

CIDOC 2016 CHAIR'S REPORT / RAPPORT DU PRESIDENT <i>Monika Hagedorn-Saupe</i>	2
THE ICOM GENERAL CONFERENCE IN MILANO: CIDOC, DOCUMENTATION STANDARDS AND DIGITAL PRESERVATION <i>Jan Behrendt</i>	4
"IT'S STILL A CLEVER CAMEL THAT GOES TO THE WELL" FIVE YEARS OF EXPERIENCES WITH THE COOPERATION OF ARCHIVES, LIBRARIES AND MUSEUMS IN ESKILSTUNA, SWEDEN <i>Susanne Nickel</i>	9
IMPLEMENTING LIDO <i>Gordon McKenna and Regine Stein</i>	15
USING CIDOC CRM AND LINKED OPEN DATA TO MODERNIZE ARTEFACTS CANADA <i>Heather Dunn</i>	26
PERSISTENT IDENTIFIER SERVICES FOR THE HUMANITIES WORKSHOP <i>Gordon McKenna</i>	32

A WORD FROM THE CHAIR

MONIKA HAGEDORN-SAUPE

Last year's CIDOC conference was part of ICOM2018 in Milan, the triannual conference of ICOM.

CIDOC had a very good meeting with more than 100 participants. As traditional with CIDOC, a mixture between plenaries, workshops and working group meetings took place. Some of the topics discussed, were: LIDO, the data harvesting format for museum data into portals; the question of long-term preservation of digital museum data as well as a panel discussion on "The UNESCO/PERSIST Guidelines for the selection of digital heritage for long-term preservation", which were launched 2016 by the UNESCO/ PERSIST Content Task Force. Thanks to the hosts CIDOC had a great support from a group of highly engaged young volunteers who supported the participants on all matters arising in the course of the conference. Due to the local arrangements – having all meeting rooms nearby – an intensive exchange was possible between the experts within CIDOC but there was also a chance for CIDOC members to connect with members from other International Committees and colleagues from other Committees could participate in one or the other CIDOC event.

In our professional excursion we went to Torino. Together with the CIDOC Board the local organization group had prepared a very interesting and stimulating day, which enabled an intense exchange on museum documentation and networking between cultural heritage institutions with colleagues from Torino. A general introduction "What is Documentation?" was followed by three discussion groups on different topics. In the afternoon the participants split in groups to visit different memory institutions, continuing the exchange of experience and with the focus on collaboration between museums, libraries, archives and related institutions.

During the AGM of CIDOC the results of the election were presented and the new board had its first meeting. Information on all board members can be found on the CIDOC website.

When you have been visiting CIDOC's website lately, you have seen that we now have a different logo. This follows the decision by ICOM for a new corporate design and new logos for ICOM and all its committees.

RAPPORT DE LA PRESIDENTE

MONIKA HAGEDORN-SAUPE

La conférence CIDOC de l'année dernière a fait partie d'ICOM2018 à Milan, la conférence triennale de l'ICOM.

Le CIDOC a eu une très bonne conférence en 2017 avec plus de 100 participants. Comme il est traditionnel avec CIDOC, la conférence a consisté de séances plénières, d'ateliers thématiques, et de réunions de groupes de travail. Les sujets discutés ont inclus : LIDO, le format de collecte de données pour les portails de musées, la question de la préservation à long terme des données numériques des musées, ainsi qu'une table ronde sur "Les lignes directrices UNESCO / PERSIST pour la sélection du patrimoine numérique pour la conservation à long terme", qui ont été lancées en 2016 par le groupe de travail UNESCO / PERSIST. Grâce au comité local d'organisation de la conférence, le CIDOC a bénéficié du soutien d'un groupe de jeunes volontaires très engagés qui ont aidé les tout au long de la conférence. La proximité des salles de réunion a permis un échange intensif entre les experts du CIDOC, et aussi entre les membres du CIDOC et les membres d'autres comités internationaux d'ICOM qui furent les bienvenus dans toutes les présentations CIDOC.

L'excursion professionnelle nous a emmenés à Turin. Avec l'aide du conseil d'administration du CIDOC, le comité local d'organisation de la conférence avait préparé une journée très intéressante et stimulante, qui a permis une discussion intense avec des collègues de Turin sur la documentation des musées et la mise en réseau des institutions du patrimoine culturel. Une séance d'introduction générale "Qu'est-ce que la documentation ?" a été suivie par trois groupes de discussion sur différents sujets. Dans l'après-midi, les participants se sont répartis en groupes pour visiter différentes institutions de la mémoire, poursuivant l'échange d'expériences, et mettant l'accent sur la collaboration entre les musées, les bibliothèques, les archives et autres institutions reliées à celles-ci.

Les résultats des élections du conseil d'administration du CIDOC ont été présentés lors de la réunion de l'Assemblée Générale du CIDOC, et le nouveau conseil d'administration y a tenu sa première réunion. Les informations sur tous les membres du conseil d'administration peuvent être trouvées sur le site web du CIDOC.

End of 2016 CIDOC together with COMCOL planned a seminar in Brazil on “Building museum collections for the future” and applied to support from ICOM’s special projects funds. We are happy that support has been granted and the seminar will take place later in 2017 in Rio de Janeiro.

In January 2017 the CIDOC board met in Berlin, Germany discussed CIDOC’s communication and social media strategy and worked on updating the website, planning future activities and especially planning the upcoming CIDOC 2017 conference, which will take place in Tbilisi, capital of Georgia, in September 2017. In March members of the board met in Tbilisi with the Local Organizing Committee of CIDOC2017 and since then we have been in continuous exchange with the colleagues in Georgia.

I am looking forward to meeting many of you at our conference later this month in Tbilisi at the Georgian National Museum.

Lorsque vous avez récemment visité le site Web de CIDOC, vous avez pu voir que nous avons maintenant un nouveau logo. Ceci fait suite à la décision de l'ICOM de créer un nouveau design d'entreprise et de nouveaux logos pour l'ICOM et tous ses comités.

Fin 2016, le CIDOC, en collaboration avec le COMCOL, a organisé un séminaire au Brésil sur le thème « Construire des collections muséales pour l’avenir » et a sollicité soutien au fonds de projets spéciaux de l'ICOM. Nous sommes heureux que le soutien ait été accordé et le séminaire aura lieu plus tard en 2017 à Rio de Janeiro.

En janvier 2017, le CIDOC s'est réuni à Berlin en Allemagne pour discuter de la stratégie de communication et des réseaux sociaux du CIDOC. Il a travaillé sur la mise à jour du site web, la planification des futures activités et surtout la planification de la prochaine conférence CIDOC 2017. En mars, les membres du bureau se sont rencontrés à Tbilissi, Georgie avec le comité d'organisation local de la conférence CIDOC2017 et depuis, nous avons été en échange continu avec les collègues en Géorgie.

Je suis impatiente de rencontrer beaucoup d'entre vous à notre conférence plus tard ce mois-ci à Tbilissi au Musée national géorgien.

THE ICOM GENERAL CONFERENCE

IN MILANO: CIDOC, DOCUMENTATION STANDARDS AND DIGITAL PRESERVATION

JAN BEHRENDT

In Milano 2016, more than 3.400 participants from 129 countries gathered in uncountable national, regional and international committees. They formed think tanks for museum and heritage issues from all over the globe. Milano 2016, that was also a remarkable museum fair with 111 exhibitors displaying what's new on the market of museum technology and services. And Milano 2016 would have been nothing without its fantastic people – like Christo and Orhan Pamuk, who gave the keynote addresses, or Suay Aksoy, who was elected new ICOM president during the session, or the many dedicated museum specialists that really shape ICOM.

Since it is physically impossible to get a complete overview of all different sessions that took place simultaneously (CIDOC, constituting but one of the 30 international committees, ran three parallel workshops most of the time), any report on Milano 2016 will be highly selective. My path through Milano 2016 has led me through various panels, workshops and presentations. Most of them were related to CIDOC, some of them belonged to other international committees that I wanted to know better in order to understand more about ICOM.

There was a great number of panels that really inspired me and gave me new impulses for my work as a curator and collection manager.

Jonathan Whitson Cloud gave us a very comprehensive and relevant overview on what documentation is all about. He started his presentation with the brief and simple sentence “A museum without documentation is just a junk shop”, offering various definitions of museum documentation afterwards and explaining a wide concept of museum documentation, which goes far beyond the traditional object documentation.

LA CONFERENCE GENERALE DE

L'ICOM 2016 A MILAN: CIDOC, NORMES DE DOCUMENTATION ET PRESERVATION NUMERIQUE

JAN BEHRENDT

À la conférence Générale de l'ICOM 2016 à Milan, plus de 3400 participants venus de 129 pays se sont réunis dans d'innombrables comités nationaux, régionaux et internationaux. Ils ont formé des groupes de réflexion autour de questions relatives aux musées et au patrimoine de par le monde. Milan 2016 avait également un salon d'exposants remarquable avec 111 exposants sur ce qui est nouveau en ce qui concerne les services et le marché de la technologie pour musée. Et Milan 2016 n'aurait rien été sans ses fantastiques participants – comme Christo et Orhan Pamuk, qui ont délivré les discours d'ouverture, ou comme Suay Aksoy, qui a été élu nouveau président de l'ICOM pendant la session, ou comme les nombreux spécialistes de musées qui font vibrer l'ICOM.

Etant donné qu'il est physiquement impossible d'avoir un aperçu complet de toutes les sessions qui se sont déroulées simultanément pendant la conférence (le CIDOC, n'étant qu'un des 30 comités internationaux de l'ICOM, a organisé trois ateliers qui étaient parallèles la plupart du temps), tout rapport sur Milan 2016 est très sélectif. Mon parcours à travers Milan 2016 m'a conduit à travers divers groupes d'experts, ateliers et présentations. La plupart d'entre eux étaient liés à CIDOC, mais certains d'entre eux appartenaient à d'autres comités internationaux que je voulais mieux connaître pour mieux comprendre l'ICOM.

Il y avait un grand nombre de groupes de travail qui m'ont vraiment inspiré et m'ont donné de nouvelles impulsions pour mon travail en tant que commissaire et gestionnaire de collection.

