Special Topics
INF XXXX
Crossing the River: Issues, Concepts and Challenges of Digitization

Instructor: Amir Lavie

Time: To be decided
Venue: To be decided

Course Description

For now and in the foreseeable future, any serious reader will have to know how to travel down two very different roads simultaneously'. (Anthony Grafton, 'Digitization and its discontents')

The goal of this course is to introduce students to the practice and cultural phenomena of 'digitization' - the conversion of recorded information to digital, machine readable, format. This will be achieved by (1) exposing students to the professional practice of digitization in an institutional environment, (2) introducing the act's deeper cultural background as a communicative act, and (3) outlining some potential social implications of dealing with a digitized document. In order to target each of these subjects the course is divided into three corresponding sections. The course will focus only on the digitization efforts of written and printed texts of cultural, historical and scholarly value, mostly within an institutional framework.

By creating this link between context, practice and critical engagement, the course will best serve the goals of the Faculty of Information – to provide students both professional and academic knowledge and offer a meeting place for information, people and technology. As current iSchool students will be at the forefront of future repositories management and will take part in the building of infrastructures for research and learning, I aim to equip them with necessary knowledge about the process and its cultural and social significance.

Another goal is for students to understand that digitization is an act that can bridge 'older' knowledge structures of cultural memory institutions with their 'newer' versions, even if those are still growing and evolving. Further, I believe digitization can be a transformative activity to some of these institutions as it enables new organizational modes of production and communication, renewed understanding of the repository and its holdings, demands new practices and skills by employees and offers new external collaborative frameworks, as well as fresh interactions with users, old and new.
The first section of the course looks at digitization as a communicative act – aiming to comparatively inspect it as part of an ongoing human commitment and effort to share information. Students will be provided with an introduction to the history of earlier communication and information technologies (assuming students do not hold previous background of the history of information), analyze those from a functional perspective and move onwards to 'digitization' as the current leading technology serving the same 'older' communicative functions. Two classes will be devoted to this section.

The second section of the course will deal with digitization as a practice. As many cultural repositories digitize their collections in order to offer the benefits of digital access, this section aims to provide students with practical tools about the process and the institutional decision making environment involved. This section will include six classes.

The third section looks at the 'post-digitized' document and focuses on the relationship of digitization to collections, libraries, collective memory practices and other circles of social activity in which the characteristics of a digitized document might alter the role and capabilities it had in its printed form. This section will be more critical in its approach and will ask questions such as what might happen to un-digitized documents or the ways in which selection and funding of digitization efforts might influence future scholarly research output. This section will include four classes.

The syllabus is built in a modular form. I strongly believe that this course could be offered, with some obvious changes of focus, in history and other humanities departments, as the knowledge it shares is valuable also to students outside of the iSchool. This is another reason I chose to divide the course into sections, each of which could be of greater interest for different disciplines. I have also tried to consider and integrate some of the remarks made by Doracic & Dahlstrom (2009) and by Tibbo (2015)1.

Some annotated references and readings suggestions are provided at the end of each class. Not necessarily all of which will be incorporated into the course – but these are the ones which guided me in producing my ideas and speaking notes and they are all helpful in conveying the messages intended for each class. I also added a short annotation and the reason why I think each reading can contribute to the course. There is a substantial amount of materials dealing with these issues and I chose to include mainly examples of 'classic' and shorter papers about this subjects. Additionally, some of the readings were generated from related sub-disciplines (such as digital preservation, digital archives etc.), and so, some concepts and subjects mentioned might overlap with other iSchool courses and students may find some of the materials familiar or related to previous learning experiences.

---

Learning Outcomes

By the end of this course students will be able to:

− **Outline** the historical and functional framework related to the process of digitization and **Describe** its impacts and preliminary expectations (dealt with in section 1 and assignment 1).

− **Identify** and **Analyze** the different concepts and terms associated with the physical and intellectual process of ‘digitization’ (class 2 and section 2).

− **Communicate and Execute** the organizational and personal steps, decisions and paths which compose the process of digitization (Section 2, assignment 2).

− **Examine and Assess** the relationship between digitization process, the digital document and the digital repository (Section 3).

− **Evaluate and Critique** past and present projects of digitization of textual artifacts and apply this knowledge to future planning of activities as professionals (Section 3 and assignment 3).

Relationship of Course Objectives to MI Program Outcomes

Digitization is a central activity and goal for many cultural memory institutions, public sector organizations and commercial enterprises. Students completing this course will understand and be conversant with the fundamental concepts, practices, and methods related to the act of digitization (Program Outcome 1). As future leaders in the field of information students will be exposed to the challenges offered by digitization with relation to the social responsibility of their profession (Program Outcome 2). The course will let students engage in aspects of digitization assessment and evaluation, thus supporting further research and expansion of knowledge (Program Outcome 3). The course will contribute to their progress towards achieving Program Outcome 5 by giving them an understanding of the application of new technological developments and an appreciation of the impact of such developments on contemporary society.

Class Format

Overall, the course is built around discussion and is not intended to be a series of lectures. I aim for most of the assigned readings to be discussed and interpreted jointly with students in class and I have tried to keep readings short and clear. Few courses exist that deal with digitization from such an angle and, as the entire field is relatively new, materials used are not set in stone and are always open to discussion. As a teacher I strongly believe in class dialogue, as well as in the need for substantial historical thinking in order to understand cultural phenomena – and even more so when dealing with the digital realm which is inherently resistant to historical analysis.
Course Materials

Textbooks/Readings
This course does not have a textbook, but throughout the course we will follow a few leading guides and 'best practices', some of which are available on-line, and which will be introduced and discussed in class 3.

Evaluation

<table>
<thead>
<tr>
<th>Assignments</th>
<th>Weight</th>
<th>Due</th>
<th>Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transferring texts between mediums</td>
<td>20%</td>
<td>Week 3</td>
<td>2-3 pages</td>
</tr>
<tr>
<td>Digitization Work-flow</td>
<td>30%</td>
<td>Week 9</td>
<td>5 pages + Class presentation</td>
</tr>
<tr>
<td>Evaluation Tool</td>
<td>30%</td>
<td>One week after class ends</td>
<td>5 pages + Class presentation</td>
</tr>
<tr>
<td>Active participation</td>
<td>20%</td>
<td>Ongoing</td>
<td></td>
</tr>
</tbody>
</table>

As the course constitutes three parts, students will be required to produce three papers during the semester. Each assignment will be related to the subjects which were covered in the corresponding section and is due one week after the end of the corresponding section.

