Deliverable 3.2

Reflection Paper on

Symposium on ‘Digital Archives and Databases as a Source of Mutual Knowledge’

May 5-6, 2016

<table>
<thead>
<tr>
<th>Work Package</th>
<th>WP3 Opportunities and Priorities for Future EU-India Research Collaboration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead Beneficiary</td>
<td>ESRC</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Margot Bezzi and Aurélie Pachkoff-Singh (APRE)</td>
</tr>
<tr>
<td>Deliverable Version</td>
<td>V.1</td>
</tr>
<tr>
<td>Grant Agreement:</td>
<td>613236</td>
</tr>
<tr>
<td>Funding scheme:</td>
<td>Coordination and Support Action</td>
</tr>
<tr>
<td>Theme:</td>
<td>SSH.2013.4.3-1</td>
</tr>
</tbody>
</table>
Table of Contents

1. EXECUTIVE SUMMARY ................................................................. 3

2. INTRODUCTION AND BACKGROUND ............................................. 5

3. SYMPOSIUM “DIGITAL ARCHIVES AND DATABASES AS A SOURCE OF MUTUAL
   KNOWLEDGE” .................................................................................. 7
   Participants of the Second EqUIP Symposium ........................................ 7
   Scope of the Symposium ...................................................................... 7
   Aim of the Symposium ........................................................................ 8
   Structure of the Symposium ................................................................ 9

4. DISCUSSIONS AND MAIN FINDINGS ............................................... 10
   Opening Plenary Session .................................................................. 10
   Thematic Sessions: Take a snapshot of the current situation, existing
   resources, capacities and gaps ........................................................ 10

5. IDENTIFICATION OF COMMON RESEARCH THEMES FOR THE FUTURE .... 16

6. CONCLUDING REMARKS AND RECOMMENDATIONS .......................... 24
   Concluding Remarks ......................................................................... 24
   Recommendation for potential research themes for future EU-India
   collaboration ................................................................................. 25

7. ANNEX A. LIST OF PARTICIPANTS .................................................. 26

8. ANNEX B. AGENDA ......................................................................... 28
1. EXECUTIVE SUMMARY

This is a report of the EqUIP symposium on *Digital Archives and Databases as a Source of Mutual Knowledge* held in Rome, Italy on 5 and 6 May 2016 organised by the Agency for the Promotion of European Research (APRE). The symposium is part of the EqUIP symposia series that aims at finding future research priorities of mutual interest between EU and India in social sciences and humanities. Another important objective of the symposia series is to facilitate networking.

Thematic areas of the symposia were defined by EqUIP member organisations beforehand. Participants were selected and invited by the organiser APRE, based on proposals given by EqUIP member organisations. The symposium gathered altogether 44 participants. They represented different organisations from all over India and Europe. Of the academic participants almost half were from India.

The symposium was organised around two plenary sessions (opening and closure) and three focused topic group discussions (parallel sessions) with the following three topics:

1. Research and technological infrastructures supporting facilitated access and sharing in SSH;
2. Digitizing Cultural Heritage;

The whole symposium was chaired by Professor Ricardo Pozzo, Head of Department Social Science and Humanities, Cultural Heritage, National Research Council of Italy. Each topic group discussion has been co-chaired by a European and an Indian expert, and moderated by EqUIP staff members.

The results of the plenary sessions and the topic group discussions are gathered in this report. The symposium produced the following five recommendations for potential research themes for future collaboration. They are:

**Enabling Data Sharing and Promotion**
- Expand formal Policy frameworks for data sharing across the data access spectrum, from open to controlled data
- Critical studies on data collections
- Survey, mapping and identification of major gaps of digital datasets, existing projects and centres of excellence to promote wider availability of data

**Data Retrieval and Repurposing**
- Create new data through data retrieval and digitisation e.g. repurposing and encoding existing data to exploit new knowledge to better understand social phenomena and tackle societal challenges (for example: tracking migration and displacement flows using administrative data or; exploiting historical sources of data to capture and monitor environmental change)
- Understanding the opportunities and quality of national and localised administrative data sources to create high quality socio-economic indicators at sub-national level

**Resource: Techniques and Tools Directory**
- Create a shared knowledge database on techniques and tools for processing and analysing data

**Good data management practices**
Promote digital literacy to foster a knowledge society by uplifting professional skills in the area of data collection, management and use of data, including soft skills linked to assessing opportunities and limitations of big data

Data Curation Infrastructure Best Practices

- Best practices in curating and providing continued access to large-scale sustainable digital archives

All above mentioned research themes are discussed in more detail in the report.

The organisers would like to thank all the participants for their valuable and important inputs for this symposium. Equip will use the results of this symposium when designing a common research funding call and other future activities. More information on future calls and other activities can be read from EqUIP web pages www.equipproject.eu.
2. INTRODUCTION AND BACKGROUND

The EU-India Platform for the Social Sciences and Humanities (EqUIP) brings together research funding and support organisations in Europe and India in order to develop a stronger strategic partnership for multilateral research collaboration. Europe and India have strong historical links, but its research collaboration relationships are relatively new. EqUIP is linking and building upon successful partner relationships developed at the European level, using the NORFACE and HERA European Research Area Networks (ERA-NETs) as a basis for expanding interactions with India. As in those networks, the EqUIP platform is supporting Social Science and Humanities research funding agencies across Europe and India to build a stronger strategic partnership, to increase opportunities for networking and dialogue amongst researchers, and to explore ways of working to enable future joint research programming. A key element of the EqUIP work will be to identify opportunities and priorities for future research collaboration.

Research themes of interest to partners that would benefit from an EU-India perspective/collaboration for further exploration were identified through an initial scoping exercise with EqUIP partners, and further refined in consultation with an expert group. Five broad thematic areas (see Table 1) were refined and are outlined in full in the Scoping Report on Existing Collaboration and Future Interests and Opportunities. These broad themes will be further explored and developed through a series of symposia events with the research community. Five thematic symposia will be organised as part of the EqUIP project activities, to inform development of these themes and widen academic networks. The final symposium will bring together the findings from these events. Discussions from each symposium will be captured in a ‘Reflection paper’ with the aim of guiding future international research collaboration initiatives.

<table>
<thead>
<tr>
<th>Table 1: EqUIP Priority Themes</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inequalities, Growth and Place/Space</td>
<td>19-20 October 2015</td>
</tr>
<tr>
<td>Digital Archives and Databases as a Source of Mutual Knowledge</td>
<td>5-6 May 2016</td>
</tr>
<tr>
<td>Sustainable Prosperity, Wellbeing and Innovation</td>
<td>9-10 June 2016</td>
</tr>
<tr>
<td>Social Transformations, Cultural Expressions, Cross-Cultural Connections and Dialogue</td>
<td>28-29 June 2016</td>
</tr>
<tr>
<td>Power Structures, Conflict Resolution and Social Justice</td>
<td>13-14 October 2016</td>
</tr>
<tr>
<td>Final Summary Symposia</td>
<td>25-26 October 2016</td>
</tr>
</tbody>
</table>

The aim of this series of Symposia is to facilitate expert discussions which will develop recommendations to the EqUIP partners about priority areas and opportunities for future joint India-Europe research activities. The symposia series provide opportunities to:

- stimulate the networking of existing collaborative projects working in areas of mutual interest;
- create new networks to identify and explore the state of the art research needs a thematic area;

1 Scoping Report On Existing Collaboration And Future Interests And Opportunities
• strengthen the production, use and communication of existing research findings to policy makers and practitioners;

• share experiences of the challenges of undertaking collaborative research between Europe and India, and how these challenges can be overcome;

• and most importantly, to identify broad research priorities for possible future collaborative initiatives between research funders.