Jonathan Whitson Cloud nous a donné un aperçu très complet et pertinent de ce en quoi consiste la documentation. Il a commencé sa présentation par la phrase courte et simple «Un musée sans documentation est juste une sucrerie», offrant ensuite diverses définitions de la documentation de musée en prenant le

concept à son plus large, en allant bien au-delà de la documentation des objets, qui est la notion traditionnelle qui vient à l'esprit quand on évoque la documentation muséale.



Unfortunately, museum documentation is often misperceived as simple object documentation, often lacking the necessary resources or the adequate priority in museum work. That is why Nicolas Crofts organized a workshop dedicated to museum directors with the appealing title "How to capitalize on museum documentation." Nicolas provided a broad survey on why it makes sense to invest in museum documentation – and how to get valuable returns on your investment.

The workshop helped museum directors to identify:

- "dormant" resources of documentation and to estimate their value;
- "hidden jewels" of a museum (such as objects, people, places, events, materials, techniques, terminology or bibliography) with the help of documentation;
- documentation weak points,

and to purposefully plan digitization efforts.

My own contribution to Nicolas' workshop was a panel on crowd intelligence, social media and documentation. I spoke about how, in my museum, we are gaining information about poorly documented artifacts with the help of social media, how 'folksonomies' can help to complement search engine proficiency of databases, and

Malheureusement, la documentation de musée est souvent perçue comme une simple documentation d'objets, ce qui a la conséquence néfaste de souvent la dépourvoir des ressources nécessaires ou de la priorité adéquate pour qu'elle crée un impact positif sur le travail des musées. C'est pourquoi Nicolas Crofts a organisé un atelier dédié aux directeurs de musées avec le titre attrayant «Comment capitaliser sur la documentation du musée». Nicolas a fourni une vaste étude sur la raison pour laquelle il est logique d'investir dans la documentation muséale – et comment obtenir des retours précieux sur votre investissement.

L'atelier a aidé les directeurs de musées à identifier:

- Les ressources "dormantes" de la documentation et à estimer leur valeur;
- Les «bijoux cachés» d'un musée (tels que les objets, personnes, lieux, événements, matériaux, techniques, terminologie ou bibliographie) à l'aide de documentation;
- Les points faibles de la documentation;

Cet atelier a également discuter comment planifier les efforts de numérisation.

Ma propre contribution à l'atelier de Nicolas a été un panel sur l'intelligence des foules, les médias sociaux et

how visitor involvement can be increased with documentation.

Dorothee Haffner gave a presentation of her impressive project for a comprehensive textile thesaurus, linking many of the existing platforms and resources. Her work might become a model for all multilingual and international thesaurus developments in the future.

Probably the most inspiring debate I participated in during Milano 2016 took place during the meeting of the working group for digital preservation, chaired by Susanne Nickel. The working group has already done excellent work contributing to the development of the "UNESCO/ PERSIST Guidelines for the selection of digital heritage for long-term preservation." Nevertheless, the meaning of a digital artifact, compared to digital documentation, seems to be unclear to many of the decision makers involved.

la documentation. J'ai parlé de comment, dans mon musée, nous obtenons des informations sur des artefacts mal documentés avec l'aide des médias sociaux, comment les «folksonomies» peuvent aider à compléter la compétence des bases de données et la façon dont la participation des visiteurs peut être augmentée avec la documentation.

Dorothee Haffner a fait une présentation de son impressionnant projet de thésaurus textuel complet, reliant plusieurs des plates-formes et des ressources existantes. Son travail pourrait devenir un modèle pour tous les développements multilingues et thésaurus internationaux à venir.

Probablement le débat le plus inspirant auquel j'ai participé pendant Milano 2016 a eu lieu lors de la réunion du groupe de travail pour la préservation numérique présidé par Susanne Nickel. Le groupe de travail a déjà fait d'excellents travaux contribuant à l'élaboration des "Directives UNESCO/ PERSIST pour la sélection du patrimoine numérique pour la préservation à long terme". Néanmoins, la signification d'un artefact numérique par rapport à la documentation numérique semble peu claire À de nombreux décideurs impliqués.



I chose to summarize in this report some of the ideas we developed during the debate, both to promote this important topic and to give an idea of the excellent discussions we had:

When we're talking about the preservation of digital artifacts, we're talking about potentially uncountable different kinds of files that are being created in our everyday lives. We talk about files that are not just

J'ai choisi de résumer dans ce rapport quelques-unes des idées que nous avons développées au cours du débat, à la fois pour promouvoir ce sujet important et pour donner une idée des excellentes discussions que nous avons eu:

Lorsque nous parlons de la préservation des artefacts numériques, nous parlons de différents types potentiellement innombrables de fichiers qui sont créés

copies of an analog reality (like the scan of a photo-negative strip), but that are originals in their very own, digital nature (like the photo taken with a digital camera that has been digital since the moment it was taken).

When we're talking about digital artifacts, we're also talking about e-mails that have been replacing handwritten letters for several decades now. We may be talking about the system files of our cars that register technical data from all engine components that can be evaluated. Or we may just as well be talking about the system files of a combat aircraft that participates in a military operation in one of today's crisis areas.

All these files have something in common: They are witnesses of our time, valuable sources for historical research or even proof for criminal investigations. They were intentionally or indirectly created by human beings. And they were created with the help of underlying operating systems, bearing the risk of becoming useless if the operating system, or the program displaying the file, is outdated. These original files may be lost due to poor storage conditions or data corruption. They may be manipulated unintentionally by other software or intentionally, by individuals whose aim it is to eradicate the proof of what they did. Preserving digital files is likely to become a completely new discipline for conservation specialists. It is not just about copying the file from one safe, redundant storage device to another. It is about maintaining digital heritage comprehensively, including its display programs and the authenticity of use.

Eventually, the possibility to copy and convert files brings up meta-questions on originality and documentation. If you want to preserve, say, an e-mail – how are you going to do it? If you're going to create a PNG screenshot or a PDF printout, you're documenting an e-mail, but you're not preserving it. You lose the e-mail's functionality, its metadata and possibly even attachments etc. In essence, the screenshot of an e-mail is not much more than the photo of a masterpiece of art: It will show what the original looked like, but it will never be the original. So what do you need to do in order to guarantee its preservation? Will it be enough to run a historic version of a 1998 e-mail client software, or do you also need the 1998 operating system to create authenticity? Or are we actually fine with deleting e-mails and replacing them with JPG and PDF copies of them, even if, in essence, this means to delete the original artifact and to keep the documentation only? Since it may require the same technology to preserve digital artifacts and the digital documentation of artifacts, where do we draw a line between artifact and documentation?

In Milano, I learned: Discussions about digital preservation are only just beginning. Despite progress

dans notre vie quotidienne. Nous parlons de fichiers qui ne sont pas seulement des copies d'une réalité analogique (comme le balayage d'une bande photo négative), mais qui sont des originaux dans leur nature propre, numérique (comme la photo prise avec un appareil photo numérique qui a été numérique depuis le moment où il a été pris).

Lorsque nous parlons d'artefacts numériques, nous parlons aussi des e-mails qui remplacent des lettres manuscrites depuis plusieurs décennies. Nous pouvons parler des fichiers système de nos voitures qui enregistrent les données techniques de tous les composants du moteur qui peuvent être évalués. Ou peut-être parlerons-nous des fichiers système d'un avion de combat qui participe à une opération militaire dans une des zones de crise actuelles.

Tous ces dossiers ont quelque chose en commun: ils sont les témoins de notre temps, des sources précieuses pour la recherche historique ou même une preuve pour les enquêtes criminelles. Ils ont été intentionnellement ou indirectement créés par des êtres humains. Et ils ont été créés à l'aide de systèmes d'exploitation sous-jacents, risquant de devenir inutiles si le système d'exploitation, ou le programme affichant le fichier, est obsolète. Ces fichiers originaux peuvent être perdus en raison de mauvaises conditions de stockage ou de corruption de données. Ils peuvent être manipulés involontairement par d'autres logiciels ou intentionnellement, par des individus dont le but est d'éradiquer la preuve de ce qu'ils ont fait. Préserver les fichiers numériques est susceptible de devenir une toute nouvelle discipline pour les spécialistes de la conservation. Il ne s'agit pas seulement de copier le fichier d'un périphérique de stockage sécurisé et redondant vers un autre. Il s'agit de préserver le patrimoine numérique de façon exhaustive, y compris ses programmes d'affichage et l'authenticité de son utilisation.

Finalement, la possibilité de copier et de convertir des fichiers fait apparaître des méta-questions sur l'originalité et la documentation. Si vous voulez conserver, par exemple, un e-mail – comment allez-vous faire? Si vous voulez créer une capture d'écran PNG ou une impression PDF, vous documentez un e-mail, mais vous ne le conservez pas. Vous perdez la fonctionnalité de l'e-mail, ses métadonnées et éventuellement même les pièces jointes, etc. En gros, la capture d'écran d'un courrier électronique n'est pas beaucoup plus que la photo d'un chef-d'œuvre d'art: Elle montre à quoi ressemble l'original, mais cette photo ne sera jamais l'original. Alors, que devez-vous faire pour garantir sa préservation? Sera-ce suffisant pour exécuter une version historique d'un logiciel client de messagerie

being made with the UNESCO/PERSIST Guideline, the public debate outside CIDOC is still lacking clarity on what digital preservation really is and what it is meant to preserve. While some are talking about the curatorial question on which files should be collected and preserved, others are talking about the technical conservation (preservation) issues – on how to best store files and how to display the files in the future. As if this wasn't already confusing enough (having to follow two completely different strings of debate at the same time), even experienced museologists completely forget about the difference between conservation and documentation when they speak about digital artifacts – and risk to delete or convert files, destroying their digital authenticity.

If you wish to pick up on this debate, join the digital preservation working group. Meet you there.

électronique 1998, ou avez-vous également besoin du système d'exploitation 1998 pour créer l'authenticité? Ou acceptons-nous la suppression des e-mails et leur remplacement avec des JPG et PDF copies, même si, essentiellement, cela signifie la suppression de l'artefact original, même si nous conservons leur documentation? Puisque la même technologie pour préserver les artefacts numériques et leur documentation numérique peut être utilisée, où se trouve la limite entre l'artefact et la documentation?

