Bibliothèque Nationale de France:
Libraries Facing the Digital Challenge

TRINITY COLLEGE DUBLIN

nli

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE
Ambassade de France en Irlande

Bibliothèque Nationale de France:
Libraries Facing the Digital Challenge

Overview of lecture by Bruno Racine
Trinity College Dublin, September 2011
including interviews with and reflections by
Hadrien Laroche, Fiona Ross
and Robin Adams

Overview

Catherine Morris

Digitization is not the enemy of the book but there is no turning back. It is the responsibility of libraries and their regulatory authorities to commit themselves, to move forward.

Bruno Racine

Introduction

The president of *Bibliothèque Nationale de France* (BnF), Bruno Racine communicated first hand to a libraries audience at Trinity College Dublin in September 2011 the ways in which the BnF had faced the challenges of the digital age not only in relation to its building infrastructure, its own book collections but also in developing cultural and financial partnerships that have enriched the digital projects and collections of France's National Library. This document is a summary version of the lecture and attempts to capture some of the information shared by Bruno Racine during his first visit to Ireland, including the views expressed about the symposium by the three partners who co-created and participated in the event: Fiona Ross (director of the National Library of Ireland (NLI)), Hadrien Laroche (Cultural Counsellor of the French Embassy in Dublin) and Robin Adams (Head of Trinity College Library).

Context of the Symposium

Under their new partnership agreement, the NLI and Trinity co-created the Paris Symposium in an attempt to further join forces and invite the French Embassy to help them in creating a local, national and pan-European dialogue about how libraries can face the challenges of the digital age during economic recession and to explore ways in which they can create a shared interface on digitization. The symposium aimed to establish a clear framework of information about the French model of digitization policies and procedures: how are choices about material to be made? How can these huge projects be funded? How can strategic connections with pan Irish and pan European projects and protocols be brokered? And what technologies and training will be needed to make sure that the 'born digital' archives and materials digitized today will be available and accessible to users and researchers in the future?

Gallica

Mr Racine examined how the Digital Library Charter made the digitization and provision of user access a legal part of the BnF's remit. 'Gallica' is the online digital library service created by the BnF. It is the online research tool through which digital materials can be located, accessed and downloaded. The BnF publicizes and is constantly improving access to Gallica. The selection of materials for digitization impacts directly on how archives are made accessible to users through Gallica.

The library selects materials and organizes them under a series of 'themes' that include: underused collections; the long 19th century; rare visual archives and ephemera. Journals and newspapers represent two priority collections. The themed and priority collections are digitized in-house through the BnF's own funding structure. Some collections (such as visual culture, newspapers and journals) involve partnership agreements with specialist departments across the library (such as Art History) or with external partnerships.

The BnF has developed a flexible and multi-faceted strategy to deal effectively with copyright issues for living authors, library holdings that are out of print and online user downloadable access. The use of the 'creative commons' copyright, partnerships with private publishing houses, electronic co-publishing deals, and collaborative working partnerships with contemporary journals and newspapers form part of this strategy.

Harvesting the Web and Preservation of the Digital Library

The 2006 Internet Legal Deposit Act extended traditional legal deposit laws and requirements to publications and information published or embedded in media produced or distributed on French territory. This does not apply to corporate or for profit documentation: those documents legally required by the library and which are harvested into their collections are those aimed at an audience and that serve a public purpose.

Audience/user research: From 2006, the BnF researched potential 'user' groups and 'user methodologies' that have informed their harvesting and access policies. The choices of materials that are harvested from the web (outside of the legal deposit acquisitions) are based on the findings of three 'user' groups: 'researchers', 'professionals' and 'average users'. This research resulted in three recommendations: (i) greater levels of transparency and better communication with users and the broader public; (ii) activate new communities of 'users' through traditional networks of amateur and academic researchers; (iii) develop a specialist set of remote user tools on Gallica that will serve a professional target audience.

Bulk harvesting and crawls: since 2004 the BnF has undertaken a series of experimental 'crawls' through the internet to harvest and record data on French domain sites. These crawls were initially organized around a series of thematic based events (such as general elections). The harvesting and regular Internet crawls are now a major feature of the library's operations. Much of the early work on Internet crawls and the library, in association with an external organization the Internet Archive, carried out harvesting. Since 2009, a permanent unit has been established within the Library's Legal Deposit Department to oversee the daily operations.

