

AccessionIndex: TCD-SCSS-X.20180907.002

Accession Date: 7-Sep-2018

Accession By: Trevor Peare

Object name: History of computerisation of TCD Library's 1872 Printed Catalogue

Vintage: c.2005

Synopsis: The long personal campaign by Prof.J.G.Byrne, Dept.Computer Science.

Description:

The Old Library of Trinity College, built 1712-1732, is famous worldwide as one of the most impressive libraries in the world.

Many thanks to Trevor Peare, Keeper (Systems) in TCD Library, 1990 - 2003, for the description of the history of the computerisation of TCD Library's 1872 Printed Catalogue below.

Background to the 1872 Printed Catalogue

The Library's collections date from the foundation of the University in 1592 and the first list of books records around fifty printed works in 1601, the year a first Library Keeper was appointed. [1]

A hundred years later, in 1720, the Librarian had as his principle duty the production of a catalogue of printed books in the Library. This listed around 12,000 volumes.

Following the completion of a new Library building (now The Old Library), in 1732, a new catalogue was prepared in six volumes to record the volumes transferred to the new building. This catalogue was annotated as new books were added to the collections and had become totally unwieldy a hundred years later by the time James Henthorn Todd was appointed Librarian in 1832.

Todd modernised the cataloguing rules and new books were catalogued on paper slips filed in drawers. In 1837, Todd set about preparing a new catalogue, first by having slips prepared for the books already in the Library. The first volume of this great work was published in 1864 and the last of the nine volumes in 1887, over 50 years after the project started and eighteen years after the death of Todd in 1869. It included all books in the Library received up to 1872.¹

Sets of the catalogue volumes were sent to the great libraries of the world, including the Library of Congress.²

¹ Books received after 1872 were catalogued on a separate sequence of slips and an Accessions Catalogue [2] was commenced. This was a Guard-Book catalogue with typed catalogued entries pasted into volumes of blank pages. It was closed in 1964 having reached 436 volumes. The conversion of the records in the Accessions Catalogue to the Library's online Stella Catalogue commenced in 1996 and is nearing completion (2019). A digital version of the Accessions Catalogue was posted on Flickr in 2014.

² In the 1990s, the sets of the Printed Catalogue used in the Library were becoming very worn and following a search for additional copies, a commercially available microfilm set made from the Library of Congress copy was identified. The Microfilm was purchased and printouts were used to replace distressed volumes in the Library reading rooms. The Microfilm itself was used in the Early Printed Books Department of the Library to conserve the printed volumes.

Professor John Byrne

“Professor John Byrne (1933-2016) was father of computing in Ireland.” [3]

“He was responsible, directly or indirectly, for many of the extra-ordinary successes of the Irish computing sector. In his five decades at Trinity College Dublin, during most of which he was Head of its Computer Science Department, he established many new courses and fostered leading-edge research. He simultaneously lobbied and campaigned to make the nascent Irish software industry a priority for the Irish Government and for its agencies. In direct consequence, several thousand undergraduates and postgraduates were trained in various aspects of the intricacies of computing and digital engineering from hardware to software to algorithmic analysis to business applications to design to human interfacing. During his leadership, many international companies chose to base operations in Ireland not least because of the excellent quality of its computing graduates. Likewise, a number of significant Irish start-ups were founded by former students and staff of the Department.” [3]

The OCR Project

“John was a bibliophile –he had a vast collection of books covering all sorts of subjects. He had a particularly good eye for rare books of historical significance especially those relating to mathematics and computation. While a postgraduate student at Imperial College London in 1959, he purchased an item for £2 10s; another copy of the same item sold for \$170,500 in New York in 2008.” [4]

In 1986 Trinity College Dublin hosted the then highly prestigious International Federation for Information Processing Annual Conference (IFIP) and Prof. Byrne decided that it would be appropriate to stage a special exhibition in the Long Room entitled *“Computing through the ages”*. He put an enormous amount of work into curating the Exhibition, sourcing material from the College Library and the British Museum as well as from his own personal collection [3]. He set about trying to locate early mathematical publications in the Library, not an easy task as there was no on-line catalogue available for books acquired prior to 1872. There was only a printed one which was largely organised by author -- with many of the entries in Latin, although with his classical education that didn't present a problem to him.