À Milano, j'ai appris que les discussions sur la préservation numérique ne font que commencer. Malgré les progrès accomplis dans le cadre de la ligne directrice UNESCO / PERSIST, le débat public à l'extérieur du CIDOC manque toujours de clarté sur ce que la préservation numérique est réellement et ce qu'elle est censée préserver. Alors que certains parlent de la question de conservation sur laquelle les fichiers devraient être recueillis et conservés, d'autres parlent des questions techniques de conservation (préservation) – sur la meilleure façon de stocker les fichiers et comment afficher les fichiers à l'avenir. Comme si cela n'était pas assez confus (il fallait suivre deux séries de discussions tout à fait différentes), même les muséologues expérimentés oublient complètement la différence entre la conservation et la documentation quand ils parlent d'artefacts numériques – et risquent de supprimer ou de convertir des fichiers, détruisant leur authenticité numérique.

Si vous souhaitez reprendre ce débat, rejoignez le groupe de travail sur la préservation numérique. Nous serions heureux de vous rencontrer.

“IT’S STILL A CLEVER CAMEL THAT GOES TO THE WELL” FIVE YEARS OF EXPERIENCES WITH THE COOPERATION OF ARCHIVES, LIBRARIES AND MUSEUMS IN ESKILSTUNA, SWEDEN¹

SUSANNE NICKEL

Background²

In the beginning (in the late 1990ies) there was the idea of localizing library, archives and museums together in order to give more and better access to the history of Eskilstuna. All archives (even the non-municipal ones) resident in Eskilstuna should be in the same place, all staff in the same office and one doorway to the historical sources should be created.

In 2009 the archives resident in Eskilstuna moved to the so called “House of archives” in the city centre next doors to the library and in 5-10 minutes walking distance to the City museum and the Art museum. Here the Municipality archives, the archives of the Swedenfinns, the archive of Swedish Children’s drawings, the archives of industries and commerce, the archives of popular movements, the municipal music archive and the archive of the City museum share a storage and an open office landscape.

Two years later in march 2011 the archives, libraries and museums resident in Eskilstuna opened a shared information centre in the venues of the library to make archival material and information from all institutions available to the public. This included also a website (<http://eskilstuna.se/eskilskallan>) a journal and events like lectures on local historical topics or celebrating the national day of the archives. The information centre provides:

- Books (on local history)
- Maps
- Databases
- Microfilm
- Genealogic studies
- Information desk

«ET LE CHAMEAU EST TOUJOURS INTELLIGENT D’ALLER AU PUITS» FIVE ANNEES D’EXPERIENCE DE COLLABORATION PARMIS LES ARCHIVES, BIBLIOTHEQUES, ET MUSEES DE ESKILSTUNA, SUEDE.¹

SUSANNE NICKEL

Historique²

L’idée originelle (dans les années 1990) était de placer les archives, bibliothèques, et musées au même endroit pour donner un plus grand et meilleur accès à l’histoire d’Eskilstuna. Il était estimé que toutes les archives de Eskilstuna (même les archives non municipales) devraient être au même endroit, que tous le personnel devrait être dans le même bureau et qu’une seule porte d’accès aux sources historiques devrait être créée.

En 2009 les archives d’Eskilstuna déménagèrent à la «Maison des archives» au centre-ville, près de la bibliothèque et à 5-10 minutes à pied du musée de la ville et du musée d’art. Là les archives municipales, des Swedenfinns, des dessins d’enfants Suédois, de l’industrie et du commerce, de mouvements populaires, musicales de la ville et du musée de la ville partagèrent un entrepôt et un bureau.

Deux ans plus tard, au mois de Mars 2011, les archives, bibliothèques, et musées de Eskilstuna ouvrirent un centre d’information commun dans la bibliothèque pour fournir accès aux informations et aux archives de toutes les institutions. Cette nouvelle opération inclut aussi un site Internet (<http://eskilstuna.se/eskilskallan>), un journal, et des événements comme des lectures sur des sujets historiques locaux ou une célébration d’un jour national des archives. Le centre d’information met à disposition

- Des livres (sur l’histoire locale)
- Cartes géographiques
- Bases de données
- Microfiches
- Etudes généalogiques
- Bureau de renseignements

¹ Lecture given at the CIDOC offsite day in Turino during the 24th general Conference of ICOM 2016.

This paper is a continuation of the paper from the CIDOC conference in Sibiu 2011 (http://network.icom.museum/fileadmin/user_upload/minisites/cidoc/ConferencePapers/2011/Susanne_Nickel_full_paper.pdf)

² For more detailed background information, cfr. paper from Sibiu.

The information centre was (and still is) open to the public 6 days a week and from October to March even on Sundays. The information desk was manned 5 days a week for four hours each day by staff from all participating institutions on an alternating scheme.

The effects we expected from this centre and the shared office and storages were:

- More and better knowledge about the other participating institutions
- Shared knowledge
- New ideas
- Shared rules and routines
- More cooperation
- New input from questions we otherwise hadn't got

For the public we hoped to give them a better service in their research. They only need to go to one point (the information centre) and get an answer from all participating institutions (instead of going to different places to meet the different institutions)

Situation five years later – lessons learned

Clear design

Soon after opening the information centre we realised that the visitors didn't understand that our services were a bit different from the rest of the library. Most of the questions we got were related to the ordinary library services and many of them we couldn't answer. To make the difference clear we reworked the design of the information centre which now better indicates that here is something new. The re-launched information centre was opened in October 2015.

Less is more

In the beginning 17 people from all participating institutions manned the information desk. But we soon saw that this was too much. Our visitors got very different information and with different quality. So we reorganized and reduced the staff to 7 people. That gave more equal service to the public and increased the quality of information our customers got. Due to economic reasons we were forced to cut down the services in the information desk. From September 2016 the opening hours for the information desk are Monday – Wednesday 10-12 am and 1-3 pm. And only staff from three participants is on duty.

Training matters

People serving in the information desk needed to be

Le centre d'information était (et est toujours) ouvert au public 6 jours par semaine du mois d'octobre au mois de mars, même les dimanches. Le bureau de renseignements était ouvert 5 jours par semaine pendant 4 heures chaque jour par du personnel des institutions qui participent à cette initiative sur un emploi du temps alterné.

Les impacts anticipés du centre, et du bureau de renseignements et entrepôt partagés étaient:

- Une plus grande et meilleure connaissance des autres organisations participants à cette initiative
- Des connaissances partagées
- De nouvelles idées
- Des règles et routines communes
- Plus de coopération
- Plus de retours de questions qui n'étaient pas demandées ultérieurement

Nous espérons fournir au public un meilleur service pour leurs recherches. IL était seulement nécessaire de s'adresser à un seul point d'information (le bureau de renseignements) pour recevoir une réponse de toutes les institutions (au lieu d'aller à différents endroits pour rencontrer les différentes institutions).

Bilan cinq années plus tard – leçons apprises

Un design clair

Peu de temps après avoir ouvert le bureau de renseignements, nous avons réalisé que les visiteurs n'avaient pas compris que nos services étaient un peu différents par rapport au reste de la bibliothèque. La plupart des questions qui nous étaient posées étaient liées aux services traditionnels de la bibliothèque et nous ne pouvions répondre à beaucoup d'entre elles. Pour rendre la différence plus claire nous avons retravaillé le design du bureau de renseignements et maintenant il indique plus clairement qu'il offre quelque chose de nouveau. Le second bureau de renseignements ouvrit en octobre 2015.

Moins est plus

Au début 17 personnes des 3 institutions travaillaient au bureau de renseignements. Mais nous avons vite réalisé que c'était trop de personnes. Nos visiteurs recevaient des

trained in search tools (databases, web services etc.) from all institutions. In the beginning we trained all staff serving in the information desk (17 people) in all relevant databases and web services, visiting the storages of the archives and of the library etc. Now we only need to train new staff or when search tools and equipment (for example the micro film reader) are changed.

Communication and understanding

Good information flow between all involved is necessary. For this one needs a workflow and a good understanding by all how the different institutions are working. This is also an important part of the training. Monthly meetings (not more than 1 hour) for organization and questions have been a good way of working. We also have a space on the internal web (intranet) where we have all the guiding documents, practical information (for example how to use a specific database), and the schema for manning the information desk. Even the non-municipal parts of our cooperation have access to this space.

Be aware of differences

One of the biggest fears in the beginning was that the participating institutions might lose their identity and their specialist knowledge. After five years of cooperation we can see that this is not the case. We have learned to appreciate the special knowledge of all participants and are really aware of the differences. This shows up especially in registration and using the different databases. Museum staff is used to one way of doing it and archivist of another and librarians use a third method. It took some time to learn that this influences even the way of searching in the different databases. Now I as a museum professional have got used to search in the archival databases (and mostly I find what I want, because I have the right expectation), but it was very difficult for me in the beginning.

Digitization

At the same time, we started the information centre we also began with a digitization project. Thanks to this we could particularly increase the amount of digitized photographic material from 6500 entries for the City museum and 4827 entries for the Municipality archives to 34 404 (City museum) and 8714 (Municipality archives). A part of these entries is available online. After a change of the database system in 2016 we will be able to publish even more entries on the web in 2017. The rules and routines for handling and digitizing the photographic material which were developed during the start of the digitalization project are now part of the daily work both at the City museum and the Municipality archives.

informations très différentes et de qualité variable. Nous avons alors réorganisé et réduit le nombre de personnes au bureau de renseignements à 7. Cela eu pour effet de donner au public un service plus similaire avec une meilleure qualité. A cause de raisons économiques nous avons été forcés de réduire le service du bureau de renseignements. Depuis septembre 2016 les heures d'ouverture du bureau de renseignements sont du lundi et mercredi de 10h00 à 12h00 et de 13h00 à 15h00. Et seulement 3 membres du personnel sont de service.

La formation est importante

Le personnel de service au bureau de renseignements a eu besoin d'être formé aux outils de recherche (bases de données, service Internet, etc.) de toutes les institutions. Pour commencer nous avons formé tout le personnel de service au bureau de renseignements (17 personnes) sur toutes les bases de données et tous les services Internet, y compris des visites des aires d'entrepôt des archives et de la bibliothèque. A présent nous n'avons seulement besoin de former les nouveaux membres du personnel ou quand les outils de recherche et l'équipement (par exemple le lecteur de microfiches) changent.