Section 1 assignment: Produce a short paper on a potential challenge in transferring texts from medium to medium and offer possible solutions to the problem (2 to 3 pages, no need for theoretical background, just open thinking, with both mediums to be chosen by the students, even if not related to the course, i.e. Students can choose to work on adaptations such as book to movie, comics to movie, movie to TV etc.). Some analysis is expected with regards to the specific qualities, capabilities and barriers of each technology and the methods to overcome them. (20% of final grade).

Section 2 assignment: Develop a work-flow for a digitization project of a library of your choice (could also be the student's own library, or any collection they are knowledgeable about). The workflow has to consider the challenges offered by this specific collection. The deliverable is to hold at least 5 pages and be arranged according to the subjects we talked about in class. A class presentation will be also required. (30% of final grade).

Section 3 assignment: Prepare a critical evaluation tool for a mass-digitization project. Use some of the tools we learnt about in class in order to evaluate the project but also try and envision the future needs of users and whether future challenges to preservation and access are
addressed in the project you are evaluating. A class presentation and discussion will be required. (30% of final grade).

**Participation** – 20% of final grade.

---

**Guidelines for Assignments**

“All written assignments for this course must be submitted in person, on paper, and handed in at the start of lecture. All assignments should be written as clearly and cleanly as possible (i.e. watch the typos, grammar, hanging sentences, etc.), in a formal but accessible academic language. The “look and feel” should be professional (i.e. no crumpled papers or faded printing).

The required format is as follows:
- Typed, 1.5 space, 11 or 12 point font, one-inch margins, page numbers in the upper or lower right hand corner. Double-sided printing is fine, as long as it’s legible.
- Align paragraphs in a standard way and avoid superfluous indentation.
- The document must be stapled together – no loose pages, no paperclips.
- No cover page required, but be sure to include your name & student number on page 1.
- Use of footnotes/endnotes is permitted.

**Referencing**

“The American Psychological Association (APA) citation style is the most commonly used in academic writing in the social sciences. I recommend that you use APA for this course, as it’s good to get used to the style you’ll likely be using over the course of your graduate career (and beyond). That said, if you think you have a valid *professional* reason for using another style, you are invited to come talk to me at least two weeks before the assignment is due and request that an exception be made. Permission to use referencing styles other than APA will be granted on a case by case basis, but only to students who make arrangements in advance. The key here is that quotes and sources must be properly and consistently cited, using:

(a) in-text citation (including author name(s), year and page number); *and*
(b) a full list of references or bibliography at the end of your paper.

This is a necessary component of academic writing, as well as a good safeguard against inadvertent forms of plagiarism.”

**Acceptable Secondary Sources**

“As graduate students, you will be expected to use a majority of academic (i.e. peer reviewed) sources when writing assignments. Students are very much allowed, but not at all limited, to use course readings and other sources referenced in lectures in their own papers. Additional sources and relevant journals that are recommended by the instructor are also acceptable. However, students are strongly encouraged to track down those resources that are best suited to their specific area of interest or inquiry, rather than rely too heavily on those provided in class.
For cutting edge information, news, announcements, etc., popular press articles are of course acceptable. But these should be used to supplement or update rather than replace peer reviewed sources, and should never be used to explain a theoretical concept. They should also come from credible, verifiable sources, who have the credentials (whatever these may be) to back up their claims. Often these articles point to underlying scholarly articles in peer reviewed journals or conferences, students are encouraged to pursue.”

Late Papers
“Unless a formal extension has been negotiated with the instructor in advance of the due date, late assignments (defined here as an assignment submitted after the deadline) will be penalized by one full letter grade per week (e.g. from A to A-), for a maximum of two weeks. After that point, late assignments will no longer be accepted. Furthermore, late papers will not receive detailed feedback or comments.”

Extensions
“Extensions on assignments within the term must be negotiated in advance, and may require supporting documentation (e.g. doctor’s note). Students must email requests for extensions to the instructor at least 24 hours prior to the due date. Exceptions will only be made in extenuating circumstances. Extensions beyond the end of the term in which a course is taken are subject to the guidelines established by the School of Graduate Studies (Which can be found here: http://www.sgs.utoronto.ca/informationfor/students/track/extsn.htm).”

Grading
“Grading for this course will follow the iSchool’s official Guidelines to Grade Interpretation of letter grades, as well as the University’s policy on Graduate Grading and Evaluation Practices. These sources define grades in the A range as “excellent” and grades in the B range as “good.” Please refer to the guidelines for detailed descriptions of these categories. Assignments will be graded and returned within 2 weeks of submission.

Ground Rules
“Each student in this course is responsible for keeping up with the course materials, which includes (all) the required course readings, as well as topics, debates, and concepts discussed in class. Students are expected to attend lectures and to take their own lecture notes. You are expected to participate in class discussions, and are encouraged to use your laptops/mobile devices during class to look up relevant information that will contribute to the discussion in a meaningful way. At all times, however, remember to be respectful of the instructor and of your classmates – turn your phone function off, turn off the sound on your computer, and be sure not to browse any websites that may be offensive or illegal, or that might be deemed irrelevant to the task of taking this course. Students should arrive on time and are expected to stay for the duration. If you miss a class, you are responsible for obtaining any information or materials given in class, either from your classmates or online. Unauthorized recording of the lectures is not permitted.”
**Students with a Disability or Health Consideration**

“Students with diverse learning styles and needs are welcome in this course. If you have a disability or health consideration that may require accommodations, please feel free to approach the instructor and/or the Accessibility Services Office (http://www.accessibility.utoronto.ca/) as soon as possible. The Accessibility Services staff is available by appointment to assess specific needs, provide referrals and arrange appropriate accommodations. The sooner arrangements are made - the quicker we can assist you.”

**Writing Support**

“The SGS Office of English Language and Writing Support provides free writing support to graduate students. Services are designed for both native and non-native speakers of English, and include non-credit courses, single-session workshops, individual writing consultations, and online resources. Students are encouraged to use these services as needed.”