He described the immediate motivation for the development of the digital version of the catalogue in his chapter in the book on the Library edited by Bill Vaughan (on page 217 of [6]):

“I was looking for an article on Pascal’s calculating machine in the Encyclopédie. Rather than call for all twenty-eight volumes one by one, I was granted permission to go to the Long Room to find the volume I wanted. Finding it quickly, I could not resist having a look at other shelves. I came across a book [in Latin published in Padua] by [Balthasar] Capra, and on looking it up in a life of Galileo found that Galileo considered it a plagiarism of one of his books. Galileo was so incensed that he brought Capra before the university court in Padua, which ordered that all 483 copies be destroyed. The copy on the shelves of the library was one of thirty that escaped”.

It occurred to him that there were probably many more such treasures to be found in the Library and so the project to produce a digital version of the 1872 catalogue was born. This was no routine project. The catalogue listed approximately 165,000 books in 9 volumes with a total of 5,374 pages in at least 14 different languages [6]. The original printed catalogue took almost 40 years to complete, the online version using principally student labour (Glynn Anderson, Andrew Collins, Brian Cullen, Brendan Culligan, Ruth Clarke, Donnla nic Gerailt and Mark Fynes) [7, 8, 9, 10, 11, 12] under Prof.Byrne's supervision took less than five years. The languages of the entries included those using the Roman alphabet but also Hebrew, Syriac, Russian, Arabic and Greek with variations in typeface and in font size. This fact together with the particular layout of the catalogue required a bespoke Optical Character Recognition System to be developed. Once the catalogue had been scanned and converted into text, other students converted it into a searchable database with a simple user interface ultimately enabling online access opening up the Catalogue worldwide. One particularly nice feature is that once the reader locates a book of interest, a screen showing an image of the entry in the original catalogue is displayed. The project has been justifiably described as "*a labour of love on the part of the late John Byrne*" [13] (p. xxiv n.11) and he continued to work on enhancements until shortly before his death."

The OCR Project and the Library

Prof.Byrne acknowledged the assistance he had from Library staff with the exhibition and then later with the project. He particularly remarked on the assistance given by Vincent Kinane in Early Printed Books for the exhibition and bringing the chance discovery in the late 1980s of an almost complete set of unbound sheets of the Printed Catalogue to his attention. This was found in a Front Square basement store, when it was being cleared of Ordinance Survey maps to establish a new Map Library in the University. The sheets were collated by the Library's Conservation Laboratory and made available to Prof.Byrne. [14].

In his own article written in 2004 [6], there is a comprehensive description of the approach taken with the computerisation of the catalogue records and the difficulties with OCR at that time. Over the years, John personally edited entries checking against an image of the original catalogue record in the Printed Catalogue. The data for each record included the analysed text along with a reference to the location of the image of the record in the Catalogue. The Shelf mark (also known as the Call Number or location) was resistant to total transcription accuracy with OCR but as that was vital for location an item, he did not use the OCR data for the Shelf mark, but instead linked to the image in the Catalogue.

He was very conscious, because of his own experience of books "hidden" in the Catalogue, that he did not add to the problem with inaccurate Shelf marks.

The 1872 Printed Catalogue online

The access system was developed within the Computer Science Department on stand-alone computers using a graphic interface and was transferred to Windows PCs with a workstation provided to the Library, where it was immediately put into use. The computerised catalogue allowed for Author, Title, Language and keyword searching – compared with the listing in the Printed Catalogue by Author only.

The early interface was also a subject for research within the project and the access was developed over the years on stand-alone workstations as the Windows operating system was revised. An internal networked version was available about 2006 and a full online version on the Internet in 2008.

Over the following years, Prof.Byrne continued to edit the data and to improve the public interface right up to the weeks before his death in 2016.

The online catalogue has replaced the use of the original volumes of the Printed Catalogue in the Reading Rooms as it provides enhanced access points while still allowing readers to see an image of the original entries in the Catalogue. [15]

The Future

After Prof.John Gabriel Byrne's death in 2016, all development work on the Catalogue ceased and the Library became increasingly concerned with ongoing access as it had become dependent on the system. The database continued to be hosted on a server in the Computer Science Department.

A copy of the data sets were transferred to the Library in 2018 and work has started to include the computerised records in the general online "Stella Catalogue" which covers printed and electronic works. This is part of an ongoing program which is completing the conversion of the Accessions Catalogue, digitising Trinity's dissertations and theses, and incorporating other printed collections into the "Stella Catalogue".

Trevor Peare, December 2019.