Communication et compréhension

Un bon échange d'informations entre tous les partis présents est nécessaire. Cette initiative requiert un processus et une bonne compréhension par tous du mécanisme de chaque institution. C'est aussi une part importante de la formation. Des réunions mensuelles (pas plus d'une heure) ont été une bonne approche. Les guides, informations pratiques (par exemple comment utiliser une base de données), et les horaires de service du personnel sont postés sur notre intranet. Ces documents sont accessibles même par le personnel non municipal de la collaboration.

Faire attention aux différences

Au début nous avions peur que les institutions qui y participaient ne perdent leur identité et leur savoir spécialisé. Après 5 ans de collaboration nous pouvons constater que ça n'est pas le cas. Nous avons appris à apprécier le savoir spécialisé des tous les participants et nous connaissons nos différences. Ceci est particulièrement apparent dans l'enregistrement et l'utilisation des différentes bases de données. Le personnel du musée est habitué à une façon de le faire, les archivistes sont habitués à une autre et les bibliothécaires utilisent une troisième méthode. Il a fallu un certain temps

Effects

The questions we get at the information desk have changed. At the beginning there were many questions like: What did the street I live in look like 50 or 100 years ago? Where was the school my grandparents went to? Such questions have decreased due to digitalization. Maps, address books and photos are available on the web. Questions today are more difficult and broader, so that even we from the staff need some time to help people to the sources where they may find an answer. The research shall be done by the user. Our ambition therefore is to help customers finding the answer by themselves (not giving the answer).

Factors of success

What made our cooperation successful? The following factors have been of importance:

The politicians and the heads of the institutions

The politicians were always on our side. They wanted us to cooperate. Therefore, they agreed to move the municipality archive from the jurisdiction of the city management to the culture and leisure department. The heads of the institutions were positive (even from non-municipal institutions). They saw a chance to increase the benefit for the public and allowed the staff to deal with the idea.

Staff burning for the idea and never give up

Besides of the heads of institution and the politics one also needs staff who is convinced that a cooperation is useful. People burning for the idea and not giving up even if it takes a long time to realize it are necessary. Without my colleague from the library Björn Lindwall the cooperation of archives, libraries and museums in Eskilstuna had never been realized. He was excellent in selling the idea to everyone (staff, politicians, public).

Time

Good things come to those who wait. So we took the time we needed. It was a long process from the first ideas 1999, to the realization of the shared information centre and the website in 2011. But we didn't stay there. Since that we have developed and fitted our services to user requirements.

Geography

All in one place or not even far away is important. By localizing the staff and the material together and creating one physical and one digital doorway we made it easy for people to access the archival material and use our services. Placing the information desk in the library made it easier for many to find their way to us. In Sweden libraries are the cultural institution with the

pour apprendre que cela influence même la façon de chercher dans les différentes bases de données., En tant que professionnel du musée, je suis maintenant habituée à chercher dans les bases de données archivistiques (et surtout je trouve ce que je veux parce que je sais quoi attendre), mais au début c'était très difficile pour moi.

Numérisation

En même temps que nous avons lancé le centre d'information, nous avons également commencé un projet de numérisation. Grâce à cela nous allions notamment pouvoir augmenter la quantité de matériels photographiques numérisés de 6500 à 34 404 pour le musée de la ville, et de 4827 à 8714 fichiers pour les archives de la municipalité. Une partie de ces matériels est disponible en ligne. Après un changement du système de base de données en 2016, nous pourrions publier encore plus de notices d'objets sur le web en 2017. Les règles et les processus de manipulation et de numérisation du matériel photographique qui ont été développées au début du projet de numérisation font maintenant partie de la routine quotidienne aussi bien au musée qu'aux archives.

Impacts

Les questions que nous recevons maintenant au bureau de renseignements ont changé. Au début, il y avait beaucoup de questions comme: De quoi avait l'air la rue dans laquelle je vis il y a 50 ou 100 ans? Où était l'école à laquelle mes grands-parents sont allés? Ce type de questions a diminué en raison de la numérisation des documents. Cartes, carnets d'adresses et photos sont maintenant disponibles sur le web. Les questions que l'on nous pose aujourd'hui sont plus difficiles et plus large, du fait que même le personnel du bureau a besoin de temps pour amener les gens aux sources où ils peuvent trouver une réponse. Notre modus operandi est que la recherche doit être effectuée par l'utilisateur. Notre ambition est alors d'aider nos utilisateurs à trouver la réponse par eux-mêmes et non pas de leur fournir la réponse).

Facteurs de succès

Qu'est ce qui a fait notre coopération réussie? Les facteurs suivants ont été importants:

Les politiciens et les chefs d'institutions

Les politiciens ont toujours été de notre côté. Ils voulaient que nous coopérions. Par conséquent, ils ont convenu de transférer les archives municipales de la juridiction de la ville à celle du département

lowest threshold.

Conclusion

After five years of cooperation I see that it is necessary not only to work together on a digital level but even on a personal level. We are providing our services for human beings (not computers) and there are still human beings making entries in databases and digitizing material. Therefore it is worth to learn more about the people working with it and their customers to increase the quality of the services to the public (regardless of the institutions). I want to close with a citation from my paper from 2011:

"[...] we must not forget the personal level. By meetings and cooperation between experts of different types and from different intuitions we create knowledge and we can provide better while personal and individualized service to our customers. As I see it we even in the future need both forms: the digital/virtual and real/personal to meet the needs and wishes of our users in the best way."

With the experiences made during the last five years and some adjustments in the future I am sure that the cooperation between archives, libraries and museums in Eskilstuna will continue to be and become even more appreciated by our public.

de la culture et les loisirs. Les chefs des institutions ont soutenu le projet (même ceux des institutions non-municipales). Ils ont reconnu la chance d'augmenter le bénéfice pour le public et ont permis au personnel de développer le concept.

Un personnel enthousiaste et persistant

Outre les chefs d'institution et les politiciens il faut que le personnel soit convaincu que la coopération est utile. Il est en effet nécessaire que le personnel soit enthousiaste et persistant pour le concept car le projet peut mettre beaucoup de temps à réaliser. Sans mon collègue de la bibliothèque Björn Lindwall la coopération des archives, des bibliothèques et des musées à Eskilstuna n'aurait jamais pu être réalisé. Il a été un porte-parole hors norme auprès du personnel, des politiciens, et du public.

Temps

De bonnes choses arrivent à ceux qui savent attendre. Nous avons donc pris tout le temps dont nous avons besoin. Ce fut un long processus depuis les premières idées en 1999 jusqu'à la réalisation du centre d'information partagé et du site internet en 2011. Mais nous ne nous sommes pas reposés sur nos lauriers. Depuis, nous avons développé et adapté davantage nos services aux besoins des utilisateurs.

Géographie

La proximité est importante. En localisant le personnel et le matériel ensemble et en créant une porte physique et une porte numérique, nous avons facilité l'accès aux documents d'archives et à l'utilisation de nos services. Placer le bureau d'information dans la bibliothèque a facilité pour beaucoup leur chemin vers nous. En Suède, les bibliothèques sont l'institution culturelle ayant le seuil le plus bas.

Conclusion

Après cinq années de coopération, je constate qu'il est nécessaire non seulement de travailler ensemble au niveau du numérique, mais même au niveau du personnel. Nous fournissons nos services à des êtres humains (pas à des ordinateurs) et il y a encore des êtres humains qui entrent le contenu des bases de données et qui numérisent les documents. Par conséquent, il est important d'en apprendre davantage sur les personnes qui travaillent avec ce contenu et leurs clients afin d'améliorer la qualité des services offerts au public (quelles que soient les institutions). Je veux

terminer avec une citation de mon article de 2011:

«[...] nous ne devons pas oublier l'aspect humain des choses. Les connaissances sont créées par des réunions et la coopération entre des experts de différents types et des intuitions différentes, et cela nous permet d'offrir un meilleur service à nos clients car personnalisé et individualisé. De mon point de vue, à l'avenir nous aurons besoin des deux aspects: le numérique/virtuel et réel/personnel pour répondre aux besoins et aux souhaits de nos utilisateurs de la meilleure façon.»

Vues nos expériences ces cinq dernières années et des ajustements à l'avenir, je suis sûre que la coopération entre les archives, les bibliothèques et les musées d'Eskilstuna continuera d'être et de devenir encore plus appréciée par notre public.



Eskilskällan information center. Photo: Eskilskällan



Eskilskällan display case. Photo: Eskilskällan

IMPLEMENTING LIDO

GORDON MCKENNA AND REGINE STEIN

MISE EN OEUVRE DE LIDO

GORDON MCKENNA ET REGINE STEIN



Introduction

Implementing LIDO is short guide for those wishing to export their content and associated metadata to aggregators (e.g. Europeana, Deutsche Digitale Bibliothek and Culture Grid) using the LIDO (Lightweight Information Describing Objects) XML schema. The guide was created within the framework of the AthenaPlus³ project which was funded between 2013 and 2015 by the European Commission, and was one the last outputs to be published.

The guide came out of the project's Deliverable 2.2 – *Survey and description of existing mapping models to LIDO*. Originally the deliverable was going to contain mappings between others' XML representations describing their objects, but the research showed that such XML representations did not exist.

Introduction

Implementing LIDO (*La mise en œuvre de LIDO*) est un bref guide pour ceux qui souhaitent exporter leur contenu et les métadonnées associées vers des agrégateurs (par exemple, Europeana, Deutsche Digitale Bibliothek et Culture Grid) en utilisant le schéma XML LIDO (Lightweight Information Describing Objects). Ce guide a été créé dans le cadre du projet AthenaPlus³, qui a été financé entre 2013 et 2015 par la Commission européenne et a été l'un des derniers résultats à publier.