**Academic Integrity**

“The iSchool has a **strict zero-tolerance policy** on plagiarism, as defined in section B.I.1. (d) of the University’s *Code of Behavior on Academic Matters*. Before you embark on your first writing assignment, please make sure that you:
- Consult the University’s site on Academic Integrity: [http://www.utoronto.ca/academicintegrity/](http://www.utoronto.ca/academicintegrity/)
- Acquaint yourself with the *Code* and Appendix “A” Section 2; [http://www.governingcouncil.utoronto.ca/policies/behaveac.htm](http://www.governingcouncil.utoronto.ca/policies/behaveac.htm)
- Review the material you covered in Cite it Right;
Schedule of Lecture Topics and Readings

Section I (Classes 1 – 2): Digitization as part of the ongoing human effort to share information and knowledge (Digitization as a Communicative act):

class 1 – An Introduction to the History of Documentation / Information / Communications

This class, and the course as a whole, will begin with a short 'performance' – going to the INFORUM, scanning a printed document (a grades transcript or another official Uni. paper) and emailing myself the file. Back in class we will explore the various elements embedded in this act: its main purposes (information accessibility and outreach; preservation), understand the technologies which render it feasible (scanner, the web, textual encoding, display monitors etc.), the people and the functions involved in the process (the source of the communication / the producer of information, the channel, the audience, the institutional setting - and how it can also be operated as a large scale business process) and the possible social and political implications of this act (what can happen if this information is shared, manipulated, lost, what if it is shared without my consent, what if I do not want it to be found etc.).

We will summarize the discussion by stating that digitization is just one aspect of a larger social-technical change taking place around us, and as the ones responsible for creating and maintaining the knowledge infrastructures of generations to come, we have a duty to explore its attributes and qualities.

The above mentioned discussion will help to introduce the course and its administrative requirements, the syllabus, the assignments and other demands, procedures and expectations.

2nd part of class: A brief introductory summary to the 'history of information' - the human need for knowledge transfer in societies and the constant advancement of mediums of communications and their effects on messages, audiences, identity, preservation methods and knowledge of the past – as a way to overcome the obstacles of space and time.


This is of course a huge subject and we will only look at these examples from a functional point of view. It is important to help students understand the nonlinear form of this technological progress and some of the problematic aspects of the comparison between the above mentioned 'information technologies'.
Readings for this section:

Most readings for this class will focus on 'distant' technologies in order to help students grasp the idea that information technologies, no matter which, where or when, has impact on social structure and cultural practices and as such, the 'digital revolution' we experience is not a unique phenomena.

This is a selection of some chapters and ideas from Innis. Despite not being written in the most accessible way, it demonstrates the importance of technology to political and social outcomes, introduces the 'Toronto school' of communications and can also serve as an helpful tool in criticizing the out-dated theory of ‘technological determinism’.

A short and very 'technology deterministic' effort to analyze the influence of digital communications on society. Again, like Innis, it can serve to both show how much social structures are affected by communication technologies, but also, how insufficient is the 'raw' 'technological determinism' explanation.

Another classic, included here because it tries to discuss oral cultures not as a phenomena which 'disappeared' after the introduction and progress of writing mechanisms but rather as a function which can gain new dimensions and new power through modern means of communications.

Yet another classic. Eisenstein introduces the problematic aspect of 'proving' the theories of technological determinism and measuring change and effect. On top, printing seems to be the closest 'revolution' to the 'digital revolution' in its perceived effect on society and culture and can therefore serve as a good case study for class discussion.

Excellent book. This introduction is useful in trying to define and categorize the various information technologies developed and used during early modern times and the challenges faced in so doing.

An early yet very 'mature' contemplation about the prospects of change and revolution that will be coming from the introduction of the 'web'. This was a time of great and sometimes utopian excitement and Nunberg delivers a sobering review, claiming social change related to technology does not necessarily follow a straight or predictable line and that predictions teach us more about the present, or on the time in which they were written, than on the future that awaits us.
A more poetic look at the idea and the philosophical concept of the 'library'. It will be interesting to analyze the story and Borges's metaphor of the library while thinking of today's tools and capabilities.

Some other recommended books (clear and accessible reviews) relating to this huge subject:

Nunberg, Geoffrey and Duguid, Paul. "The History of Information" A course from the University of California (Berkeley) iSchool. (https://www.youtube.com/watch?v=0PjN4pfdPsAk).

Class 2 – Technologies of Copying, Reproduction and Representation; Transformation of Texts between Mediums and some 'Digital' Concepts

We will continue the review of information technologies from last week but this time focus on technologies of copying, reproduction and representation (hand copying, Photography, fax and Microfilm – technologies that carry resemblance to digitization through their expected verisimilitude, trustworthiness and authenticity). We will comparatively review these technologies, their origins, evolution and economic and social characteristics and analyze them as carriers of social expectations - not always in tune with their actual capabilities. We will also talk about the roles and influences of these technologies on memory institutions.

Important points to make: changes and developments within the technology (we still use the title of 'photography' for a process completely different from the one which took place hundred years ago); their relationship and dependency on other technologies; the economic factors underlining the process of production in each of these technologies and its impact on their social roles.

Readings:

An illuminating discussion about documentary photography and notions of truth and representation. Emerling is using photographs from some key 20th century events and traumas to discuss the problematic and paradoxical elements of accepting photography as a documentary genre. Also touches on issues of aesthetics and ethics of photography as a medium.


These four sources are all speaking about microfilms and provide a glimpse to the expectations regarding knowledge and social change reflected through the introduction of a new information technology. I analyzed this subject in a paper from last year and I will make a point that to an extent, what matters is the human endeavour and expectation for better communication, rather than the capabilities of the new technology, whether microfilm or digitization.


Another poetic look at the subject, again from Borges (who was a librarian as well as a writer). This short story speaks about the impossible task of creating an exact copy or reproduction of a work of art.

2nd part of class: We will discuss the concept of 'Textuality' and how everything around us can be seen as a text - not just written works. Text is present in itself, in connection to other texts (context) and in relation to the reader. We will talk about this from a semiotics point of view – using the concepts of signs, coding and decoding. This review will lead us to the question of encountering the same content in different mediums and the ways this encounter creates different texts.

There are plenty of examples to discuss. Some include adaption and translation and some carry the notion of representation and reproduction: (1) Book / movie adaptations with examples such as Raymond Carver's short stories vs. Robert Altman's 'Short cuts' movie or 'Lord of the rings' as a book by Tolkien and as a movie by Jackson; (2) Robert Burns' poem – ‘my love is like a red red rose’ vs. A picture of a rose; (3) A real Christmas tree vs. A plastic tree; (4) The same radio show encountered via radio waves or heard through the web.

