst women?
- Can we use the census to trace the victims of industrial accidents named in the press - both backwards and forwards in time and to map them spatially?

Communities Lab

This lab focuses the interactions within and between the distinct communities contributing to the project as a whole. It has both operational and research missions. We will explore the many contributing academic specialisms represented in the project, reflect on the challenges of an 'uncompromisingly collaborative research philosophy' and develop a model for truly radical interdisciplinary collaboration. We will broadly communicate and share recommendations for success learned. Working closely with the other labs, it will amplify the impact of research, tutorials, workshops, code and documentation produced within the project. Furthermore, we will engage with academic partners and the public, working with individuals and communities to create common understandings of their shared histories through an exhibition, training opportunities and participatory tasks. Participatory aspects will embody and seek to extend best practices in crowdsourcing, citizen history and human computation.

Key questions:

- How can we contribute to the existing scholarship on a) interdisciplinary collaboration, and b) outreach and engagement?
- When working in ways that challenges disciplinary norms, how do we communicate the findings, benefits and outcomes to our home communities in a compelling way?
- How do we meaningfully include public participation within a complex, rapidly evolving research project, particularly given licensing and copyright constraints on public access to source collections?
- How do we best document the process of developing practices that will undoubtedly unravel over a long time, encounter tensions and failure, and only arrive at solutions through many iterations?

- How does this form of radical collaboration contribute to discourse around the future of work? So often these discussions occur in relation to industry, but what does it mean for academia and cultural heritage organisations?
- How should research methods and processes from history, public engagement and data science that shaped the project be represented when communicating about the project as a whole?
- How can we integrate data science and public engagement models of 'success' within a human computation system?

Integration, Infrastructure and Interfaces

The Integration, Infrastructure and Interfaces (3I) group has both an operational and a research mission.

Operationally, the group develops and maintains the storage and computing infrastructure for the Living with Machines project. It also contributes to the development and deployment of data science and machine learning solutions, either developed independently or in partnership with laboratories, in the event that they need to become part of the project infrastructure (e.g. Named Entity Recognition). Furthermore, it facilitates the use of the project data and infrastructure.

The 3I group research agenda focuses on the study and application of computational methods for the project. Members of 3I participate in research conducted in the laboratories. Typically, at least one 3I member will be involved in each laboratory. The group also conducts research on general methodological topics of interest to the project (e.g. active and transfer learning, inspectability). Finally, the group strives to spot and suggest opportunities for sharing computational methods across the project.

The 3I group's ambition is to have an impact in the following ways:

- By guaranteeing the necessary level of support to the research activities of the Living with Machines project.
- By developing a large-scale and AI-driven infrastructure for digital research in the humanities.
- By taking active part in the research effort, firstly to inform the development of the said infrastructure, and also for its own sake.
- By facilitating methodological connections across laboratories

Impact Plan

As the UK's first large-scale synergy between data science, artificial intelligence research, and the arts and humanities, we see it as our mission to drive innovation, set standards that can be replicated, build research capacity in this area, and catalyse new research. This can be broken down into eight areas of impact. For

qualitative and quantitative assessment of our ongoing progress towards this impact, see section 5 below

1.   Provide a model for an effective collaborative environment, that facilitates interdisciplinarity, and supports and promotes all members.

There are relatively few projects that are working across so many disciplinary boundaries and at this scale. One of the legacies we want to leave is a set of recommendations for facilitating this process in the future, by sharing lessons learned both from our successes and our failures. Our contention is that in projects like this, the research process can be as important as the outcomes (i.e. development of methods and research findings), and we therefore want to make this accessible by working in radically open ways, and by publishing our recommendations for effective methods of interdisciplinary collaboration. There is a lack of such self-reflexive work currently available, which we will rectify. In addressing this lack, we hope to help establish norms by which future projects will share and document their process.

2.   Develop generalisable tools and code components embedded into an infrastructure that can be adapted for and inspire future interdisciplinary research projects in the areas of data cleaning and historical text processing, computational semantic analysis, and machine learning.

The research engineering contribution to the project will address the task of creating a general and comprehensive data model for the data and metadata of the project, including process and human or machine generated (meta)data. This data model will be mapped and implemented in an agile data workflow, allowing for multiple entry points to the (meta)data, from domain specific query languages to graphical interfaces. These outcomes will be developed in view of generality (i.e. their applicability to other projects), and by documenting the reasoning and research motivating them.

While contributing to marshalling and integrating all the data sources of interest to the project, several machine learning solutions will be developed. These include approaches to ingest, clean, enrich, interlink and analyse data. The project will not only use or develop solutions as appropriate (machine learning research), but also integrate them into the project's infrastructure (machine learning engineering) and in so doing make them available to the broadest public possible, within and outside the project.

3.   Develop new computational techniques to marshal the UK's rich historical collections (digitised and born-digital), to enable new research questions to be posed of holdings.

While individual researchers have made great use of discrete datasets based on digitised collections, very few projects internationally have attempted to apply computational techniques to historical collections at huge scale. One legacy of this project will be a documented, accessible and reusable body of computational tools and techniques that can be applied by other institutions with digitised historical collections. Another will be increased understanding of the benefits of the methods developed by the project as a whole within the cultural heritage sector, delivered through case studies in publications and presentations.

