Digital Humanities @ Oxford 2022 Programme

Contents

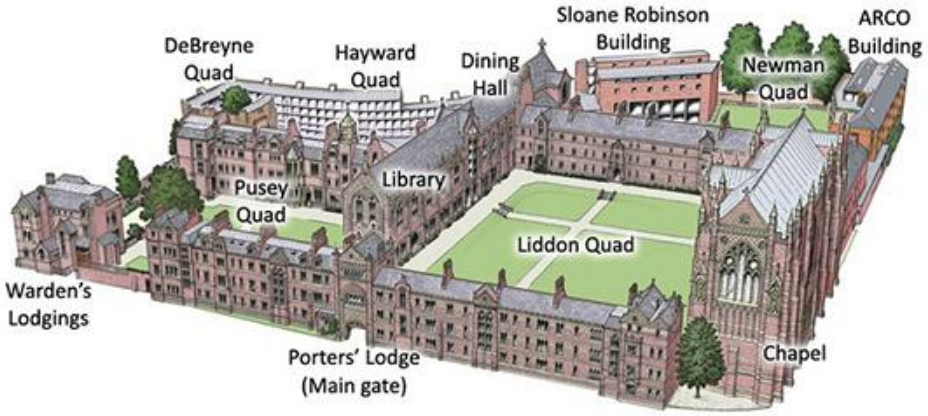
Contents 2
Maps of Keble 5
Timetable 7
Helpdesk 8
Introduction to digital humanities 9
Monday 11th July 9
Tuesday 12th July 11
Wednesday 13th July 12
Thursday 14th July 13
Friday 15th July 14
When Archives become digital 16
Monday 11th July 16
Tuesday 12th July 17
Wednesday 13th July 17
Thursday 14th July 18
Friday 15th July 20
Applied Data Analysis 21
Monday 11th July 21
Tuesday 12th July 22
Wednesday 13th July 23
Thursday 14th July 24
Friday 15th July 25

Linked data 27
Monday 11th July 27
Tuesday 12th July 28
Wednesday 13th July 29
Thursday 14th July 29
Friday 15th July 30
Crowdsourced research 31
Monday 11th July 31
Tuesday 12th July 32
Wednesday 13th July 33
Thursday 14th July 34
Friday 15th July 34
Humanities data 36
Monday 11th July 36
Tuesday 12th July 37
Wednesday 13th July 37
Thursday 14th July 38
Friday 15th July 39
Virtual strand 41
Monday 11th July 41
Tuesday 12th July 42
Wednesday 13th July 44
Thursday 14th July 44

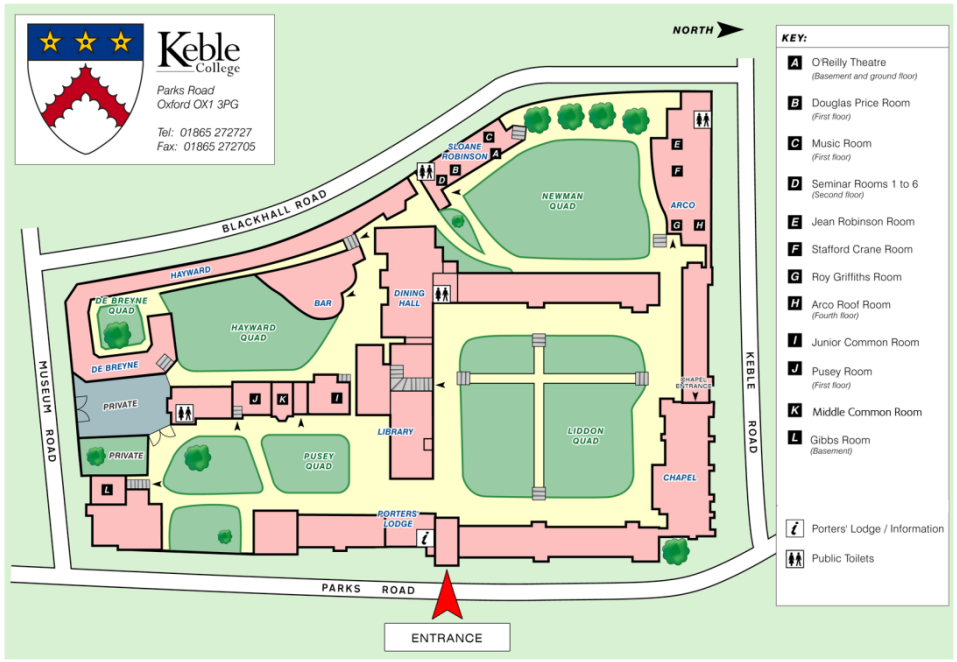
Friday 15th July 45
Text2Tech 47
Monday 11th July 47
Tuesday 12th July 47
Wednesday 13th July 48
Thursday 14th July 48
Friday 15th July 49
Introduction to TEI 50
Monday 11th July 50
Tuesday 12th July 50
Wednesday 13th July 51
Thursday 14th July 51
Friday 15th July 51
Digital Cultural Heritage 53
Monday 11th July 53
Tuesday 12th July 53
Wednesday 13th July 54
Thursday 14th July 54
Friday 15th July 55

Maps of Keble

Labelled illustration of Keble College



Overhead site map of Keble College showing building and room locations


Keble College
Parks Road
Oxford OX1 3PG
Tel: 01865 272727
Fax: 01865 272705
NORTH
KEY:
A O'Reilly Theatre (Basement and ground floor)
B Douglas Price Room (First floor)
C Music Room (First floor)
D Seminar Rooms 1 to 6 (Second floor)
E Jean Robinson Room
F Stafford Crane Room
G Roy Griffiths Room
H Arco Roof Room (Fourth floor)
I Junior Common Room
J Pusey Room (First floor)
K Middle Common Room
L Gibbs Room (Basement)
Porters' Lodge / Information
Public Toilets
BLACKHALL ROAD
HAYWARD
DE BREYNE QUAD
HAYWARD QUAD
BAR
DINING HALL
SLOANE ROBINSON
NEWMAN QUAD
ARCO
DE BREYNE
PRIVATE
PRIVATE
LIBRARY
PUSEY QUAD
LIDDON QUAD
CHAPEL ENTRANCE
CHAPEL
PORTERS' LODGE
MUSEUM ROAD
KEBLE ROAD
PARKS ROAD
ENTRANCE

Timetable

	Mon 11th	Tue 12th	Wed 13th	Thu 14th	Fri 15th
9.00-10.00	Opening keynote	Session 1	Session 1	Session 1	Session 1
10.00-10.30	Break				
10.30-11.00	Session 1	Break	Break	Break	Break
11.00-12.00		Session 2	Session 2	Session 2	Session 3
12.00-13.00	Lunch				
13.00-13.30		Lunch	Lunch	Lunch	Lunch
13.30-14.30	Session 2				
14.30-15.30		Session 3	Session 3	Session 3	Session 3
15.30-16.00	Break	Break	Break	Break	Break
16.00-17.00	Session 3				Closing keynote
17.30-18.00				DiSc Debate	
18.00-18.30	Welcome drinks				
18.30-19.00		Walking Tour			
19.00-19.30			Banquet dinner		
19.30-21.00					

Helpdesk

If you need help during the summer school you can

- Visit the registration desk in the Sloane Robinson building, Keble college
- Contact us on the Slack #helpdesk channel : https://app.slack.com/client/T03KJKDP1M1/C03K43NJ3LP
- Email us on events@eng.ox.ac.uk

If you need help about your accommodation at Keble College, please visit the Porter's Lodge