Readings:


In my opinion the best and clearer example of semiotic analysis. Barthes presents the subject in an helpful and constructive way, helping the prospects that arise from the application of semiotics to cultural phenomena to be very intellectually stimulating.

There are many introductions to semiotics. This one is not too complicated and includes definitions, traditions and models while introducing De-Saussure and Peirce and even discussing analog vs. digital signs.

3rd part of class: We will dedicate the last part of class to review some important concepts that have to be built upon for the rest of the course. I find a need to better define some of the 'digital' related concepts as it is not always clear what does 'library', 'archive' or 'curation' mean once used with the 'digital' signifier before them. The main source for concept definitions is *The Dictionary for Library and Information Science* by Joan M. Reitz and the definitions are attached as an appendix to this syllabus.

**Concepts to be reviewed:** Digital Information; Digitization and Mass Digitization; Digital Archive and Digital Archives; Digital Collection; Digital images; Digital library; Digital repository; Retro Conversion and Digital Representation; Markup and Markup Language; Optical character recognition (OCR); Intelligent Character Recognition (ICR); Digital preservation; Digital Surrogate; Digital Curation.
Section II (Classes 3 – 8) Digitization as an organizational act / as a business practice:

The aim of this section is to review and analyze the main steps of the institutional digitization process. Students are expected to understand that the process is both chronological and synchronous (activities and decisions which are on-going throughout the project) and that the process can greatly vary across (1) different kinds of organizations (archives, libraries; governmental or communal) (2) different magnitudes of budgetary constraints; (3) different documentary sources, different user groups and different project goals.

class 3 – Introduction to Digitization and an Analysis of Mass-Digitization case Studies

We will introduce the process of digitization and speak about its early history, major projects, challenges and promises. I will follow the Terras historical overview, introduce the Alston, Kelly and Grafton papers and will end the discussion with Abby Smith. We will also review some of the foundational literature and guides about digitization and the different audiences they approach.

Guides and Literature:

_The NINCH guide for Good Practice_ (2002); _The NISO Framework of Guidance for Building Good Digital Collections_ (2004). Two 'authoritative' guides for professionals planning and building collections of digitized sources. Each tries to provide a context, planning scheme and practical advice for different stakeholders related to the process. Both are very structured and are great tools for instruction and learning.

_Leggett, Elizabeth. Digitization and digital archiving: A practical guide for librarians._ (2014). _Rowman & Littlefield: Lanham._ Another guide, this time with a focus on preservation aspects, dealing also with other, non-textual, kinds of media.


Two guides published at the UK, more or less at the same time, offering a framework for establishing and operating a digitization project.

_Bor Ng, Kwong & Kucsma, Jason. (Ed.). Digitization in the real world._ (2010). _New York: Metropolitan New York Library Council._ A publication by the New York Library Council – trying to analyze more than thirty small and medium size digitization projects case studies supported by grants from the library council. Some very interesting case studies are mentioned and the overall tone of the book emphasizes the great difference in goals, methods and outcomes of the process, depending on the organizations and collections involved.
Mass digitization: implications for information policy (2006). A publication by the NCLIS, aimed at policy and decision makers trying to consider, identify and analyze, in reaction to the Google initiative of mass-digitization, key issues for a national (US) information policy.

Cohen, Dan & Rosenzweig, Roy. Digital History. (2006). Philadelphia: Uni. Of Pennsylvania Press. An informal guide aimed at historians who want to use the digital realm as a source / tool / presentation space for their work. Less professional than other guides but great in providing the specific focus for historians. The Rosenzweig centre for history and new media from the George Mason University is very active in providing also other tools for the digital humanities. Cohen now serves as the CEO of the 'digital library of America'.

These are just a few of the publications trying to provide an overview of the process of digitization and we will refer to more specific chapters or parts in the coming classes.

Readings related to this section:

Probably the only 'history of digitization' currently available – trying to look at the last thirty years and provide insights into the development, changes and future trends of the process of digitization within the landscape of various institutional efforts.

Kelly, Kevin. 'Scan This Book'. New york times, May 14, 2006
http://www.nytimes.com/2006/05/14/magazine/14publishing.html?_r=3&
A very positive outlook on the possibility of creating a world library of all 'texts' ever published through digitization. Ofcourse, it is mainly about quantity and not about quality. But also about the coming 'death' of the book as a printed form in order to find place for the next generation – the linked, and tagged, text. The article is full of language of freedom and equality of access.

Alston, Robin. 'the Battle of the books'.
(http://hnet.msu.edu/cgibin/logbrowse.pl?trx=vx&list=halbion&month=9309&week=&msg=/iIvhBzQ0uIis53JmkjkUw&user=&pw=)
A 1993 talk given by a British librarian, contemplating the dangerous implications of transferring paper libraries to 'electronic warehouses'. Very well written and thought provoking.

An early attempt to contemplate on value and effects of digitization efforts. Very clearly and humbly written, it touches foundational questions such as what is unique in digital information, digitization as a tool for access and for preservation and the promises and challenges of each of these roles, the question of authenticity with relation to digital documents and the relationship between the forms of the document which is present in both mediums.

(http://www.newyorker.com/magazine/2007/11/05/future-reading)
An examination of digitization written by a well known historian It aims to show, in a similar fashion to this course, digitization as a tool of knowledge management. Grafton also introduces some problematic aspects which we will encounter later on – messy character recognition,
copyrights restrictions, the selection process, the need to cooperate and coordinate efforts across institutions, the problematic reliance on search tools and the scholarly need to understand both the printed and the digital forms of the text.


A confident and very polemical overview of mass digitization efforts in France in face of the Google books project. Sometimes self-contradicting and sometimes not very clear in his message but always thinking ahead and articulating just how high are the stakes with regards to future research and access to knowledge.

2nd part of class: A structured analysis, done in small groups and afterwards presented to the rest of the class, of some major mass-digitization projects. I will provide background about the history, goals and current standing of each project and will let students analyze and present to class some other aspects of these projects according to their own research into the points listed below.

3. Gutenberg Project ([http://www.gutenberg.org/wiki/Main_Page](http://www.gutenberg.org/wiki/Main_Page))
4. Library of Congress – American Memory ([http://memory.loc.gov/ammem/about/index.html](http://memory.loc.gov/ammem/about/index.html))
7. EEBO ([http://eebo.chadwyck.com/home](http://eebo.chadwyck.com/home))

The analysis will be based on following points:

Dates of initiative / completion; Source materials (textual / non textual); funding and sector (Business / non profit / government); access level (full text / catalogue); searchability; selection criteria (subject based / everything); cooperation (network / single organization); goals; value; public acceptance; Users; project documentation.