Discussions from each symposium have been captured in a ‘Reflection paper’ for each event with the aim of guiding future international research collaboration initiatives.
3. SYMPOSIUM “DIGITAL ARCHIVES AND DATABASES AS A SOURCE OF MUTUAL KNOWLEDGE”

Participants of the Symposium

This report outlines discussions at second EqUIP Symposium on Digital Archives and Databases as a Source of Mutual Knowledge held in Rome on May 5-6, 2016 organized and hosted by the Agency for the Promotion of European Research (APRE). Fifty participants, nominated by partners from India and Europe, accepted the invitation to attend the symposium; 44 attended the symposium (see Annex B for a full list of participants). The organisers ensured a good balance between Europeans and Indians participants and of Social Sciences and Humanities experts in each group. All participants were willing to share their views, drivers, approaches and experiences at the symposium, and contributed to recommendations to the platform.

Scope of the Symposium

The challenge

The pervasiveness of digital technologies has led to transformation of processes in all sectors, including research. The cornerstone of this thematic area is enabling researchers to access and utilise the data they want to in order to carry out excellent research. This challenge requires the political will to make data available, the appropriate technology to archive, curate, and store them in a repository that is easily accessible, and the capacity and capability of researchers to be able to use the data they hold. Data repositories should be established in such a way as to enable open access where appropriate, guarantee the retention of local knowledge and preservation of sources. It also requires innovation in methodology and skills to handle large datasets and new forms of data.

During the initial EqUIP scoping exercise, the following issues were raised for this theme for further exploration:

- How to create nodal digital centres for archives and ensure (open) access to data sources for Indian and EU researchers? Creating both a general centre that provides open access to researchers and an online repository of available databases to which open access is ensured (including enabling mutual data sharing and access between archives in the EU and in India).
- How digitisation of resources supports different languages and data? Digitalisation/Digitisation refers to the process of converting something from analogue (incl. image, video, text, sound), into digital formats.
- The value of preserving digitised content: Accessibility and durability of digitised content (e-resources) is vital in order to safeguard the use by future generations of researchers.
- Changing research methods and methodologies needs and challenges: Cross-cutting issues related to the development of new tools, methodologies and strategies for research, including the use and storage of new forms of data (open data, and big data, such as administrative data held by
governments, business data, transport and travel data, data generated by mobile phones, GPS location technologies, internet interactions, sensors etc.) shall be explored, to ensure that researchers have the capacity and capability to utilise data in the ways they would like to.

Priorities flagged by the EqUIP expert as relevant to this theme were:

- Historiography and Historical Methods
- Historical Databases and Oral History Archives
- Archives and Digital Humanities
- Information Infrastructures on Social Sciences and Humanities
- Storing, curating and making available ‘new’ data (e.g. administrative data, business data, ‘born digital’ data)
- Equitable access to data

This initial scoping suggested that a dual focus for discussions was needed, both around the development of infrastructure (including data and digitization), and the use of these resources by the research communities. Due to the dual emphasis on infrastructure and use, the event aimed to explore these issues, allowing for focussed sector and cross-sectoral discussion and networking.

**Aim of the Symposium**

The symposium aimed to explore the impact of the progressive and pervasive digital transformation of society in the context of social science and humanities research, analysing the whole range of dimensions where the digitisation and digitalisation have a weight in re-shaping the future of research, taking into account the international perspective.

In particular, the symposium aimed to gain:

- A full and nuanced view of research and infrastructure capacities, opportunities and challenges connected to digitalisation/digitisation for social science and humanities research, and the identification of the main barriers to digital archives and databases in the EU and India.
- Clear recommendations for EqUIP partners and funders on which future priority areas for joint India-EU research activities and collaboration in this area, ensuring research excellence and without duplicating national activities.

In particular the main questions the Symposium aimed to answer were:

- What are the research and infrastructure capacities?
- What are the opportunities and challenges connected to digitalisation for social science and humanities research?
- Which are the main barriers to a holistic approach to digital archives?
- What are the main factors to consider relating to the different geographic areas (e.g. India and Europe)?

This meant discussing the pervasiveness and impact of the digital transition across the following societal levels:

- Policy, cultural and legal framework (commitment to transparency, openness, sharing).
• ICT technical assets and requirements (systems’ capacity and interoperability) for accessing, sharing and correctly storing and preserving data in a digitised form.

• Methods and tools to digitise different cultural assets (in different languages or forms - images, video, and voices) in a way that is suitable for research and sharing.

• New digitally-based research methods (digging to the power of data – open data, big data).

Structure of the Symposium

The symposium was organised around two plenary sessions (opening and closure) and around three group focused discussions (parallel sessions). Participants were invited to choose their preferred small group discussions to attend, focused on the following topics:

1. Research and technological infrastructures supporting facilitated access and sharing in SSH (topic 1)

2. Digitising (Multi) Cultural Heritage (topic 2)

3. New digital research methods for social sciences and humanities in a digitalised world (topic 3)

Each parallel session was organised as follows and aimed to:

- **Take a snapshot of the current situation**, existing resources, capacities and gaps. Each participant was asked to give a one-minute statement on the research and/or infrastructure capacities in Europe and India related to the specific topic (1, 2 or 3) highlighting gaps and opportunities.

- **Identify research priorities**. Prior to the symposium, participants were asked to submit in advance research priorities to be discussed during the event. The organisers gathered a brief outline of each research priority and distributed it to the symposium participants before the event. During the discussion, participants were invited to give a short oral presentation (max. 5 minutes) on the priorities they previously submitted, following this structure:
  • A description of the research priority
  • The rationale for the research priority and why it is important
  • The main actors in Europe and India
  • Existing or emerging cooperation
  • Why this is a common challenge and which are the opportunities
  • Main barriers/challenges

A final selection of maximum five priorities was conducted for each group, which were presented in the final plenary session.

- **Widening our perspectives on the opportunities and challenges of increased availability and use of digital resources and databases for SSH research**. The discussion was conducted through a Disney method technique, which meant imagining a desired situation, then identifying the challenges and difficulties hindering achievement of this desired situation, and finally combining the two first steps to design a realistic approach to achieving the best possible results while taking into account the existing challenges.
4. DISCUSSIONS AND MAIN FINDINGS

Opening Plenary Session

The Symposium was opened with a plenary session which comprised an introduction by the Symposium’s convenor Professor Ricardo Pozzo, (Head of Department Social Science and Humanities, Cultural Heritage, National Research Council of Italy) who introduced the 5 principal EU infrastructures with legal status of ERIC (European Research Infrastructure Consortium) relevant for SSH:

- CESSDA (Council of European Social Science Data Archives)
- European Social Survey
- SHARE (Survey of Health, Aging and Retirement in Europe)
- CLARIN (Common Language Resources and Technology Infrastructure)
- DARIAH (Digital Research Infrastructure for the Arts and Humanities)

The European Union monitors these infrastructures through the European Strategy Forum on Research Infrastructures (ESFRI).

