INKE (Implementing New Knowledge Environments) is a research project that aims to help bridge the gap between the book-based tradition of research in the humanities and the new affordances brought by the computer and the World Wide Web.
Historical context

In 1440, when Gutenberg began tinkering with movable type, few people could have anticipated that his invention would spark off one of the most important and widely felt revolutions in the world's history. By making books affordable and easily available, the printing press unleashed the full power of the word. In less than half a century, historians estimate that between fifteen and twenty million books were printed in Europe for a population, at that time, of approximately one hundred million (Febvre & Martin, 1958). This invention had incommensurable effects on all aspects of life, causing an explosion of knowledge, fostering a new relationship with religion, as well as precipitating the triumph of the individual (see, for example, McLuhan, 1962; Ong, 1977; Eisenstein, 1979; Chartier, 1995).

When the computer appeared in the late 1940s, it was a machine whose crude interface was designed for doing statistics. Some scholars, however, quickly grasped its possibilities for searching through vast numbers of words and for establishing concordances and indexes. The first “electronic book” was thus a 56-volume index to the works of Thomas Aquinas, compiled between 1949 and the mid-1970s by the Jesuit Roberto Busa, who is revered today as the first digital humanist. For many years, though, the computer could not be considered a potential substitute of the book, since it did not even have a screen.

Following an idea proposed in 1945 by Vannevar Bush, Ted Nelson foresaw the affordances of the computer for enhancing the activity of reading through “hyperlinks,” a term he proposed in 1965. The hyperlink can be seen as an extreme development of the table of contents, as well as of the index and the footnote, three components of the scholarly book that give readers the option to go directly to the place that is of interest to them, rather than following the lengthy and sometimes tiresome developments of a linear thread. It is worth noting that none of these components existed at the birth of the codex. Rather, the table of contents and the index appeared in the Middle Ages and became relatively common by the thirteenth century, whereas the marginalia, which first appeared on critical editions in the Alexandrian period, gave way to the footnote, which “in its modern form seems to have been devised in the seventeenth century, as part of an effort to counter scepticism about the possibility of attaining knowledge about the past” (Grafton, 1994, p. 53). The hyperlink, first implemented by Douglas Engelbart in 1966, without any contact with Nelson, became part of high-end word processing systems, notably IBM’s Generalized Markup Language (GML) that morphed into SGML in 1986. It made its appearance in the consumer market with HyperCard, which came installed on the Macintosh as of 1987.

The hyperlink can connect not only internal parts of a book but any object, be it an image, a sound, or a video. It thus abolishes the difference between the book as a container and the library, making all books virtually the components of one big book. As such, this technology was able to bridge telecommunications and computer technology and, someday, to make digital texts available from everywhere, transcending the limitations of distance. With this technology, the world was thus getting closer to the old dream of the universal library that Arthur C. Clarke had reactivated in a quotation by Wernher von Braun in the May 1970 issue of Popular Science magazine, and that Ted Nelson had pursued in Xanadu during twenty years.
in a “quixotic quest” (Wolf, 1995). That dream finally became a reality with the World Wide Web, developed by Tim Berners-Lee and Robert Cailliau between 1989 and 1992. With this worldwide system of interlinked hypertext documents, books attained the ultimate form of portability, that had long been an elusive goal: the Babylonian tablets, heavy, breakable, and cumbersome, were at the lowest level of portability; the papyrus scroll was a big improvement, being light and extensible to the exact length of a chapter or a book; it was superseded, however, by the codex, which represented during two millennia the best form of portability, thanks to its small size, its high capacity for text, and its handiness. Still, nothing illustrates better the awkwardness of travelling with a collection of bulky codices than a Persian legend told by Alberto Manguel:

In the tenth century, the Grand Vizier of Persia, Abdul Kasem Ismael, in order not to part with his collection of 117,000 volumes when travelling, had them carried by a caravan of four hundred camels trained to walk in alphabetical order. (1996, p. 193)

Immediacy of search, hyperlinking technology and ubiquity of access are such valuable tools for a reader that the digital environment soon became the first place most people visited in their quest for reading material. This necessarily affected the way people read. Indeed, by facilitating some aspects of the activity of reading and discouraging others, technology insidiously shapes readers’ expectations and behaviour. While the personal computer was already common in the 1990s, many users still routinely chose to circumvent its limitations by sending long texts to the printer. This was clearly not a good solution in the long-term; either the text had to be adapted to the limitations of the technology, or the technology had to be adapted to the text and the demands of the readers, the latter clearly being a more sensible solution.

Some observers quickly saw the long-term implications of the personal computer becoming the method of choice for writing and communicating. Jay Bolter (1991), among others, enthusiastically embraced the new paradigm, deciding to get rid of his collection of printed books and do his reading on the computer. Most people, however, did not rejoice at the announced disappearance of the book – a device that had been so important for the intellectual development of millions of users for half a millennium. In his Gutenberg elegies, Sven Birkerts (1994) expressed the sadness and anxiety raised by the mere idea of this. Many people, though, retained the hope that print would still accompany us for many years. In 1996, Umberto Eco was convinced that “[b]ooks will remain indispensable not only for literature, but for any circumstance in which one needs to read carefully, not only to receive information, but also to speculate and to reflect about it” (p. 300). Indeed, at that time, the desktop computer was definitely unfriendly for long reading sessions, and most notably for the reading of novels, as Stephen King learned from his experience with The Plant, a novel made available through paid instalments during a few months in the summer of 2000, which he stopped writing when he saw the number of subscribers dwindling dramatically.