We will continue with a discussion on possible solutions to questions that arise from the transformation of printed materials to digital format and especially the act of digitizing a book. We will emphasis the advantages of access, preservation and the ability to create new collections and new knowledge structures versus the challenges of digital formats being manipulable, machine-dependent, rapid technological obsolescence, the high expectations of funders and researchers and the problem of loss of context. **Examples for themes to be examined in the class discussion:** The textuality of the material book: Ownership, book covers, marginalia and underscores (also of library books), book reviews clipped to the pages; Book related institutions - the book store; the library and its visibility; what if there is a mistake in the book which is to be
digitized (is it about the content or just about copying, which is also a moral question); Are all books equal and if not, in what way do they differ (holy books; medieval manuscripts; mass produced books; author's copy; signed and dedicated copy), non textual information (images, tables) in books etc. As a case study example we can also look at [http://www.johnfoxe.org/](http://www.johnfoxe.org/) - the book of martyrs project by the University of Sheffield.

---

**Class 4 – Planning, Work-flow and funding / budgeting a digitization project**

This class will focus on the two above mentioned fundamental aspects of a digitization projects.

**Planning and Work-flows:** Planning is the initial step in each organizational activity. We will talk about what does planning mean, about the different levels and participants of the planning process, about the different subjects that have to be considered as part of a digitization project planning, about the importance of understanding available resources and capturing cost models, about the human resources factor, the equipment, about the need to document decisions and about the importance of knowing the cultural and social background of the organization and collections that are about to be digitized, as well as familiarization with external initiatives and other internal organizational projects, past, present and future. Also, we will be looking at some publicly available digitization work-flow models: (University of Michigan, Scottish national archives, Harvard Uni., University of Santa Clara).

**Readings:**

'Digitization and digital archiving' – Pp. 187 – 204.

Another breakdown of issues and subjects that comprise the steps needed for implementing a successful digitization project. This is not a planning guide per se but the structure of this paper can be used as a good guide for planning ahead.

This is a survey of the planning process by academic libraries aiming to digitize collections. The authors look at issues of staff knowledge and creation of new positions (and sometimes units), at cross organizational communications and coordination and at other responsibilities and actions involved in the ongoing process of digitization.

A short paper suggesting a 'layered' approach to digitization, or, in other words, digitization as an on-going and continuous project, moving from images of the text to encoding as an ongoing process.

The paper is looking at various objections and restrictions to digitization projects and offers practical solutions to trying to overcome them.

**Case Studies:**


A comparison between Google Books and the Open content alliance – focusing mainly on copyright issues and on metadata.

A description of an 'early' digitization project, involving a unique and well known manuscript at the British Library. Most of the technical complexities encountered are less relevant today, but most of the functional and philosophical debates and challenges are still relevant.

**Funding & budgeting:** Funding and budgeting are a crucial part of the planning effort. We will review some funding guidelines and funders expectations as they appear on major funders public websites and will also speak about other sources of funding (both internal and external), the difference between private and government organizations with regards to funding; about project cost models (cost per unit created; other cost factors); about long term funding and about sustainability of the project; about funders expectations and the need for marketing and communications and the challenge of budgeting in a world of rapid changing technologies (not just do we need to digitize but also do we need to digitize now ?)
Readings:


Even if a bit outdated, an interesting take on trying to measure financial costs vs. value of digitization projects.

Case Studies:
Looking at guidelines of two major funders.

Class 5 – The Selection Process and Copyrights

The selection / criteria process: This is one of the most important aspects of digitization – how to select which materials to digitize, taking into account that funding and timing are in limited supply. We will speak about selection as part of the planning process (what are the aims of the project and how it guides selection; who are the users and what are their needs). Beside identification we will also speak about prioritization according to pre-defined criteria such as size, condition, uniqueness, costs, rarity, demand, work-flow requirements, added research value, reaching new audiences, copyrights restrictions and whether the chosen materials are already available in digital format.

Readings:


An helpful and practical guide for the selection process, even though Gertz (academic librarian at Columbia University) tends to focus more on technical aspects and less on assessment of intellectual value.

A report about the implementation of a selection model developed by the Harvard library towards the digitization of another collection (at the Uni. Of Indiana). The author approves of the model yet raises the need for a 'local' version due to the different needs and customs of each institution.
Case Studies:

USC digital library selection criteria – a short questionnaire, consisting of 19 questions, aiming to serve as a selection tool to support budgeting and decisions for digitization within USC.

A collection of selection criteria statements from various University libraries digitization projects.

Copyrights and Privacy Considerations: The subject of copyrights is related to selection criteria yet is also a world within itself. We will speak about what is 'copyright' and what is 'privacy' – concepts which are differently understood across different periods and different cultures. We will review the different stakeholders (libraries, publishers, authors, users and collectors) and their different needs and goals and about copyright law as a tool which aims to find a way to protect both the creators and the public interest. On a more concrete level - what does an institution need to do before starting to digitize and how to determine the status of the object, how do we define an original work, what are the layers of distinct rights, what do we do with the issue of international resources, the time frames of copyrights, licensing and 'fair use' and copyrights management technologies.
We will also briefly speak about Digital asset management with regards to the product – what is it, why do we need it, how to think about and develop one.

Readings:


A very thorough guide to the subject. For this course we will probably only need the introduction and chapter 1 which gives context, concepts and background to the subjects (Pp. 1 – 12) and chapters 9 and 10 (Pp. 103 – 124) that deal with digital rights management systems.

An historical account of Canadian copyright law, starting from the origins in the United Kingdom and moving onwards to a discussion of the unique features of the Canadian law vs. other models.

Class 6 – Equipment, Text Encoding and Metadata

These three subjects cover the main technical range of the process. Returning to the subject of different approaches to digitization (image capture through scanning, text keying, OCR) we will
talk about the attributes of scanners and of softwares that can be used and explore the rapid pace of technological development and the need for offering benchmarks for quality of resolution and compression – as related to users needs.

Howard Besser & Jennifer Trant. (1995). *Introduction to imaging*. The Getty: Santa Monica. An introduction to digital imaging explaining terms and concepts which are still relevant, even if technological capacities are more advanced since the days this guide was written.