4.   Enrich and interlinked data holdings for the British Library, to add additional context and value to content.

To date, historical records including newspapers, serials, census records, monographs, and other records exist in isolation. For serials, metadata is available only at the title level and does not reflect the content of articles, text, or images on pages. Where additional progress has been made, this has largely been through painstaking manual additions - but even where these have been done, they are not reflected in the metadata available through major collecting institutions, such as the British Library. An important contribution of the project will be to enrich and interlink the information across many of these resources. This will provide an enduring legacy even as it changes the capabilities, systems, and workflows of collecting institutions.

While the cultural sector is excited about the potential of crowdsourcing and AI-based data, it has found it difficult to operationalise and use the metadata these methods can create. The task of ingesting enriched metadata created through crowdsourcing or computational tools is still challenging for most collecting institutions. In addition to creating examples for workflows usable by other institutions, we will begin to address the challenges that mixing 'gold standard' traditional cataloguing with crowdsourced or software-led metadata creates for users of discovery interfaces.

5.   Generate new historical perspectives on the effects of the mechanisation of labour and associated changes on the lives of ordinary people during the long nineteenth century.

For more than a century, historians have been interested in investigating the impact of industrialisation and mechanisation on the working patterns and living standards of ordinary people. Scholarly interest in this area still remains very lively. Working within traditional historical paradigms, however, much of this work is local and very piecemeal in nature. This large interdisciplinary project, harnessing the skills of data scientists as well as historians, computational linguists and others permits us to address some of these questions at scale, and also allows us to pioneer new approaches. Linguistic analysis of the newspaper corpus, for example, allows us to ask how the creation and proliferation of new machinery altered human perceptions around machines and associated themes such as progress and modernity.

Meanwhile, our analysis of the demographic data contained in the census will provide new insights on a raft of questions surrounding mobility, migration, family, and work.

6.   To build capacity in the wider academic and GLAM communities in the use of digital methods and software.

If our aim is to develop data sets, tools, code, and infrastructure that can be adapted for and inspire future interdisciplinary research projects, it is vital to make sure that the communities in which we work have the technical capacity to make use of them. For this reason, we are planning to provide training in how to use our data sets, tools, code, and infrastructure via workshops and the development of teaching materials and documentation. The immediate aim of such activities is to create a larger community of scholars able to traverse disciplinary divides and push forwards new forms of interdisciplinary research, and who will be able in turn to influence scholarship and deliver training in their own institutions. Reaching GLAM - gallery, library, archive and museum - practitioners will require a different strategy than reaching academics, but running events which both can attend could lead to fruitful synergies.

But there is also a potential longer-term impact of building this capacity: it goes beyond skills to effect an attitude-shift towards collaboration, sharing data, co-publication, and the credibility given to these practices in humanities disciplines both at an institutional (e.g., how these activities are weighted for promotion) and national level (e.g. REF policy, and grant capture within funding councils).

7.   Catalyse future research.

We expect that this ambitious project will inspire and generate spin-off projects, both from within the project team, and through the data sets, tools, code, and infrastructure that we will make available. Aims 2-4 show the desire to give back to the community data and tools that can picked up and repurposed for other research questions and data sources. Moreover, aim 6 shows how we are seeking to expand the number of people able to make use of such resources.

8.   Advance public understanding of digital research in the humanities and how it can change and enhance our understanding of history.

The outreach and engagement strategy (outlined at 6.1 and 6.2) seeks to influence the public understanding of digital research. An exhibition, a popular book, and a programme of press and broadcast outputs will communicate new findings generated by our research and the innovative means by which they were reached. We will involve the public in our research not simply as end consumers, but also as co-producers. Building on the British Library's expertise in crowdsourcing and digital scholarship and emerging paradigms of 'human computation', we will deliver a public

participation and engagement programme with regional heritage partners. This online and event-led programme will enhance project data while providing meaningful access to public domain project data and research questions. We believe that such processes, because of the active engagement they require (as opposed to passive consumption) are more likely to affect knowledge acquisition and understanding.

Section 3 – Work plan

3.1 Delivery milestones

YEAR ONE (1 September 2018 – 31 March 2019: i.e. Q2-3 and Q4)

Q2-3

- Project team recruited
- Project Charter in place
- Project management systems and documentation processes in place
- Project risk register in place
- Project collaboration and communication processes in place
- Project repository for documentation, management, reporting, software, data in place

Q4

- Project team in place
- Project processes documented
- Data quality assessment process in place
- Digitisation strategy agreed
- First digitisation completed
- Usage rights secured for data to be ingested
- Initial datasets quality assessed
- Initial datasets available for use
- Initial digital infrastructure designed, documented, and in place
- Labs established
- Research agenda refined
- Initial linguistic pre-processing of data
- Communications Plan in place
- Blog posts published

YEAR TWO (1 April 2019 – 31 March 2020)

Throughout:

- Communications delivered according to plan

- Lab experiments undertaken and documented
- Research findings presented at symposia and conferences
- Engagement activities delivered as planned
- Usability and UX methods (competitor analysis, usability tests, user needs analysis etc.) undertaken and results integrated into future work
- Usage rights secured for data to be ingested

Q1

- Visual identity of project defined and implemented
- Data management plan finalised
- Release approach and plan developed for infrastructure, data, and code

Q2

- Year 2 datasets in place and available for use
- Crowdsourcing Plan in place
- Engagement Plan in place

Q3

- Year 2 data quality assessed
- First data visualisations produced and shared
- Project contribution featured in Infographics Exhibition

Q4

- Year 2 digitisation completed
- Digital infrastructure refined
- Regional Exhibition Programme in place
- Research findings presented at symposia and conferences
- First datasets shared
- First research results published (with code and data where applicable)

YEAR THREE (1 April 2020 – 31 March 2021)