As it was evident that the desktop was too cumbersome and lacked the mobility and ease of use of the book, engineers began working on e-book readers, such as the Rocket eBook released in 1999, which offered some of the new affordances of the digital text. But these machines did not gain many followers, and most readers of
novels still preferred printed books to digital ones. In spite of its apparent external simplicity, the codex is well adapted to reading, thanks to the high-definition of its typography, the layout of its double page, as well as the facility of leafing through it and annotating interesting passages (see for example Vandendorpe, 2009). Moreover, for many people, it was not clear, in fact, if the book should be replaced at all or just left to collect dust, since we had entered a new “mediasphere” (Debray, 1992, p. 286), where the unit of information was no longer the book as a container, but the flux formed by the specific web of relationships one can trace among millions of nodes of information. This question, however, was definitely settled in December 2004, when Google announced that it would partner with five major libraries in order to scan and make available online some fifteen million books within a decade. To this huge number, one must add the books digitized by various other programs around the world – some of them established years before, others promptly following Google’s initiative – be they national, privately funded, or collaborative, such as Gallica, Europeana, Library of Congress, Hathi Trust, Million Book Project, Gutenberg, and Wikisource, etc.

From then on, it was clear to everyone that the chasm between the printed book and the digital world would eventually be bridged and that the entire trove of information ever produced by the human mind since the invention of writing – books, codices, scrolls, and Babylonian tablets – would someday be available in digital form. As a result, new interfaces and workspaces, as well as new methods of working collaboratively would be needed for scholarly research to move entirely onto the screen and make use of the affordances of the computer and the Web.

The genesis of INKE
When discussions first started in 2004 about a research project initiated by Ray Siemens and incorporating other Canadian scholars working in the digital humanities, there was a consensus that the most important issue was to rethink the book in the context of the digital environment. After many discussions between the main researchers, and some preliminary meetings, notably in Vancouver and Victoria (see Lynne Siemens, “From Writing the Grant to Working the Grant”, in this volume), a first draft was submitted to the Social Sciences and Humanities Research Council of Canada (SSHRC) in 2006 under the title HCI-Book – Human-Computer Interaction/ Interface with the Book – (see Ray Siemens, 2011). In the course of the following years, a new name, better adapted to the reality of this endeavour was chosen: INKE: Implementing New Knowledge Environments. The abstract of the final submission gives a good idea of the scope of this research project:

The pace of that change has created a gap between core cultural and social practices that depend on stable reading and writing environments, and the new kinds of digital artifacts – electronic books, being just one type of many – that must sustain those practices into the future. Our project will bridge this gap by theorizing the transmission of culture in pre- and post-electronic media, by documenting the facets of how people experience information as readers and writers, by designing new kinds of interfaces and artifacts that afford readers new abilities, and by sharing those designs in online prototypes that
implement new knowledge environments for researchers and the Canadian public. (Siemens et al., 2008, p. 10)

The goal was not only to import the affordances of the book on a computer, but to create the best possible space for working on information and producing knowledge, devising tools for manipulating text and implementing interfaces suited to a variety of tasks. This implies a wide-range reflection on the new possibilities that the computer offers for the “use of text in dynamic reading environments where the reader is capable of controlling and modifying the format and content of the text as part of standard interaction with it” (Siemens et al., 2009).

The challenge facing our enterprise has been complicated by the rapid developments occurring in reading devices and display technologies. Following the lead of the Sony Reader (2006) that used e-ink technology to create a paper-like reading experience, Amazon released the Kindle in the fall of 2007, just when INKE was submitting its research program. Various other similar devices promptly appeared on the market (see in this volume “Electronic Environments for Reading,” by Koolen, Garnette and Siemens). In 2010, to make matters still more complicated, Apple bet on quite a different technology; instead of emulating the closed environment of the printed book, it launched the iPad, whose shiny glass screen takes (almost) full advantage of the affordances of the web, within a tightly controlled operating system (OS). Many tablets would subsequently appear on the market, competing with the Apple OS in order to offer a comparable user experience, supported either by Windows, Linux, QNX or, most frequently, Android, the free OS designed by Google. Thanks to these breathtaking developments, and the subsequent publicity, the idea of reading long books on mobile devices is today widely accepted, to the point that libraries have begun lending books for electronic devices. The pace of development is not slowing, however, and large screens with retina-display definition (more than 300 ppi) are close on the horizon, opening new spaces for knowledge environments. Keeping pace of these developments in academic research certainly poses a particular challenge to INKE and those working in this field.

The methodology of research also poses particular problems. Contrary to experimental sciences, where “the proof of the pudding is in the eating,” our goal is not to build new hardware but to design prototypes. Ray Siemens discusses this methodology in his article titled “Implementing New Knowledge Environments: Year 1 Research Foundations” (in this volume). This article also gives a detailed discussion of the mandates and activities of the research teams. Initially, there were four groups, but in 2010, as discussed in Lynne Siemens’ paper submitted at the Kyoto gathering, these groups have been reorganized into three: Interface Design (ID), Textual Studies (TS), and Modelling/Prototyping (M/P).

**The birds of a feather gatherings**

In a project of this scope, researchers have to be in contact quite frequently in order to keep the synergy of the community at its highest level. Emails and videoconferences are important for keeping up the pace and the collegiality, but they are not sufficient. People have to engage in face-to-face meetings in order to feel fully integrated in the team and exchange on all aspects of the research.
It is also necessary to exchange with the larger community of digital humanists because external scholars bring fresh viewpoints and novel ideas, giving the members of the teams precious feedback. With this in mind, annual international conferences have been held under the theme “Research Foundations for Understanding Books and Reading in the Digital Age.” The first one took place in Victoria, BC (October 2009), sponsored by INKE and SSHRC.

**Victoria, 2009**

**THEORETICAL FRAMEWORKS**

At the Victoria conference, Ray Siemens gave an overview of the first year of funded research for INKE and reiterated the objectives and methodologies of the various teams.

In “Beyond Remediation,” Alan Galey presents the philosophy and the goals of the Textual Studies team, whose field of research encompasses both the history of the book and understanding how digital media will affect reading. A lot of work has already been done in this field over the last sixty years, but with different approaches on both sides of the Atlantic. While the French school of thought gives a lot of attention to the social aspect of book reading and book publishing, the Anglo-American counterpart is more interested by the materiality of the book and bibliographic study. The TS team aims to bridge the gap between these two approaches and examine long-term continuities and discontinuities without falling into the oversimplifications that have marred this field of research. It also aims to connect the study of print and digital environments in order to develop a technical vocabulary for describing the salient features of electronic artifacts. The team will also compile a knowledge base, titled *Architectures of the Book*, or *ArchBook*, which aims at being “the most comprehensive online reference for the history of the book,” as well as a medium for promoting research on the future of the book.