Bulk harvesting is conducted all year round 24 hours a day by robots: these machines are bulk harvesting but not blind harvesting: their crawls are tightly controlled by rules and settings. Within this new internal unit, the BnF has built a fully integrated workflow based on virtual architecture and the integration of open source software agreed by the International Internet Preservation Consortium (IIPC). Since 2004, over 200 terabytes and 13 billion files have been archived from the web. All web archives are immediately indexed by the library for a search within Gallica by URL and date. While permissions are not required from publishers to harvest websites, access is restricted. Due to data protection, users can only access data collated through live web crawls inside the library on designated computers. The BnF currently focuses its efforts on harvesting rich media (e.g. video material), e-books and e-journals but this means that other forms of critical social cultural material remains uncollected.

Safe long-term storage and preservation of the BnF web archive and digital materials: The size of the web archive and the extensive variety of file format types impose new preservation strategies that require more than 'bit stream' preservation. The emulation of old web platforms and the documentation of 'odd' or irregular formats will be necessary to keep the archive accessible in the long term. So far, the BnF has made its web archive a major collection track within its digital repository datacenter that is currently under construction. Mr Racine commented that for real long-term preservation to be guaranteed, digital documents will have to be printed on high quality paper and stored in purpose-built repositories located in the French countryside.

Accessibility and Public/Private Partnerships

The preservation of data harvested from the web is a major task that requires the BnF to activate and connect with a national, European and global support network and to work with organizations such as the International Internet Preservation Consortium. The BnF aims to use the resources of this network to lobby for joint software development, advocacy, and a collaborative collection building. The BnF further strengthens its support network by sharing technology, information, data and digital archives: Gallica is harvested into 'Europeana', the library also supplies the French digital network, the BnF plays an active role in the World Digital Library and supports numerous digitization programmes with other national libraries.

National Partnerships, resource sharing and training: Since 2008, the BnF has partnered with public libraries, higher education and research libraries, libraries under the authority of government and parliament, libraries with diverse status (such as associations, and foundations) and archive and museums libraries. The aim is to disseminate and promote digital resources as broadly as possible and to achieve the long-term preservation of the new archives through sharing the costs, knowledge, technical skills and human resources. The BnF aims to share its digitization programmes, skills and resources with regional associated libraries and with thematic networks of libraries.

Public Private Partnerships: In 2010 the French Government raised a National Loan of 35 billion ('the grand emprunt national') to fund the country's strategic priorities in culture and education over the period to 2030. Following the recommendations of the 'Commission on Strategic Investment Priorities and the National Loan', 750M was set aside for the digitization of cultural materials through Public/Private partnerships. This programme of work is to be implemented not through the library alone, but through the creation of multiple partnerships with private business. In order to promote and to deliver this new programme of work a new formal agreement has been reached between French right-holders, the BnF, the French Ministry of Culture and the Prime Minister (through the government investment commission).

The BnF has forged partner agreements (undersigned by the Ministry of Culture) with several private contractors to carry out digitization work on two of its collection areas: Printed Collections and Special Collections. A new set of high quality standards has been defined as operational guidelines for this new partnership project: the documents will be digitized with a high level of OCR accuracy (ranging between 98.5% and 99.9%) and at least 10% of the documents from the Printed collections will be provided in an e-Publishing version compatible with e-readers. The principal aims attached to the digitization of special collections are to generate digital copies of treasures and masterpieces from the graphic collections. The preservation as well as the communication of the treasures is also of paramount importance.

Two private contractors have also been commissioned to carry out two strands of data work: (i) digitization using original documents (680,000 images are being created by in-house workshops that address specifically fragile original documents that cannot be moved off-site), and (ii) digitization of microfilms and reproductions (to generate 916,000 images, work that is carried out off site in Germany).