- Communications delivered according to plan
- Lab experiments undertaken and documented
- Research and methodology articles drafted and submitted
- Research findings presented at symposia and conferences
- Engagement activities delivered as planned
- Technical white papers published
- Infrastructure, data, code and documentation shared, with regular updates
- Book projects scoped
- International conference launched

YEAR 4 (1 April 2021 – 31 March 2022)

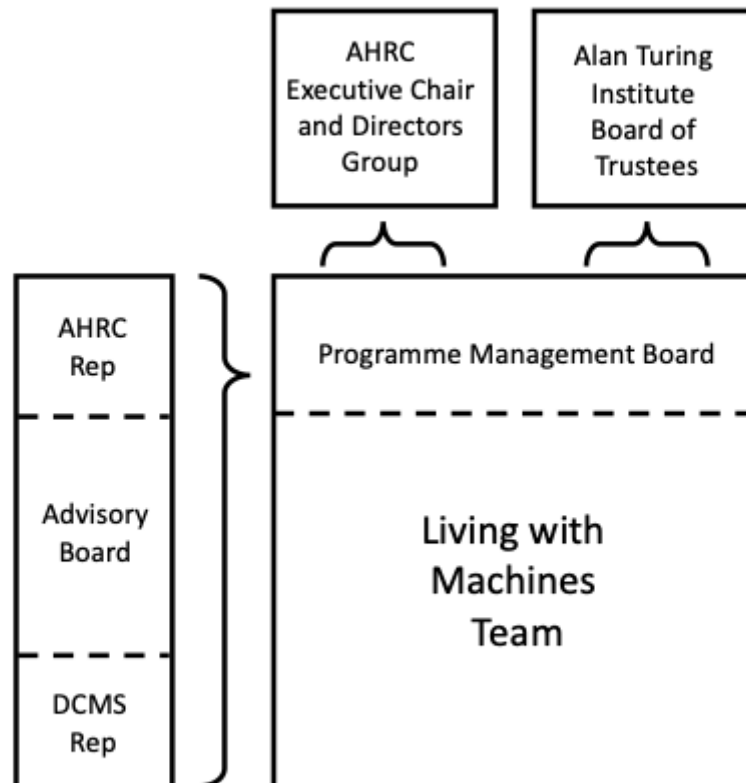- Communications delivered according to plan

- Lab experiments undertaken and documented
- Research and methodology articles drafted and submitted
- Exhibition hosted
- International conference held
- Book proposals drafted and submitted
- Technical white papers published
- Infrastructure, data, code and documentation shared, with regular updates
- Research findings presented at symposia and conferences
- Engagement activities delivered as planned
- Plan in place for long-term use, access, and preservation of all outputs
- Ingest of project data into strategic systems within the British Library piloted
- Multiple bids submitted for complementary research projects at various scales, as well as follow-on funding

YEAR FIVE (1 April 2022 – 31 March 2023)

- Communications delivered according to plan
- Research and methodology articles drafted and submitted
- Research findings presented at symposia and conferences
- Technical white papers published
- Infrastructure, data, code and documentation shared, with regular updates
- Books-length projects contracted
- Plan implemented for long-term use, access, and preservation of all outputs
- Engagement activities delivered as planned
- Best practice recommendations for research collaboration generated and published
- Ingest of project data into strategic systems within the British Library documented and delivered

3.2 Governance, management and advisory framework

```
┌──────────────┐  ┌──────────────┐
│     AHRC     │  │  Alan Turing │
│Executive Chair│  │  Institute   │
│ and Directors │  │   Board of   │
│     Group     │  │   Trustees   │
└──────────────┘  └──────────────┘
       ⎵                ⎵
┌────────────┐ ┌────────────────────────┐
│  AHRC      │⎱│                        │
│  Rep       │ │ Programme Management   │
│            │ │        Board           │
├ ─ ─ ─ ─ ─ ─┤ ├ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┤
│            │ │                        │
│ Advisory   │ │      Living with       │
│ Board      │ │      Machines          │
│            │ │        Team            │
├ ─ ─ ─ ─ ─ ─┤ │                        │
│  DCMS      │ │                        │
│  Rep       │⎰│                        │
└────────────┘ └────────────────────────┘
```

The Project Management Board is at the centre of the project's governance structure. Comprised of the team members with Investigator status and the Research Project Manager, it will meet quarterly to oversee the delivery of the project. Twice a year it will consult with the Advisory Board, and it will report on a quarterly basis to the AHRC's Executive Chair and Directors Group and The Alan Turing Institute Board of Trustees.

Project Management Board

The Alan Turing Institute's Living with Machines Project Management Board is comprised of team members with investigator status, a director from The Alan Turing Institute (Alan Wilson in the first instance), and the Research Project Manager. The Board has two broad aims:

1. To provide oversight and ensure the delivery of high quality research within the legal and financial terms, and research objectives agreed between The Alan Turing Institute, UK Research and Innovation (UKRI), and Arts and

Humanities Research Council (AHRC) for the project entitled Living with Machines (henceforth LwM); and

2. To steer the LwM team on research direction in order to deliver the project proposed and agreed by UKRI and AHRC.

The Project Management Board will meet quarterly. Following each meeting the Research Project Manager, on behalf of the Project Management Board, will draft a report to be submitted to the AHRC's Executive Chair and Directors Group and The Alan Turing Institute's Board of Trustees.