In “From Writing the Grant to Working the Grant,” Lynne Siemens reflects on the nature of academic collaboration and the factors that may contribute to its success. She presents a flowchart of the administrative structure made up of an executive committee, an advisory board and partners committee, and a research area group committee, overseeing the four sub-area research groups. The charter and administrative documents, abundantly discussed by the members of the core administrative group between 2005 and 2008, cover a number of issues, notably a process for the inclusion of new partners and researchers, an authorship convention, an intellectual property clause, protocols for data and document storage, etc. These documents are not static, however, and will be revisited on a yearly basis.

Craig Saper discusses a curious device designed around 1929-1930 by Bob Carlton Brown, an American polygraph who lived in many places and knew personally some of the major avant-garde artists and writers of the period. Brown is known today for “The Readies,” an electric machine displaying a text as a single continuous line, flowing either to the left or to the right, whose pace could be accelerated by the reader, supposedly allowing the reader to absorb many thousands of words per minute. An online simulation of this machine is now available at www.readies.org. The process of recreating this machine in digital form led Saper to reflect about databases and
interfaces, and to realize that a new field was open to simulation: the reading situation itself.

Brown’s conception of reading would appear naïve, or at best a mere provocation, in light of the cognitive conception of reading, supported by decades of psychological research, which informs Jean-Guy Meunier’s presentation of an outline of CASTR (Computer-Assisted Expert Text Reading). For Meunier, the most important aspect of reading does not reside in the linear decoding of one word after another, but rather in the capacity of the reader to extract the meaning of a sentence and integrate it into their own knowledge base. In order for a computer program to help in expert reading, it must be able to work at both the micro-level, facilitating search in the dictionary or the encyclopaedia, and at the macro-level, extracting the structure of the text. The program CASTR has been designed to meet precisely these goals.

Building on the experience he gained redesigning Representative Poetry Online and lemmatizing Lexicons of Early Modern English, Marc Plamondon, an INKE collaborator, suggests that it would be much more useful to build databases for text based on lemmata rather than word entries. This method would provide a “more seamless integration with available digital tools and reference sources, such as period dictionaries, modern dictionaries, pronunciation dictionaries, and encyclopaedias.” It would also “allow for greater control on the users part of the selection and activation of tools and resources while reading.”

**Collaborative tools**

James MacGregor et al. make a presentation of tools for reading and online annotation. Since 2001, PKP (Public Knowledge Project), a partner of INKE, already has provided a set of “reading tools” for OJS journals, but these fall short of covering the social needs of today’s scholars, notably as concerns the three strategies identified by Ray Siemens: creating an online identity, allowing for clear connections and communications between identities, and managing research material. The eventual redevelopment of the OJS platform would offer the reader tools for highlighting, note making, and linking, as well as a workspace for carrying through these activities. This is contingent, however, on the appropriate funding of the Synergies project by SSHRC.

In order to avoid the pitfalls of collective authorship, Rama Hoetzlein has devised Quanta, an offline experimental piece of software able to track the authors of every individual sentence, allowing different levels of knowledge to co-exist in a same entity of text. A reader can filter out unwanted authors or groups of authors, just as we normally do in real life. If installed on a collaborative platform, this system would restore the authority of the individual and create an “author-centric knowledge organization.”

Serina Patterson advocates for integrating into educational websites the features of web 2.0 that the young generation now enjoys so thoroughly. First of all, young people want to be active participants in digital media. Older modes of interaction do not attract them because in the online environment “interaction is no longer simply an exchange between two people, but a method of influencing and changing the content in significant ways.” She is thus designing a website devoted to promoting books

Interface
A current and efficient way for academics and graduate students to start a specific line of research is to follow the sources most commonly cited in that field. This process, called “chaining,” is the focus of “Drilling for Papers in INKE” by Stan Ruecker, leader of the ID (Interface design) team. The prototype created during the first year of research is adequately called The Paper Drill. It starts from a seed article, asks the user to indicate how many levels deep they wish to go, then makes “the system traverse the available metadata and articles to produce a summary report of the authors and articles most cited starting from that seed, as well as links where possible back to the articles.” In order to be effective, the set of articles selected for the chaining process must be consistent, which is easy to achieve with a database of journals like Synergies. Ruecker is also experimenting with a visual representation of the summary report, which could be animated.

In “Re-conceiving ‘Watching the Script,’ ” Jennifer Roberts-Smith presents an interface designed by the same ID team in order to enhance users’ interactions with a theatrical text. In its first version, ‘Watching the Script’ presented four vertical panels on the screen: (a) the “Overview panel,” representing the entire text in microscopic format on a single column; (b) the “Reading view,” showing portions of the text in a readable format; (c) the “Stage view,” separating the text into speeches and associating them with coloured dots representing the characters that deliver them. Although this interface could serve as an aid for students and a visual mnemonic for actors, it had the fundamental flaw of considering the continuous utterance of a single character as the basic unit of a theatrical text. Going back to Aristotle, the team decided to make action the basic unit of its interface. The new version “offers a centrally-placed visualization of a line of action in the form of a horizontal image analogous to a timeline or a number line,” with characters built in Lego-like blocks, whose actions can be defined by the user and associated with words. This interface appears to be extremely well conceived and could not only help professional actors and directors, but also provide students with a dynamic reading experience as well as a very relevant sense of play.