Social cohesion, co-creation and mutual benefit were raised as important principles for the development of new forms of digital interactions. Co-creation in this context refers to the capacity of the public, of users and customers to contribute to and shape a (digital) service. In this context it was suggested that there was a need for a change in mindset around the development and use of digital resources and data infrastructures, which could be understood in this light as forms of social and cultural innovation. Three fundamental questions were raised in this context:

1) Does this innovation solve a problem?  
2) Does it save costs?  
3) Does everybody accept it?

After the opening session, the six chairs (three from the EU and three from India) introduce the three discussion topics.

Thematic Sessions

Take a snapshot of the current situation, existing resources, capacities and gaps

**TOPIC 1 - Research and technological infrastructures supporting facilitated access and sharing in SSH**

This topic focused on the level of digital infrastructure and capacities supporting digital humanities research. The issues highlighted relate the current capacity and/or needs in terms of online repositories, mutual access (also from remote areas), and knowledge sharing. This included reflection on the cultural dimension of open science and open processes as prerequisite of knowledge sharing and effective access.
Digital Divide

- The digital divide and inequality of access, needs to be acknowledged, especially in India.

Policy framework for sharing

- An improvement was seen to be needed in the policy framework at the government level (particularly pressing in India) to allow proper data sharing - data acquisition, access, and preservation – and create trusted institutions able to sensitively manage data.
- Interoperability shall include also aspects such as privacy, sensitive data treatment and aggregation. Therefore, this should be the responsibility not only of data providers, but also of policy makers, in a perspective where ICT support, not drive, the process. Thus, interoperability not only refers to at the technological level – to make data reusable and useful for other purposes – but it also relates to the legal and cultural level in order to embrace the culture of collaboration, openness and sharing and creating the basis for a evidence-based governance.

Archives

- Creation and sustainability of large-scale digital archive with a strong humanities research focus.

TOPIC 2 - Digitizing (Multi) Cultural Heritage

This topic focused on the definition of standards, models and tools for the digital representation, curation and valorisation of cultural artefacts and objects in a multicultural perspective/context (including oral history sources and ethnographies, census and other survey data, archaeological data, music, and data derived from written or spoken language. The main issues highlighted in this discussion were around:

Researchability

- The improvement of the overall quality and researchability (in terms of search) of digitised material and the need to better exploit and improve on what has already been digitised prior to digitising more in order to avoid the risk of creating a large but unusable digital archive
- Due to the variety of collections, platforms and access requirements, finding and retrieving what has been digitised is a major problem. This makes it very hard to know where the gaps are and what is missing from digitised corpora. The combined treatment of born-digital data and of digitised content should be prioritised to enable better use of existing resources. This includes, for example, a better use of metadata and metalinguistic tags for better cataloguing. This could maybe be helpful also to reconcile different collections with different levels of cataloguing.\(^3\)

Digitisation of natural language including minority languages

- The need for archives to reflect the multi ethnic/cultural/lingual nature of cultures.
- Another challenge is in closing the gap between Natural Language Processing (NLP) tools and the existing datasets, in particular for Indian languages. The Indian government has been working in this field; the CEDAC Pune is a center of excellence on NLP and it has developed many tools for Indian languages used in digitisation initiatives\(^4\).

---

\(^3\) e.g. Holocaust archives project, which uses NoSQL to work with metadata.

\(^4\) Some highlights on the current situation in India: National Museums have been active in India on manuscript digitisation, but no concerted digitisation efforts have been undertaken in other areas. The Digital Library of India, after a mass digitisation project, now contains about 1 million books. However there are shortcomings in the retrieval process. Therefore, in 2008 the “National digital preservation programme” was launched with the aim to gather international best practices in these fields and to propose a
• Participants questioned how far digital environments support minority languages’ preservation and how digitalisation could help in preserving cultural and linguistic diversity (multilinguality)? This is a central issue for India, where some tribal languages are under threat. A database and a technological support could be helpful, also for the several Indian universities that have started teaching these native languages.

Integration of old and new digitised content
• It is necessary to find ways to combine different kinds of content (digitised books, images, sounds, and so on) in rich multinational, multilingual and multisensory datasets, able to combine together existing digitised content and new digitisation in order to fill gaps.

Cultural heritage and preservation
• Digitisation is a way to preserve and protect cultural heritage, including intangible heritage, archaeology, and cultural artefacts. How will it be possible to increase the number preserved monuments in a viable and pragmatic way, also according to international best practices, in order to cover as much as possible all Indian cultural heritage? Currently, some 3,500 monuments are officially looked after in India, but many are not receiving the same level of attention as others by public authorities. How should we resource taking care of them? Better protection of intangible heritage is also necessary.

Access rights to collections
• Access and rights – which collections are digitised, and which are not? Which collections are openly accessible, or/and can be put online? Access to outsourced libraries and archives in India, especially privately-owned collections, remains difficult, and support on this front is needed. Some of these data could not be released online, but storage on CD cannot be considered either, since it will deteriorate quickly – in this case access issues are linked to preservation issues.

Community-based actions
• Networking should be interpreted for a better (digital) engagement of communities through co-creation, co-curation, and crowdsourcing:
  • Crowdsourcing could help the cataloguing of manuscripts that are still not well identified. Communities shall also be engaged as co-creators and partners.
  • Digital engagement, collaborative methodologies in co-creation of digitised resources, and the showcase of good practices can together support urgent challenges on preserving tangible and intangible heritage.
  • A networking approach is also needed at national and international level. Cooperation is the best way to share ideas and digitisation skills and learn from each other’s successes (and failures) in records management. Cultural heritage in this context should be considered world heritage and sharing heritage and archives are of mutual benefit for Europe and India.

national policy for digital archiving. After the final report in 2010 the implementation started through the establishment of a center of excellence in digital preservation and the creation of a standards committee on e-archiving procedures.

5 e.g. Linguistic Data Consortium in India has digitised documents for 24 languages, but nothing before 1990.
TOPIC 3 - New digital research methods for social sciences and humanities in a digitalized world (current situation)

This topic aimed at exploring new digitally-based methods, protocols and tools to advance the use and interpretation of digital material and cultural heritage, taking into account new types of data (for example, social, geo-location data), historical survey, and census data.

The main issues highlighted in this discussion were around:

**Data quality and management**

- Standards, Metadata and Analytics: participants highlighted the need to enhance metadata to facilitate research, utility, address inconsistencies in discoverability, to improve data quality and management, and increase awareness of digitised material.

- The need for stronger semantic data structure was also highlighted, together with the need for more powerful computational tools.

- The need for strengthening data evaluation at each stage of treatment in order to guarantee the quality of knowledge resources, was also stressed.

- The necessity of an ontology modelling process, allowing different researchers coming from different perspective to create the correct ontologies before mapping, and getting something that is more than the simple sum of the single elements. It is important to make available and accessible the proper tools to be applied to these new semantic data sets (from semantic reasoners to visualisation tools). Running projects shall be linked to these new methods in order to ensure sustainability.

- The need for sharing knowledge on methods and tools for analysis and data mining was stressed.

**Research methodology (discoverability, analysis, re-use)**

- The role of hyperlinking is changing the research process, as well as semantic computing connected to open data methodologicals.

- The need to develop standards and tools (such as specific computational methods) that meet the needs and demands of social sciences and humanities, and facilitate both the discovery, the analysis and the use/reuse of digitised contents in large datasets.