The full Terms of Reference for the Project Management Board, including details of its membership, are outlined in **Appendix 1.**

Advisory Board

The Advisory Board supports the Project Management Board and the project team. It has three broad aims:

1. To advise on the strategy for LwM in order for it to achieve its overall objectives, acting as a critical friend in regard to the overall shape and academic direction of the research programme;
2. To provide review and commentary on presentations and proposals presented by the LwM Project Management Board
3. To offer contextual advice about the wider UK research landscape and identify opportunities for consultation or collaboration with other researchers and projects.

The Advisory Board will meet twice per year. Minutes of each Advisory Board meeting will be taken and reported to the AHRC's Executive Chair and Directors Group and The Alan Turing Institute's Board of Trustees.

The full Terms of Reference for the Advisory Board, including details of its membership and representatives from AHRC and DCMS, are outlined in **Appendix 2**.

Reporting

The quarterly reports to the AHRC's Executive Chair and Directors Group and The Alan Turing Institute's Board of Trustees (following the Project Management Board meetings) will use the AHRC's standard format for a highlight report, providing:

- Progress summary: key achievements since last report and planned activities for next reporting period;
- A financial overview for the reporting period, and spend to date;
- Key risks (see 3.3 below);

- Key Delivery Milestones;
- Overall RAG [red, amber, green] status

3.3 Risk register

A full risk register will be maintained for the project. In the first instance, during the start up phase, the risk register will be populated by holding a series of pre-mortems with researchers from other large collaborative projects. The pre-mortem is a managerial strategy in which a project team imagines that a project or organisation has failed, and then works backward to determine how to avoid or remedy that failure. These strategies of risk identification can be used to populate the risk register with a list of perceived risks which can subsequently be categorised and quantified as per the UKRI's risk register template, both in terms of the probability of it occurring and the impact it would have.

We will start every Programme Management Board by consulting the risk register, and updating it with risks that have been closed, new ones that have recently been opened or need to be opened, and developing a strategy for the risks identified as most urgent. Our Agile project management framework provides a clear way of actioning these strategies, by producing tasks (or tickets) for the backlog and assigning relevant team members to those tasks. The ticketing system means that there is a clear trail of documentation left behind so that we can ensure that the risk register can be fully updated following each iteration.
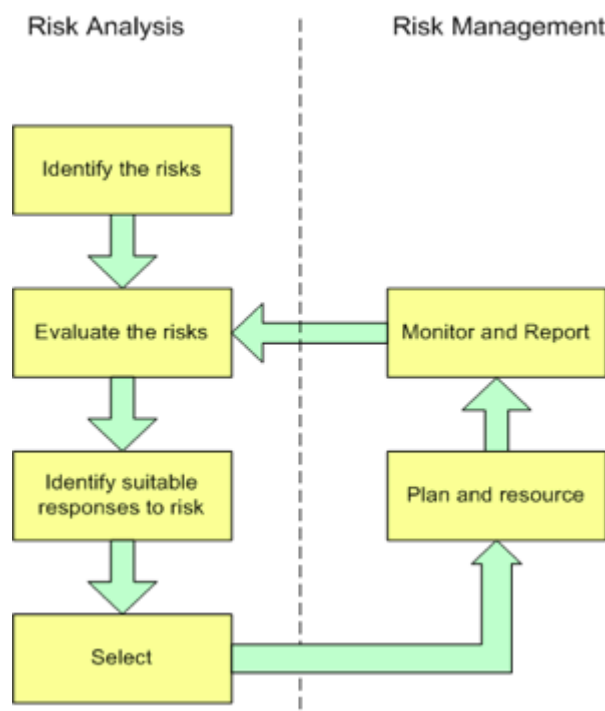
Section 4 – Finance profile

[Most of this section has been redacted but we have left an outline of each directly incurred job for reference. Other costs included directly allocated posts (named investigators), travel and subsistence; corpus building, computing equipment and infrastructure.]

Staff – Directly Incurred Posts

- **Two Postdoctoral Research Associates in History [**…] to provide historical expertise in research labs, drive research agenda and deliver outputs.
- **Postdoctoral Research Associate in Digital Humanities […]** to provide expertise in processing textual data, to contribute to research labs, drive research agenda and deliver outputs.
- **Postdoctoral Research Associate in Computational Linguistics […]** to contribute to research labs especially Lab 2: Language, drive research agenda and deliver outputs.
- **Research Project Manager […]** to manage project delivery through scheduling of and attendance at meetings and labs, to manage relationships with stakeholders, undertake project documentation, oversee the budget, and report to the AHRC.
- **Digital Curator [**…] to create, promote, and integrate digital collections, datasets and tools for academic and public users.
- **Rights Assurance Manager […]** to develop relationships with content and right holders relevant for the LwM project, and to negotiate, cost, document and clear rights.
- **Research Software Engineer […]** to ensure the delivery of digital outputs, to integrate tools and services into online interfaces, foster a technical and researcher community based on these tools and services, and ensure their legacy following the end of the grant.
- **Data and Content Manager […]** to ensure that digitised data and content is available for storage and advanced analysis, including leading the development and implementation of relevant workflows, data aggregation, transfer, cleaning, description and mapping of different formats and processes.

Section 5 – Monitoring and evaluation

The monitoring and evaluation framework outlined below employs best practice as outlined in the AHRC guide, [Understanding Your Project: A Guide to Self Evaluation](#).

5.1 Monitoring

Our monitoring framework will be determined by our KPIs, as outlined under 5.2. We will establish a monitoring tool/framework that allows for easy recording of data to provide evidence that our project goals are being met. Data will be collected annually (at a minimum) against the KPIs at activity level and recorded on Researchfish. Years are counted according to the financial year April-April, in line with our funding. The records created to fulfil our monitoring framework will be stored safely and confidentially in observance of data protection laws.