Public Private Partnership funded digitization: The BnF has set up a special department in the library to foster new relationships with business and to connect collections with private funding. The library is therefore able to connect digitization projects with targeted businesses and offer to match any private investment with government funds. The PPP model in large part aims to grant access to copyrighted, orphan collections and out of print books. A call for partnerships was launched in 2011 and aims to connect private business investment to 12 areas of the Library's collections. Key areas for legal concern relate to: the commercial rights on digital files and the framework for agreed protocols on access to newly digitized archives. The BnF has created evaluation criteria for standards that private partners have to sign up to. While the BnF has set up an internal department to promote PPP, this initiative is still at a very early stage and it is not yet clear how effective or viable private sponsorship of the public digitization project will be.

Sharing SPAR (Scalable Preservation and Archiving Repository): SPAR is a system that has been designed and developed by the BnF to preserve its digital material. It stands as a benchmark in the world of sustainable digital information. The cost of implementation so far amounts to 3.3M. The objectives include getting a return on investment and sharing this tool with other public institutions and private partners in an attempt to distribute maintenance costs.

Commercialization of out of print book collections: another dimension of the BnF's PPP strategy relates to the digitization and exploitation of 500,000 out-of commerce 20th century books published in France and now only accessible in the BnF collections. The aim of the project is to generate a new commercial exploitation and value for these books. The market for the books will operate through the

Internet where the books will find new commercial value. It is the duty of the private partners to ensuring a fair remuneration of right-holder.

Conclusion

Mr Racine ended his talk by suggesting that new educational research and cultural possibilities will emerge by promoting the digitization of local archives and by working in partnership with a broad range of national cultural institutions and private businesses. The BnF hopes to generate the same kind of momentum at a regional level as from within its own collection by working with the agreed regional partners, all of whose activities include digital referencing and promotion of local cultural heritage. The Digital Library Charter, the new PPP model of investment (whereby the BnF pays 50% of the costs and provides training and technology to national partners), and involvement of the Ministry of Culture have all combined to generate a national topic-based digitization programme in legal science and cultural heritage.

Perspectives

Interview with Hadrien Laroche

As the Cultural and Scientific Counsellor in the French Embassy in Ireland, what did Bruno Racine's visit to Ireland and his talk to the libraries community mean to you and to the French Embassy?

Well, first of all we are very keen to cooperate with every aspect of Irish culture – this means art, audio visual, the film industry, education, sciences – every aspect of Irish culture. But of course, the written cultures of the book, the knowledge, the memories, and libraries are most important for us.

You took the initiative to invite Bruno Racine, whom I told you I knew, and of course, the French Embassy was immediately ready to support this initiative. The connection between culture and universities is very important for us. We try to implement very strong cooperation with all the universities in Ireland and we look for opportunities to develop projects and support initiatives across all disciplines – the sciences, culture and economics, languages. For us, intellectual debate is of paramount importance and we have a special mandate for this: to bring as many French writers, artists, scholars, philosophers, scientists to Ireland. This is not only to enable them to speak to new audiences, but also to initiate dialogue within a new community of peers that can help develop discussion and ideas. We promote and create different events throughout the year to help link and connect those people. Inside this programme, the project and event that you initiated was of course one of the most important for us.

What can you tell us about Bruno Racine's work in France?

I know Bruno Racine for quite a long time now. He was director of Centre Pompidou in Paris when I first met him. He is both a civil servant and a writer: a very interesting combination. Since he arrived in the BnF he has done a wonderful and very impressive job particularly in driving forward the digitization aspect of the libraries agenda. As you know, he wrote a book on the topic of digitization and Google and so he has clearly been one of the leading intellectuals in this field who is ready to debate the implications of technology, memory and libraries today .

Could you say something about French national culture today and the digitization of heritage? You as a writer of fiction and philosophy must have used the BnF as a researcher yourself. What does digitization mean for writers? What does it mean to digitize cultural

heritage? It is clear that the French government have really prioritized this and given it a major role in contemporary French Cultural life. As a writer could you just reflect on some of the implications of that[1].