The TCD Library was an early user of computing in Trinity College Dublin, see [16].

The homepage for this catalog is at: <https://www.scss.tcd.ie/SCSSTreasuresCatalog/>
Click 'Accession Index' (1st column listed) for related folder, or 'About' for further guidance.
Some of the items below may be more properly part of other categories of this catalog,
but are listed here for convenience.

Accession Index	Object with Identification
TCD-SCSS-X.20180907.002	History of computerisation of TCD Library's 1872 Printed Catalogue. The long personal campaign by Prof.J.G.Byrne, Dept.Computer Science, c.2005.
TCD-SCSS-X.20180907.001	History of TCD Library computing, Trinity College Dublin. The evolution of TCD Library computing, c.1968.

References:

1. Peter Fox, *Trinity College Library: a History*, Cambridge University Press, 2014.
2. Trinity College Dublin, *Accessions Catalogue*, see:
<https://www.tcd.ie/library/collections/accessions.php>
Last viewed Nov-2019.
3. Ed: Chris Horn, *Professor John Byrne: Reminiscences: Father of Computing in Ireland*, ISBN-10: 1520696841, ISBN-13: 978-1520696843, 2017.
4. Jane Grimson, *Memorial Discourse on Prof.J.G.Byrne*, presented on Trinity Monday, 2018, Trinity College Dublin, see:
<https://scss.tcd.ie/SCSSTreasuresCatalog/literature/TCD-SCSS-V.20180409.001/TCD-SCSS-V.20180409.001.pdf> (Last viewed 3-Dec-2019).
Also see the text of the Discourse:
https://www.tcd.ie/Secretary/FellowsScholars/discourses/discourses/2018_John%20Gabriel%20Byrne.pdf (Last viewed Nov-2019).
Also see the video of the Discourse:
<https://www.scss.tcd.ie/SCSSTreasuresCatalog/literature/TCD-SCSS-V.20180409.001/TrinityMondayDiscourse2018-JohnGabrielByrne-OfficialVideoAndAudio-20180409-1130.mp4> (Last viewed Nov-2019).
5. W. E. Vaughan, *The Old Library Trinity College Dublin 1712-2012*, Four Courts Press, 2013.
6. John G. Byrne, *The Trinity College Dublin 1872 Online Catalogue*, In: Ed: Simone Marinai, Andreas R. Dengel, Proc.Document Analysis Systems VI, 6th International Workshop, DAS 2004, Florence, Italy, September 8-10, 2004, Berlin, Springer-Verlag, 2004.
7. Glynn Anderson, *Computerising a library catalogue using optical character recognition*, M.Sc. Thesis, Trinity College Dublin, 1993 [Library Shelf mark: THESIS 2842].

8. Brian Cullen, *A policy based approach for assuring Internet based services*, M.Sc. Thesis, Trinity College Dublin, 2004 [Library Shelf mark: TX-1-456].
9. Ruth Clarke, *User-oriented access to a multilingual database*, M.Sc. Thesis, Trinity College Dublin, 1995 [Library Shelf mark: THESIS 3882].
10. Mark A. Deegan, *The development of a searching and retrieval system for the 1872 Printed Catalogue*, M.Sc. Thesis, Trinity College Dublin, 1994 [Library Shelf mark: THESIS 3193].
11. Mark Fynes, *A remote administration and system monitoring solution for Windows 95*, M.Sc. Thesis, Trinity College Dublin, 1999 [Library Shelf mark: THESIS 4775].
12. Donnla nic Gerailt, *Dictionary characteristics in cross-language information retrieval*, Ph.D. Thesis, University of Cambridge, 2003.
13. T. R. Jackson, T.R. (Ed), *Frozen in Time - the Fagel Collection in the Library of Trinity College Dublin*, Lilliput Press, 2016.
14. Trevor Peare, *personal email to Jane Grimson, 24-Nov-2017*, see transcript in the related folder in this catalog.
15. Trinity College Dublin, *Introduction to the 1872 Printed Catalogue*, see: <https://www.scss.tcd.ie/misc/1872catalogueB/>
Last viewed Nov-2019.
16. Trinity College Dublin, *History of TCD Library computing*, see: <https://scss.tcd.ie/SCSSTreasuresCatalog/hardware/TCD-SCSS-X.20180907.001/TCD-SCSS-X.20180907.001.pdf>
Last browsed to on 7-Sep-2018.

See the associated documents in the related folder in this catalog.