The kinds of data we will be gathering are as follows.

1. Data from full team on research activity (annually). Method: census/survey. Responsibility: each team member to report; Research Project Manager (RPM) to ensure reporting complete and results recorded in our monitoring tool. Example of data to be gathered:

- Conferences papers given, location, audience
- Speaking invitations
- Publication details
- Citation of your work
- Any media coverage
- Funding bids submitted

2. Data from the project website and GitHub repository (at least annually). Method: analytic tools/counters (e.g. Google Analytics), and use of registration for mailing list. Responsibility: RPM. Example of data to be gathered:

- Number of blog posts published
- Number of visits to the project website and individual pages/blog posts (with data on links from locations such as Twitter)
- Number of tracked public engagements with open source project code repositories (issues raised by non-project staff, `stars', and `forks'.)

3. Data on citation and reuse. Method and responsibility: collected annually through team survey (see above), and systematic searches carried out by RPM.

- Number and locations of citations
- Status of citations, determined from journal impact factor/status in field, etc.
- Downloads of code
- Details of reuse of code

4. Data on social media, journalism, broadcast and other communications. Responsibility: RPM in collaboration with the British Library and The Alan Turing Institute Comms teams.

- Details (date, location, URL etc.) of all media coverage
- For newspaper/magazine coverage: size of readership (to help estimate number reached)
- For radio and TV broadcast outputs: viewer statistics
- For Twitter: number of retweets and likes.

5. Data on academic events we host, including conferences, workshops, premortems and fellowship programme. Methods: registration and questionnaires. Responsibility: RPM

- Number of attendees
- Names and affiliations of attendees
- Feedback on the events including usefulness, lessons learned etc.

6. Gather data on engagement activities, including the exhibition, workshops, and crowdsourcing activities etc. Method: registration/booking forms, visitor books, and questionnaires (systematic sampling). Responsibility: RPM.

- Count of people on entry
- Data on attendees such as gender, age, postcode (to determine geographical reach), nature of interest in the project
- For crowdsourcing activities: patterns of usage (single time users, repeat users, and expert-level users); number of questions and comments from participants
- Feedback on the events including usefulness, lessons learned etc.

7. Data on collaborative process. Method: through team usage of GitHub (which we are using for project management, storing code, documentation etc.). Responsibility: self-documentation through use of ticketing system, creates a log of all activities undertaken, time ticket open, when closed; RPM minute-taking in fortnightly Sprint Review/Planning meetings (justification or deeming a ticket closed and discussion of items).

8. Data on reach of teaching initiatives. Method: registration/booking forms, questionnaires (systematic sampling), use of download statistics. Responsibility: RPM.

- Number of participants
- Data on backgrounds (HEI, GLAM institutional affiliation, other), gender, and geographical reach (postcode)
- Count of downloads of teaching materials

9. Data on spin-off projects and funding bids, including development of any IP or patents filed, etc. Method: maintain a log, and folder of submitted paperwork.

Responsibility: each team member to report; Research Project Manager (RPM) to ensure reporting complete.

5.2 Key performance indicators

The KPIs are organised in terms of nine overarching goals, which are derived from the project's desired areas of impact (see above). These are then broken down into indicators on which it is possible to collect quantitative data. For the gathering and assessment of contextual qualitative evidence, see 5.3-4.

[Here we have redacted specific figures against KPIs. KPI figures were grouped into columns for Year 2, 3, 4 and 'Year 5 and beyond'.]

Goal	KPI
1. To be recognised nationally and internationally for research excellence and innovation. NB. Because we are at least four different publication cultures the lead time on publication varies considerably. For this reason we will be looking at the submission rates rather than the date of publication. This also includes publication of conference proceedings (the route most esteemed in computational linguistics, for example)	High quality publications in respective fields and in interdisciplinary venues Number of conference papers given; separate national and international events attended, and national and international speaking invitations External attendees hosted through workshops, data study groups, fellowship programme, or at the project conference,
2. To enable an effective collaborative environment that facilitates interdisciplinarity, and supports and promotes all members.	Developed and published recommendations on best practice for integrating crowdsourcing and human computation with academic research questions based on complex digitised collections Other indicators: Publications and other outputs (e.g. visualisations) authored by team members from two or more different disciplines; publications placed in interdisciplinary venues PDRAs and members of team on fixed term contracts successfully transitioned into suitable employment post-project, e.g. into permanent academic/research posts or permanent industry posts
3. To develop generalisable tools, code, and infrastructure that can be adapted for and inspire future interdisciplinary research projects.	Open source code and documentation released via GitHub for initial components of project infrastructure, including linguistic pre-processing, data workflow and modelling, storage and computing architecture, machine learning components and data usage