In an emergency dial 999 for the emergency services

Introduction to digital humanities

The sessions will take place in the O'Reilly Theatre

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	Lorna Richardson
10:00-10:30	Refreshment break	
10:30-12:00	What is Digital Humanities? This session is an introduction to the Introductory strand. We know Digital Humanities can be daunting – it's hard to know what it is, which tools to use, and what skills you need to learn before embarking on your first (or third, or fifth!) DH project. This session will give you an introduction to the DH project lifecycle as a starting point for planning your dissertation, thesis or next big research project. We will work together to share the kinds of data, methods, and research questions we encounter in our own disciplines, and then expand on how you can manage your DH data effectively to give you results in the short-term, and your DH project a lengthy life as a resource for other scholars.	Megan Gooch and David De Roure
12:00-13:30	Lunch	
13:30-15:30	What are Digital Editions and Why? In this 2-part workshop you will learn about the theory of digital critical editions – why we make them, and the unique challenges that face editors in the digital world. Electronic Enlightenment (EE) is a pioneering digital resource for the study of the letter and its place in the rich writing cultures of the 'long' eighteenth century. Based on the best critical editions, EE benefits from collaboration with a large network of university presses, editors, scholars, and students who have contributed texts, and continues to grow by adding more archive/print editions and commissioning born-digital correspondences.	Professor Andrew Kahn Dr Jack Orchard Mark Rogerson

	The first part of the presentation will consider what the migration of critical editions to digital platforms offers the researcher; the relationship between the digital as a replication of the print edition and the status of the digital edition; how digital resources can enhance scholarly work and what the challenges are for electronic scholarly editors in an editorial environment that is more fluid than bibliographic editing; and why dynamic databases can enable a more collaborative critical engagement of reader with editor than has been possible previously. In the second half of the presentation, the technical and content editors of Electronic Enlightenment will take attendees through the process of building the digital edition of a letter from page to publication – including metadata capture and management, html tagging, constructing biographies and adding geolocation data, and displaying scholarly annotation in a digital format.	
15:30-16:00	Refreshment break	
16:00-17:00	The Pandemic's Digital Turn – excavating socio-political affordances of digital memorialisation under COVID-19 This lecture assesses the digital turn of memorialisation in the pandemic, focusing on the politics of inclusion/exclusion evident in the design and use of the digital memorial, Remember Me. The view that "the digital" offers more pluralistic, community-focused, or democratic, routes for memorialisation has become entrenched in the pandemic. This lecture takes a critical approach to the "digital turn" of memorial work under COVID-19. I argue that, whilst the pandemic has increased acculturation, this is largely because it has also recalibrated the interfaces between cultural/digital infrastructures and memorialisation, rather than generating 'from below' memorial practices. Pandemic impact research highlights ethical and methodological challenges facing digital humanities researchers, increasing imperatives to assemble various qualitative tools to conduct multifaceted analyses, and pose humane questions of digital phenomena. I analyse the UK's most significant digital memorial – Remember Me (April 2020) – a website in remembrance of dead of COVID-19. I approach the memorial not as a "corpus" of grief discourse to be analysed computationally, but as a locus to excavate socio-political artefacts of grief from the pandemic's different	Dr Eleanor O'Keeffe

	temporal and political phases. I thus suggest how digital humanities can inform our handling of the digital afterlife of this memorial and others in the future.	
Tuesday 12th July		
Time	Session	Speaker(s)
09:00-10:30	From Clay to Bytes, Ancient Cuneiform in the Age of AI The oldest script in existence, cuneiform, was invented around 5500 years ago by the people who lived in the marshes of Southern Iraq. The Mesopotamians recorded their administrative transactions, lexical lists, epic narratives, and official texts by impressing the characteristic wedge-form signs on clay. The cuneiform corpus, composed of around half a million recovered artifacts spanning more than three millennia is a unique set of primary sources which opens up the pathway to humanity's shared cultural heritage. Digitizing these sources started early, in the seventies. Progress was not linear but approaches in digital Assyriology are now at the forefront in areas of low resource languages optical character recognition (OCR), processing, and translation, using a variety of techniques, including machine learning, while also pioneering in applications of linguistic linked open data. But these sorts of technological advances would not have been possible without the patient and sustained data collection to create digital surrogates of these artifacts often difficult to access and of which some are now lost or destroyed. Aside from the careful curation of catalog entries (metadata) and the meticulous preparation of transcriptions, digital Assyriology shines through the techniques developed to provide stunning images of cuneiform tablets, cylinder seals, and other inscribed artifacts, both in terms of capture and processing. In this presentation, we will give particular attention to ingenious imaging techniques, and have a look at machine learning applications to OCR, machine translation, and semantic analysis of cuneiform languages, while grounding these topics in the history of the discipline	Émilie Pagé-Perron
10:30-11:00	Refreshment break	
11:00-13:00	Introduction to relational databases This session looks at what a relational database is, and when and why it might be	Meriel Parick and Pamela Stanworth

	helpful to use one. It introduces some basic database concepts, and works through the process of planning one - with some hands-on exercises to help start putting the theory into practice. We also look at some challenges posed by the sort of data often used in humanities projects, and how these might be addressed.	
13:00-14:30	Lunch	
14:30-15:30	Introducing Linked Data Linked Data is a method for publishing information online. It relies on representing information through the capture of specific data points (such as people, places, and events) and the relationships between them as an interconnected graph. In this session, we will look at examples of such graphs, and evaluate the strengths and weaknesses of this methodology for knowledge representation and information publication for use in Digital Humanities projects	Terhi Nurmikko-Fuller
15:30-16:00	Refreshment break	
Wednesday 13th July		
Time	Session	Speakers
09:00-10:30	Machining the archive When is an archive digital? How does it relate to its non-digital forebear? What happens in a search? How do these questions affect how archives are used, how scholarship happens, how materials mean? These and other topics form the basis of a lecture & discussion considering some of the foundational questions of digital archival practice.	Andrew Cusworth
10:30-11:00	Refreshment break	
11:00-13:00	As open as possible, as closed as necessary: An introduction to open access publishing, open data and copyright	Ruth Mallalieu, Sarah Stewart, Chris Morrison
13:00-14:30	Lunch	
14:30-15:30	An introduction to computer vision tools for the digital humanities: How to Search, Compare, Classify and Annotate your images Computer vision has made rapid progress in recent years: images are now as readily	Dr. Giles Bergel, Dr. Abhishek Dutta, Dr. David Pinto and Prasanna Sridhar

	searchable as text is in web search engines. In this presentation, we will introduce software tools that enable researchers to organise and search large collections of images instantaneously - by allowing search queries based on images (such as a building or a book illustration) or categories (such as “gothic-architecture” or “birds”). We will demonstrate how these tools are being used in many projects within humanities disciplines such as art and book history; film studies; archaeology and literature. Attendees will leave the session knowing how to match, compare, classify and annotate many kinds of images. Since these tools are open-source, researchers can freely use them for any purpose. Attendees will have the opportunity to book an appointment to get these tools installed on their personal laptop computer, or will be provided with instructions for doing so themselves.	
15:30-16:00	Refreshment break	
Thursday 14th July		
Time	Session	Speakers
09:00-10:30	Statistics for Digital Humanists Some digital humanities approaches involve numerical tools and techniques which are not always familiar to humanities scholars. This workshop provides an introduction to some of the statistical and numerical ideas used in DH, introducing some core concepts and providing a basis for further study as desired. There will be self-paced exercises mid-session which will simply require a web browser to access an online tool. This introductory workshop is ideal for students who have not previously had an intro to statistics course, or wish to find out more about quantitative approaches in humanities.	David De Roure
10:30-11:00	Refreshment break	
11:00-13:00	An Overview of Digital Musicology and an Introduction to Music Encoding A wealth of music and music-related information is now available digitally, offering tantalising possibilities for digital musicologies. These resources include large collections of audio and scores, bibliographic and biographic data, and performance ephemera -- not to mention where music occurs within other digital materials. With	Dr Kevin Page and Dr David M. Weigl