- Characteristics that were repeatedly stressed concerned: Optical Characters Recognition, Optical Handwriting Recognition and Multilingual computing as they are important in handling the very inherently diverse multilingual digitised datasets likely to be used in EU-India collaboration.

- Handwritten and textual based information shall be ideally exploited in an automated way (automatic translation, machine learning algorithm and self-organizing maps).

- Archives shall comprise audio/video/multimedia (not only textual archives).

**Interoperability and re-use**

- The need of adequate protocols and standards to allow interoperability was stressed. This would include the development of a common semantic framework, with a strong common vision of the use-reuse of the digital archives ecosystem.

- A better integration among stakeholders that should improve capacity to talk each other is necessary. This should include work to enable the integration of diverse sources including user-generated content: potentially fostering collaborative work, while mindful of the challenges this presents for evaluation and trust.
• It is important to make research data usable and useful for the society. Therefore, institutions shall aim not only to make data and information available to users and stakeholders, but also to create appropriate archiving processes that allow a meaningful integration and re-purposing of data for new knowledge creation in an open data approach.

Data preservation
• Long-life maintenance of our resources and data preservation issues is central in order to ensure their usefulness and accessibility. Long term preservation ensures stability, persistence and autonomous digital objects.
• Definition of relevant evaluation and measurement, KPIs (key performance indicators) for infrastructures and mechanisms allow long-term sustainability of infrastructures (locally defined KPIs, not standard ones)

Access and Open Access
• Practical problems around open access were stressed by participants, particularly around standards in publication and plagiarism which can be the negative side of greater data availability.
• Access to image and video data shall be made available, with appropriate descriptions allowing proper use of them.
• Shape discourse around enhanced scientific publications; what is needed to achieve this?
• Understanding technologies is a priority, and at the EU level priority shall be given to small-scale projects.
• Open access depends on the willingness to constitute open archives (mindset).

Cross-cutting issues derived from group discussions
The following is a set of issues that emerged from the discussions as recurring and shared concerns, across all the three topic discussions.

Impact of Data infrastructure and uses
• Societal impact of research. Digitisation was described as an enabler—in addressing global challenges, by creating and allowing the exploitation of new data in innovative ways. For example, digitisation of historical collections can become significant as longitudinal datasets can inform research on global challenges like climate change and environment, food security, public health, and migration over longer periods. Equally data coming from other sectors (such as the industrial sector) could be of use for SHH research.
• Understanding the impact of the digital world and digitised content on cultural and social life. Human science shall drive cultural change using new technologies as a leverage and enabler (for example big data and high performance computing), balancing any technology-driven approach. Humanities and Social Sciences, more than other disciplines, shall elaborate the moral, ethical and philosophical underpinning of the digitalised and digital world, reflecting about the ways it is changing culture and social life. Academics and professionals from this sector can help facilitate a public debate on these issues. Multinational collaboration needs to address the very different

For example: data and information related to migrations flows can be re-used, if properly extracted, for measuring short-term current migration challenges. Likewise, historical data on environmental change can be assembled in order to create database and inform current observations on climate swing.
assumptions and ethical positions that can have an influence when dealing with digitalisation of cultural heritage. This challenge requires:
— removing obstacles in Universities (curricula, training) to create positions for digital humanities to be influential in the development of the digital turn;
— Identifying socio-economic indicators (at the sub-national/regional level) that are able to capture and measure the economic impact and value of culture. This means being able not only to measure the monetary impact directly linked to the development of a cultural economy, but it also relates to the capacity to catch the wider and multi-layered complexity of cultural impact embracing concepts such as well-being, access to education, employability, innovation rates, etc.

Digital Literacy and engagement

• Improve digital literacy, data literacy, and shared skills and techniques in order to promote innovative and excellent research, and to promote evidence-based policy making and build a sustainable and inclusive societies. This include innovating university curricula in order to align the skills that our training and academic system is providing to young people/researchers aligned with the needs of a modern knowledge society, based on continuous learning and open access. This includes:
  — increasing people’s ability to draw on big data, linked data, and digitised content, and including greater awareness of issues like IPR (Intellectual Property Rights) and open access, i.e. issues that affect digitalisation and sharing of digitalised contents;
  — training of highly specialised IT people in parallel with infrastructure creation, building also the framework for their employability;
  — create positions for digital humanities to be influential;
  — creation of learning resources for outreaching and outscaling (i.e. making them available to a wider number of people).

• For selected research priorities, networking events with key actors shall be held to forge concrete proposals.

Cultural issues

It was highlighted how the uptake of new technological resources, as well as the adaptation of methodology and processes, are tightly linked to the existence of a proper cultural and value environment, that is conducive for innovation and includes all societal actors (government, private sectors, citizens). For innovation and new models to thrive, the following aspect were identified:

• Innovative mindset and capacity to recognize opportunities (versus a natural reluctance to change);
• universal access to infrastructure (broadband; interoperable interfaces), and open access to contents’
• commitment towards values such as openness, accountability and transparency;
• a legislative and regulatory enabling framework that does not hinder innovation
5. IDENTIFICATION OF COMMON RESEARCH THEMES FOR THE FUTURE

At the end of the first day each group chair provided an overview of research themes identified. Where there was overlap these were combined. Research themes were discussed using the guidelines below:

- A description of the research priority;
- The rationale for the research priority and why it is important;
- The main actors in Europe and India;
- Existing or emerging cooperation;
- Why this is a common challenge and which are the opportunities;
- Main barriers/challenges

A final selection of priorities from each group was then presented and discussed during the final plenary session.

**Topic 1: Research and technological infrastructures supporting facilitated access and sharing in SSH**

The following research themes were developed:

**Expand the Policy framework for data sharing**

- **Description:** the general objective is to improve data sharing policies, through the development of data acquisition, access preservation and sustainability policies, particularly focussed on publically funded data creation (government, admin, academic).
- **Rationale:** the expansion of data sharing will lead to wider data availability. This operation should be at government level, taking into account the needs to do it with greater transparency; security of sources, citation, credit, etc.; mutual drivers.
- **Actors:** ICSSR, CLARIN, CESSDA, SERISS, DwB (some good existing EU work on this area).
- **Challenge:** may take time to implement legal and ethics issues. It would need a timeline for expected release.

**Best practices in looking after of large-scale sustainable digital archives**

- **Description:** objectives are to share technologies and methodologies to manage trusted digital records and information, that would be free at the point of use and to include digital sensitivity review of public records.
- **Rationale:** there is a concrete risk that digital records and information will not exist or be trusted or maybe will be permanently closed.
- **Actors:** National Archives India, EU National Archives. ICSSR, Digital Preservation Coalition, CLARIN, CESSDA, ESRC UK Big Data Network, Digital Repository of Ireland, other actors.
- **Challenge:** the scope is defined data deemed to be of historical value, including large-scale and complex data.
Understanding the opportunities and quality of administrative data sources to create high quality socio-economic indicators at sub-national level

- **Description:** There is a poor quality data at local level due to lack of reach of rural populations: in fact, administrative data are collected at local and rural level in different formats by different players. The objective is influence a more systematic data collection (e.g. village level data collection by tablets), or a further digitisation of information to build up a database.

- **Rationale:** Exploit the potential and quality of sub-national admin data sources to develop quality indicators for policy and planning.

- **Actors:** ICSSR, MOSPI, NSIs, DwB Europe.

- **Challenges:** To collect statistical data (Possible that UN Statistical groups are looking at it).