'Equipment for digitizing and editing archival materials' In: *Digitization and digital archives*. Pp. 129 – 143.


A clear introduction to the technical practices and choices, methods and equipment involved in digitization projects.

The second part of class will deal with character encoding methods (ASCII, Unicode) and the need for markup (what is XML, TEI and EAD and their respective qualities); examples of different projects and the related decision making: for example, the difference in digitizing medieval manuscripts compared to records of a foreign affairs ministry.

*Digitization and Digital Archives, Chapter 4 'Storing Text' Pp. 37 – 49.*


A LOC guide about textual file formats and the requirements to adhere to in order for the LOC to consider collection.

In the third part of class we will speak about what is metadata and why is it so important; the types of metadata; the main schemas (Dublin Core; EAD; LIDO, METS), the need for interoperability and choosing the right scheme that fits our needs.


These are all short introductions to the subject, explaining just how important metadata is in the digital environment for retrieval, description, administration and preservation of records.
Class 7 – Distribution and Access, Preservation and Curation

This class will be dedicated to some 'post-production' subjects. We will examine the distribution of the product and decisions related to user-access – or in other words, to the 'library' side of a 'digital library': How do users gain access to the objects – free or paid, do they need to login? do we collect their details and if so which details? Can the project be accessed through other portals? Will it be distributed through portable media, internet or both – and the pros and cons of each, can users download, tag or edit materials within the collections? And the way these decisions relate to the 'digital divide'.

The NINCH Guide Pp. 109 - 115

This paper touches on the challenges faced by an academic library in addressing the user demands and expectations after a digitization project.

The preservation and 'digital curation' aspects of the project are complex subjects and we will explore the dilemmas and challenges involved in the preservation of digitized and digital items. We will refer to some guidelines for digital preservation (the SAA 'statement on the preservation of digitized reproductions'; Columbia Uni. 'policy for preservation of digital resources'; 'Digitization and archives' by the Canadian council of archives; the LOC guideline etc.) and speak about the concepts of integrity, fixity, authenticity, preservation metadata and the challenges of migration and emulation.


Trying to define what is digital preservation and what are the implications of digital preservation for cultural heritage. Conway makes a distinction between 'creating digital products worth preserving (digitization for preservation) and investing in the long-term preservation of digital assets (digital preservation), warning that most institutions need to decide on funds allocation between the two.

The paper deals with the selection and preparation processes in an academic library environment but also touches on preservation of source materials after being digitized.

A good introduction to the subject of preservation, speaking about the characteristics of digital information and aiming to guide preservation efforts in asking the fundamental question of what, why and how to preserve.

Class 8 – Quality Assurance and Project Evaluation

We will discuss ideas and methods to evaluate and measure the different components of a digitization project, focusing on the challenging aspects of measuring success, value and impact. We will also explore the subjects of quality control and quality assurance in digitization projects (how do we define and measure mistakes, completeness, the concepts of legibility and fidelity; how to define benchmarks, how do we sample and what to do if the product is not up to the standards we requested).


A study trying to assess the Google books digitization project (through sampling) and finding about 1% of materials with what is termed as 'severe error' and asking questions about what is to be done about the findings.

A discussion that can also be read as part of the planning class but chosen to appear in this section as the authors dedicate part of the chapter to ways of studying users, separating types of involvement and commitment to the projects as tools for evaluation.

A clear overview about the main questions and methods of trying to measure the various consequences of the creation of a digital collection.

'Managing digitization activities' – Pp. 151 – 156. Offering a couple of case studies (from Johns Hopkins and Syracuse Universities) regarding their assessment methods.

A Quantative study trying to assess usage and working patterns of hisotrians looking and citing digital collections in articles from AHR between 2001 and 2010. I didn't find any other research which is trying to do this very important task. The author makes an important distinction between intensity (amount of digital records used) and extensity (how widely was a digital record used).

Joyce Chapman and Samantha Leonard. (2013). 'Cost and benefit of quality control visual checks in large scale digitization of archival manuscripts'. Library Hi Tech, 31, 3. Pp. 405 – 418. A case study description asking whether time spent on QA is well spent or whether it is better used for more digitization activities. As always, the answer depends on the kind of materials involved and on the project goals.


Peter Botticelli, Patricia Montiel- Overall & Ann Clark. 'Building Sustainable digital cultural heritage collections'. (2012). In: The Memory of the World in a Digital Age: Digitization and Preservation Conference Proceedings, Luciana Duranti and Elizabeth Shaffer. (Eds.). Pp. 1205 – 1218. An Interesting paper, offering to use 'cultural competence' as a tool for evaluating digitization efforts – by looking at the history and background of the institution as the main ingredient for the assessment of the project.

Section III (classes 9 – 12): Digitization as a social and political act

The goal of this section is to contemplate the broader and longer-term implications of digitization through exploring the place of digitized texts in repositories and some of the potential social and political consequences of the ability to engage with digital texts.

Class 9 – Digitization and digital repositories

This class will explore the digitized document as part of a collection and as part of an institutional environment. Documents need to be understood with regards to a specific context and usually also with regards to the specific collection in which they are held and as such, the digital format creates expectations related to the new 'abilities' of the document and of the repository. We will inspect these subjects through a review of the history of the ideas of collection and preservation and the various manifestations of libraries, archives, repositories and collections as related to the digital realm.


Both of these encyclopedia entries provide short, authoritative and concise backgrounds about the institutions which historically carry the task of preserving and documenting information in society. Reading these (I assume students are already familiar with these entries) in the context of the course will also help contextualize the subject.


This two 'semi-encyclopedic' papers look at the changes taking place in the roles and functions of libraries and archives in face of digitization and the phenomena of digital collections. Both offer a clear perspective on the changing scenery and practices involved.

An interesting paper about the ongoing human fantasy of creating a 'total' or 'universal' library and the way in which the concept of 'digital library' became a fertile ground to some of these ongoing utopian notions.


Good definitions and discussion on the vagueness of the term 'digital library' especially as it is now being used also by corporations and other for-profit organizations. Rikowski also speaks about existing libraries building digital collections (hybrid libraries), about e-books and discusses some case-studies of digitization projects.


A paper by one of the chief advocates for non-profit, massive scale, digital library. This is part of a series of papers from NYRB, offering his world-view and following the steps in the design and development of the DPLA (digital public library of America).