	such large and wide ranging opportunities come new challenges in methods, principally in adapting technological solutions to assist musicologists in identifying, studying, and disseminating scholarly insights from amongst this data deluge. The first half of this session will give a whistle stop overview of some of the tools, methods, and results from digital musicology, including many introduced over five years of workshops at the Digital Humanities at Oxford Summer School. The second half of the session will focus on the Music Encoding Initiative (MEI) -- a community-driven machine-interpretable representation of music notation -- through a guided tour of online editing tools including an opportunity for attendees to try their own hand at music encoding on their laptops (please bring headphones!). The session will culminate in a demonstration from the TROMPA project combining music encoding and live performance.	
13:00-14:30	Lunch	
14:30-15:30	Does humanities 'do' reproducible research? The reproducible research movement encourages researchers to ensure that others can repeat the same analysis multiple times with the same results. But what does reproducibility look like in the context of humanities projects?	David De Roure
15:30-16:00	Refreshment break	
Friday 15th July		
Time	Session	Speakers
09:00-10:30	Building Digital Collections: Choices, Challenges and Opportunities This workshop will explore the decision-making and core considerations, processes, and technologies that go into building and delivering digital collections of different sorts – be they based on image, text, metadata, born-digital or audio-visual data – online in a way that is user-focused and sustainable. Areas covered in the session will include data types; collection scoping; rights and reuse triage; cataloguing, metadata and other Special Collections requirements; conservation assessments; imaging and metadata standards; the International Image Interoperability Framework (IIIF); sustainable technology choices; digital content creation, ingest and	Judith Siefring, Andy Irving and Tim Dungate

	management; Digital Bodleian and other core Bodleian platforms; digital collections lifecycle – retiring and archiving content. The session will be structured as a presentation on these general issues, including an in-depth look at some technology-related topics, followed by a short case study demonstrating how some of the choices and challenges are made and addressed in reality. Time will then be set aside for questions and discussion	
10:30-11:00	Refreshment break	
11:00-13:00	If technology was the answer, what was the question? Looking towards a digital future for libraries and museums During the Covid 19 pandemic libraries and museums, like many cultural organisations, pivoted to digital delivery, and researchers pivoted to digital tools and digitised resources. This accelerated shifts that was already in progress, but it also exposed gaps and inequities in digital capability and provision. What can we learn from this for libraries and museums, and for the digital humanities? This workshop will use strategic foresight approaches to think together about what the future might hold, after the pandemic.	Dr Katie Eagleton
13:00-14:30	Lunch	
14:30-15:30	The end of the introduction This session will be a chance for you to reflect on your week on the Introduction strand and consider where you might continue in your research or professional development	Megan Gooch and David De Roure
15:30-16:00	Refreshment break	
16:00-17:00	Keynote Lecture	Julianne Nyhan

When Archives become digital

Seminar room 5

Monday 11th July		
Time	Session	Speaker
09:00-10:00	Opening keynote	
10:00-10:30	Refreshment break	
10:30-12:00	You, me, & the future of history Introductions to each other, and to the week's course. This session will comprise of a short introductory lecture, time for questions, and a chance to introduce yourself and your interests to the class, so that we can gain the full benefit of a week in each other's company, and set a discussion in motion about what it means when archives become digital.	Andrew Cusworth
12:00-13:30	Lunch	
13:30-15:30	The language games we play: data, metadata, and our view of the world In this two part session, we look more closely at some of the terminologies we use in relation to digital archives in order to understand some of the core concepts at play. What is an archive? When is it digital? How do we describe collections and artifacts? What is data and how do we model it? Understanding these matters a little better will lay the groundwork for the week, its activities, and its ideas. In the second part of the session we start the word of the next session, 'All the world's a stage'.	Andrew Cusworth
15:30-16:00	Refreshment break	

16:00-17:00	All the world's a stage, And all the men and women merely players In this workshop, we will set ourselves a challenge for the week: Using a range of (relatively) free tools, we will create an online platform and introduce a data and metadata structure into which we can introduce our collection. As the week goes on, we will build on this in order to create a digital archive of 'When archives become digital'.	Andrew Cusworth
Tuesday 12th July		
Time	Session	Speaker
09:00-10:30	Digital Archival Literacy (Title & abstract TBC) In this seminar, Professor Helle Standgaard-Jensen will introduce the topic of Digital Archival Literacy and some of the key pedagogical, ethical, and knowledge challenges that face us in the creation and use of digital archives	Helle Strandgaard Jensen
10:30-11:00	Refreshment break	
11:00-13:00	An introduction to data (Pt 1, Title & Abstract TBC) In this two-part workshop, Neil Jefferies will introduce some of the fundamental principles of data to us.	Neil Jefferies
13:00-14:30	Lunch	
14:30-15:30	An introduction to data (pt 2)	Neil Jefferies
15:30-16:00	Refreshment break	
Wednesday 13th July		
Time	Session	Speaker

09:00-10:30	Machining the archive When is an archive digital? How does it relate to its non-digital forebear? What happens in a search? How do these questions affect how archives are used, how scholarship happens, how materials mean? These and other topics form the basis of a lecture & discussion considering some of the foundational questions of digital archival practice.	Andrew Cusworth
10:30-11:00	Refreshment break	
11:00-13:00	From archival collection to digital asset A visit to the imaging studios at the Bodleian Libraries will provide an opportunity to see a world-class digitisation pipeline at its most practical level, and to ask questions of those doing the actual work.	Bodleian imaging studios
13:00-14:30	Lunch	
14:30-15:30	Perfect pragmatism This session will focus on situating projects within the spectrum of digital archival practices in relation to establishing best possible practice within available means. Giving play to a shortlist of potential project paradigms from academic archival research through to digitisation by national institutions, it will address the thorny matters facing those involved in digital archival scholarship, including digitisation quality, metadata creation, and the presentation and management of digital archives.	Andrew Cusworth
15:30-16:00	Refreshment break	
Thursday 14th July		
Time	Session	Speaker

09:00-10:30	Archive of tomorrow: Health Information and Misinformation in the UK Web Archive In this workshop, we will hear about a born digital archive project relating to contemporary public health. The Bodleian Libraries are a partner in the Archive of Tomorrow project, a 14-month pilot which curates UK Web Archive materials documenting contemporary public health discourse. These materials are being used to explore the challenges of developing research collections which reflect the nature of information and misinformation in circulation online. Areas of interest to the project include approaches to collection development, options for metadata, computational analysis, ethics and rights issues, building a research network across a range of disciplines, and making recommendations to make web archives more representative, inclusive and open for health research. After a briefing on the project and its aims, students will be assigned to small breakout groups where they will be asked to respond to a series of questions around some of the central challenges of the project.	Cui Cui
10:30-11:00	Refreshment break	
11:00-13:00	Critical Mass or Critical Mess As the term 'big data' suggests, data are often most useful and informative at scale. Yet, as we work with datasets of all sizes, the noise of messy, uneven, and inconsistently processed data can introduce interference into our work. In this practical workshop, we will think about data, its consistency and quality, how that affects our work, and how we can improve things using a combination of self-aware quality control and data-cleaning tools such as OpenRefine.	Andrew Cusworth
13:00-14:30	Lunch	
14:30-15:30	Making a museum's collection accessible online In this session, Aruna Bhaugeerutty, Head of Digital Collections at the Ashmolean Museum, will introduce her team and some of their projects, and some of the issues they encounter and deal with in their work. With time for discussion of what happens when museum collections become digital.	Aruna Bhaugeerutty