	The above developed into a blueprint for a general AI-powered research infrastructure for digital humanities projects
	Algorithms developed to answer the project's research questions - see Goal 6.
	Code accessed, used, and repurposed
4. To develop new computational techniques to marshal and linguistically process the UK's rich archival collections (digitised and born-digital), to enable new research questions to be posed of the holdings.	Open source code and documentation made available via GitHub
	Articles published on our methods, their benefits and their findings, in the area of Natural Language Processing, data mining, and machine learning
	Access to source code, documentation, and our measures reused
5. To build capacity in the wider academic and GLAM communities in the use of digital methods and software.	Development and delivery of training schools to teach methods and software developed on project
	Number of people who attended
	Development and public release of teaching documentation
6. To generate new historical perspectives on the effects of the mechanisation of labour on the lives of ordinary people during the long nineteenth century	Papers and panels presented at conferences outlining the substantive historical potential of the project collaboration
	Research articles published on individual and smaller-scale new insights
	Major new historical perspectives developed into research monographs and popular history books
	Significant reach (in terms of audience size and geographical spread) of publicly disseminated narratives via broadcast and other engagement activities (see Contextual Evidence at 5.3-4)
7. Enriched and interlinked data holdings for the British Library, to add additional context and value to content	Percentages of created data successfully ingested into strategic British Library discovery systems
8. To catalyse future research.	Funding bids submitted by members of project team for spin-off projects, and follow-on funding
	Projects funded for members of project team

	Spin-off projects developed and funding bids submitted on related work by external individuals/groups who have participated on the project through fellowships, data study groups, or training schools
9. To advance public understanding of digital research in the humanities and how it can change and enhance our understanding of history	Significant number of people involved in our engagement activities (including crowdsourcing) Exhibition well attended Significant reach via press and broadcast (see Contextual Evidence at 5.3-4)

5.3-4 Ongoing contextual evidence and evaluation

As the AHRC guidelines on Self-Evaluation make clear that when evaluating the success of a project "the emphasis is likely to be on numerical data but depth of understanding is important at this stage. Qualitative data can be crucial in explaining what lies behind your quantitative data."

The need for both numbers and narrative is demonstrated by evaluation models such as The Kirkpatrick Model, which enumerates four levels of potential impact:

- reaction – the initial response to participation
- learning – changes in people's understanding, or raising their awareness of an issue
- behaviour – whether people subsequently modify what they do
- results – to track the long-term impacts of the project on measurable outcomes.

In terms of reaction, we would like to set objectives regarding perceived levels of enjoyment and usefulness. We are also keen to establish records of what the various audiences of our project learned. However, tracking and measuring changes in behaviour, as the AHRC guide observes:

> is resource-intensive: you'll need to know what the baselines were and will need some sort of ongoing contact to monitor change. You might rely on self-evaluation, but you may want independent verification. Either way, you will need resources and expertise capable of delivering this sort of evidence

Or these reasons we will need to ascertain our ability and commitment to gathering that kind of data. The gathering of long-term qualitative data on results is even more resource intensive, and so in this case we will rely on the quantitative KPIs.

The qualitative data needed to ascertain reaction, learning and behaviour changes, in particular the contextual evidence gathered from anecdotes and examples, is

essential for understanding the true drivers behind the numerical KPIs outlined above. Our monitoring and evaluation framework will provide a space to pull out useful narratives that explain both our successes and our failures, and what we can learn from them.

There are different solutions for gathering contextual evidence for the different participants and audiences. The following are some of the strategies we intend to, or aspire to, employ. Several of them lead on from the monitoring mechanisms described above.

1. Collaboration and the team. This is a really important part of the project for us. In addition to the quantifiable indicators listed under KPIs, how can we demonstrate how our research practices have changed and become more radically collaborative?

- Self-reflexive blog posts: all team members are producing blog posts at the beginning of the project about their background, and what they're hoping for from the collaboration. We will follow up (minimum) with mid-project and end-project posts on what those team members have learned from the process. In addition we will have an ongoing series of blog posts called 'how we collaborate' which document the measures we are using to facilitate our collaborations, the successes, the failures, and lessons learned.
- A more aspirational method that we are considering: asking an ethnographer specialising in lab culture to come and watch our process at intervals, and report on their observations.

2. The wider academic/GLAM circle around the project, participating through conferences, project fellowships, data study groups etc.

- Observation of event
- Use opening session/discussion to establish opinions and what is known at the beginning the event, and use this at conclusion to structure conversation about what has been learned.

3. Members of HEI or GLAM who participate in training events.

- Short questionnaire at time of registration to establish baseline knowledge/expectations
- Short questionnaires at the end of any training events can discover what participants learned, what they found most useful, and what they found least useful
- We might also consider following up with participants one or two years later to see if and how they had made use of any training since the event

4. Public engagement activities other than the exhibition would employ the same approach as 3 above.

5. Public attending the exhibition

- Develop evaluation questions to ask visitors before viewing the exhibition with British Library experts
- Comment/visitor books and exit questionnaires
- Build in opportunities for staff to engage with visitors

In reporting our results to the AHRC we will base it on 'Annex 3 – A sample template for reporting to funders' found in the AHRC guide on Self-Reporting. As mentioned above, we have also built into the process a mechanism for publicly reporting on aspects on our project through our blog. As well as self-reflective posts on the process of collaboration we will regularly update the blog with posts on all our events, reflecting on the difficulties and failures as well as our successes.

Section 6 - Communications and engagement

6.1 Communications strategy

We have experienced and active Communications teams supporting us both within The Alan Turing Institute and the British Library, to help field approaches from broadcast and media outlets, and to identify key audiences and opportunities to communicate our work.

An important aspect of our project communication will take place through the project website. We have been working with a branding agency to develop our website and key messages for launch in April 2019. The will, initially, be our main space to share work in progress through blogposts, to announce publications and other project outputs (code, documentation), and opportunities to engage with the project.

Our communications strategy will reflect our emphasis on agility and on reflecting and sharing lessons learnt from our 'radical collaboration'. Influenced by the 'week notes' common in digital projects and teams, we aim to 'work in public', sharing our progress via blog posts and prototypes as we proceed. Updates and case studies will be written for different audiences, including policy makers, cultural heritage and academic practitioners, and the public.