Exploring suitable data and techniques for creating indicators of migration and displacement (flows)

- **Description:** to assess and collate suitable data sources for measuring these concepts to capture systematic measures and create a database (it could include surveys and mobile data sources).

- **Rationale:** there are no good measures in this field in EU or India, and it occurs to get better data on people flows from migration and from displacement. This has policy implications for management of migration/displacement and, for this purpose a comparative study would be useful.

- **Actors:** Demographers, Census units EU and India.

- **Challenges:** which data sources?

Exploiting historical sources of data to capture and monitor environmental change

- **Description:** to gather high quality and reliable data on climate change from historical sources from the past to build up accurate pictures (e.g. using historical ships logs of gardeners’ diaries to measure weather and climate). For this reason, the visualisation outputs would be useful.

- **Rationale:** to create a historical evidence base across EU and India; to demonstrate patterns and enable predictions.

- **Actors:** ICSSR, Met offices across EU, National archives and records offices, Science Academies, Climate data centres.

- **Challenges:** identifying trusted sources of data.

Shared knowledge database on techniques and tools for processing and analysing data

- **Description:** to create a shared knowledge base on methods and tools for processing, cleaning and analysis of data across the whole SSH portfolio (structured to unstructured) and to create exemplars and user journeys of methods and tools with real data types (e.g. through videos).
• **Rationale:** a knowledge exchange on techniques and tools for particular data types.

• **Actors:** NSSO, MOSPI, CESSDA, CLARIN, DARIAH, key universities in the field, other research methods networks.

• **Challenges:** How to define data types and how to bring tools to the data (tools via shared data platform) as well as data to known tools.

**TOPIC 2 - Digitizing (Multi)Cultural Heritage**

The following research themes have been developed:

**Survey of existing projects and centres of excellence**

How to make it easier for people to know what data is available to them? It is essential to capture good practices and lessons-learnt, and inform future projects and practice.

**Intangible heritage of Delhi, holistic approach**

There is the need to create models and complex and nuanced development agenda focused on the holistic dimension of digitalisation, taking into account also skills, memories, foods, and so on, as well as economics/livelihoods; modernity and tradition, and relationship between past and present.

**World Heritage sites – coordinated management plans to avoid duplication**

There are often issues with duplication of work and implementation of plans already completed, even in the most well-known sites. The question is how to prepare databases that support national and regional heritage management. A more integrated connection between management agencies to facilitate best practise is required, taking into account a broad range of approaches, considering cultural and natural heritage together with a multi-disciplinary approach.

**Co-creation and co-design of projects, databases, and digital outputs.**

The question is how make sites multi-sensory and technological, encouraging more people to access cultural heritage sites and contribute to cultural heritage archives. Geo-based technologies and engagement throughout project lifetime; participatory society and quadruple-helix approach (where government, industry, academia and civil participants work together); ground-level perspective, and society as part of the design and creation process.

**Co-creation, leveraging on IoT technologies in UK**

The project should explore the impact of “new materiality”, such as the potential of the IoT and a wider use of 3D printing to preserve and engage, especially in a perspective of tourism and
economic valorisation of areas. For instance: hear about objects, and also put your memories into them; walk around and hear stories associated, for example, with an area.

**Critical studies on data collections**

Research priority, since it is an established concept in Europe, represents a meta-process, a lens through which we can look at things. Digitisation risk to reinforce old canonicities regarding priorities. Critical studies, looking at collections of data and how they were collected and fit together, analysing as if text collections are necessary, might be useful. On this path is also the speech of whom recalled the importance to offer data to the research and to take into account other kind of data (e.g. them from industrial sector).

**Informed policy decision through open data**

Linking good practises around data to social sciences and humanities, encouraging policy decisions to be made more and more on the basis of good data. This is relevant for both EU and India, since data are pervasive in both. This shall go hand in hand with capacity-enhancement of researchers, on how to create, use, and share linked data, through appropriate standards.

**Open source analytic tools for social science data**

A further development of tools for OCR, multi-lingual search, multi-lingual computing, and multi-lingual topic modelling, is required. Some work in this direction already began (e.g. under DEITY and at CDAC Pune). These tools should help not only with analysis, but also on discoverability and indexing, at full-text level and at metadata level. Open source software development would be preferable, and challenges regarding costs of development, adaptation and ongoing support have to be addressed.

**Reflection of cultural and linguistic diversity**

There should be deeper global-scale studies on linguistic diversity and on linguistic minorities since there are many languages not enumerated in census and that are endangered. The question is how databases (in particular the existing ones) can be used in this area by students and researchers, allowing to treat languages and cultures as living entities, rather than as static or “dead”. JNU is working on endangered languages for a long time, but it only collected information since 1990 on major languages, and separately on minor languages: standardisation and collection of the right data at the right granularity are the main challenges in this topic.

**Sanskrit digitisation**

The scope is to build a Sanskrit “tree bank” out of the digitised information for facilitating study of
ancient languages. An automatic processing of ancient Indian languages has progressed in Europe in recent years, with major resources resulting (Sanskrit Library and another in Gotingen), but actually there are no NLP tools for that, no basic tools to take the work with these corpora forwards. Indian projects are working on computationally with Sanskrit (C-DAC Pune, also some work in Chennai), but only computer scientists (and not linguists) are involved in these projects.

Climate-related data

Creation of long-run datasets for major challenges like climate, food, water, health, by looking to historical collections and finding new methodologies for bringing these together with more recent datasets. The United Kingdom could be helpful in this challenge, thanks to its strong scientific expertise in the topic, also if there is a problem regarding archival materials, unfortunately dispersed in various archives that requires a “traditional” archive-digging to identify the right material and create data sets from digitised material.

Creation of a shared database on colonial periodical press

This project is important because periodicals were the space where democracy was created and nurtured, and the work will give the opportunity to provide a way of thinking about the colonial and post-colonial world. Currently national projects exists (the Union catalogue at CSIR represents a huge resource with millions of records that should be checked), but a Europe-wide approach is now needed. Moreover, new techniques to speed up indexing (e.g. OCR, summarising, topic modelling, parse page to create item-level structure for each issue) could be used in order not to limit metadata to title-level, but including also need item-level or article-level metadata. The NewsPlan mapping project for microfilm with British newspapers provides a good pre-digitisation model, that could be revised and rolled-out more widely. Moreover, a network analysis of contributors to newspapers and periodicals (using data mining techniques to look at mobility of people and information) could be implemented.

Future-proofing of technology

The challenge is how to know if digitised content is going to be accessible and useful in the future, concerning the digital preservation and the need to ensure interoperability. This requires the development of open and robust standards, software sustainability, documentation of Interconnectivity, resolution of cyber security issues, cloud computing and cloud storage. It should be verified how connectivity could vary in different kinds of institutions.