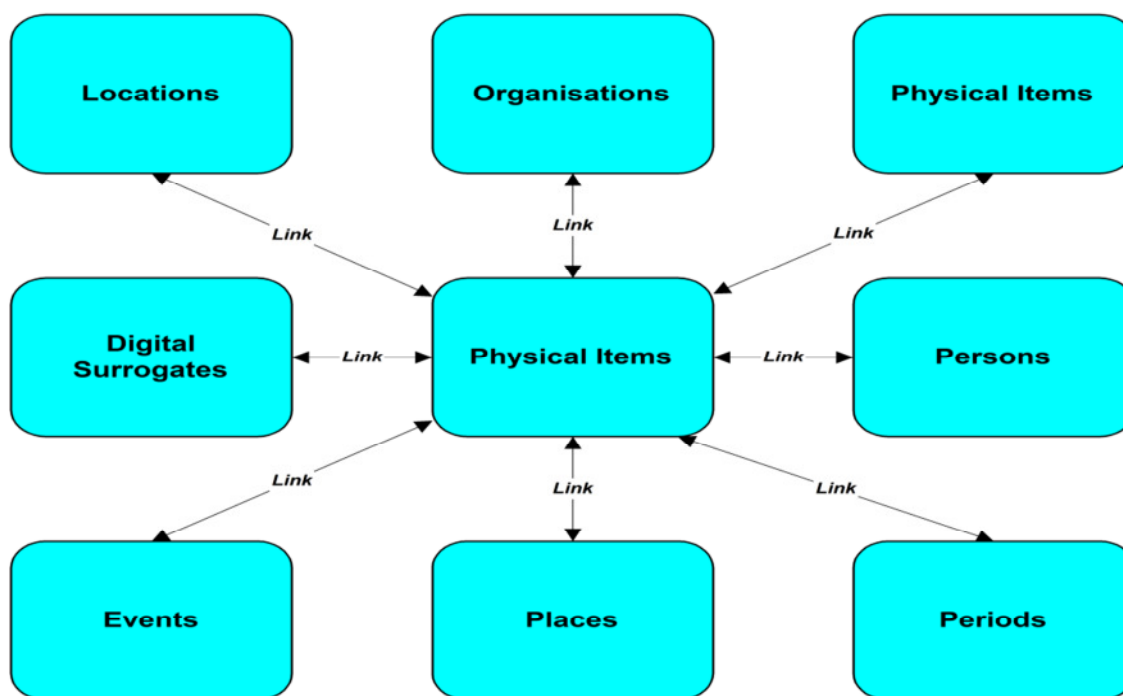
Le guide est le résultat du projet Deliverable 2.2 – *Enquête et description des modèles courants de data mapping à LIDO*. À l'origine, l'objectif allait contenir des data mappings entre d'autres représentations XML décrivant leurs objets, mais la recherche a montré que de telles représentations XML n'existaient pas.

³ See: <http://www.athenaplus.eu>

Therefore, in order to make the work useful it was decided to develop advice for people already using LIDO, and for those who have no experience of the schema but would like to consider using it.

A short history of LIDO

Cultural heritage objects do not exist in isolation. They are linked to many other entities. Some of them are shown below:



Possible relationships between different types of cultural heritage entity

If the metadata is held and published in a simple data model (e.g. similar to Dublin Core) a large proportion of the potential richness will be lost. In particular actors with roles and events with different dates were especially difficult to distinguish. The disquiet with this simple data model, articulated by partners and others, led to the development of an XML schema within the framework of the ATHENA project.

The schema, was developed, starting in 2008, from a number of existing metadata standards, and is compliant with the relevant sector ontology. It was the result of the joint collaborative efforts of the international stakeholder communities which created those standards. The purpose of the work was to create a common standard for contributing cultural heritage information and content to portals and other aggregation repositories.

The story of LIDO began, before the ATHENA project, with

Par conséquent, afin de rendre le travail utile, il a été décidé d'élaborer des conseils pour les personnes qui utilisent déjà LIDO, et pour ceux qui n'ont aucune expérience du schéma, mais voudrait envisager de l'utiliser.

Une courte histoire de LIDO

Les objets du patrimoine culturel n'existent pas en isolation. Ils sont liés à de nombreuses autres entités. Certaines d'entre elles sont indiquées ci-dessous:

Relations possibles entre les différents types d'entités du patrimoine culturel

Si les métadonnées sont conservées et publiées dans un modèle de données simple (par exemple similaire à Dublin Core), une grande partie de la richesse potentielle sera perdue. En particulier, les acteurs ayant des rôles et des événements avec des dates différentes étaient particulièrement difficiles à distinguer. L'inquiétude suscitée par ce modèle de données simple, articulé par des partenaires et d'autres, a conduit au développement d'un schéma XML dans le cadre du projet ATHENA.

Le schéma a été développé à partir de 2008 à partir d'un certain nombre de normes de métadonnées existantes et est conforme à l'ontologie pertinente du secteur d'activité. Le schéma a été le résultat des efforts conjoints de collaboration des communautés d'intervenants internationaux qui ont créé ces normes. Le but des travaux était de créer une norme commune pour la contribution de l'information et du contenu du

the J Paul Getty Trust and ARTstor who developed *CDWA Lite* schema (Categories for the Description of Art)⁴ in the USA. The scope of this was extended by the Working Group Data Exchange of the German Museum Association's development of *museumdat*⁵. This generalised the schema to include being able to describe non-art objects, e.g. social history, natural science, and technology, and to include the need to support multilingual records.

It was agreed that rather than keep *CDWA Lite* and *museumdat* separate, to create a single schema. In addition, input from the greater community of cultural heritage information and technology professionals was sought. Part of this effort was to secure compliance of the new schema with CIDOC CRM (ISO 21127)⁶ as sector standard ontology. A working group was established for the development of what became LIDO.

Resulting from the report on existing standards applied in European museums, it was concluded by partners in the ATHENA project that any metadata format for ATHENA would also have to meet the needs of SPECTRUM⁷. As a result, ATHENA decided to join the LIDO working group to further development that would integrate SPECTRUM requirements into what became the published LIDO schema.

The result of all this work: *LIDO – Lightweight Information Describing Objects Version 1.0*, was delivered to the sector during the CIDOC meeting at the ICOM triennial conference in November 2010 which took place in Shanghai, China.

Since its publication LIDO has been used in many projects, portals, and initiatives:

- European related projects (MIMO, Digitising Contemporary Art, Judaica Europeana, Linked Heritage, Partage Plus, Europeana Inside, EuropeanaPhotography, LoCloud, Europeana Food and Drink and AthenaPlus);
- Kansallinen digitaalinen kirjasto;
- Deutsche Digitale Bibliothek (DDB);
- BAM portal;
- Bildindex der Kunst und Architektur;
- eMuseum network;
- digiCULT;

⁴ See: http://getty.art.museum/research/publications/electronic_publications/cdwa/cdwalite.htm

⁵ See: <http://www.museumdat.org/index.php?ln=en>

⁶ See: <http://www.cidoc-crm.org>

⁷ See: <http://www.collectiontrust.org.uk/spectrum>

patrimoine culturel aux portails et autres entrepôts d'agrégation.

L'histoire de LIDO a commencé, avant le projet ATHENA, avec le J Paul Getty Trust et ARTstor qui ont développé le schéma *CDWA Lite* (Catégories pour la Description de l'Art)⁴ aux Etats-Unis. Le champ d'application de ce projet a été élargi par le développement de *museumdat*⁵ par le Groupe de Travail sur l'Echange de Données de l'Association Allemande des Musées. Ceci a généralisé le schéma pour pouvoir être capable de décrire des objets qui ne sont pas des objets d'art, comme les objets d'histoire sociale, des sciences naturelles, et de la technologie, et de pouvoir générer des notices d'objets multilingues.

Il a été convenu, que plutôt que de garder *CDWA Lite* et *museumdat* séparés, de créer un schéma unique. En outre, on a cherché à recueillir les commentaires de la communauté des professionnels de l'information et de la technologie du patrimoine culturel. Une partie de cet effort consistait à assurer la conformité du nouveau schéma avec CIDOC CRM (ISO 21127)⁶ en tant qu'ontologie standard du secteur. Un groupe de travail a été créé pour le développement de ce qui est devenu LIDO.

Résultant du rapport sur les normes existantes appliquées dans les musées européens, les partenaires du projet ATHENA ont conclu que tout format de métadonnées pour ATHENA devrait également répondre aux besoins de SPECTRUM⁷. En conséquence ATHENA a décidé de rejoindre le groupe de travail LIDO pour un développement qui permettrait d'intégrer les exigences SPECTRUM dans ce qui est devenu le schéma LIDO publié.

Le résultat de tout ce travail: *LIDO – Informations Légères Décivant les Objets Version 1.0*, a été délivré au secteur lors de la réunion du CIDOC lors de la conférence triennale de l'ICOM en novembre 2010 qui a eu lieu à Shanghai en Chine.

Depuis sa publication, le LIDO a été utilisé dans de nombreux projets, portails et initiatives:

- Projets liés à Europeana (MIMO, Digitising Contemporary Art, Judaica Europeana, Linked Heritage, Partage Plus, Europeana Inside, EuropeanaPhotography, LoCloud, Europeana Food and Drink and AthenaPlus);
- Kansallinen digitaalinen kirjasto;
- Deutsche Digitale Bibliothek (DDB);
- Portail BAM;
- Bildindex der Kunst und Architektur;

- Yale Center for British Art;
- WissKI;
- Deutsche Forschungsgemeinschaft;
- Museum Data Exchange Project;
- Culture Grid.

- Réseau eMuseum;
- digiCULT;
- Yale Center for British Art;
- WissKI;
- Deutsche Forschungsgemeinschaft;
- Projet d'échange de données sur les musées;
- Culture Grid.

There is also a CIDOC Working group on LIDO⁸.

Despite all this activity support for the implementation of LIDO was largely within the framework of the projects and initiatives, e.g. workshops for project members, therefore it was decided that resource was needed to help those who were not able to benefit from this. The guide is that resource.

Scoping the guide

As part of the scoping the authors carried out research based c1600 questions on the implementation of LIDO that were asked of project helpdesks over a period of five years. Here a tool (MINT⁹) developed by the National Technical University of Athens (NTUA) was used to map in-house XML to LIDO, and then transform it to Europeana's EDM (Europeana Data Model). About 45% of the questions asked were about the MINT itself and were not covered in the guide.