In “The NT2 Hypermedia Art and Literature Directory,” Bertrand Gervais, a partner and a collaborator of INKE, presents a project aimed at cataloguing, describing, and evaluating hypermedia art. The growing body of artistic works created for the web offers the basic features of being both iconic and textual, with eventual components manipulated by software. Due to the hybrid nature of these objects, an interdisciplinary team has been assembled, involving specialists of literature, semiotics, art history, ludology, and cinema. Since the beginning of the project, over 3,000 works have been treated in the “Répertoire” (directory). Each hypermedia work is catalogued according to an extensive set of key words. Quite predictably, interactivity has been the most difficult characteristic to describe, necessitating no less than 39 components in order to accurately report the user’s experience. Every work is documented with screen captures, and in many cases with video captures, of the navigation through the hypermedia, which will allow long-term preservation and study of works that are for young people along with out-of-class discussion (http://sd62onlinelibrary.ca/) according to three criteria identified by Eliza Dresang as trend-setting in young adult literature: connectivity, interactivity, and accessibility.
essentially ephemeral. The NT2 website also offers a discussion forum, “Délinéaire,” and video-interviews with artists/authors exploring the creative process as well as technical aspects of their productions. Being a partner of INKE, the NT2 project will be useful for documenting some particular characteristics of reading in electronic environments.

The title of Jon Bath’s article, “Book design still matters in the digital age,” sounds like a manifesto that is far from being universally accepted. When looking at web pages and e-books, one has to agree with Bath that most of them are “sorely lacking” in comparison with the standards of traditional book typography. A similar situation happened before, in a sense, since “most of the 17th and 19th centuries were periods of lacklustre, or downright abysmal, book printing [in England],” due notably to a break with tradition and a concern for making money. With the rise of advertisement, there has been a battle between conventional typography, whose aim is to be transparent, and the completely opposed ethos of graphic design, which draws attention to itself. While the former has maintained its empire in the double page of the book, the latter dominates the single page format of posters and advertisements and has thus invaded the web pages of would-be artists eager to express themselves. One can only hope with Jon Bath that someday “electronic reading interfaces can, and will, be improved by paying attention to traditional typographical practices.”

Text encoding

In “Sound, Ink, Bytes,” Øyvind Eide presents his work on a series of interviews made in the 1740s by Major Peter Schnitler, who travelled extensively in northern parts of Norway and Sweden/Finland in order to gather information about the people and border areas. The manuscripts of the transcripts were recently published in print and were useful in settling border issues after Norway left the union with Sweden in 1905. In the 1990s, the print edition was digitized and encoded in XML and the element structure converted to TEI. Eide is now working on the internal history of these documents and will focus on the geographical aspects.

A number of other papers were presented at the Victoria gathering that do not appear in this volume, namely:

Ethan Hawkley, “Where’s Walden?: Searching, Indexing, Reading and Living in the Digital Age”

Martin Mueller, “An English Diachronic Digital Annotated Corpus (EDDAC)”

Sharon Oviatt, “Designing Interfaces that Stimulate Ideational Superfluency”

Brett Hirsch, “Mark the Play: Electronic Editions of Shakespeare and Video Content”

Patrick Juola, “Guessing at the Content of a Million Books”

Although they were not members of INKE, and chose to publish elsewhere, their presentations were highly appreciated and a valuable contribution to the conference.

**The Hague, 2010**

The second gathering was held in The Hague in December 2010, in conjunction with the Text & Literacy conference and sponsored by the National Library of the Netherlands, the Department of Book and Digital Media Studies of Leiden University, INKE, and SSHRC. The main topic was “Textual Methodologies and Exemplars.”

**Collaborative writing and editing**

Peter Boot made a presentation about Verhalensite, a Dutch website where a vibrant community of writers can post poems and stories or comment on the pieces submitted. In the course of ten years, this site has registered 48,000 texts and some 350,000 comments, thus constituting a corpus quite relevant for research of online community writing. One of the first exploratory findings is the strong correlation between the number of comments writers make on other people's creations and the number of comments their own creations receive. This correlation suggests that “giving comments might […] be a conscious strategy” for people in search of fame. Studying the frequency of words used in the comments and comparing it to a general corpus, Boot found a very high frequency of terms of greeting, followed by technical literary terms. As for words referring to the thematic content, “loneliness” and “sad” are the most frequent. In a preliminary conclusion, Boot suggests that “the role of the common reader in determining reputation may become more important,” displacing former institutions of literary fame.

This radical shift of the locus of influence, from the institution to the web, is a topic that Claire Clivaz from Lausanne University documents with great detail in “Homer and the New Testament as Multitext in the Digital Age.” The starting point of her reflection was the launch in late 2010 of an independent online edition of the Greek New Testament. This version could be made freely available by returning to 19th century editions and thus bypassing all the information provided by a century of research and discoveries. This is a clear example of the “deregulation of the scholarly critical edition.” It is also at odds with the philosophy governing the Homer Multitext, a project that plans to present the textual transmission of the Homeric text rather than produce a unique edition. According to Clivaz, the reason for the divergence between two close fields of research lies in the characterization of Christianity as “a religion of the Book,” a topos invented in the 19th century. Another important shift brought about by the Internet is the transformation of “the familiar boundaries of scholarship,” as evidenced by Arabic websites displaying Greek manuscripts of the New Testament and studies showing that “the Gospels were first translated into Arabic in either the sixth or early seventh century.” This also offers Western scholars an opportunity to closely examine often-ignored manuscripts. Clearly, digital culture is redefining our ways of thinking about a whole field of research.
The digital environment is even redefining the way we think about knowledge and authority, as Janneke Adema argues quite eloquently in a paper titled “On Open Books and Fluid Humanities.” There is a deep chasm between the opportunities offered by the new tools and the traditional practices of scholarly publishing. On the one hand, we want to have easy and constant access to information, to be able to mine it at will and use it interactively in collaborative settings; on the other hand, we are still publishing scholarly monographs that cling to an out-dated model of authority and to the idea that knowledge can be stabilized. In order to overcome this contradiction, Adema makes the case for “the possibility to reuse, adapt, modify and remix material” and explores in depth the concepts of “modularity” (see Manovich, 2008), of the author being replaced by a “selector” in a sampling environment (see Navas, 2010), and of the printed book giving way to the “fluid text” (see Bryant, 2002). While recognizing that these are only halfway solutions and that “just like stability, fluidity is an ideal type,” Adema concludes her Derridean critique of the present situation by proposing that these experimental concepts may help to reinvent the scholarly knowledge dissemination system.