**TOPIC 3 - New digital research methods for social sciences and humanities in a digitalized world**

The following research themes have been developed:
**Produce new data (i.e.: telephony data) and exploit available datasets (extract relevant information) to better understand social phenomena (i.e.: migrations) and tackle societal challenges**

- **Priority:** Produce and exploit new forms of data
- **Description:** The use of existing datasets and newly created datasets (ranging from TELCOs and other agents actors within big data landscape) in order to use, manage and extract patterns from these data, and meaningful information and face societal challenges (migration, etc.): this is an innovative approach in using, re-using and extracting data to study and to better understand societal phenomena such as mapping dynamic populations using mobile telephone data.
- **Rationale:** Understanding mobility in broad sense, migration and diaspora. With different impact for two very different contexts: EU and India. From the methodological point of view the challenge is to find ways to extract useful information out of the available collected metadata, to allow for innovative research processes and sound preservation/access policies. The goal is to start capturing now the stream of information that will be used in 20 years’ time to fully understand migration flows. Technological infrastructure for communication and innovation.
- **Actors in EU and India:** There are disconnected activities and projects. Estonia, Naples. There are Indian collaborations in some projects.
- **Common challenge:** Data access issues, need for new analysis methods to be developed; big-data approach to foster methodological innovation and elaborate extraction and archival strategies
- **Main barriers / challenges:** data access; new methods of analysis needed (big-data); Knowledge transfer; how and what we extract and what to archive or lose forever.

  Difficulties related to this priority relates to overcoming regulatory boundaries, cultural and languages differences should be taken into account. Also the different contexts: for example the availability high band-width, universal access to digital ecosystem.
- **Notes:** Citizen science; Consider several dimension, not only technological

---

**Build interoperability (governmental, technical, etc.) to support evidence based governance**

- **Priority:** Build interoperability framework at various level (not intended as merely technical interoperability), ranging from governmental, institutional and administrative interaction to support “evidence-based” governments and governance, and enhance societies’ inclusiveness, by involving communities on this process.
- **Description:** Interoperability here is intended as the study of the context of data production, not merely considering only the technological one, but addressing also the layers of governmental and administrative interoperability. Privacy, safety and safeguard, as well as limitations are among the main drivers.
- **Rationale:** Switch from the database as-a-static-object to a knowledge-base-to-support-decisions approach. Databases, form a technological point of view, could be seen as an active set of rules, which needs limitations. Contextualise data priority and adapt it cross-cuttingly. The technological-driven track is only one of the context that has to be considered, the relevance of the context here
is bigger than the focus on the tools.

- **Actors in EU and India**: Existing cooperation, in Europe there is a lot of debate for evidence-based policy making.

- **Existing/emerging collaborations**: Wikimedia foundation, Centre for Internet & Human Rights (Berlin), Tactical Technologies (Berlin), Centre for Internet and Society (Bangalore), Centre for Media Communication and Governance, NLSUI (New Delhi), Digital Cultures Research Lab (Lüneburg), The Alternative Law Forum (Bangalore), Global Voices Online.

- **Common challenges / opportunities**: cross fertilization of Social Sciences and Humanities, promote new forms of evidence-based policies and governmental decisions, and transform knowledge into lobby.

- **Main barriers / challenges**: A merely technological dimension is not the solution. This process requires tackling various issues: government’s lack of accountability and transparency; lack of adequate legislative and regulatory frameworks (for example on DRM), or even access of legislation in certain areas; lack of interfaces; lack of innovation mindset (e.g. cultural reticence in using open data) and of a truly enabling framework for this innovation potential to really unleash.

- **Notes**: Getting evidence based governance would require generation of new learning resources, translating these resources for multi-stakeholder outreach, upscaling existing networks for knowledge transfer, building open and free publicly owned infrastructure, and participatory and public frameworks for adequate regulatory and legislative structures.

---

Transform already available digital assets (i.e.: archival collections, including various document types), into exploitable sources of information, through the application and further development techniques (i.e. OCR and other tech.). Get the contents of manuscripts and printed books available; foster the development of new forms of appropriation, use-reuse of data (i.e.: publications, translations, other remixes)

- **Priority**: We have a lot of digital assets (from archives, collection in libraries) but we should try to find a way to shift from data directly in to knowledge. This can be done through various scanning technologies – such as intelligent character recognition and writing recognition, and others - in order to extract the knowledge hidden in these available collections. The aim is to get contents out of digitised manuscripts, documents, printed-books and making it available to foster the development of new forms of appropriation, use, reuse, remix, of those data (for instance for publications, translations and other formal/informal remix of these information.

We shall find an answer to the following questions: “What do we do with 10.000 miles of archival documents?” “What to do with a Million Books?”7. The answer is to find a way to make archival (at large, including various document types) collections, getting their contents through the application of OCR to manuscripts and printed books.

- **Description**: Currently, traditional and digital humanities are not methodologically equipped to cope with the big-data turn: humanities and ICT shall work together to find a common ground and starting from the specific needs of different disciplines.

- **Rationale**: Reach a good level of terminological precision (i.e.: define “archives”, “documents”

---

7 http://www.dlib.org/dlib/march06/crane/03crane.html
according to different contexts etc.), represent contents and interpretations in a separate way, provide semantic interoperability, enable users to develop their own ontologies on top of the contents, and the necessity to foresee a common holistic semantic framework encompassing the above, also including interfaces (user interaction and experience) and multilinguality (and multiculturality) are among the goals.

• **Challenges / opportunities:** Convert digital material into readable text, easily and permanently available. Reflect on the selection criteria: should we limit this activity to a “canon” (selection) of available archival collections? We need new forms of publications (i.e.: publication as a discourse), but we should consider the fact that each kind of digital publication has its own challenges (i.e: archiving hyperlinks: do you archive only the documents or also the hyperlinks?).

• **Existing and emerging cooperation:** 5 or 6 research centers in Europe; one important one in the Netherlands, particularly for archival documents. No info known about India.

• **Common challenges and opportunities:** Reading manuscripts and documents through ICR/OCR means facing the same problems in EU and India (multilinguality, multiscrupturality, necessity to cope with various scripts in the same document, etc.): trying to address both EU and India problems would produce a more complete and overall solution (an example discussed deals with a Persian literature project which could act as a basis for this framework to be scalable and exportable in other contexts).

• **Main barriers / challenges:** High Performance Computing and Big Data Challenge; Influence the IT process, not react at it

• **NOTES:** Provide semantic interoperability developing a common holistic (multilingual and multicultural) semantic framework; represent the content and let users develop their own ontologies; consider the problems of User Interfaces.

---

*Promote Digital literacy, to align the skills to the needs of the knowledge society*

• **Priority:** Promote digital literacy

• **Description:** Digital skills for humanities: 1) scenario, 2) actors and stakeholders, 3) mapping of skills (supply and demand for skills and gaps). Instructional / education design. E-learning - blended learning - life-long learning (keep the set of skills aligned with what we need).

• **Rationale:** align the skills to the needs of the knowledge society

• **Actors in EU-INDIA:** universities, research agencies, territorial stakeholders. Indian institutional perspective.

• **Existing and emerging cooperation:** Erasmus project platform

• **Common challenges and opportunities:** new way in which we keep up to date our university system. This is not following the way society is changing. It deals with alignment between education system and society.

• **Barrier and challenges:** conservative approach.
6. CONCLUDING REMARKS AND RECOMMENDATIONS

Concluding Remarks

The following remarks below summarise the core issues raised by participants that emerged from the topic discussions:

- Open access and mutual sharing of international data and information requires a cultural shift towards the values of collaboration, sharing and openness, as well as the compatibility of policy and legal frameworks including privacy issues, data security and intellectual property, in order to be able to release data online. For a comprehensive international collaboration on research, interoperability is crucial, and not only at the technological level – to make data reusable and useful for other purposes – but also at the legal and cultural level;

- Digitisation of information brings with it several challenges-related to:
  - The nature of items digitised: due to the multi ethnic/cultural/lingual nature of cultures, digitisation methods shall be adequately represented, not only through images and texts, but also video, sounds and natural voices, including natural languages. These different types of materials require different digitisation methods, and multinational, multilingual and multisensory datasets which combine together existing digitised content and new digitisation. It has been emphasised that Optical Characters Recognition, Optical Handwriting Recognition and Multilingual computing are important in handling diverse multilingual digitised datasets likely to be used in EU-India collaboration.
  - The integration of digitised contents with digital-born contents: The combined treatment of born-digital data and of digitised content should be prioritised to enable better use of existing resources.