Well, I have some personal memories of the French national library. I finished my second novel in this beautiful profound reading room in the BnF. I like to go there. I like to work there. When I arrived in Dublin, thanks to the Ambassador, I had the privilege of renovating my offices in Kildare Street and I made a point to pick the same floor tiles as the BnF – they are distinctive red brick and sandy colour. My model was the BnF! The most important thing that I want to say is that Ireland is strongly involved in culture and Irish people always find a means to survive and to live and sometimes to win. During the Famine, during the War of Independence and now during the economic crisis, culture was and is always a way for a country to re-invent itself and artists ask and answer the critical questions of the moment. I remember when Enda Kenny, for his first visit abroad, met with Obama in Washington on St Patrick's day he said that culture was the means by which Irish people have always reinvented themselves. So we, the French embassy and the French people support the Irish in their belief that in this time of crisis, poets are needed and books are needed. Books and poets and intellectuals sow food for thought, they give meaning to the times and the crisis through which we are living. And artists are essential for the economy of the country. This is why we are trying to support Irish people and the development of cultural European links in every aspect of education, including the production of books and digitization within libraries.

Do you think of the contemporary connections between Irish writers and France as historically rooted?

Yes, absolutely. Especially here in Trinity where there is a tradition of educational exchanges that starts with Samuel Beckett going from Trinity to Ecole Normale Supérieure (ENS) in Paris to take up a post as a lecturer. This is a model of exchange that we aim support by bringing over specialists from France to engage with universities in contemporary Ireland. There are many forms of educational reciprocity and very good connections between Trinity, the universities in Ireland and French institutions like the ENS or others. Now, thanks to this event and the initiatives that will follow, we have developed good working links between the National Library of Ireland and the BnF.

What did you think about the economic model that Bruno Racine has developed at the BnF?

I was fascinated when Bruno Racine, during his talk, spoke about those places that are maybe virtual – the library almost as a virtual home to keep the virtual memories of the book. Digitization is about memory. But you have to be able to access and 'remember'

---

[1] Bruno Racine, Google et le nouveau monde, 2010.

this memory. To construct a place, a virtual place maybe, to contain and keep this 'memory of the memory'. Do you know that the BnF was built on the exact place where the Nazis created a camp in Paris in 1942 ? A very peculiar camp in fact, known as 'camp d'Austerlitz"' or 'Galeries d'Austerlitz', a camp of objects where all the furniture and clothes and whatever you think of looted by the Nazis in the apartments of the deported jews were sorted out, labeled by Jewish prisoners and then railway to German families in Germany. Memory is important, indeed.

The other aspect of the digitization programme is, of course, money. I understand that it is very unique of Bruno Racine and the French system to have linked in the digitization of national cultural heritage with private industry to help with this enormous job. In the beginning it was an economic war between the state and Yahoo and Google for dominance of all these virtual library projects. But Bruno Racine was the first to really make a change and an intervention in France where the dominant culture was always for the state to finance culture and to separate culture from the private sector. Unlike the United States, connections between the private and public sectors are completely new in France. Racine was the first to try to find a way to engage the private sector in this public project. It is very important for the people of France to achieve the digitization and availability and dissemination of the cultural heritage. There is now money coming from the private sector to support this project. It is not easy but it is working. Also the partnership between Google and BnF is a new step forward. All of this offers good examples that were brought to the fore in the discussion with Trinity and the National Library at the Paris Symposium and subsequently.

Interview with Fiona Ross

Why was it important to hear Bruno Racine talk in Dublin and address the National Library of Ireland, Trinity College Dublin and the broader libraries community of Ireland?

We tend to be very domestic and nationally focused here in Ireland and there is nothing like having the opportunity to hear someone from another country talk about the same issues and challenges albeit in a different context and on a different scale. Having someone of the calibre and stature of Bruno Racine from a country that leads the way in honouring and valuing culture and cultural heritage was both inspirational on a high level, big vision point of view but also from a pragmatic and practical point of view. His openness, his honesty, his ability to share quite detailed information with us was very valuable in proving that there are models out there that can be deployed to make the quantum leap we are trying to make in terms of cultural heritage.

What for you are the benefits of having partnerships with other libraries such as Trinity College?

Ireland's a very small country and we tend to think of ourselves as being much bigger than we are. Yet when you actually look at the demographics it is amazing to me that we have as many libraries as we have and that we have as many legal deposit libraries as we have and that we have as many varied collections. As we look to this era of economic difficulties, the silver lining to that is definitely the sense that everyone is willing to collaborate. What people would have done on their own in the past they are now willing to do together. Having a connection not just with the Trinity's library but with other university libraries has been absolutely fundamental to help us understand the challenges that we as libraries face in the 21st century: our relevance to our users and our readers; our challenges in terms of our collecting remit both for our readers, students and other users; and our relevance to our citizen tax payer who supports us on a day to day basis. It is only by sharing that knowledge and experiences that we can build a sense of how we may move forward.