15:30-16:00	Refreshment break	
Friday 15th July		
Time	Session	Speaker
09:00-10:30	The Archive and the Digital, a panel discussion In this session, we will welcome Pip Willcox, Head of Research at National Archives, and Jane Winters, Professor of Digital Humanities at School of Advanced Studies, to participate in a panel presentation and discussion, sharing their experiences of being at the forefront of the intersections of Digital Humanities and Archives.	Pip Willcox, Jane Winters
10:30-11:00	Refreshment break	
11:00-13:00	Let's get meta[data] Drawing together elements from earlier in the week, we will work to curate a digital archive of the week's course by building an online collection from media including tweets, photographs, lecture slides, and notes.	Andrew Cusworth
13:00-14:30	Lunch	
14:30-15:30	At the end of the day, you're another day wiser ... in this final session, we will try to mop up any last questions, and take a little time to reflect on the week, what we have learned, and where we might go next …	Andrew Cusworth
15:30-16:00	Refreshment break	
16:00-17:00	Closing Keynote	

Applied Data Analysis

Seminar room 6

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening Keynote	
10:00-10:30	Refreshment break	
10:30-12:00	(L1.1) * Presentation of the strand, objectives and schedule * Introductions * Examples of data analysis applications	Matteo Romanello Giovanni Colavizza
12:00-13:30	Lunch	
13:30-15:30	(L1.2) Setting up systems Python refresher Data carpentry: * Data formats (XML) & Input/Output * Read/Write The Python data analysis stack (pt. I)	Matteo Romanello
15:30-16:00	Refreshment break	

16:00-17:00	Options (attendees may pick one from the below): * Catching-up: assistance is provided to clarify any issue from the previous classes or in setting-up your Python environment. * Exercises/project: exercises or mini-projects will be provided for practice. Alternatively, attendees can bring their own mini-project to the class and work on it, individually or with others.	Matteo Romanello
Tuesday 12th July		
Time	Session	Speakers
09:00-10:30	(L2.1) Tidy data, part I: * Basic concepts of (proper/tidy) data modelling	Matteo Romanello
10:30-11:00	Refreshment break	
11:00-13:00	(L2.2) Tidy data, part II: * Manipulating data with pandas * Select-apply-transform	Matteo Romanello
13:00-14:30	Lunch	

14:30-15:30	Options (attendees may pick one from the below): * Catching-up: assistance is provided to clarify any issue from the previous classes or in setting-up your Python environment. * Exercises/project: exercises or mini-projects will be provided for practice. Alternatively, attendees can bring their own mini-project to the class and work on it, individually or with others. * Lectures at Text to Tech: attend the invited lectures given as part of the Text to Tech strand	Matteo Romanello
15:30-16:00	Refreshment break	

Wednesday 13th July

Time	Session	Speakers
09:00-10:30	(L3.1) Descriptive statistics Exploratory data analysis (part I)	Giovanni Colavizza
10:30-11:00	Refreshment break	
11:00-13:00	(L3.2) Exploratory data analysis (part II)	Giovanni Colavizza
13:00-14:30	Lunch	

14:30-15:30	Options (attendees may pick one from the below): * Catching-up: assistance is provided to clarify any issue from the previous classes or in setting-up your Python environment. * Exercises/project: exercises or mini-projects will be provided for practice. Alternatively, attendees can bring their own mini-project to the class and work on it, individually or with others. * Lectures at Text to Tech: attend the invited lectures given as part of the Text to Tech strand	Matteo Romanello
15:30-16:00	Refreshment break	
Thursday 14th July		
Time	Session	Speakers
09:00-10:30	(L4.1) Joker session: Pandas' recap/revision/dive-in	Matteo Romanello
10:30-11:00	Refreshment break	
11:00-13:00	(L4.2) Data visualization: * Primer on good and bad data visualization practices * Mapping	Matteo Romanello
13:00-14:30	Lunch	

14:30-15:30	Options (attendees may pick one from the below): * Catching-up: assistance is provided to clarify any issue from the previous classes or in setting-up your Python environment. * Exercises/project: exercises or mini-projects will be provided for practice. Alternatively, attendees can bring their own mini-project to the class and work on it, individually or with others. * Lectures at Text to Tech: attend the invited lectures given as part of the Text to Tech strand	Matteo Romanello
15:30-16:00	Refreshment break	

Friday 15th July

Time	Session	Speakers
09:00-10:30	(L5.1) Social network analysis	Giovanni Colavizza
10:30-11:00	Refreshment break	
11:00-13:00	(L5.2) Constructing datasets for research	Matteo Romanello
13:00-14:30	Lunch	

14:30-15:30	(L5.3) * Communicating data analysis results * Best practices about publishing datasets, licensing issues, reproducibility, data repositories * Q&A and farewell	Matteo Romanello Giovanni Colavizza
15:30-16:00	Refreshment break	
16:00-17:00	Closing keynote	

Linked data

Seminar room 2

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	
10:00-10:30	Refreshment break	
10:30-12:00	Introduction to Linked Data - theories and practice This session is an introduction to Linked Data and to the workshop itself. It begins with an overview of the course and provides the participants with the opportunity to outline their aims for the workshop as well. We will discuss the practical requirements of setting up, completing, and mainintaining a Linked Data project and contextualise Linked Data by comparing it to other alternatives such as relational databases and XML. We will discuss jargon, acronyms, the triple (RDF) data model, and look at examples of knowledge graphs. We will discuss how the components of instance level RDF, ontologies, SPARQL queries, and triplestores all of which form part of a cohesive workflow.	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
12:00-13:30	Lunch	
13:30-15:30	Practical session: Engaging with Existing Linked Data Participants will explore Linked Data through the four different UIs available for DBpedia: DBpedia's own resource page, OpenLink Faceted Browser, OpenLink Structured Data Editor, and the LodLive Browser.	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
15:30-16:00	Refreshment break	
16:00-17:00	Exploring Eurovision Demo session of a recent project, which has published Eurovision data as Linked Data. Participants will be introduced to the data, and we able to explore the Eurovision RDF.	Dr Terhi Nurmikko-Fuller; Dr David Weigl

Tuesday 12th July		
Time	Session	Speakers
09:00-10:30	Confidence in Cultural heritage Data To understand and develop new strategies for Cultural Heritage data, we need to critically understand its history and development. In this session we will be discussing the original modernist theories of information and data and its purpose; the first uses of computers and data in cultural heritage organisations, and the dominant reasons for its production; and the extent to which those purposes, if at all, have changed as many museums openly publish their previously internal collection data? The session will examine current challenges to cultural heritage data, particularly in relation to diversity and inclusion, raising issues that can be discussed within the group and addressing questions about the role data could and should be playing in society particularly given its growing general influence within digital information infrastructures. Finally, what role could Linked Data have in addressing new strategies for Cultural Heritage data?	Dominic Oldman Dr Diana Tanase Cristina Giancristofaro
10:30-11:00	Refreshment break	
11:00-13:00	Case Study - The Institutional use of Linked Data This second session, following on from some of the discussions in the first session , examines a case study of an organisation using Linked Data to address practical and strategic objectives addressing both internal and external processes. Linked Data is often perceived as a tool for technical people, but its real strength, we argue, is in linking people and processes. The National Archives Collection Care Department recently implemented the ResearchSpace to create a knowledge base system to support its Collection Care Department and its strategic objectives. The nature of graph based systems linked to expert knowledge is that they can evolve and change over time responding to new questions and therefore are more useful than relational databases in enabling knowledge oriented work.	Dominic Oldman Dr Diana Tanase Cristina Giancristofaro
13:00-14:30	Lunch	