- Digitisation is less relevant in absence of a complete and effective retrieval process. In order to trigger the full potential of data-based research, we shall develop standards and tools (such as specific computational methods) that meet the needs and demands of social sciences and humanities, and facilitate both the research, discovery, retrieval, analysis and the use/reuse of digitised contents in large datasets. In particular, digitised content shall share a compatible and meaningful ontology modelling process. Attention shall also be given to the semantic data structure and to a better use of metadata and metalinguistic tags for effective cataloguing.

- The importance of such interventions goes beyond the traditional boundaries of Social Science and Humanities (SSH). Attention shall be given on studying the impact of the digital world and digitised content on cultural and social life. SSH shall drive cultural change using new technologies as a leverage and enabler (e.g. big data and high performance computing). SSH shall elaborate the moral, ethical and philosophical underpinning of the digitalised and digital world, reflecting about the ways it is changing culture and social life.
Recommendation for potential research themes for future EU-India collaboration

The table below summarises the five research themes recommended by participants for future EU-India research collaboration.

<table>
<thead>
<tr>
<th>1. Enabling Data Sharing and Promotion</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Expand formal Policy frameworks for data sharing across the data access spectrum, from open to controlled data</td>
</tr>
<tr>
<td>• Critical studies on data collections</td>
</tr>
<tr>
<td>• Survey, mapping and identification of major gaps of digital datasets, existing projects and centres of excellence to promote wider availability of data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. Data Retrieval and Repurposing</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Create new data through data retrieval and digitisation e.g. repurposing and encoding existing data to exploit new knowledge to better understand social phenomena and tackle societal challenges (for example: tracking migration and displacement flows using administrative data or; exploiting historical sources of data to capture and monitor environmental change)</td>
</tr>
<tr>
<td>• Understanding the opportunities and quality of national and localised administrative data sources to create high quality socio-economic indicators at sub-national level</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• Create a shared knowledge database on techniques and tools for processing and analysing data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Good data management practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Promote digital literacy to foster a knowledge society by uplifting professional skills in the area of data collection, management and use of data, including soft skills linked to assessing opportunities and limitations of big data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>5. Data Curation Infrastructure Best Practices</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Best practices in curating and providing continued access to large-scale sustainable digital archives</td>
</tr>
</tbody>
</table>
## 7. ANNEX A. List of participants

<table>
<thead>
<tr>
<th>First Name</th>
<th>Last name</th>
<th>Organisation</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jagdish</td>
<td>Arora</td>
<td>INFLIBNET Centre</td>
<td>India</td>
</tr>
<tr>
<td>Sandra</td>
<td>Ataide Lobo</td>
<td>CHAM - Portuguese Centre for Global History</td>
<td>Portugal</td>
</tr>
<tr>
<td>Margot</td>
<td>Bezzi</td>
<td>APRE</td>
<td>Italy</td>
</tr>
<tr>
<td>Sudhanshu</td>
<td>Bhushan</td>
<td>National University of Educational Planning and Administration (NUEPA)</td>
<td>India</td>
</tr>
<tr>
<td>Francesca</td>
<td>Cantone</td>
<td>Federico II Napoli University</td>
<td>Italy</td>
</tr>
<tr>
<td>Arun Kumar</td>
<td>Chakraborty</td>
<td>RRRLF, Kolkata - National Mission on Libraries</td>
<td>India</td>
</tr>
<tr>
<td>Upendra</td>
<td>Chaudhury</td>
<td>ICSSR</td>
<td>India</td>
</tr>
<tr>
<td>Fabio</td>
<td>Ciotti</td>
<td>University of Roma Tor Vergata</td>
<td>Italy</td>
</tr>
<tr>
<td>Louise</td>
<td>Corti</td>
<td>UK Data Archive - University of Essex</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Shefali Sushil</td>
<td>Dash</td>
<td>National Informatics Centre</td>
<td>India</td>
</tr>
<tr>
<td>Emiliano</td>
<td>Degl'Innocenti</td>
<td>University of Florence</td>
<td>Italy</td>
</tr>
<tr>
<td>Catherine</td>
<td>Eagleton</td>
<td>British Library</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Tomaž</td>
<td>Erjavec</td>
<td>Dept. of Knowledge Technologies, Jožef Stefan Institute</td>
<td>Slovenia</td>
</tr>
<tr>
<td>Ilaria</td>
<td>Fava</td>
<td>Italian National Research Council</td>
<td>Italy</td>
</tr>
<tr>
<td>Raman</td>
<td>Ganguly</td>
<td>Vienna University Computer Center</td>
<td>Austria</td>
</tr>
<tr>
<td>Sanjay</td>
<td>Garg</td>
<td>National Archives of India</td>
<td>India</td>
</tr>
<tr>
<td>K.S</td>
<td>James</td>
<td>Population Research Centre, Institute for Social and Economic Change</td>
<td>India</td>
</tr>
<tr>
<td>Valerie</td>
<td>Johnson</td>
<td>The National Archives</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Manija</td>
<td>Kamal</td>
<td>ESRC</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Ajay</td>
<td>Khare</td>
<td>School of Planning and Architecture</td>
<td>India</td>
</tr>
<tr>
<td>Mahesh</td>
<td>Madhukar</td>
<td>ICSSR</td>
<td>India</td>
</tr>
<tr>
<td>Jyoti Ranjan</td>
<td>Majumdar</td>
<td>M/O Statistics &amp; Programme Implementation</td>
<td>India</td>
</tr>
<tr>
<td>David</td>
<td>Martin</td>
<td>University of Southampton</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Reena</td>
<td>Marwah</td>
<td>ICSSR</td>
<td>India</td>
</tr>
<tr>
<td>Nafees</td>
<td>Meah</td>
<td>Research Council UK</td>
<td>India</td>
</tr>
<tr>
<td>AGK</td>
<td>Menon</td>
<td>INTACH Delhi Chapter</td>
<td>India</td>
</tr>
<tr>
<td>Usha</td>
<td>Munshi</td>
<td>Indian Institute of Public Administration</td>
<td>India</td>
</tr>
<tr>
<td>Appasamy</td>
<td>Murugaiyan</td>
<td>Ecole Pratique des Hautes Etudes</td>
<td>France</td>
</tr>
<tr>
<td>Jussi</td>
<td>Nuorteva</td>
<td>The National Archives of Finland</td>
<td>Finland</td>
</tr>
<tr>
<td>Aurélie</td>
<td>Pachkoff-Singh</td>
<td>APRE</td>
<td>Italy</td>
</tr>
<tr>
<td>Name</td>
<td>Affiliation</td>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Marco</td>
<td>Passarotti</td>
<td>Università Cattolica del Sacro Cuore, Italy</td>
<td></td>
</tr>
<tr>
<td>Riccardo</td>
<td>Pozzo</td>
<td>Department Social Science and Humanities, Cultural Heritage, National Research Council of Italy, Italy</td>
<td></td>
</tr>
<tr>
<td>Andrew</td>
<td>Prescott</td>
<td>University of Glasgow, United Kingdom</td>
<td></td>
</tr>
<tr>
<td>Archana</td>
<td>Saad Akhtar</td>
<td>Aga Khan Trust for culture, India</td>
<td></td>
</tr>
<tr>
<td>Massimiliano</td>
<td>Saccone</td>
<td>Italian National Research Council, Italy</td>
<td></td>
</tr>
<tr>
<td>Babu</td>
<td>Sebastian</td>
<td>State Institute of Education Technology, India</td>
<td></td>
</tr>
<tr>
<td>Nishant</td>
<td>Shah</td>
<td>Leuphana University, India</td>
<td></td>
</tr>
<tr>
<td>Peter</td>
<td>Stockinger</td>
<td>Communication et Formation Interculturelles, INALCO, France</td>
<td></td>
</tr>
<tr>
<td>Gregor</td>
<td>Strle</td>
<td>ZRC SAZU Research Centre of the Slovenian Academy of Sciences and Arts, Slovenia</td>
<td></td>
</tr>
<tr>
<td>Manfred</td>
<td>Thaller</td>
<td>Cologne University, Germany</td>
<td></td>
</tr>
<tr>
<td>Sukhadeo</td>
<td>Thorat</td>
<td>ICSSR, India</td>
<td></td>
</tr>
<tr>
<td>TP</td>
<td>Vijay</td>
<td>Kannad University, India</td>
<td></td>
</tr>
<tr>
<td>Vania</td>
<td>Virgili</td>
<td>Italian National Institute for Nuclear Physics, Italy</td>
<td></td>
</tr>
<tr>
<td>Jerca</td>
<td>Vodušek Starič</td>
<td>MIZS - external expert, Slovenia</td>
<td></td>
</tr>
<tr>
<td>Martin</td>
<td>Warnke</td>
<td>University of Lüneburg, Germany</td>
<td></td>
</tr>
<tr>
<td>Alison</td>
<td>Weir</td>
<td>ESRC, United Kingdom</td>
<td></td>
</tr>
<tr>
<td>Tamara</td>
<td>West</td>
<td>University of Birmingham, United Kingdom</td>
<td></td>
</tr>
<tr>
<td>Matthew</td>
<td>Woollard</td>
<td>University of Essex - the UK Data Service, United Kingdom</td>
<td></td>
</tr>
<tr>
<td>Gonçalo</td>
<td>Zagalo</td>
<td>FCT - Fundação para a Ciência e a Tecnologia, Portugal</td>
<td></td>
</tr>
<tr>
<td>Xenia</td>
<td>Zeiler</td>
<td>University of Helsinki, Finland</td>
<td></td>
</tr>
</tbody>
</table>
8. ANNEX B. Agenda