**Interface**

As entire libraries are now moving online, there is an urgent need to help users easily find their way through vast collections of works. In his paper titled “Interfacing the Collection,” Daniel Sondheim presents the findings of the Interface Design (ID) team after examining the best practices in the field. One of the most basic interfaces to a vast collection of items is a search engine such as Google, which can be supplemented by a “Page Preview.” A “Timeline” helps in presenting the findings historically, while the “Wonder Wheel” arranges the findings visually. Other innovative methods of display are in use by alternative search engines, such as Redz and SearchCube, or the now defunct Viewzi. The search can eventually be refined with filters, as with Google Search and YouTube, or categories, as with Google Books. With these tools, users can also design and manage their own directories or playlists. The most innovative interface to large corpora, according to Sondheim, is the Perseus Digital Library, which offers a map of places mentioned, a timeline, lists of documents by author, as well as micro-level overviews; its only limitation is that it does not support user input. Building on these observations, the team, led by Stan Ruecker, has been working on an experimental prototype of a collection interface controlled by the user called JiTR (Just in Time Research), which allows the user to create and manage collections for research and study. The most innovative aspect of JiTR will be to introduce “structured surfaces,” which “can help users understand the organization of items in a collection and allow a user to interact with the collection.” The results of a search could also be visualized as a structured surface, inspired by Florence Nightingale’s rose diagram, or as a treemap with a cloud of dots. The user then superimposes another level of data onto the infographic, much as one puts pins in a map, in order to begin constructing an argument about the data. All these examples are very promising.

In much the same vein, and with the same modular architectural approach as JiTR, Texttiles 2.0, presented by Alejandro Giacometti (“Showcase Browsing with Texttiles 2.0 and BubbleLines”), another member of the ID team, is a powerful tool for browsing and managing collections of text files. Applying the principles of rich-prospect browsing laid out by Ruecker in 2003, it provides a meaningful representation
of each item in a collection as well as controls to manipulate the display. Various experimental interfaces have been developed according to these principles, notably the delegate browser, which allows reviewing details on attendees met at an academic conference. As a complement to Texttiles, BubbleLines represents a series of texts by horizontal lines of relative length; it displays on each line, in the form of a bubble, the places where a queried term appears, the bubbles being greater when there are more occurrences. This tool may be very helpful for getting a synthetic view of a huge corpus and, at the same time, being able to locate the places where particular words are most frequent.

For Brent Nelson and the Textual Studies group, it is not at all certain that the hyperlink is better at linking texts than “old ways for linking.” Examining the chain-linking referencing system for the Bible published in 1908 by Frank Charles Thompson, he shows quite convincingly that this system easily beats its digital adaptation for the iPhone, thanks to the ability offered by the print edition to follow links both forward and backward. Moreover, this system provides “smart linking” by declaring the nature of the targeting link, and enables the contextual association of links. These are challenges worthy of being emulated in the digital world. One must keep in mind, however, that the Thompson system can beat an automated search system thanks to: (a) the intense semantic coding performed on the Bible by many thousands of scholars and (b) a uniform numeration of verses, which provides a level of granularity not ordinarily seen in other corpora.

In “E-readers are for Reading,” Joost Kircz reminds us that “every medium has its own characteristics which makes it the preferred medium for certain types of reading.” There is no doubt that an e-reader can be adequate for reading a novel, but one may wonder if it could be a substitute for the textbook. In order to test this question, Kircz did an experiment with three groups of students who were assigned the same textbook either in print, on a laptop, or on the IREX e-reader. The results show that the e-reader was mostly dropped by the users because it was “too slow in starting-up and browsing and ergonomics were not ideal.” In order for the e-reader to compete, digital textbooks need to be completely reorganized with hyperlinks and a well-designed interface. Even if the iPad was not part of this experimentation, the last recommendation certainly also applies to it.

In her article titled “Reconfiguring Narrative,” Sharon Webb makes a similar observation and points to the necessity of reconsidering “how we present scholarly narratives” in writing. Instead of the linear thread of classic writing, we have to embrace the hypertext technology and accept the eventuality that users could jump on a link and leave our texts behind.

**Literary studies**

In "Utilizing Digital Technology in Literary Research," Sonia Howell presents a reflection on interdisciplinary research in the field of literature aimed at developing “a new critical method.” Taking two novels, one Irish and the other one English, both centred on a patient having suffered a traumatic experience, she asked the subjects of the experiment to mark “what they considered to be indicators of cultural context and trauma.” The resulting mark up can be visualized, which may lead to interesting
discussions between different critics of these works and help them to become conscious of their tacit assumptions in evaluating a novel.

**User Studies**

In “New Theories and Methods for Screen-Centred Interfaces,” Sheila Petty presents the objectives and the methodology of a pilot study she intends to do with her colleagues in order to determine “how media content is influenced by the device on which it is presented, from cognitive, cultural, and aesthetic perspectives.” This experimental study will be carried out over eighteen months and will involve 80 participants, with four types of devices (computer screen, iPad, iPhone, mobile phone) and four types of content (poetic, visual, and language texts, as well as hybrid multimedia fragments) being tested. There is no doubt that the user experience is greatly influenced by the media, and that we need to have more information on the cognitive processes at play. Led in parallel to and independently of INKE, this research may be useful to the INKE's ID and TS teams.

**Administration**

In an article titled “Understanding Long Term Collaboration: Reflections on Year 1 and Before,” Lynne Siemens presents the findings of interviews conducted with seven individuals coming from three INKE's teams, including graduate research assistants, researchers, and administrative leads. All of them highlight the advantages that flow from a collaboration while also pointing to the challenges inherent to disciplinary differences and interpersonal issues.