A key dynamic that came from the Paris symposium was a focus on the digital. We clearly have major physical library buildings and infrastructures for book collects and material archives in the National Library and at Trinity. But what would you see as the key digital challenges for the National Library? And is it possible within this economic climate for you as Director of the National Library to meet any of those for the Irish nation?

I think we have absolutely no choice. If I learned one thing in the last eighteen months it is that the future is online and the future is digital. And so there is absolutely no way that we can ignore that opportunity or challenge (depending on which perspective you take on a given day). The value of seeing the model that the BnF in France has in place, particularly the way it is funded and the Public Private Partnership component that is built into it, is an inspirational model to prove that there are models out there that work. But the BnF is only

one of many models that we can see across Europe. There is no doubt that the National Library of Ireland's core driving ethos now is that it has to be a 21st Century Library. That vision is being led by one word and that is 'accessibility' – which, of course, technology allows you to do online in ways that could never have been conceived of before.

Within the discussions that we have started with Bruno Racine, the French National Library and Trinity, are there clear objectives that you have in mind? Are there any particular archival projects or training programmes that you would like to see delivered within this pan European context of cultural diplomacy?

The BnF certainly is a model and it gives us an example of how things might be done. Bringing it back closer to home in thinking about ways in which Trinity Library and the National Library have and do collaborate: we would do so and have done so at the level of collections. But there is a clear opportunity that we would look at and we've had preliminary conversations about: and that is around the digitization of our respective newspaper collections. There is an example of two fantastic, unique collections – of the newspapers held at the NLI and then complimentary collections held at Trinity, which we know that there is huge interest in both domestically and internationally, for the academic and the populous in general. And that's an area where by working together we could jointly digitize the newspaper collections of Ireland. And indeed, the extent that they exist in other libraries, we would look to bring them in. But primarily, given our collecting remit: it is the NLI and Trinity that have the big collections here for our newspapers.

Could you just say something about the storage? Clearly that is a major issue for both libraries. Is there a joint vision being forged around the issue of storage and space?

Absolutely. And again, going back to my previous point. In the past we all looked to develop storage solutions on an individual institutional level. But that is no longer relevant in this day and age certainly given the size of this country. And so the library at Trinity and the National Library and University College Dublin (and indeed, the other university libraries and institutions such as the Royal Irish Academy and the National Archives) have all conceived of a collaborative storage facility. This will be one big storage facility built to the highest international standards into which these rare, unique and important and endangered collections would be housed in perpetuity. That also allows you a huge level of activity around what would be called 'the ingest' if this facility were ever it to be built. And indeed, it is a vision, which we must all hold to. Digitization would be a huge part of this plan whereby you may have a digitization facility or factory as it were on site at this location. So as materials move from their homes at an institutional level (be they inside the libraries at Trinity or here at the National Library) they will be handled once, digitized, and then put on a shelf for ever but then available on line forever as well.

Ireland's just had a recent budget in December 2012 that took two days to deliver to the public. Can you reflect on the financing now for the year to come of the National Library? Can you give us some figures and reflect on the implications in terms of staffing, facilities, and acquisition and obviously in this extraordinary vision that you have for the digital collections?

Anyone answering those questions will be able to attest to the very difficult and at times very depressing economic circumstance we find ourselves in. The national Library will have seen essentially a 40 per cent cut in its funding between 2008 to 2012. That funding situation will only get worse as we go forward. We have a grant allocation of 7.1 million and little over 6 million of that 7.1m is allocated to staff costs and fixed costs. So this budget leaves us with a tiny amount of money to play with in terms of the type of outsourcing that you do day-to-day at a library such as people to move stuff around; box making; cleaning; and the dream of making some kind of quantum leap in cataloguing by bringing in some of the outside contract services that we have been able to do over the last number of years. What it means from a budget point of view is that the National Library will have to operate within its own staff cohort. So what the people here can do, will be done. That obviously is not enough for a National Cultural Institution given the scope of our vision and our aims and our aspirations both domestically and internationally; it is not enough in terms of research and academia; or indeed for cultural tourism and family history. So we will have to partner and we will partner. We will build on the great relationships that we have developed over the last couple of years both domestically and internationally as well as with the other libraries and other cultural bodies and universities. We will partner with the commercial world also. And we will have to partner and do deals with others in order to do anything substantive over the next five to ten years. But we shouldn't be afraid of that because such partnership collaborations offer us a great opportunity as well. Sometimes when you have nothing, it is at that point in time when you are able to make some of the inspired decisions.