More polished research outcomes will be communicated through press and broadcast programmes. Building on the broadcast experience of Co-Is Griffin and Lawrence, and the British Library's extensive media connections, the project will seek avenues to broadcast its findings to the widest possible audience. These programmes would raise the profile of the project and set the basis for meaningful engagement in outreach activities outlined below. We will aim to tell stories that are not only about the past, but also about how the past can inform our futures.

6.2 Engagement strategy

'Living with Machines' exhibition

The UK national newspaper collection, held by the British Library, is one of the greatest of its kind in the world. Spanning more than three centuries, it comprises local, regional and national newspapers from across the UK as well as many overseas titles. The National Newspaper Building, located at the British Library's northern site at Boston Spa, West Yorkshire, opened in 2015, was purpose-built to provide the ideal environmental conditions in which to store millions of old newspapers – many of which are in a fragile state. It is one of the British Library's strategic goals to increase the visibility of this collection with local communities. The Library is in constant dialogue with local communities about the significance of this heritage resource being located in West Yorkshire and how it can be more actively used by local people. It has an active partnership with the Leeds City Council and Leeds Central Library.

It is proposed for this project that we would hold a Leeds-based 'Living with Machines' exhibition, specifically curated around the project. It will be developed by the British Library and local institutions. It will tell a story that will distil key project findings that are suitable for an exhibition format, but also co-create unique content in collaboration with key partners. The exhibition will not simply tell people about research, but will linked to our participatory initiatives and developed with local stakeholders.

In addition, other departments in the British Library will produce a small exhibition at St Pancras in 2022-23, looking into 400 years of news and newspapers in the UK with emphasis on news readership. British Library staff are collaborating with the exhibition team to explore links with our research findings.

Crowdsourcing and engagement programme

As mentioned above, we will seek to engage the public in crowdsourcing initiatives designed to enhance the project data while providing meaningful access to underlying public domain collections and datasets. This will be designed in the initial stages of the project and it might include elements such as data cleaning, name disambiguation/tracing life histories, or the verification of place names and other entities found with computational tools.

The nature of this digital infrastructure will enable us to integrate public engagement through methods such as crowdsourcing that simultaneously enhance the datasets. Digitised newspaper content is particularly popular with the genealogy communities that we will seek to involve. The British Library already has a substantial experience of such engagement through the communities tagging images on the British Library 1 million image Flickr collection, the In the Spotlight playbills transcription project, and the Georeferencer for geotagging maps.

We will explore options for coordinating this activity with members of the Living Knowledge Network - a nationwide partnership between the British Library and 21 major libraries in cities and towns across the UK, which combines local expertise and national organising power to share knowledge, resources and activities. The current partners are: Birmingham City Council, Bournemouth Borough Council, Edinburgh Council, Libraries Unlimited, Glasgow Life, Kirklees Council, Suffolk's Libraries Industrial and Provident Society (IPS), Leeds City Council, Liverpool City Council, Manchester City Council, Newcastle City Council, The Preston Harris Library, National Library of Scotland, National Library of Wales, First for Wellbeing, Norfolk County Council, Northern Ireland Library Authorities, Portsmouth City Council, Reading Borough Council, Sheffield City Council, Wakefield Council, Middlesbrough Council and Hull City Council.

As well as enhancing the project, this will have an impact through enhancing digital literacy as we will be able to inform participants through the Living Knowledge Network about key issues arising from data-driven research, thus expanding their awareness of the usefulness but also limitations of digital resources. The project will encourage critical evaluation of digital content and enable deeper public understanding of news sources in our age of social media and 'fake news'.

'Living with Machines' contributions to doctoral and learning programmes

We will arrange a series of training events, partnering with existing doctoral and postdoctoral training opportunities offered by the British Library and academic consortia to teach colleagues within higher education and the GLAM sector about the tools and infrastructure we have developed for cleaning, linking, analysing and sharing our data. We consciously do not want to limit these training opportunities to early career scholars. Changing attitudes to the value of data-driven research in the humanities requires an ideological buy-in from colleagues at all levels.

From these events we would develop documentation and training materials to release with our tools and data models, so they can be used in external workshops and other research environments, to provide self-guided training and tutorials. This documentation will also be used as the basis for a series of methods articles.

Appendix 1: Terms of Reference – The Living with Machines Project Management Board

Summary

The Alan Turing Institute's Living with Machines Project Management Board has two broad aims:

1. **To provide oversight and ensure the delivery** of high quality research programme within the legal and financial terms, and research objectives agreed between The Alan Turing Institute, UK Research and Innovation (UKRI), and Arts and Humanities Research Council (AHRC) for the project entitled Living with Machines (henceforth LwM); and
2. To steer the LwM team on **research direction** in order to deliver the project proposed and agreed by UKRI and AHRC.