**Kyoto, 2011**

In November 2011, the third “Birds of a feather” gathering was held at Ritsumeikan University in Kyoto, in conjunction with the Second International Symposium on Digital Humanities for Japanese Arts and Cultures (DH-JAC2011). The meeting was sponsored by the Japanese Association for Digital Humanities, the Digital Humanities Center for Japanese Arts and Cultures at Ritsumeikan University, the Implementing New Knowledge Environments (INKE) research group, and the Social Science and Humanities Research Council of Canada (SSHRC). As noted by Mitsuyuki Inaba in his welcome words, this University was particularly relevant for our own endeavour in digital humanities since the name “Ritsumeikan” attributed to the Confucian philosopher Mencius means “the Place to Establish Your Destiny”.

Kozaburo Hachimura gave a talk on “Digital Archiving of Intangible Cultural Properties,” which presented his projects aimed at the digitization and preservation of cultural heritage. In order to facilitate research on intangible objects, like Noh and Kabuki theatre, Hachimura has devised a system for capturing actors’ body motions and recording the data in a database, which will allow search and comparisons of specific motions across various dances. A system called Labanotation will allow the editing of Noh motions and their subsequent representation in digital animated sequences. In another groundbreaking project, Hachimura will make a Digital Museum of Gion Festival Yamahoko Parade. Coupled with the environment of the Virtual Kyoto developed by Prof. Yao’s group, this project will give users the opportunity to experience virtually what it is like to participate in this important festival.
Working on text-mining, Tomoji Tabata presented the drawbacks of keyword-based analysis. Instead of keywords, he proposed a “Random Forests” approach, consisting of a machine learning system building a large number of classification trees from a database. In a set of 47 works (23 by Charles Dickens and 24 by his contemporary Wilkie Collins), Random Forests reliably identified Dickens’ style with 95% accuracy.

**Collaborative archiving and editing**

One of the great affordances offered by the computer in a web environment is the possibility of abolishing physical distance and immediately linking people. It is no surprise therefore that six of the presentations discussed various ways of working collaboratively on texts. Masahiro Shimoda presented the challenges facing the Buddhist Research Knowledge Base (RBIB), an international project involving over 70 scholars from 30 universities whose goal is to digitize important Buddhist works in Chinese, Indic, and Japanese characters, as well as to build a Digital Dictionary of Buddhism and a Bibliographical Database.

In his presentation of the Shelley-Godwin archive, Neil Fraistat said texts are “massively addressable objects,” meaning that, if properly edited, the position in text can be queried at different levels. Texts should no longer “lie in silos but be scalable and e-verted,” or turned inside out, interoperable, layered, modular, multimodal, dynamic, scalable, and sustainable. In a not so distant past, scholars had to rely on the mail in order to collaborate on a project; now they can interact daily, by email, voice, or video, exchange files or, better still, work on a wiki or a common file located in the “cloud.”

Christian Wittern, from the University of Kyoto, proposed an “architecture for active reading” that would allow the collaborative editing of texts. He followed a model that already existed in various software communities who, many years ago, adopted a “distributed version control system” (DVCS) where there is no central authority, but instead an organic network of branches. This system relies on a piece of software that the members of the community install on their local computer. One of the most popular today is the program GIT, designed and developed in 2005 by Linus Thorwald for the Linux kernel development. This program is able to track and record the changes made to a specific version, and gives any participant the choice of implementing the change on their machine or not. A program may thus evolve in many branches and offer various flavours, adapted to the needs of a specific task or a particular hardware. Wittern suggests that this model could easily be adopted by scholars working on a same critical edition. Each scholar’s annotations and corrections could be merged and selectively shared among the members of the community. The various branches could also be converted to a “flat” version for the public. This system would be particularly useful for canonical texts existing in a variety of translations; a German user, for example, could choose to download a master text in Chinese and its translation in German. Moreover, GIT keeps track of who made the changes. This is particularly useful in the humanities, where the signature still bears a lot of value, and, in some cases, a distinct level of trust and credibility.

Constance Crompton and Ray Siemens, both from the INKE team, explore other alternative ways of harnessing the social resources of the web for a critical edition. They are working on a tripartite system edition of the Devonshire Manuscript. The first one
will be a scholarly edition of this manuscript that will reside on the Iter Community website, complete with introduction, notes on the text, contributor biographies, witnesses, and indices. This text will be static and available in PDF format. A second edition, also on the Iter Community website, will incorporate the static edition in wiki format, thus allowing for hyperlinked content and the scholarly apparatus. The third instantiation of this edition will be distributed between Wikipedia, Wikiversity, and Wikibooks, where the various fragments may be edited by anyone. After a while, the group will compare the different editions in order to arrive at a final copy.

In “Wikisource and the Scholarly Book,” Christian Vandendorpe, an associate member of the ID team, took a different approach to the same problem. Instead of creating new environments from scratch for working on canonical texts, he proposed to build on Wikisource. This multilingual portal already incorporates 58 languages, allowing the reader of some books to easily shift from one translation to another. This feature would be extremely useful for studies on the reception of a canonical text across various cultures. Scholars could be discouraged, however, by the restrictive policy of Wikisource with regard to annotations. In order to allow annotations and comments on specific passages of books, Wikisource would need to break with two of Wikipedia’s overarching principles, namely the exclusion of original works and the obligation of adopting a neutral point of view. Even if the efficiency of these two principles has been proven in the redaction of the encyclopaedia, they are clearly not adequate for a reading space like Wikisource where a large part of the interest stems from the juxtaposition of various points of view on subjective matters. If comments and annotations were allowed, however, it would add to the costs of administration because Wikisource would have to open a private space for each user in order to be able to filter derogatory comments. Here, the Quanta software, devised by Rama Hoetzlein (see above), would be very helpful.

Jon Saklofske, an INKE associate researcher, takes another approach to the question of annotations—or, in a more general way of speaking, user-generated content—by proposing a dynamic articulation between an open archive, in the style of NINES, and a digital edition. In a paper titled “Fluid Layering: Reimagining Digital Literary Archives through Dynamic, User-Generated Content,” he proposed making a distinction between a digital edition, which is a primary source, and a digital archive, which he defined as an “active, dynamic public record, an information commons that substantiates a foundational collection of primary texts with a continuous aggregation of critical contexts and conversations that grow from that foundation.” He applies this distinction to the work he is presently doing on William Blake, suggesting that the existing William Blake Archive should be re-imagined as “a centre for critical engagement, user-generated dialogue, argument, commentary and response.”