14:30-15:30	The ResearchSpace system (http://www.researchspace.org) In this last session we explore the design vision behind the ResearchSpace system. ResearchSpace, is a semantic and contextually driven platform designed to allow humanities, cultural heritage and community experts to build expert knowledge bases and integrate knowledge processes. We examine the principles and methodology behind its design, reveal how Linked Data is used to drive flexible user interfaces for authoring data, and examine some of the challenges for using Linked Data in collaborative environments.	Dominic Oldman Dr Diana Tanase Cristina Giancristofaro
15:30-16:00	Refreshment break	

Wednesday 13th July

Time	Session	Speakers
09:00-10:30	Ontologies - in theory In this session, we discuss the purpose and functionality of ontologies, and look at some well-known examples of ontologies used in Digital Humanities projects and the GLAM sector.	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
10:30-11:00	Refreshment break	
11:00-13:00	Ontologies - in practice In this session, participants will develop an ontological structure of their own in groups, capturing the knowledge within a given sample dataset, or a dataset of their own (if applicable).	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
13:00-14:30	Lunch	
14:30-15:30	Implementing ontologies Participants will implement their ontologies in Protege, demonstrate their model to the rest of the group, and export it as .TTL in preparation for tomorrow's activitites.	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
15:30-16:00	Refreshment break	

Thursday 14th July

Time	Session	Speakers

09:00-10:30	RDF production In this session, particpants will be introduced to Web-Karma, which they will use to produce RDF based on the ontology they designed and implemented the day before. They will also have a chance to edit and finalise their ontology.	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
10:30-11:00	Refreshment break	
11:00-13:00	SPARQL - in theory & in practice In this session, participants will be introduced to the SPARQL query language. They will also have an opportunity to practice SPARQL queries using SPARQL Playground. At the end of this session, participants will upload the RDF triples they produced yesterday into an instance of Blazegraph.	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
13:00-14:30	Lunch	
14:30-15:30	SPARQL your own RDF In this session, participants will be able to practice writing queries over the RDF they produced yesterday.	Dr Terhi Nurmikko-Fuller; John Pybus; Graham Klyne
15:30-16:00	Refreshment break	

Friday 15th July		
Time	Session	Speakers
09:00-10:30	Introduction to Linked Data and the Linked Art data model to connect cultural heritage collections data In this session, we will introduce you to the Linked Art data model as well as some of the processes involved in connecting cultural heritage collections data using Linked Art, namely, transformation, reconciliation and visualisation. We will introduce JSON-LD, Open Refine, Python and Jupyter Notebooks in preparation for the hands-on workshop with the code notebooks.	Tanya Gray Dr Kevin Page
10:30-11:00	Refreshment break	

11:00-13:00	Connecting cultural heritage collections data with Linked Art – transformation, reconciliation and visualisation In this session, we will support you in a hands-on session looking at the processes involved in transforming cultural heritage data to Linked Data. We will use code notebooks to transform, reconcile, and visualise cultural heritage data using the Linked Art data model.	Tanya Gray Dr Kevin Page
13:00-14:30	Lunch	
14:30-15:30	Data standards in the Digital Humanities (panel discussion) In this session, we invite panellists to share their experiences with data standards in the Digital Humanities. The panellists will also be available to answer questions put forward by audience members.	Tanya Gray Dr Kevin Page
15:30-16:00	Refreshment break	
16:00-17:00	Closing Keynote	

Crowdsourced research

Seminar room 4

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	
10:00-10:30	Refreshment break	
10:30-12:00	Who We Are + Intro to Crowdsourcing - Introductions and an overview of the workshop - A brief presentation examining the history of crowdsourced research in the Humanities, and an overview of the field in its current state - Discussion: motivations for creating a crowdsourcing project: why are we all here?	Dr Samantha Blickhan
12:00-13:30	Lunch	
13:30-15:30	Translating a Research Goal into Project Design - Discussion: how to identify a research question; best practices for interrogating your data	Dr Samantha Blickhan

	- Intro to project design - SWOT analysis + SMART goals - Articulating values Zooniverse Project Builder I - Intro to the Zooniverse, the world's largest platform for crowdsourced research - Overview of the Zooniverse Project Builder (https://www.zooniverse.org/lab), the main tool we'll be using this week to prototype our projects	
15:30-16:00	Refreshment break	
16:00-17:00	Zooniverse Project Builder II - Finish overview of project builder and go over any questions - Small group activity in which participants review existing crowdsourcing projects from multiple perspectives and discuss their strengths and weaknesses	Dr Samantha Blickhan

Tuesday 12th July

Time	Session	Speakers
09:00-10:30	Subject Sets + Manifests - Learn how to upload data to the platform - Learn how to include metadata with your datasets, and what opportunities that metadata can afford you on the platform - Data management: how to prepare your project data - Trial and error: participants will upload test datasets provided by the instructor, as well as their own test data	Dr Samantha Blickhan
10:30-11:00	Refreshment break	
11:00-13:00	Intro to Workflow Development + Paper Prototyping - How to use the available resources from the platform to achieve your project goals. - Consider how to translate research questions into tasks. - Recreating example workflows to understand how the tools work.	Dr Samantha Blickhan

	- Explore some best practices for creating workflows that are approachable for a wide audience, yet which also produce quality data.	
13:00-14:30	Lunch	
14:30-15:30	Work Session I - Participants work on projects, with a focus on workflow development	Dr Samantha Blickhan + Zooniverse team members supervising
15:30-16:00	Refreshment break	

Wednesday 13th July

Time	Session	Speakers
09:00-10:30	Exporting Project Data - Learn about working with project results, including how to export data from the Project Builder - Overview of the data export format - Available options for Zooniverse data management - How results influence project design - What to do when your data isn't optimal	Dr Samantha Blickhan
10:30-11:00	Refreshment break	
11:00-13:00	Project Review and Revision - The importance of project testing and feedback - Group activity to provide feedback on prototypes: participants will engage in a small-scale version of the Zooniverse beta review process to provide and receive feedback on their prototypes - Discuss best practices for evaluating and implementing feedback	Dr Samantha Blickhan
13:00-14:30	Lunch	
14:30-15:30	Work Session II - Participants work on projects, with a focus on iteration and editing workflows	Dr Samantha Blickhan + Zooniverse team members supervising
15:30-16:00	Refreshment break	

Thursday 14th July		
Time	Session	Speakers
09:00-10:30	Launching a Project - Checklist for project launch (platform requirements) - PR and media strategies - Communicating with volunteers	Dr Samantha Blickhan
10:30-11:00	Refreshment break	
11:00-13:00	Creating Resources for Volunteers - Learn how to write a good Tutorial and how to make the best use of additional platform resources like the About page, Field Guide, Help Text - Recognition and attribution: the ethics of volunteer labor	Dr Samantha Blickhan + Zooniverse team members supervising
13:00-14:30	Lunch	
14:30-15:30	Work Session III - Participants work on projects, with a focus on creating resources for supporting volunteers and finishing prototypes in advance of tomorrow's share-out	Dr Samantha Blickhan + TBD
15:30-16:00	Refreshment break	

Friday 15th July		
Time	Session	Speakers
09:00-10:30	In For the Long Haul: Maintaining a Project - How to maintain a healthy project over a period of months, or even years - Strategies for continued engagement and communication - Tips for communicating impact and getting institutional buy-in - Considering resource needs and developing funding strategies	Dr Samantha Blickhan + TBD
10:30-11:00	Refreshment break	
11:00-13:00	Presentation of Projects and Group Discussion - Participants present their prototypes to the group	Dr Samantha Blickhan