Terms of Reference

- To provide oversight and ensure the delivery of a high quality research project within the legal and financial terms, and research objectives agreed between The Alan Turing Institute, UK Research and Innovation (UKRI), and the Arts and Humanities Research Council (AHRC) for the project entitled Living with Machines
- To review the project's research plan and operations at regular intervals (as determined by the PI) or when there is a particular requirement for refreshing its content, e.g. when presented with options for expanding, elaborating or otherwise changing the project as specified in the delivery plan
- To identify new opportunities for collaborative research and translation opportunities in keeping with the aims of the project as articulated in the business case
- To work with the PI to advise on the content and translation of the research generated by the project, in line with its original proposal and delivery plan
- To help develop and encourage collaborative and impact-oriented research within the project
- To help ensure that the budget is distributed to meet the objectives of the project as articulated in the spend profile for each of the five years of the project, as submitted in the Je-S application
- To identify and address risks, documenting them through the risk register (see below)
- To develop and measure Key Performance Indicators (KPIs) in dialogue with the AHRC
- To determine current and future priorities safeguarding the long-term relevance of the project

- To develop further investment plans beyond the timescale of this project

The terms of reference for the LwM Project Management Board will be reviewed annually.

Ways of working

The LwM Project Management Board meets quarterly.

Members are expected to help the LwM project and The Alan Turing Institute within which the project is delivered to develop and encourage an effective research culture with delivery of research, economic and/or societal impact at its core.

Minutes of each Project Management Board meeting will be taken and made available upon request. A report based on these minutes will be sent to the AHRC's Executive Chair and Directors Group and The Alan Turing Institute's Board of Trustees.

Membership

Representation will normally include the Principal Investigator (Ruth Ahnert, as Chair); the Research Project Manager (Andre Piza), the Co-Investigators, and one of the directors from The Alan Turing Institute (currently Alan Wilson. An alternate can be sent in the absence of the nominated representative where necessary. Meetings will be regarded as quorate only if at least 50% of members are present, and at least 50% of the partner institutions are represented.

Members:
- Dr Ruth Ahnert (Queen Mary University of London), LwM Project Management Board Chair and Principal Investigator, LwM
- Dr Giovanni Colavizza (The Alan Turing Institute), Co-Investigator, LwM
- Dr Adam Farquhar (British Library), Co-Investigator, LwM
- Professor Emma Griffin (University of East Anglia), Co-Investigator, LwM
- Dr James Hetherington (The Alan Turing Institute), Co-Investigator, LwM
- Professor Jon Lawrence (University of Exeter), Co-Investigator, LwM
- Dr Barbara McGillivray (The Alan Turing Institute/University of Cambridge), Co-Investigator, LwM
- Andre Piza (The Alan Turing Institute), Research Project Manager, LwM
- Dr Mia Ridge (British Library), Co-Investigator, LwM
- Professor Sir Alan Wilson, (The Alan Turing Institute), Director, Special Projects, and Co-Investigator, LwM

In attendance: Other staff, secondees or grantees, as required by the LwM Project Management Board Chair

Appendix 2: Terms of Reference – The Living with Machines Advisory Board

Summary

The Advisory Board's role is:
1. **To advise on the strategy** of The Alan Turing Institute's Living with Machines Project (henceforth LwM) in order for it to achieve the objectives agreed between The Alan Turing Institute, UK Research and Innovation (UKRI), and the Arts and Humanities Research Council (AHRC), and;
2. **T**o **provide review and commentary** on the work, proposals and reports shared by the LwM Project Management Board

Terms of Reference

- To advise on the strategy for LwM in order for it to achieve its overall objectives, acting as a critical friend in regard to the overall shape and academic direction of the research project;
- To review the LwM Project Management Board's research plan and operations at regular intervals (as determined by the LwM Principal Investigator) or when there is a particular requirement for refreshing its content, e.g. when presented with options for expanding, elaborating or otherwise changing the programme as specified in the LwM Delivery Plan;
- To work with the LwM Project Management Board to identify new funding opportunities, and key areas for the development of spinoffs;
- To offer contextual advice about the wider UK research landscape and identify opportunities for consultation or collaboration with other researchers and projects;
- To act as ambassadors on behalf of LwM, attending key events where possible and networking with key players/audiences on its behalf;
- To support the project in its impact-generating activities, offering advice on the development of the communication plans; supporting LwM in the dissemination its research output to key audiences and potential users of the research; offering pathways into user groups;
- To read any circulated reports and offer comments.

The terms of reference for the LwM Advisory Board will be reviewed annually.

Ways of working

The LwM Advisory Board meets twice per year.

Members are expected to help the LwM Project Management Board and The Alan Turing Institute within which the project is delivered to develop and encourage an effective research culture with delivery of research, economic and/or societal impact at its core.

Minutes of each Advisory Board meeting will be taken and reported to the AHRC's Executive Chair and Directors Group and The Alan Turing Institute's Board of Trustees.

Membership

The Advisory Board will not normally consist of more than twelve members. Members are appointed for the duration of the grant, but should a member need to resign, additional members can be appointed to retain the balance of expertise represented. Representation will normally include an appointed Chair; a representative from the AHRC; a representative from the Department for Digital Culture, Media and Sport; and a representative from Findmypast. An alternate can be sent in the absence of the nominated representative where necessary. Other interested parties may be invited to attend, with the prior knowledge of the Chair and in consultation with LwM's Principal Investigator.

Members:
- Professor Martin Daunton (University of Cambridge), Chair
- Dr Melodee Beals (Loughborough University)
- Dr Tessa Hauswedell (University College London)
- Professor Edward Higgs (University of Essex)
- Professor Tim Hitchcock (University of Sussex)
- Mr Thomas McGill (FindmyPast)
- Professor Andrew Prescott (University of Glasgow)
- Professor Tom Rodden (ex officio, DCMS)
- Professor David de Roure (University of Oxford)
- Dr James Smithies (King's College London)
- Professor Roey Sweet (ex officio, AHRC)
- Professor Melissa Terras (University of Edinburgh)
- Professor Jane Winters (School of Advanced Study)