A paper suggesting that rather than a threat to the 'brick and mortar' library – digital information offers an opportunity for rejuvenation of libraries as a place to provide free access and guidance to the new plethora of available information.


Even if not fully understanding Conway's claims, this is an interesting paper discussing the archival qualities of digitized documents (mainly books) and the need to treat them as such, i.e. Understanding their life-cycle and provenance. This paper turns the class discussion away from libraries to archives but is also useful in bridging some of the qualities and concepts discussed so far.

**Class 10 – the social life of the document / the work of art in the age of digital reproduction**

This class is a continuation of class nine and we will discuss some of the possibilities which the digitized document can offer the citizen, the scholar and any other 'knowledge-seeking' individual. We will start the class with discussing Walter Benjamin's 'the work of art in the age of mechanical reproduction', explain its importance in intellectual history and offer students to think about 'the work of art in the age of digital reproduction' through talking about differences and resemblances of 'mechanical' and 'digital' and the knowledge infrastructures they can offer. Other subjects to discuss are as follows:

- The ability to re-unite collections which are geographically scattered: 'Foreign devils on the silk road' ([http://idp.bl.uk](http://idp.bl.uk)); 'the Cairo Genizh' ([http://cudl.lib.cam.ac.uk/collections/genizah](http://cudl.lib.cam.ac.uk/collections/genizah))
- The ability to create collections and private digital repositories through linking.
– The political implications of providing access to information (the notion of the information wants to be discovered - the Yanovitch leaks - http://yanukovychleaks.org/documents)

– Networked collections: Using the digitized document on social media (OJA) and on other websites.

– What can happen to libraries and archives as organizations and as social spaces (cooperation and collaboration? Convergence? Relationship with users, with the nation state, with private funds, is it still the same organization?) what might happen to smaller archives and libraries who do not have the funds for digitizing?

– what might happen to documents which will not be digitized and the ways in which selection and funding of digitization efforts might be influential for future scholarly research output.

Readings:

Looking at issues of digitization related to special collections in academic libraries – associating issues of digitization with changing economic, political and cultural climates which make it ever harder to build and sustain special collections.

A paper looking into collaboration methods of some European organizations around the subject of digitization. Offers some good background on options, possibilities and challenges.

An interesting paper on a growing phenomena (especially with some of the post-colonial discourse it can offer). Punzalan reviews some attempts at reunification and offers some models for future projects as well as identifying areas for future research.

An effort to re-unite an early medieval collection, involving a cooperation between some geographically (and culturally) distant libraries and research centres. The paper optimismlly concludes with the creation of a joint catalogue.


Another Interesting article about the political opportunities offered by digitization of documents – combined, of course, with the ability to widely disseminate them. The case study of 'the Yanukovitch documents' can be discussed in class, as well as the obvious Snowden / Wiki-leaks case.

Using archeology in its focouian sense, Mak is trying to criticlly engage with the act of digitization at a specific site (the EEBO project) in order to produce insights on contemporary political and culutral preferences. She tries to capture the development of the site and the decisions that shaped it according to users, funding and political chocies and refers to the digitized items as 'artefacts of the modern day'.


An analysis of the research about collaboration between archives, libraries and museums, identifying the core ethical differences between them. The paper proposes digital curation as the “change agent” which will bring the prophesied and anticipated convergence.

**Class 11 – Digitization, Digital History and Collective Memory**

We will explore the concept of 'collective memory' as an on-going and highly active 'battle ground' of various players and forces. I will try to situate digitization projects within this context by using case studies of digitization efforts related to the historical event of the holocaust.

I will provide some background on the holocaust as an historical event and ask some questions with regards to its definition and boundaries (can it be defined ? dates of beginning and ending, geography, what does it include - only murders ? Or can we look at fascism ? At antisemitism ? At the rise of the nation-state ?); talk about the available historical sources and the way they multiplied along the years, about the vast amount of institutions and collections and the challenges this deluge produces to the researcher (languages, geographically disperse, dealing and analyzing an abundance of information), its relation to historiographical trends (opening of archives, discovery of new sources) and the different genres involved (research, fiction, memories, photos, oral testimonies, museums).

We will then be looking at some digitization efforts, the different narratives produced by each instituton and the ways in which this narrative is manifested in the digitized collection and the possible effects it will have on future memory practices.

The sites which will be inspected:

- Yad vashem / The holocaust memorial Museum (Washington) / The Shoah foundation/ The international tracing service / The Jewish historical Institute in Poland / The Wiener Library (London, UK) / EHRI

**Readings:**


A critical look at the ways 'cultural memory' is shaped by the design, structure and availability of digital libraries. Dalbello speaks mainly about the loss of context and narrative coherence and of selecting the sources better adapted to visual display.

A contemplation, done mainly through case studies, of relationship between historical artifacts (whether textual or material – but with a focus on museum artifacts) in their digitized form and in their material form. Maybe a bit too optimistic in nature but interesting and offers some good points to think about.


Another study into effects of digitization on 'collective memory'. Again, offers a much wider definition of 'digitization', related not just to the act in itself but also to the social usages of the new form of documents and the way they alter existing 'collective memory' practices by creating new linkages between the 'digital' and the 'material'.


Prescott is talking about the different practices of literary scholars and historians with relation to the digital text as the way it allows us to further investigate and doubt the record / text. Digitization will allow the reader to double check the historian and so change existing knowledge power structures. And help discuss the essence of the sources rather than taking them for granted before making assumptions. It is well written and argued but something is missing with regards to the continued need for historians as mediators and meaning-makers.


A short online paper by the chief conservator of the US Holocaust museum in which she shows, through a case study, the difficult choices and resulting limitations in trying to digitize some of the materials and sources held by the museum (in this case – a diary written during the war).


A unique effort to try and analyze the ethical perspectives of approaching a digital archive of testimonies holding more than 100,000 hours of materials. Presner contemplates ways for historians to approach the vast amount of available data but still respect the individual testimony as the core historical unit of research.

Class 12 – Site Visit, Wrap up and Guest lecture

The last class will re-visit the main topics of the course and will aim to accommodate a visit to a digitization site and/or a guest lecture by a practitioner or scholar who could highlight some extra subjects and talk about their personal experience. Again, this class aims to show just how the fascinating process of digitization can allow an institution to re-examine its own goals and
working methods in order to stay active, relevant and attuned to changing models of relationships between texts, collections, repositories and user groups.