	- What did you find easy? Difficult? - What might you do differently in the future?	
13:00-14:30	Lunch	
14:30-15:30	Wrap-Up Session: Questions & Feedback - An opportunity to ask practical questions and give feedback on the course	Dr Samantha Blickhan
15:30-16:00	Refreshment break	
16:00-17:00	Closing Keynote	

Humanities data

Pusey Room

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	
10:00-10:30	Refreshment break	
10:30-12:00	Introduction A critical review of humanities data approaches After an introduction to what's coming up during the rest of the week, Neil Jefferies will provide a critical review of humanities data approaches. Choosing the correct approach for your data can have a significant impact on the success, or otherwise, of your research, and this talk will encourage you to critically evaluate all standards and practices much as you would evaluate your scholarly sources.	The convenors Neil Jefferies
12:00-13:30	Lunch	
13:30-15:30	Open data, Creative Commons, and community collections Description forthcoming	Rowan Wilson and Ylva Berglund Prytz
15:30-16:00	Refreshment break	
16:00-17:00	Corpus linguistics Before the terms 'big data', 'data science', 'evidence-based research', 'digital humanities', 'distant reading' or 'text mining' were current, there was 'corpus linguistics', in which a digital, data-driven approach was developed to text analysis and the study of language usage. Whether it is considered as a set of methods or an academic sub-discipline, corpus linguistics continues to thrive and grow as an approach to research in linguistics and related fields. Nowadays, there are a wide range of tools - applications, interfaces, environments and platforms of various types which are available to do corpus linguistics. This presentation will offer a tour of these tools, what you can and can't do with them, and offer some suggestions and tips on how to choose among the sometimes bewildering array of options.	Martin Wynne

Tuesday 12th July		
Time	Session	Speakers
09:00-10:30	OpenRefine for tabular data OpenRefine is a powerful tool for working with messy data. It can help you explore large datasets, clean them, and transform them from one format into another.	Owen Stephens
10:30-11:00	Refreshment break	
11:00-13:00	Computer vision and machine learning This session will explore computer vision - or the extraction of data from images. Although digital humanists have a rich set of tools for searching and analysing text, images have until recently been less amenable to processing. The session will showcase tools developed by Oxford's Visual Geometry Group in collaboration with researchers in the humanities working in fields including book and art history, film, music and photography. Participants will gain an understanding of the state-of-the-art in such computer vision operations as image comparison, classification and search; of how humanists are currently using these methods; and how they can apply them to their own materials. We will use web demos and (optionally) user-installable software. No coding is required.	Giles Bergel, David Pinto, and Prasanna Sridhar
13:00-14:30	Lunch	
14:30-15:30	From project to preservation What happens to your data when your project is complete? This session provides an overview of archiving and data management from the perspective of institutional repositories.	David Tomkins
15:30-16:00	Refreshment break	

Wednesday 13th July		
Time	Session	Speakers
09:00-10:30	Framing digital objects with context and provenance Cultural and historical objects derive a lot of their meaning and interpretation from the contexts in which they are created and subsequently experienced. When digital surrogates or born-digital artefacts are created, it is important that this contextual information is also represented in the digital domain. This talk will explore the nature	Neil Jefferies

	of context and provenance (which can be seen as a historical series of contexts), and consider how they might be modelled digitally.	
10:30-11:00	Refreshment break	
11:00-13:00	Framing digital objects with context and provenance (cont.) Continuation of earlier session.	Neil Jefferies
13:00-14:30	Lunch	
14:30-15:30	Reproducibility for the humanities The reproducible research movement encourages researchers to ensure that others can repeat the same analysis multiple times with the same results. But what does reproducibility look like in the context of humanities projects?	Dave De Roure
15:30-16:00	Refreshment break	

Thursday 14th July

Time	Session	Speakers
09:00-10:30	Taylor Digital Editions This session will give an overview of the Digital Editions course taught to students at the Taylor Institution Library. It demonstrates how to create, store, preserve and publish your digital objects for free!	Emma Huber
10:30-11:00	Refreshment break	
11:00-13:00	Personal digital archiving and preservation This session introduces methods and tools for preserving your personal digital archive. Sequential steps will be introduced for forensically examining the contents of digital files, preparing documentation and best practice for storage. Participants can choose to bring their own digital files to work on or to use the sample files provided.	Edith Halvarsson
13:00-14:30	Lunch	
14:30-15:30	The Quill Project The Quill Project is an interdisciplinary digital-humanities project that examines the negotiation of legal texts. This session will discuss problems of representing these kinds of processes, designing interfaces for data-entry, collaboration with undergraduate students, and visualisation design.	Nick Cole
15:30-16:00	Refreshment break	

Friday 15th July		
Time	Session	Speakers
09:00-10:30	Scripting for specialised audio and text data Working with specialised language data often means either selecting a targeted sample from a larger, more general dataset, or building that sample from scratch. In this workshop, we will cover both ends of constructing a simple corpus. On the general-to-specific side, the first half of the session will focus on HTML scraping using Python to extract and annotate text data from digitised transcriptions. The second half of the session will introduce the general procedure of building and querying a specialised speech corpus. We will focus on how to perform forced alignment of speech and text, with hands on exercises.	Winnie Smith and Chenzi Xu
10:30-11:00	Refreshment break	
11:00-13:00	Introduction to relational databases This session looks at what a relational database is, and when and why it might be helpful to use one. It introduces some basic database concepts, and works through the process of planning one - with some hands-on exercises to help start putting the theory into practice. We also look at some challenges posed by the sort of data often used in humanities projects, and how these might be addressed.	Meriel Patrick and Pamela Stanworth
13:00-14:30	Lunch	
14:30-15:30	Wrap-up session An opportunity for an informal discussion about key issues in the world of humanities data and talking points arising from the week's presentations. What questions do those working with humanities data need to consider, and what special challenges (and opportunities) do humanities researchers face? How can data-driven humanities research best be harnessed to produce good scholarship? There will also be an opportunity for attendees to share details of work they are looking to undertake with humanities data, to think about their next steps, and to tap in to the knowledge and experience of their fellow delegates.	The convenors
15:30-16:00	Refreshment break	
16:00-17:00	Closing Keynote	

Virtual strand

Online

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	David De Roure
10:00-10:30	Refreshment break	
10:30-12:00	What is Digital Humanities? This session is an introduction to the Introductory strand. We know Digital Humanities can be daunting – it's hard to know what it is, which tools to use, and what skills you need to learn before embarking on your first (or third, or fifth!) DH project. This session will give you an introduction to the DH project lifecycle as a starting point for planning your dissertation, thesis or next big research project. We will work together to share the kinds of data, methods, and research questions we encounter in our own disciplines, and then expand on how you can manage your DH data effectively to give you results in the short-term, and your DH project a lengthy life as a resource for other scholars.	Megan Gooch and David De Roure
12:00-13:30	Break	
13:30-15:30	No session	
15:30-16:00	No session	
16:00-17:00	The Pandemic's Digital Turn – excavating socio-political affordances of digital memorialisation under COVID-19 This lecture assesses the digital turn of memorialisation in the pandemic, focusing on the politics of inclusion/exclusion evident in the design and use of the digital memorial, Remember Me. The view that "the digital" offers more pluralistic, community-focused, or democratic, routes for memorialisation has become entrenched in the pandemic. This lecture takes a critical approach to the "digital turn" of memorial	Dr Eleanor O'Keeffe

	work under COVID-19. I argue that, whilst the pandemic has increased acculturation, this is largely because it has also recalibrated the interfaces between cultural/digital infrastructures and memorialisation, rather than generating 'from below' memorial practices. Pandemic impact research highlights ethical and methodological challenges facing digital humanities researchers, increasing imperatives to assemble various qualitative tools to conduct multifaceted analyses, and pose humane questions of digital phenomena. I analyse the UK's most significant digital memorial – Remember Me (April 2020) – a website in remembrance of dead of COVID-19. I approach the memorial not as a "corpus" of grief discourse to be analysed computationally, but as a locus to excavate socio-political artefacts of grief from the pandemic's different temporal and political phases. I thus suggest how digital humanities can inform our handling of the digital afterlife of this memorial and others in the future.	