Reflections by Robin Adams

The Library of Trinity College has collections that reach back to the 3rd Millennium BC up to the electronic records of contemporary life and scholarship. It continues to acquire material in traditional print and manuscript formats and enjoys the benefits of Legal Deposit from Ireland and the United Kingdom. It collections now include around five million items in printed books, maps, manuscripts and music. The acquisition, description, preservation and storage of such large collections are a significant logistical burden. Like many similar libraries, the Library is facing a crisis of accommodation in providing

appropriate facilities for the storage of very large volumes of lesser-used print material, in order to provide services for emerging research and learning requirements. The partnership for a combined storage facility with the National Library of Ireland and the Library of University College Dublin is offering the real prospect of a facility, which can provide accommodation for the next generation on a national scale. On this foundation can be built a number of co-operative national initiatives such as conservation, digitization and processing.

Making such material available to students and scholars of the College remains a primary responsibility, together with fulfilling its role as a national and international research destination. With over 12,000 visits annually by external readers, the number of readers from the home community is almost matched by those visiting from outside.

To this traditional range of duties is more recently added the challenge of creating, acquiring, preserving and making accessible sources of information in electronic form. The Library presently makes available research material in the form of electronic journals, databases and other reference resources and more than 60 per cent of its information budget is now expended on electronic products. Such external content provides Trinity College scholars with access to global resources.