According to William Bowen, from the University of Toronto, the same impulse to rethink the role of big archives in the context of the social transformation of the web is at work at the Iter Gateway to the Middle Ages and Renaissance. Created in the mid 1990s as a bibliographical database, the Iter Gateway holds a collection of more than one million records and is now evolving toward a Knowledge Base, allowing personalized searches in all types of documents, and connecting metadata to the relevant objects. Bowen also envisions going a step further within the next ten years...
and morphing the Iter Gateway into an Iter Community that would better serve the
needs of its affiliated members by offering additional services, such as blogs, wikis,
online polls, collaborative editing, collegial research area groups, et cetera.

**ANALYSIS AND ENVIRONMENT**

In her “Analysis of Silent Cinema and Benshi Narration in Digital Humanities,”
*Kyoko Omori*, from Hamilton College, studied a collection of recordings of *benshi*—
live performers who, in Japan, stood in a spotlight adjacent to a movie screen and
performed three distinct roles: plot narration, character dialogue (*of all characters*),
and impromptu commentary on the movie, actors, and story. Omori aims to provide a
semiotic analysis of voice and space in modernity and to map the relationship between
media and personhood. In order to disseminate and complement her research, Omori
is creating a website whose interface will provide multimedia case studies of Japanese
movie theatre experiences and allow scholars from different institutions to join and
collaborate in the research. Of particular interest is a case study of *The Cabinet of Dr.
Caligari*, accompanied by the *benshi* performance of a top performer in the 1920s.

*Harvey Quamen*, from the University of Alberta, proposes a theoretical discussion on
the “limits of modelling” in the digital humanities. He takes as a starting point Willard
McCarty’s assertions that “analytical computing in the humanities is modelling” and
that modelling is “the continual process of coming to know by manipulating things.”
Such presuppositions lead Quamen to fear that “the Humanities might succumb to
easy temptations to become more ‘scientific’ in order to legitimate [its] presence on
university campuses.” In opposition to McCarty, Quamen underlines the merits of
the database and advocates for a symbiotic relationship between the database and the
modelling, arriving at the conclusion that, “when the model seems impossible, build a
database.”

In order to have a better grasp of the functionalities that a digital interface should
embody, it is necessary to have a full understanding of the characteristics of printed
books. This is precisely the endeavour of *Richard Cunningham* and INKE’s Textual
Studies team, who propose a reflection on the architecture of the book. Starting from
the assessment that “the pre-digital book has a world of features very few people ever
see—despite looking at some or all of them daily, typically for years,” Cunningham’s
team plans to produce an ArchBook, i.e. an online, open-access reference resource of
textual features throughout the long history of the book, in order to advance

the state of knowledge about pre-digital textuality so that informed
choices can be made about what textual features to retain, which to
reject, which to modify, and, perhaps, which to invent to make digital
textuality the best possible environment it can be for readers and users
of text in the Humanist tradition.

*Daniel Paul O’Donnell*, from the University of Lethbridge, examines how new
technology affects reader’s understandings of digital objects. He starts with various
cases of media enhancing or modifying photos in order to make them look prettier or
to support an argument. In many cases, the editorial process is now affecting objects
that were until recently thought of as independent realities. A similar transformation is
taking place with the “Spatial Turn” in geography due to the introduction of GIS. In the second part of this reflection, O’Donnell wonders how to avoid falling into the same predicament with his own project of realizing a digital edition of a collection of objects belonging to the Anglo-Saxon “Visionary Cross” tradition.

Takaaki Kaneko, from Ritsumeikan University, is working on the digitization of another type of objects: printing blocks of the Edo period. This material is difficult to handle and very fragile. Yet, its close examination is essential in order to study in a systematic manner the variants of a book in relation with its initial prints. The digitization of an archive of five thousand printing blocks will allow researchers to study this fragile material in collaboration and to incorporate the data into Japanese bibliographical studies. In his presentation, Kaneko details the methods of digitization and the structure of the metadata.

New interfaces for reading
The interface is a crucial component of any book, and it is still more important in scholarly editions since they have much more complex and hierarchically organized content than the classical novel. When working on a critical edition, for example, a scholar may need to question the text in many ways in order to compare the versions of various manuscripts, ascertain the meaning of a specific word, or examine the interpretations offered by various commentators, etc.

Stan Ruecker proposes to address these multiple concerns by prototyping a set of tools among which the scholar will be able to choose at will. The first prototype is a dynamic table of contents that will support not only the dynamic insertion by the reader of XML-marked passages, but also “bubblelines” or frequency graphs. A second prototype consists of “a series of small embedded text analysis tools that the reader can activate while reading.” The Citation Visualization System will produce a set of interactive visualizations of citations, based on the metadata they contain about time, place, and their function in the argument. The problem with tools, however, is to ensure that they are known and adopted by the public for whom they have been built. Strategies for addressing this issue include discussions with stakeholders, user experience studies of the prototypes, and collaborations with INKE research partners.

In his presentation titled “The Face of the Scholarly Corpus and Edition,” Geoffrey Rockwell, also a member of the ID team, presented a topology of corpora, resulting in three main classes: linguistic, artifactual, and literary. Differences between print and web instantiations are highest in linguistic corpora because they use complex databases and their digital form looks more “like complicated search engines than traditional texts.” Using the Internet Archive Wayback Machine, Rockwell also documents the evolution of the interface of Perseus, from its first release on a CD-ROM to the present, underlying the important changes that have occurred over twenty years due to changes in the mediating technology, the intended audience, the mission of the scholarly work, and the culture of the media.