Day 1

Plenary session

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.00 – 9.30</td>
<td>Registration and welcome coffee</td>
</tr>
<tr>
<td>9.30 – 9.40</td>
<td>Welcome words</td>
</tr>
<tr>
<td>9.40-9.50</td>
<td>Introduction to Equip</td>
</tr>
<tr>
<td>9.50-10.10</td>
<td>Digital Humanities for Social and Cultural Innovation</td>
</tr>
<tr>
<td>10.10-10.20</td>
<td>Symposium objectives and organization</td>
</tr>
<tr>
<td>10.20-10.40</td>
<td>Introductions to the topics by the Indian and European co-chairs</td>
</tr>
<tr>
<td>10.40-11.00</td>
<td>Topic 2: Digitizing(^8) (Multi)Cultural Heritage</td>
</tr>
</tbody>
</table>

---

\(^8\) Digitization is defined as the material process of converting individual analogue streams of information into digital bits. In contrast, we refer to digitalization as the way in which many domains of social and economic life are redefined around digital communication and infrastructures.
Centre, India

11.00-11.30  Coffee break and networking


Developing new digitally-based methods, protocols and tools to advance the interpretation of digital material and cultural heritage

Dr Vania Virgili, Expert in research infrastructures for social sciences and humanities, Italian National Institute for Nuclear Physics, Italy

Dr Sanjay Garg, Deputy Director, National Archives of India, India

---

Resources, capacities and gaps

11.50-12.10  Selected examples in India and Europe:

ICSSR Data Service: a national level data repository service for the social sciences

Professor Sukhdeo Thorat, Chairman, Indian Council for Social Science Research, India

DARIAH: Pan-European Digital Research Infrastructure for Arts and Humanities

Dr Fabio Ciotti, DARIAH (Digital Research Infrastructure for Arts and Humanities), Head of Virtual Competency Centre (VCC), Italy

---

Parallel sessions

12.10-12.20  Tour de table in each group

12.20-13.30  Taking a picture of the current situation: existing resources, capacities and gaps.

What are the research and/or infrastructure capacities in Europe and India related to your specific topic (1, 2 or 3)? What are the gaps and complementarities?

All participants will be asked to give a 1 minute ‘pitch’. It will be followed by a group discussion moderated by the topics co-chairs.

13.30-14.30  Lunch

---

Parallel sessions

14.30-17.00  Identification of research priorities.

What are the future priorities for EU-India research funder and researcher
collaboration in this area – ensuring research excellence (this means where collaborations between researchers in India and Europe add value and don’t duplicate national activities)?

Moderated discussion based on participant’s input prior to the symposium.

19.30 Networking Dinner at Hotel Diana, Via Principe Amedeo, 4

Day 2

<table>
<thead>
<tr>
<th>Time</th>
<th>Session/Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.00-9.05</td>
<td><strong>Second day welcome</strong></td>
</tr>
<tr>
<td></td>
<td>Aurélie Pachkoff-Singh, APRE</td>
</tr>
<tr>
<td>9.05-10.00</td>
<td><strong>Parallel sessions</strong></td>
</tr>
<tr>
<td></td>
<td>Widening our perspectives: Opportunities and challenges of greater digitalisation for SSH research</td>
</tr>
<tr>
<td></td>
<td>What are the opportunities and challenges presented through greater digitisation for social science and humanities research? Which are the main barriers to a holistic approach to digital archives?</td>
</tr>
<tr>
<td></td>
<td>Interactive discussion moderated by EqUIP partners</td>
</tr>
<tr>
<td>10.00-10.30</td>
<td>Conclusion of group discussions and key outcomes</td>
</tr>
<tr>
<td></td>
<td>Summary of discussions and agreement on key outcomes to be presented to the plenary</td>
</tr>
<tr>
<td>10.30-10.45</td>
<td>Coffee break</td>
</tr>
<tr>
<td>11.00-11.45</td>
<td><strong>Plenary session</strong></td>
</tr>
<tr>
<td></td>
<td>Inputs and reporting from different groups</td>
</tr>
<tr>
<td>11.45-12.15</td>
<td>Final discussion</td>
</tr>
<tr>
<td>12.15-12.30</td>
<td>Closing remarks</td>
</tr>
<tr>
<td></td>
<td>Professor Riccardo Pozzo, Head of Department Social Science and Humanities, Cultural Heritage, National Research Council of Italy</td>
</tr>
<tr>
<td>12.30</td>
<td>Lunch</td>
</tr>
</tbody>
</table>