A greater strategic challenge is posed in providing electronic access to the rich and in many cases unique holdings accumulated by the Library over four centuries of collecting. It is these collections which give the Library its standing as an internationally recognized destination for scholars and provide Trinity's academics with the opportunity to undertake research not possible elsewhere. As a strategic change, these collections and the services around them have now been brought together as the Research Collections Division, with the aim of facilitating research across different formats and providing a large scale of operation for planning, funding and development. Under this structure, the Manuscripts and Archives Research Library, the Department of Early Printed Books and the Music and Map collections will operate as a co-coordinated unit.
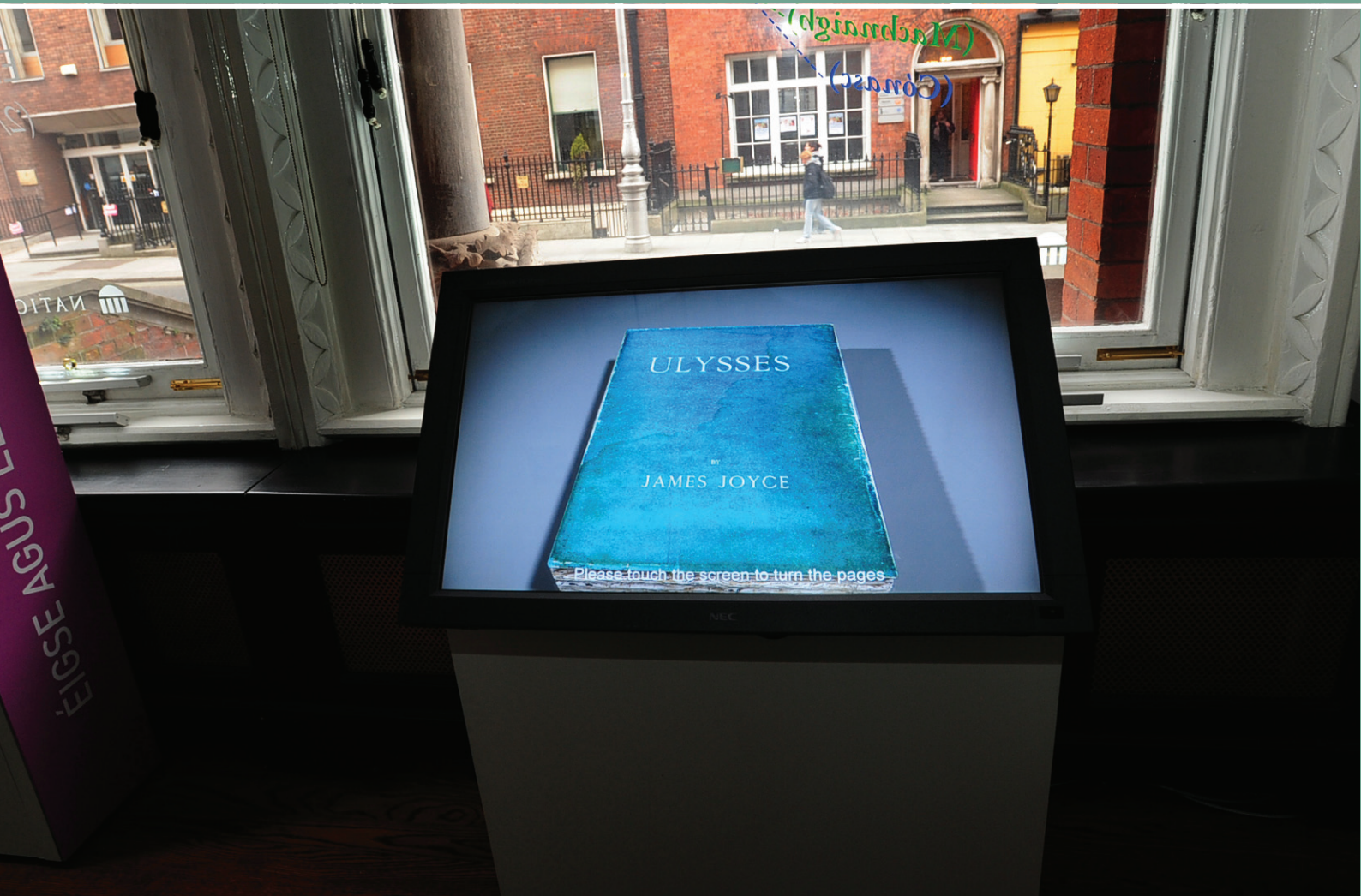
The creation of digital surrogates to print, manuscript and photographic material is an expensive and labour intensive process, as is the essential work of preliminary stabilization of the originals through conservation and the generation of appropriate descriptive metadata.   Access to digital versions of material opens up new avenues of investigating, teaching and research and the College has identified the digital humanities as an area for strategic investment, with the appointment of a lecturer in the discipline.

To date, the Library has addressed the challenge of digitization through a series of discrete initiatives aimed at digitizing collections of material such as maps, or imaging partial collections as an element of the production of a commercial package. Trinity library's Digital Resources and Imaging Service (DRIS) was established in 2008 to undertake a programme of work determined by the needs of the Library and the academic community, and it has completed some important work, contributing to the Europeana programme. At present, the DRIS unit is small and its funding is limited, so the production of electronic content is inevitably restricted, thus frustrating plans to open up new areas of research and enhance access to Library treasures as rapidly as possible to the widest community.

The scale of the challenge requires a large-scale response and the opportunity to work with the National Library of Ireland presents real opportunities for both institutions in fulfilling the need to digitize their collections at a macro level. There is considerable technical and bibliographic expertise within the Libraries but not the scale of operation to undertake digitization at sufficient throughput.

The example provided by the Director of the Bibliothèque Nationale de France on the programme of digitizing France's literary and historical archive is an enormously encouraging stimulus to the introduction of mass digitization of collections. It has demonstrated that government commitment to the cultural record is an essential requirement if a country is to preserve and make available its history and recorded thought in a format which has become ubiquitous. Trinity College Library is following up on the presentation by Bruno Racine by strengthening its links through the Cultural Co-ordinator and arranging exchange of experience visits to the Bibliothèque Nationale. The Library has particularly rich collections of French language publications. In particular, its 17th century drama is probably one of the strongest outside France and it looks forward to co-operative ventures with colleagues in France.