In an article presenting a survey of tables of contents across the last three centuries, Brent Nelson, from the TS group, divides their functions into four categories:
summarize, locate, visualize, and conceptualize. A table of contents is a malleable form that can be distributed across a whole book. While being the front door to the contents, the table can be complemented by an index, which can be considered as a back door situated at the end of the book. Following this historical study, Nelson presents the “Dynamic Table of Contents,” developed with Stan Ruecker, Milena Radzikowska, and Mark Bieber (ID team). It is described as a “digital re-imagining of these two forms of summary apparatus (the TOC and the index), placed in conversation with each other in an interactive system that allows the reader to dynamically add to and subtract from the table of contents what are essentially index items – actually material that has been encoded with semantic-level XML.” This tool will certainly be very useful because it is always visible at the side of the text and provides readers with a large range of affordances when working with long texts. Visualization tools, presently under construction, will also be integrated into this dynamic table, which should still further enhance the reader’s experience.

**Jon Bath**, from the TS team, proposes experimental interfaces for reading illustrated books. Challenging “the idea of reading as solely the mental act of translating letterforms into language, and of the book as merely a container for the written word,” he gives some very telling examples of the shortcomings of the classic interface displaying one page at a time. This form of display can be particularly misleading in the case of illustrated books because these books were conceived of as opening on a double page, which led the illustrator to eventually introduce a play between the illustration on the left page and the one on the right page. For that reason, the British Library has designed a double-page interface that gives the reader a quasi-book experience. Another possible solution is the “Rich Prospect Browsing” that allows the reader to see a whole collection of pages simultaneously on the screen and eventually filter the pages according to a specific range of page numbers or a set of metadata. A visualization tool also allows the reader to explore all the co-occurrences of metadata in the same book. Bath is very conscious of the fact that digital tools must take full advantage of the new environment offered by the computer and “go beyond replicating pre-digital processes and displaying the answers to pre-determined questions.” He thus plans to develop a framework that will allow the users to add their own metadata to a text. That would give the user a lot of flexibility and make reading an operation of co-construction between author and reader.

**Susan Brown**, in collaboration with the INKE research group, presents a paper titled “From CRUD to CREAM: Imagining a rich scholarly repository interface.” It consists of research on the components found essential in the interface of a dynamic collaborative scholarly repository in order to support the creation, enhancement, and re-use of available texts as well as eventual operations of analysis of derivative collections. The acronym “CREAM” stands for “create, read, enhance, analyze, manage.” The research involved a group of users who were asked to organize a set of 94 activities into functional categories. Quite predictably, the participants did not reach a consensus but there was a high level of agreement on some aspects of the categories “create,” “manage,” and “analyze/search.” As a conclusion, Brown considers that her research offers a useful starting point on thinking about the enriched functionality a scholarly repository requires.
An interface actualizes a preferred set of actions that a user will make in order to manipulate data. In many cases, these sets of actions are very much dependent on the hardware. This is why new e-book readers have appeared in the last few years under the form of light machines that users can easily take with them and hold in any position. In an article titled “Electronic Environments for Reading,” Corina Koolen, Alex Garnette, and Ray Siemens propose an annotated bibliography of pertinent hardware and software. The group recognizes that “people seem willing to read on an LCD-screen, as long as the device has an aesthetically pleasing design and is portable,” a finding generally considered valid and that “has implications for the development of digital information environments.” They also stress the wide difference of support by the industry for the scholarly way of reading in comparison with the social and leisurely ones:

reading with the aim of building knowledge means that many levels of handling information need to be supported: […] flexible documentation organization, multi-document and text navigation, information triage, annotation, and possibly, the inclusion of a writing space.

This is followed by a survey of commercial e-reading devices, namely, the Sony Reader (2006), Kindle (2007), Nook (2009), Kobo (2010), iPad (2010) and Android smartphones.

ADMINISTRATIVE DIMENSION
In her article titled “Firing on all cylinders: Progress and Transition in INKE’s Year 2,” Lynne Siemens explores the nature of collaboration within INKE – a seven year multidisciplinary project with 35 active researchers, as well as postdoctoral fellows, graduate research assistants, and partner organizations across four countries. According to her report, the second year was characterized by “forward research progress, positive relationships, and transitions and challenges related to human resources, team restructuring and partner institutional policies.” She ponders the benefits of collaboration against individual research and concludes in favour of team research. She makes clear, however, that this also entails many challenges, notably the possible conflicts and confusion arising from the fact that teams consist of people coming from different disciplines, with different languages and methodologies. A still more challenging issue is to ensure the knowledge transfer between former graduate students and new recruits, as well as to retain highly skilled researchers and post-doctoral students.

It is worth noting that the format adopted for the presentations was quite different in Kyoto. Instead of reading their articles, every participant had to send it in advance so that most of the time for the presentation could be dedicated to a discussion of the papers. This type of discussion was very useful in enhancing the level of interactions between the participants.
Conclusion

As one can see by this rapid survey, the body of research produced by INKE is already significant and it was time that the dissemination process shifted into high gear. The new digital environment has brought a profound shift in attitudes at the university level, going from text-based to a hybrid of text and visual elements, from solitary to collaborative, and from narrowly disciplinary to interdisciplinary.

While the book was an interface for discourse, and the library catalogue an interface for books, we now need interfaces in order to navigate the wealth of information available on the web and to mark out access paths to specific cognitive tasks. Visualization is a key element of the interfaces presented because it helps synthesize in a single map huge amounts of data, giving the user a bird’s eye view of a mass of relations between various objects.

From the first gathering to the third one, we can also see a growing interest for collaborative tools and platforms. This is one more sign that the traditional metaphor of the ivory tower is becoming irrelevant; rather, scholars work more and more in teams and take advantage of the open nature of the web. Collaborative research is not easy, though, and necessitates adequate tools and dedicated spaces where people can interact, as well as administrative protocols and a new conception of authorship.

This collaborative spirit also tends to soften disciplinary borders. Being naturally multidisciplinary, the digital humanities can play a major role in widening the range of questions we traditionally ask. Moreover, the web brings a wind of change inside our traditional sphere of inquiry, the screen texts and images from other cultures or fields of research, as shown notably by the presentation of Claire Clivaz. All in all, the Web is the ultimate serendipity engine.

References