About 5% of the questions related to general XML issues covering: Entity references (e.g. &); Not following schema rules; Illegal characters in URLs (particularly '&'); Use of Unicode; and URLs not working. All of these issues were dealt with by giving the implementers additional general information about how XML worked. The rest of the questions asked were on LIDO itself. Topics, in order of number of questions were: LIDO structure; Europeana; Multilingual record; Links; Rights; Actors; Dates; Events; Holding institution; IDs; Terminologies; Multiple elements; Mandatory elements; Inscriptions; and Subject. The conclusion from these was that it would be useful to 'pre-empt' these issues by covering these topics in the guide. However it is realised that there is still a role for a helpdesk supported by real experts.

Il existe également un groupe de travail du CIDOC sur le LIDO⁸.

En dépit de toutes ces activités, le soutien à la mise en œuvre du LIDO a été en grande partie dans le cadre des projets et initiatives, p. Ex. Ateliers pour les membres du projet, il a donc été décidé que des ressources étaient nécessaires pour aider ceux qui ne pouvaient en bénéficier. Le guide est cette ressource.

Définition du guide

Dans le cadre de l'examen de la portée, les auteurs ont étudié environ 1600 questions sur la mise en œuvre de LIDO qui ont été demandés aux bureaux d'aide du projet sur une période de cinq ans. Ici, un outil (MINT⁹) développé par l'Université Technique Nationale d'Athènes (NTUA) a été utilisé pour le data mapping d'XML interne à LIDO, puis le transformer ensuite en EDM d'Europeana EDM (Europeana Data Model). Environ 45% des questions posées portaient sur le MINT lui-même et n'étaient pas couvertes dans le guide.

Environ 5% des questions relatives aux questions XML générales couvrent: les références d'entité; le manquement à suivre les règles de schéma; les caractères illégaux dans les URL (en particulier '&'); l'utilisation d'Unicode; et les URL qui ne fonctionnent pas. Toutes ces questions ont été traitées en donnant aux implémentateurs des informations générales supplémentaires sur le fonctionnement de XML. Le reste des questions posées concernait LIDO lui-même. Les sujets, par ordre de nombre de questions, étaient

les suivants: Structure de LIDO; Europeana; notice multilingue; Liens; Droits; Acteurs; Dates; Événements; Organisme propriétaires; IDs; Terminologie; Éléments multiples; Éléments obligatoires; Inscriptions; Et Sujet. La conclusion de cette étude fut qu'il serait utile d'anticiper ces questions en couvrant ces sujets dans le guide. Cependant, on s'aperçoit qu'il existe toujours un rôle pour un bureau de renseignements soutenu par de vrais experts.

⁸ See: <http://network.icom.museum/cidoc/working-groups/lido>

⁹ Metadata INTeroperability. See: <http://mint.image.ece.ntua.gr/redmine/projects/mint/wiki/Wiki>

What the guide contains

The guide contains the following sections:

XML basics and LIDO

The guide covers:

- Tag-names;
- Attribute-names;
- Namespaces.

It advises the organisation implementing LIDO, and particularly those persons carrying out the implementation should familiarise themselves about the basics of XML, and how LIDO handles these areas.

Overview of LIDO

This section covers the basics:

- What LIDO is;
- What LIDO is not;
- LIDO's design principles;
- What's in LIDO?

This gives a good grounding in LIDO and can be used to make a case to senior management for its use in an organisation.

Before starting to implement LIDO

This section is the key part of the guide, and suggests the implementer considers:

Setting your own house in order by:

- Managing their own systems to deliver metadata to the quality needed to meet their needs to provide access. One way to guarantee this, if properly implemented, is to base the system used on accepted standards, e.g. for museums *SPECTRUM* or *CDWA*.

In-house developed systems should be able to map their metadata elements to the metadata elements of these standards. If their system is not as comprehensive as one of the recommended standards it might not be possible to submit fully rich metadata into LIDO.

- Correctly using their collections management system looking at; Data entry errors (e.g. spelling mistakes); putting the wrong type of data in a field (e.g. confusing dates with periods); and inconsistency in the use of terminology. The organisation must have in place policies and procedures to minimise these issues.

Contenu du guide

Le guide contient les sections suivantes:

Bases XML et LIDO

Le guide couvre:

- Noms de tags/étiquettes;
- Noms d'attributs;
- Espaces de noms.

Il conseille l'organisation chargée de la mise en œuvre de LIDO et, en particulier, les personnes chargées de la mise en œuvre de se familiariser avec les bases du XML et comment LIDO gère ces sujets.

Vue d'ensemble de LIDO

Cette section couvre les notions de base:

- Qu'est-ce LIDO est;
- Qu'est-ce que LIDO n'est pas;
- Les principes de conception de LIDO;
- Qu'y a-t-il dans LIDO?

Cela donne une bonne base de connaissance de LIDO et peut être utilisé pour convaincre l'équipe de direction d'adopter LIDO dans une organisation.

Avant de commencer à mettre en œuvre LIDO

Cette section est la partie clé du guide, et suggère que l'implémentateur considère:

Préparer le terrain:

- Gérer leurs propres systèmes pour fournir des métadonnées à la qualité nécessaire pour répondre à leurs besoins d'accès. Une façon de garantir cela, si elle est correctement mise en œuvre, est de baser le système utilisé sur des normes acceptées, c'est-à-dire pour les musées respecter *SPECTRUM* ou *CDWA*.

Les systèmes développés à l'interne devraient pouvoir mapper leurs éléments de métadonnées aux éléments de métadonnées de ces normes. Si le système n'est pas aussi complet que l'une des normes recommandées, il pourrait ne pas être possible de générer des métadonnées LIDO riches.

- Utiliser correctement leur système de gestion des collections en payant attention aux erreurs de saisie de données (par exemple, erreurs d'orthographe); à

Looking at the interaction between LIDO and the target portal examining how various features are handled by the portal. This analysis might lead to the adjusting the LIDO exported to give the best results of display in the portal. Areas to consider include: multilingual metadata; display elements and their corresponding indexing elements, and what happens if either is not present; and the ordering of repeatable elements.

General LIDO issues

The main advice is that the organisation should have knowledge of which elements in a LIDO record hold data, and what kind of data each of these elements should hold. It should also know its own collections management system well enough to be able to map to these elements. To help in this the guide covers:

- Elements containing data;
- Language(s) of the record and elements;
- Display and indexing elements;
- Repeatable elements;
- Ordering of repeatable elements;
- Preference of repeatable elements;
- Events.

The section on 'Events' should be noted as a multilingual terminology has been developed [*non-preferred term*]:

- | | |
|--|-------------------|
| • Acquisition; | • Loss; |
| • Collecting
[<i>Collection Event, Field
Collection, Collection</i>]; | • Modification; |
| • Commissioning
[<i>Order</i>]; | • Move; |
| • Creation
[<i>Conception, Create</i>]; | • Part addition; |
| • Designing
[<i>Design</i>]; | • Part Removal; |
| • Destruction; | • Performance; |
| • Event; | • Planning; |
| • Excavation; | • Production; |
| • Exhibition; | • Provenance; |
| • Finding
[<i>Find</i>]; | • Publication; |
| | • Restoration; |
| | • Transformation; |
| | • Type assignment |
| | • Type creation; |
| | • Use. |

Rules and guidance are given for each area. For example, if the organisation wishes to have more than one language for their metadata in LIDO records, they:

- Should be aware of how multilingual records will be supported by the target portal;
- Must be able to record multilingual metadata in their collections management system (in some way, e.g. by different field names or by

l'utilisation du mauvais type de données dans un champ (par exemple, confondre les dates avec les périodes); l'utilisation incohérente de la terminologie. L'organisation doit mettre en place des politiques et des procédures pour minimiser ces problèmes.

En examinant l'interaction entre LIDO et le portail cible, nous examinons comment les différentes fonctionnalités sont gérées par le portail. Cette analyse pourrait conduire à ajuster le LIDO exporté pour donner les meilleurs résultats d'affichage dans le portail. Les domaines à considérer comprennent: les métadonnées multilingues; Les éléments d'affichage et leurs éléments d'indexation correspondants, et ce qui se produit si l'un ou l'autre n'est pas présent; Et l'ordre des éléments répétables.

Questions générales de LIDO

L'avis principal est que l'organisation doit avoir connaissance des éléments qui se trouvent dans un fichier LIDO et des données que chacun de ces éléments devrait contenir. Il devrait également connaître son propre système de gestion des collections assez bien pour être en mesure de mapper à ces éléments. Pour vous aider, le guide couvre:

- Éléments contenant des données;
- Langue (s) du dossier et éléments;
- Éléments d'affichage et d'indexation;
- éléments répétables;
- Commande d'éléments répétables;
- Préférence des éléments répétables;
- Événements.

La section sur les «événements» devrait être notée comme une terminologie multilingue [terme non préféré]:

- | | |
|-----------------|--------------------------|
| • Accession; | • Découverte |
| • Collecte | [Découvrir]; |
| • [Événement de | • Perte; |
| collecte, | • Modification; |
| Champ de | • Déménagement; |
| collecte, | • Addition partielle; |
| collecte]; | • Suppression partielle; |
| • Commission | • Performance; |
| • [Commande]; | • Planification; |
| • Création | • Production; |
| • [Conception, | • Provenance; |
| Créer]; | • Publication; |
| • Conception | • Restauration; |
| • [Design]; | • Transformation; |

parameter for a field), in order to facilitate the creation of the LIDO record.

The next sections are most of the guide, and contain guidance for every part of LIDO schema. The advice allows a mapping from a SPECTRUM *Unit of information*, which might be being used in an organisation's collections management system (CMS), gives the LIDO elements path, and some explanatory notes:

- Destruction;
- Événement;
- Excavation;
- Exposition;
- Attribution de types;
- Création de types;
- Utilisation.

Des règles et des directives sont données pour chaque domaine. Par exemple, si l'organisation souhaite avoir plus d'une langue pour ses métadonnées dans les notices LIDO, ils:

- devrait être conscient de la façon dont les notices multilingues seront prises en charge par le portail cible;
- Doit être en mesure d'enregistrer des métadonnées multilingues dans leur système de gestion des collections (d'une façon ou d'une autre, par exemple par des noms de champ différents ou par un paramètre pour un champ), afin de faciliter la création de la notice LIDO.