Contacting the Instructor & Supporting Instructor

Amir Lavie is available by email: amir.lavie@mail.utoronto.ca. Usual response time: within 2 working days.

Acknowledgement and citation. Thanks to Prof. Seamus Ross (and before that, Prof. Sara Grimes) for providing the outline for this Syllabus as a model. Some of the administrative related texts are taken directly from their earlier syllabus.
Appendix A: Definitions of Concepts to be reviewed and discussed during class # 2:

**Digital:** The informational manifestation of a series of bits; Data recorded or transmitted as discrete, discontinuous voltage pulses represented by the binary digits 0 and 1, called bits. In digitized text, each alphanumeric character is represented by a specific 8-bit sequence called a byte. The term can also refer to the wave of information technology generated by the invention of the microcomputer in the second half of the 20th century, as in the expressions "digital divide" and "digital library."

**Digitization:** The process of converting data to digital format for processing by a computer. In information systems, digitization usually refers to the conversion of printed text or images (photographs, illustrations, maps, etc.) into binary signals using some kind of scanning device that enables the result to be displayed on a computer screen. In telecommunication, digitization refers to the conversion of continuous analog signals into pulsating digital signals. The Association of Research Libraries (ARL) has endorsed digitization as an accepted preservation reformatting option for a range of materials. Also, capturing a signal in a digital form, converting content of physical media to digital format, a translation of information into bits (many possible mediums). For textual artifacts digitization can be either as an image of the original shape of the text; a full text version in a different shape, or it can have extra information (hyper links)

**Digital archive:** A system designed for locating, storing, and providing access to digital materials over the long term. A digital archive may use a variety of preservation methods to ensure that materials remain usable as technology changes, including emulation and migration. The National Digital Information Infrastructure and Preservation Program (NDIIPP) led by the Library of Congress is an example of a program aimed at preserving digital content.

**Digital Archives:** Archival materials that have been converted to machine-readable format, usually for the sake of preservation or to make them more accessible to users. Also refers to information originally created in electronic format, preserved for its archival value.

**Digital Collection:** A collection of library or archival materials converted to machine-readable format for preservation or to provide electronic access. Also, library materials produced in electronic formats, including e-zines, e-journals, e-books, reference works published online and on CD-ROM, bibliographic databases, and other Web-based resources. The National Information Standards Organization (NISO) has published *A Framework for Building Good Digital Collections* (2007).

**Digital Image:** An analog image that has been converted, usually by a scanner or digital camera, into a gridded array (matrix) of small discrete locations called *picture elements* ("pixels") that hold binary data quantifying the size in area of the location and the colour and
brightness of the image at the location. The data of which a digital image is comprised can be stored, manipulated, transmitted electronically, printed, reproduced on film, or displayed on a computer monitor or television screen.

**Digital Library:** A library in which a significant proportion of the resources are available in machine-readable format (as opposed to print or microform) and accessible by means of computers. The digital content may be locally held or accessed remotely via computer networks. In libraries, the process of digitization began with the catalog, moved to periodical indexes and abstracting services, then to periodicals and large reference works, and finally to book publishing.

**Digital Repository:** Many academic and research libraries are actively engaged in building digital collections of books, papers, theses, media, and other works of interest to the institution served, as a means of preserving and disseminating scholarly information. Usually locally authored or produced, content can be either born digital or reformatted. Access is generally unrestricted, in compliance with the Open Archives Initiative, which makes such archives interoperable and cross-searchable.

**Markup:** A system for annotating a document in a way that is distinguishable from the text. The idea evolved from the "marking up" of paper manuscripts, i.e. revision instructions by editors, traditionally written with a blue pencil on authors' manuscripts. In computing: encoding of a textual document with information external to the document, for example, to indicate the structure of the document, the type or grammatical function of words and phrases, or the way in which portions of the document are displayed on screen or page, usually accomplished by inserting tags and format codes in the text.

**Markup Language:** A predefined set of descriptors (symbols and tags) or a method of defining descriptors that are used to embed external information in an electronic text document, usually to specify formatting or facilitate analysis. The Hypertext Markup Language (HTML) used in creating Web pages. In 1998, the World Wide Web Consortium recommended the use of a simplified version of SGML known as Extensible Markup Language (XML). Most markup languages differ from databases in identifying elements within a stream of text, rather than discrete, structured data elements, but XML is capable of turning text into the equivalent of a database.

**Mass Digitization:** The conversion of texts or images to digital format on a very large scale using robotic equipment. Many publishers and other stakeholders have expressed concerns about the legal, social, economic, and other impacts of these kinds of projects and initiatives.

**Retro-Conversion** – In a general sense, converting existing textual collections into a machine readable form. **Digital representation** – transferring content (without form or image) to the digital realm.

**Intelligent Character Recognition (ICR):** An advanced form of optical character recognition, capable of learning to recognize different typefaces and styles of handwriting. Most ICR systems
automatically update their recognition databases whenever they encounter a new typeface or style of handwriting.

**Optical Character Recognition (OCR):** A process by which characters typed or printed on a page are electronically scanned, analyzed, and if found recognizable on the basis of appearance, converted into a digital character of code capable of being processed by a computer. OCR eliminates the time-consuming process of re-keying information produced in print, but results can be unpredictable if the scanned copy is imperfect or contains diacritical marks or unrecognizable characters.

**Born Digital:** Refers to materials that originate in a digital form. This is in contrast to digital reformatting, through which analog materials become digital.

**Digital Curation:** The active management, enhancement, and preservation of trustworthy digital data for the duration of its lifecycle. Digital curation entails verification of the integrity of digital data, selection of authoritative digital data for its long-term value, creation of digital objects and associated metadata, transfer of digital objects to reliable digital repositories for secure storage, provision of access to designated users, and periodic re-evaluation of digital formats to avoid obsolescence.

**Digital Preservation:** The process of maintaining, in a condition suitable for use, materials produced in digital formats, including both bit stream and continued ability to render and display the content represented by the bit stream. Some digital storage media deteriorate quickly and the digital object is inextricably entwined with its access environment (software and hardware). Also refers to the practice of digitizing materials originally produced in nondigital formats (print, film, etc.) to prevent permanent loss due to deterioration of the physical medium. Also: A formal endeavor to ensure that digital information of continuing value remains accessible and usable. Involves planning, resource allocation, and application of preservation methods and technologies, and combines policies, strategies and actions to ensure access to reformatted and "born-digital" content, regardless of challenges of media failure and technological change.