Tuesday 12th July		
Time	Session	Speakers
09:00-10:30	From Clay to Bytes, Ancient Cuneiform in the Age of AI The oldest script in existence, cuneiform, was invented around 5500 years ago by the people who lived in the marshes of Southern Iraq. The Mesopotamians recorded their administrative transactions, lexical lists, epic narratives, and official texts by impressing the characteristic wedge-form signs on clay. The cuneiform corpus, composed of around half a million recovered artifacts spanning more than three millennia is a unique set of primary sources which opens up the pathway to humanity's shared cultural heritage. Digitizing these sources started early, in the seventies. Progress was not linear but approaches in digital Assyriology are now at the forefront in areas of low resource languages	Émilie Pagé-Perron

	optical character recognition (OCR), processing, and translation, using a variety of techniques, including machine learning, while also pioneering in applications of linguistic linked open data. But these sorts of technological advances would not have been possible without the patient and sustained data collection to create digital surrogates of these artifacts often difficult to access and of which some are now lost or destroyed. Aside from the careful curation of catalog entries (metadata) and the meticulous preparation of transcriptions, digital Assyriology shines through the techniques developed to provide stunning images of cuneiform tablets, cylinder seals, and other inscribed artifacts, both in terms of capture and processing. In this presentation, we will give particular attention to ingenious imaging techniques, and have a look at machine learning applications to OCR, machine translation, and semantic analysis of cuneiform languages, while grounding these topics in the history of the discipline	
10:30-11:00	No session	
11:00-13:00	No session	
13:00-14:30	Lunch	
14:30-15:30	Introducing Linked Data Linked Data is a method for publishing information online. It relies on representing information through the capture of specific data points (such as people, places, and events) and the relationships between them as an interconnected graph. In this session, we will look at examples of such graphs, and evaluate the strengths and weaknesses of this methodology for knowledge representation and information publication for use in Digital Humanities projects	Terhi Nurmikko-Fuller
15:30-16:00	No session	

Wednesday 13th July		
Time	Session	Speakers
09:00-10:30	Machining the archive When is an archive digital? How does it relate to its non-digital forebear? What happens in a search? How do these questions affect how archives are used, how scholarship happens, how materials mean? These and other topics form the basis of a lecture & discussion considering some of the foundational questions of digital archival practice.	Andrew Cusworth
10:30-11:00	Refreshment break	
11:00-13:00	As open as possible, as closed as necessary: An introduction to open access publishing, open data and copyright	Ruth Mallalie, Sarah Stewart, Chris Morrison
13:00-14:30	Lunch	
14:30-15:30	No session	
15:30-16:00	No session	

Thursday 14th July		
Time	Session	Speakers
09:00-10:30	Statistics for Digital Humanists Some digital humanities approaches involve numerical tools and techniques which are not always familiar to humanities scholars. This workshop provides an introduction to some of the statistical and numerical ideas used in DH, introducing some core concepts and providing a basis for further study as desired. There will be self-paced exercises mid-session which will simply require a web browser to access an online tool. This introductory workshop is ideal for students who have not previously had an intro to statistics course, or wish to find out more about quantitative approaches in humanities.	David De Roure

10:30-11:00	Refreshment break	
11:00-13:00	No session	
13:00-14:30	No session	
14:30-15:30	Does humanities 'do' reproducible research? The reproducible research movement encourages researchers to ensure that others can repeat the same analysis multiple times with the same results. But what does reproducibility look like in the context of humanities projects?	David De Roure
15:30-16:00	Refreshment break	

Friday 15th July		
Time	Session	Speakers
09:00-10:30	Building Digital Collections: Choices, Challenges and Opportunities This workshop will explore the decision-making and core considerations, processes, and technologies that go into building and delivering digital collections of different sorts – be they based on image, text, metadata, born-digital or audio-visual data – online in a way that is user-focused and sustainable. Areas covered in the session will include data types; collection scoping; rights and reuse triage; cataloguing, metadata and other Special Collections requirements; conservation assessments; imaging and metadata standards; the International Image Interoperability Framework (IIIF); sustainable technology choices; digital content creation, ingest and management; Digital Bodleian and other core Bodleian platforms; digital collections lifecycle – retiring and archiving content. The session will be structured as a presentation on these general issues, including an in-depth look at some technology-related topics,	Judith Siefring Andy Irving Tim Dungate

	followed by a short case study demonstrating how some of the choices and challenges are made and addressed in reality. Time will then be set aside for questions and discussion	
10:30-11:00	Refreshment break	
11:00-13:00	No session	
13:00-14:30	Lunch	
14:30-15:30	The end of the introduction This session will be a chance for you to reflect on your week on the Introduction strand and consider where you might continue in your research or professional development	Megan Gooch David De Roure
15:30-16:00	Refreshment break	
16:00-17:00	Keynote Lecture	Julianne Nyhan

Text2Tech

Douglas Price Room

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	
10:00-10:30	Refreshment break	
10:30-12:00	Introduction to programming in Python The session provides a basic introduction to programming for digital humanities using the Python language. Among the topics covered are assignments and variables, data types, conditional statements, and reading/writing data.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
12:00-13:30	Lunch	
13:30-15:30	Introduction to programming in Python (continued) The session provides a basic introduction to programming for digital humanities using the Python language. Among the topics covered are assignments and variables, data types, conditional statements, and reading/writing data.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
15:30-16:00	Refreshment break	
16:00-17:00	Exercises and catching-up session	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni

Tuesday 12th July		
Time	Session	Speakers
09:00-10:30	Basic text processing with Python The session gives an introduction to working with text data in Python. Topics include simple regular expressions and other methods for handling text data.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
10:30-11:00	Refreshment break	
11:00-13:00	Data structures in Python This session will cover basic data structures like lists and dictionaries in Python, with practical examples.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
13:00-14:30	Lunch	

14:30-15:30	Invited talk Natural language processing for the humanities: a use case [title TBC]	TBC
15:30-16:00	Refreshment break	

Wednesday 13th July		
Time	Session	Speakers
09:00-10:30	Working with semi-structured and tabular data This session gives an introduction to working with semi-structured texts, such as XML or HTML documents. It shows how to access data via APIs and analyse the content with Pandas.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
10:30-11:00	Refreshment break	
11:00-13:00	Working with semi-structured and tabular data This session gives an introduction to working with semi-structured texts, such as XML or HTML documents. It shows how to access data via APIs and analyse the content with Pandas.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
13:00-14:30	Lunch	
14:30-15:30	Invited talk Natural language processing for the humanities: a use case [title TBC]	TBC
15:30-16:00	Refreshment break	

Thursday 14th July		
Time	Session	Speakers
09:00-10:30	Introduction to Natural Language Processing in Python This session provides an introduction to natural language processing in python, with a focus on text preprocessing, covering tasks such as lemmatization, part-of-speech tagging and named entity recognition.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
10:30-11:00	Refreshment break	
11:00-13:00	Introduction to Natural Language Processing in Python (continued) This session provides an introduction to natural language processing in python, with a focus on text preprocessing, covering tasks such as lemmatization, part-of-speech tagging and named entity recognition.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni

13:00-14:30	Lunch	
14:30-15:30	Word embeddings We offer a kind introduction to semantic analysis with word embeddings. We show how this technique can be used for humanities research, such as tracking semantic change or understanding biases in a corpus.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
15:30-16:00	Refreshment break	

Friday 15th July		
Time	Session	Speakers
09:00-10:30	Topic modelling In this session we cover how to apply topic models for understanding the content of a corpus and discover underlying trends.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
10:30-11:00	Refreshment break	
11:00-13:00	Transformers for NLP This session introduces the concept of contextualized word embeddings and shows how to use them for different tasks using the HuggingFace libraries.	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
13:00-14:30	Lunch	
14:30-15:30	Open discussion	Kaspar Beelen, Mariona Coll Ardanuy, Federico Nanni
15:30-16:00	Refreshment break	
16:00-17:00	Closing Keynote	

Introduction to TEI

Seminar room 3

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	
10:00-10:30	Refreshment break	
10:30-12:00	What is TEI and why you might want to use it The principles of XML, markup and TEI and how they relate to research questions	Yasmin Faghihi, Matthew Holford and Huw Jones
12:00-13:30	Lunch	
13:30-15:30	Exercise: Creating and editing TEI documents The basics of TEI files using the Oxygen XML editor	Yasmin Faghihi, Matthew Holford and Huw Jones
15:30-16:00	Refreshment break	
16:00-17:00	The Beckett Digital Manuscript Project. This workshop will explore how TEI has been used in a major 'genetic' editing project which aims to reconstruct the processes by which different versions of a text were created.	Dirk van Hulle

Tuesday 12th July		
Time	Session	Speakers
09:00-10:30	What to mark up and why? What are the options for TEI markup - and how to decide what markup to use	Yasmin Faghihi, Matthew Holford and Huw Jones
10:30-11:00	Refreshment break	
11:00-13:00	Exercise: A practical guide to TEI mark-up Hands on exercises and examples of TEI markup	Yasmin Faghihi, Matthew Holford and Huw Jones
13:00-14:30	Lunch	
14:30-15:30	Beyond the basics - working with the header and the body	Yasmin Faghihi, Matthew Holford and Huw Jones
15:30-16:00	Refreshment break	

Wednesday 13th July		
Time	Session	Speakers
09:00-10:30	Describing and transcribing manuscripts using TEI Using the msDesc module to encode manuscript descriptions	Yasmin Faghihi, Matthew Holford and Huw Jones
10:30-11:00	Refreshment break	
11:00-13:00	Exercise: Encoding manuscripts Practical exercises based on real life manuscripts	Yasmin Faghihi, Matthew Holford and Huw Jones
13:00-14:30	Lunch	
14:30-15:30	TEI in Taylor Editions and where digital editing can take you. This workshop will explore the journey from encoding to publication focusing on the digital editions created by students for Taylor Digital Editions (https://editions.mml.ox.ac.uk/), and will show what two students went on to do next, building on their TEI knowledge for their research	Emma Huber, Sebastian Dows-Miller and Mary Newman
15:30-16:00	Refreshment break	

Thursday 14th July		
Time	Session	Speakers
09:00-10:30	Using TEI to encode correspondence	Elizabeth Smith
10:30-11:00	Refreshment break	
11:00-13:00	Exercise: Encoding correspondence Practical examples and exercises of encoding letters with the corresp Desc module	Elizabeth Smith
13:00-14:30	Lunch	
14:30-15:30	Case study: The Darwin Correspondence Project	Elizabeth Smith
15:30-16:00	Refreshment break	

Friday 15th July		
Time	Session	Speakers

09:00-10:30	Using Xpath to query TEI	Mike Hawkins
10:30-11:00	Refreshment break	
11:00-13:00	Customising the TEI with schemas	Mike Hawkins
13:00-14:30	Lunch	
14:30-15:30	Case study: The Newton Project	Mike Hawkins
15:30-16:00	Refreshment break	
16:00-17:00	Closing Keynote	

Digital Cultural Heritage

Monday 11th July		
Time	Session	Speakers
09:00-10:00	Opening keynote	Richard Smith & Lisandra (Lia) Costiner
10:00-10:30	Refreshment break	
10:30-12:00	Introduction to Digital Cultural Heritage Introductory lecture. In this session, we will explore a variety of virtual and augmented reality applications that engage with cultural heritage objects and sites.	Lisandra (Lia) Costiner
12:00-13:30	Lunch	
13:30-15:30	Introduction to Digital Cultural Heritage (continued) Demonstrations of a variety of virtual and augmented reality applications (with Richard Smith)	Lisandra (Lia) Costiner
15:30-16:00	Refreshment break	
16:00-17:00	Introduction to Digital Cultural Heritage (continued) Cabinet (?)	

Tuesday 12th July		
Time	Session	Speakers
09:00-10:30	Introduction to Photogrammetry What is photogrammetry and its uses. Types of capturing methods.	Richard Smith & Lisandra (Lia) Costiner
10:30-11:00	Refreshment break	
11:00-13:00	A Practical Introduction to Photogrammetry [Photographing process: Demo of photogrammetry setup and how the capture process works. Explore different lighting scenarios (lightbox, natural light, interior and diffuse light). Practice capturing an object.	Richard Smith
13:00-14:30	Lunch	
14:30-15:30	A Practical Introduction to Photogrammetry (continued) Practical demo of AgiSoft. In this session we will continue working with our photogrammetry data to create a mesh and to add texture to the model. This	Richard Smith

	workshop will provide a step-by-step introduction to using the photogrammetry software Agisoft PhotoScan. The workshop will start with a practical photography session before moving on to using photographs to create 3D-models.	
15:30-16:00	Refreshment break	

Wednesday 13th July		
Time	Session	Speakers
09:00-10:30	Digital Modelling and the Reconstruction of Ancient Rome	Prof. Matthew Nicolls
10:30-11:00	Refreshment break	
11:00-13:00	Introduction to Digital Modelling in SketchUp This is an introduction to the intuitive and accessible 3D modelling software SketchUp. The aim is for participants with no previous experience of 3D work to get to grips with the software's fundamental principles and toolset, and undertake various simple 3D modelling tasks	Prof. Matthew Nicolls
13:00-14:30	Lunch	
14:30-15:30	Introduction to Digital Modelling in SketchUp (continued) In this session we will continue working with SketchUp on 3D modelling tasks	Prof. Matthew Nicolls
15:30-16:00	Refreshment break	

Thursday 14th July		
Time	Session	Speakers
09:00-10:30	A Practical Introduction to Creating Virtual Reality Applications in Unity	Richard Smith
10:30-11:00	Refreshment break	
11:00-13:00	A Practical Introduction to Creating Virtual Reality Applications in Unity (continued) We will be getting acquainted with the software Unity and setting up a basic scene	Richard Smith
13:00-14:30	Lunch	
14:30-15:30	Creating Augmented Reality Applications in Unity	Richard Smith
15:30-16:00	Refreshment break	

Friday 15th July		
Time	Session	Speakers
09:00-10:30	Introduction to 3D Printing In this session, participants will be introduced to the basics of 3D Printing, including an overview of printers, materials, software and models. Participants will learn how to choose appropriate filaments, work with slicer programs, create and implement designs. If permitted, students will also tour the Bodleian Library's 3D printing facility.	
10:30-11:00	Refreshment break	
11:00-13:00	Field Trip to a local museum to explore the use of AR in the context of a cultural institution.	
13:00-14:30	Lunch	
14:30-15:30	Kick-start your own project workshop and discussion	Richard Smith & Lisandra (Lia) Costiner
15:30-16:00	Refreshment break	
16:00-17:00	Closing Keynote	