Les sections suivantes constituent la part principale du guide, et contiennent des conseils pour chaque partie du schéma LIDO. Le conseil permet un mapping à partir d'une unité d'information SPECTRUM – qui pourrait être utilisée dans le système de gestion de collections (CMS) d'une organisation – indique le chemin d'accès des éléments LIDO et quelques notes explicatives:

SPECTRUM Unit	Notes	LIDO elements path
		<code><lido:descriptiveMetadata xml:lang="[code]"></code> <code><lido:objectClassificationWrap></code> <code><lido:objectWorkTypeWrap></code>
Object name	Object type. Optionally there can be <code><lido:conceptID></code> element containing a published URI for term.	<code><lido:objectWorkType></code> <code><lido:conceptID lido:type="URI">[Published URI for term]</code> <code></lido:conceptID></code> <code><lido:term>[Item type]</lido:term></code> <code></lido:objectWorkType></code>
	Without published URI for term.	<code><lido:objectWorkType></code> <code><lido:term>[Item type]</lido:term></code> <code></lido:objectWorkType></code>

The guide advice for object names (LIDO – objectWorkType)

In all cases the organisation must be aware where in their CMS the appropriate metadata is kept in order to successfully harvest it for a LIDO record. Also how the metadata is displayed in the target portal, and any requirements for metadata of the portal must be taken into account.

Les conseils du guide pour les noms d'objets (LIDO – objectWorkType)

Dans tous les cas, l'organisation doit savoir où dans son CMS les métadonnées appropriées sont conservées afin de réussir la récolte pour une notice LIDO. De même, la manière dont les métadonnées sont affichées dans le portail cible et toutes les exigences relatives aux métadonnées du portail doivent être prises en compte.

The areas of LIDO that the guide covers are:

Object classification area

Object work type – It is **mandatory** to provide an object/work type in a LIDO record. It is **recommended** that the term comes from a controlled vocabulary, perhaps published with a URI identifier.

Classification – An object can be classified in many ways (e.g. style, form, age, and phase), all of which can be represented in a LIDO record. The guide **recommends** that the term comes from a controlled vocabulary, perhaps published with a URI identifier. The guide also suggests classifications for various types of object.

Object identification area

Title – However not all types of object have a title, therefore in those cases it will be necessary to insert `objectWorkType` element.

Inscriptions – An object can have many different inscriptions and marks

Repository – Information about the organisation holding the object is unlikely to be held in its collections management system. The guide **recommends** that at least *Repository name* appears in the LIDO.

Display and edition – Prints and other types of multiple can have **optional** *Display and edition* information.

Description – An object can have many different types of description. These might concern different aspects of the object or be aimed at different audiences. The guide **recommends** that at least a brief description should be in the LIDO record.

Measurements – An object and its parts, where relevant, can have many measurements and non-measured technical attributes. The exact types of them are dependent on the perceived collections management and research needs of the object, and can also vary with different types of object (some examples are given). The guide **recommends** that measurements and/or technical attributes (where present), should be in the LIDO record.

Event area

The exact types of event associated with an object are dependent on the object itself and its historical context. Some information may be partial or entirely missing. An object, and its parts, where relevant, can be associated with many events, during:

Les zones de LIDO que le guide couvre sont:

Zone de classification des objets

Type d'objet/d'oeuvre – Il est **obligatoire** de fournir un type d'objet/d'oeuvre dans une notice LIDO. Il est recommandé que le terme provienne d'un vocabulaire contrôlé, peut-être publié avec un identifiant URI.

Classification – Un objet peut être classé de plusieurs manières (par exemple, style, forme, âge, et phase), qui peuvent toutes être représentées dans une notice LIDO. Le guide **recommande** que le terme provienne d'un vocabulaire contrôlé, peut-être publié avec un identifiant URI. Le guide propose également des classifications pour différents types d'objets.

Zone d'identification d'objet

Titre – Cependant, tous les types d'objets ne possèdent pas de titre; dans ce cas, il est nécessaire d'insérer l'élément `objectWorkType`.

Inscriptions – Un objet peut avoir beaucoup d'inscriptions et de marques différentes.

Dépositaire – Il est peu probable que des informations sur l'organisation qui détient l'objet soient conservées dans son système de gestion des collections. Le guide **recommande** qu'au moins le nom du Repository apparaisse dans le LIDO.

Affichage et édition – Les gravures et autres types de documents multiples peuvent comporter des informations d'affichage et d'édition en **option**.

Description – Un objet peut avoir plusieurs types de description. Celles-ci peuvent concerner différents aspects de l'objet ou s'adresser à des publics différents. Le guide **recommande** qu'au moins une brève description soit inscrite dans la notice LIDO.

Mensurations – Un objet et ses parties, le cas échéant, peuvent avoir de nombreuses mensurations et des attributs techniques non mesurés. Leurs types exacts dépendent de la gestion perçue des collections et des besoins de recherche de l'objet, et peuvent également varier selon les différents types d'objets (quelques exemples sont donnés). Le guide **recommande** que les mensurations et/ou les attributs techniques (le cas échéant) soient inscrits dans le registre LIDO.

Espace événementiel

Les types exacts d'événements associés à un objet dépendent de l'objet lui-même et de son contexte historique. Certaines informations peuvent être

- Its bringing into being;
- After that but before it became part of an organisation's collections;
- After it became part of an organisation's collections.

The exact types of event associated with an object are dependent on the object itself and its historical context. Some information may be partial or entirely missing. The guide **recommends** that at least, where information is present and relevant, to give that information in the LIDO record in the following events:

- **Production** (all);
- **Designing** (designed objects);
- **Publication** (published objects);
- **Collecting** (those collected in some way);
- **Excavation** (those from archaeological excavations);
- **Exhibition** (those that have taken part in an exhibition).

Other events can be included if the organisation thinks they are relevant.

Object relation area

Subject – For visual and textual works and objects it is possible to describe the *Subjects* of their content in this area of a LIDO record. The exact types of *Subjects* content are dependent on the object itself. The guide **recommends** that where the information is present and relevant, to give that information in the LIDO record. It can include information about or depicting:

- **Activities;**
- **Concepts;**
- **Dates;**
- **Periods;**
- **Events;**
- **Objects;**
- **Organisation;**
- **People;**
- **Persons;**
- **Places.**

In addition, if the organisation thinks they are relevant, a general description of content and where on the object the content is located can be included.

Related items – It is **optional** to relate objects to the object being described in the LIDO record. However the guide **recommends** that where the information is present and relevant, to give that information in the LIDO record. It should include information the:

partielles ou manquantes. Un objet et ses parties, le cas échéant, peuvent être associés à de nombreux événements, pendant:

- Sa mise en œuvre;
- Après cela, mais avant de faire partie des collections d'une organisation;
- Après avoir fait partie des collections d'une organisation.

Les types exacts d'événements associés à un objet dépendent de l'objet lui-même et de son contexte historique. Certaines informations peuvent être partielles ou manquantes. Lorsque les informations sont présentes et pertinentes, le guide **recommande** d'au moins les inclure dans la notice LIDO dans les cas suivants:

- **Production** (tous);
- **Conception** (objets désignés);
- **Publication** (objets publiés);
- **Collecte** (ceux collectes d'une certaine façon);
- **Excavation** (provenant de fouilles archéologiques);
- **Exposition** (ceux qui ont participé à une exposition).

D'autres événements peuvent être inclus si l'organisation pense qu'ils sont pertinents.

Zone de relation objet

Sujet – Pour les visuels et textes ainsi que les objets, il est possible de décrire les sujets de leur contenu dans cette zone d'une notice LIDO. Les types exacts de contenu des sujets dépendent de l'objet lui-même. Le guide **recommande**, lorsque l'information est présente et pertinente, que ces renseignements figurent dans la notice LIDO. Cela peut inclure des informations sur ou représentant des:

- **Activités;**
- **Concepts;**
- **Dates;**
- **Périodes;**
- **Événements;**
- **Objets;**
- **Organisation;**
- **Gens;**
- **Personnes;**
- **Des endroits.**

En outre, si l'organisation pense qu'ils sont pertinents, une description générale du contenu et où sur l'objet le

- Nature of the relationship;
- ID of the related work;
- Note describing the relationship.

Rights for work area

The guide advises that it is very important that an organisation gives **all the relevant information about the rights** affecting the use of the physical object in order to benefit the organisation, and to respect the rights of others. Also not confuse the rights affecting the use of the physical object with those of its surrogates and the metadata record describing the object. Metadata to harvest is on:

- **Rights type** (it is suggested a term for a published list);
- **Rights date;**
- **Rights holder;**
- **Credit line.**

Record area

Record ID – The guide advises that it is **mandatory** to provide an identifier for the metadata record unique in the organisation's collections management system in the LIDO record.

Record type – The guide advises that it is **mandatory** to provide a type for the metadata record being described in the LIDO record.

Record source – The guide **recommends** that the organisation should to provide some information about source, usually the organisation itself, for the metadata record being described in the LIDO record. This information must include one, or more, of the following:

- The name of the source;
- An identifier for the source;
- Link to the website of the source.

It is possible that it is not held in the CMS. In this case the organisation must have a method for getting the information into the LIDO record.

Record rights – The guide says that it is very important that an organisation gives **all the relevant information about the rights** affecting the use of the metadata record describing the object in order to benefit the organisation, and to respect the rights of others. It is also important not to confuse the rights affecting the use of the metadata record describing the object with those of the physical object itself and its surrogates.

contenu est situé peut-être inclus.

Objects associés – Relier des objets à l'objet décrit dans la notice LIDO est **facultatif**. Toutefois, le guide **recommande**, lorsque l'information est présente et pertinente, qu'elle figure dans la notice LIDO. Cela doit inclure les informations suivantes:

- nature de la relation;
- ID de l'objet associé;
- Note décrivant la relation.

Section sur les droits affectant l'objet

Le guide indique qu'il est très important qu'une organisation donne **toutes les informations pertinentes sur les droits** affectant l'utilisation de l'objet physique afin de bénéficier à l'organisation et de respecter les droits d'autrui. Ne confondez pas non plus les droits affectant l'utilisation de l'objet physique avec ceux de ses substituts et la notice de métadonnées décrivant l'objet. Les métadonnées à récolter sont:

- **Type de droits** (il est suggéré d'utiliser un terme pour une liste publiée);
- **Date des droits;**
- **Titulaire de droits;**
- **Ligne de crédit.**

Section sur la notice

ID pour la notice – Le guide indique qu'il est **obligatoire** de fournir l'identificateur unique de la notice de métadonnées du système de gestion des collections de l'organisation dans la notice LIDO.

Type de notice – Le guide indique qu'il est **obligatoire** de fournir un type pour la notice des métadonnées décrit dans la notice LIDO.

Source de la notice – Le guide **recommande** que l'organisation fournisse des informations sur la source, généralement l'organisation elle-même, pour la notice des métadonnées décrit dans la notice LIDO. Ces renseignements doivent inclure un ou plusieurs des éléments suivants:

- le nom de la source;
- Un identifiant pour la source;
- Le lien vers le site Web de la source.

Il est possible que la source de la notice ne soit pas documenté dans le CMS. Dans ce cas, l'organisation doit avoir une méthode pour obtenir les informations dans la notice LIDO.

Record metadata information – The guide **recommends** that the organisation provides a URL link to a metadata record for a physical object in a LIDO record.

Resource area

Resource ID – The guide advises that is **optional** to provide an identifier for a surrogate for a physical object in a LIDO record.

Resource representation – The guide **recommends** that the organisation provides URL links to surrogates for a physical object in a LIDO record.

Resource type – The guide advises that is **optional** that the organisation provides type information for a surrogate for a physical object in a LIDO record.

Resource relationship type – The guide advises that is **optional** that the organisation provides the relationship between a surrogate and the physical object it relates to in a LIDO record.

Resource perspective – The guide advises that is **optional** that the organisation provides the perspective shown in a surrogate for the physical object in a LIDO record.

Resource description – The guide advises that is **optional** that the organisation provides the description of a surrogate for the physical object in a LIDO record.

Resource date taken – The guide advises that is **optional** that the organisation provides information about the date of creation of surrogate for the physical object in the LIDO record.

Resource source – The guide advises that is **optional** that the organisation provides information about the source of a surrogate for the physical object in the LIDO record. It is possibly the organisation itself, but might be an outside person or organisation. This information must include one, or more, of the following:

- The name of the source;
- An identifier for the source;
- Link to the website of the source.

Resource rights – The guide says that it is very important that an organisation gives **all the relevant information about the rights** affecting the use of the surrogate for the physical object in order to benefit the organisation, and to respect the rights of others. Also it notes that the organisation should not confuse the rights affecting the use of the surrogate for the object with those of the physical object and metadata record describing it.

Droits de la notice – Le guide dit qu'il est très important qu'une organisation donne **toutes les informations pertinentes sur les droits** affectant l'utilisation de la notice de métadonnées décrivant l'objet afin de bénéficier à l'organisation et de respecter les droits d'autrui. Il est également important de ne pas confondre les droits affectant l'utilisation de la notice des métadonnées décrivant l'objet avec ceux de l'objet physique lui-même et de ses substituts.

Notice d'information des métadonnées – Le guide **recommande** que l'organisation fournisse un lien URL à la notice des métadonnées pour un objet physique dans une notice LIDO.

Section sur les ressources

ID de ressource – Le guide indique qu'il est **facultatif** de fournir un identifiant pour un substitut pour un objet physique dans une notice LIDO.

Représentation des ressources – Le guide **recommande** que l'organisation fournisse des liens URL aux substituts d'un objet physique dans une notice LIDO.

Type de ressource – Le guide indique qu'il est **facultatif** que l'organisation fournisse des informations sur le type de ressource pour un substitut d'un objet physique dans une notice LIDO.

Type de relation de la ressource – Le guide indique qu'il est **facultatif** que l'organisation fournisse la relation entre un substitut et l'objet physique décrit dans la notice LIDO.

Perspective des ressources – Le guide indique qu'il est **facultatif** que l'organisation fournisse la perspective montrée dans un substitut pour l'objet physique dans une notice LIDO.

Description de la ressource – Le guide indique qu'il est **facultatif** que l'organisation fournisse la description d'un substitut pour l'objet physique dans une notice LIDO.

Date de capture de la ressource – Le guide indique qu'il est **facultatif** que l'organisation fournisse des informations sur la date de création du substitut pour l'objet physique dans la notice LIDO.

Source de la ressource – Le guide indique qu'il est **facultatif** que l'organisation fournisse des informations sur la source d'un substitut pour l'objet physique dans la notice LIDO. La source est peut-être l'organisation elle-même, mais peut-être une personne ou une organisation extérieure. Ces renseignements doivent

Information needed by Europeana

The last area that the guide considers are aspects of LIDO that are relevant to organisations contributing to Europeana. The following information (EDM elements) are mandatory:

- europeana:dataProvider;
- europeana:isShownAt.

AND/OR

- europeana:isShownBy;
- europeana:object;
- europeana:provider;
- europeana:rights;
- europeana:type.

The guide shows how this information can be represented in LIDO.

Where can I find out more?

The guide can be downloaded at:
<http://www.athenaplusteu/getFile.php?id=559>.

For full information on LIDO and the CIDOC working group visit:

<http://network.icom.museum/cidoc/working-groups/lido/what-is-lido/>

inclure un ou plusieurs des éléments suivants:

- le nom de la source;
- Un identifiant pour la source;
- Lien vers le site Web de la source.

Droits des ressources – Le guide dit qu'il est très important qu'une organisation **donne toutes les informations pertinentes sur les droits** affectant l'utilisation du substitut de l'objet physique afin de bénéficier à l'organisation et de respecter les droits d'autrui. Il note également que l'organisation ne doit pas confondre les droits affectant l'utilisation du substitut de l'objet avec ceux de l'objet physique et des métadonnées qui le décrivent.

Information nécessaire pour Europeana

La dernière section que le guide considère sont des aspects de LIDO qui sont pertinents pour les organisations qui contribuent à Europeana. Les informations suivantes (éléments EDM) sont obligatoires:

- europeana:data Provider;
- europeana:isShownAt.

ET/OU

- europeana:isShownBy;
- europeana:objet;
- europeana:fournisseur;
- europeana:droits;
- europeana:type.

Le guide montre comment cette information peut être représentée dans LIDO.

Où puis-je en savoir plus?

Le guide peut être téléchargé à l'adresse suivante:
<http://www.athenaplusteu/getFile.php?id=559>.

Pour de plus amples informations sur LIDO et le groupe de travail du CIDOC:

[Http://network.icom.museum/cidoc/working-groups/lido/what-is-lido/](http://network.icom.museum/cidoc/working-groups/lido/what-is-lido/)

USING CIDOC CRM AND LINKED OPEN DATA TO MODERNIZE ARTEFACTS CANADA

HEATHER DUNN

UTILISATION DE CIDOC CRM ET DE DONNEES OUVERTES LIEES POUR MODERNISER ARTEFACTS CANADA

HEATHER DUNN



150 years of Canadian Art A linked open data project



Abstract

The Canadian Heritage Information Network (CHIN)¹⁰ used the CIDOC-CRM (Conceptual Reference Model) in a pilot project to demonstrate the feasibility of providing enhanced access to digital cultural and heritage content through a Linked Open Data (LOD) model. Based on the success of the pilot, CHIN will be using CIDOC CRM (among other ontologies) and LOD to modernize Canada's national repository of museum collections data, Artefacts Canada¹¹.

Artefacts Canada – History and Today

Artefacts Canada is Canada's only national repository of museum collections records and images. It began as a centralized "National Inventory" in 1972, and Canadian museums used this central database for their local collections management from the early 1970s until the mid-1990s. A set of core fields was shared, first only within the Network of museums, but later with the public on the Artefacts Canada web application. By the mid-1990s, many museums had acquired their own collections management systems and began to manage their data locally, but a core set of their data fields was retained centrally at CHIN for Artefacts Canada. Museums today send batch updates of their data (delimited text), mapped to the Artefacts Canada fields. Artefacts Canada Humanities now contains 4.3 million object records and 1.1 million images, contributed by over 500 museums and galleries across Canada.

Synopsis

Le Réseau canadien d'information sur le patrimoine (RCIP)¹⁰ a utilisé le CIDOC-CRM (modèle conceptuel de référence) dans le cadre d'un projet pilote visant à démontrer la faisabilité d'un meilleur accès au contenu culturel et patrimonial numérique grâce à un modèle de données ouvertes liées. En s'appuyant sur le succès du projet pilote, le RCIP utilisera CIDOC CRM (entre autres ontologies) et LOD pour moderniser le dépôt national de données sur les collections de musées, Artefacts Canada¹¹.

Artefacts Canada – Histoire et aujourd'hui

Artefacts Canada est le seul magasin national d'archives et d'images de collections de musées. Il a commencé comme un «inventaire national» centralisé en 1972, et les musées canadiens ont utilisé cette base de données centrale pour leur gestion locale des collections du début des années 1970 au milieu des années 1990. Un ensemble de champs principaux était partagé, d'abord seulement au sein du Réseau des musées, mais plus tard avec le public sur l'application Web d'Artefacts Canada. Au milieu des années 1990, de nombreux musées ont acquis leurs propres systèmes de gestion des collections et ont commencé à gérer leurs données localement, mais un ensemble essentiel de leurs champs de données a été retenu centralement au RCIP pour Artefacts Canada. Aujourd'hui les musées envoient des mises à jour par lots de leurs données (texte délimité), mis en concordance avec les champs d'Artefacts Canada. Artefacts Canada Humanities contient maintenant 4,3 millions de notices d'objets et 1,1 million d'images, auxquels ont contribué plus de 500 musées et galeries partout au Canada.

¹⁰ Government of Canada, Canadian Heritage, Canadian Heritage Information Network <http://canada.pch.gc.ca/eng/1454520330387>

¹¹ Government of Canada, Canadian Heritage, Canadian Heritage Information Network <http://www.rcip-chin.gc.ca/bd-dl/artefacts-eng.jsp>

The Canadian heritage community considers Artefacts Canada to be CHIN's most valuable resource. It is a showcase for local, provincial, national collections, an online presence for museums that otherwise have none, and it is the only way of searching across collections. But there are serious problems with Artefacts Canada as it exists today. CHIN does not have the capacity to perform much quality control or validation (to either the data or the mapping), and there are problems with the data model, architecture, and data provisioning process. The Artefacts Canada interface is outdated, and the functionality is low. There is little incentive for museums to contribute data.

The Ideal Solution: Linked Open Data

Ideally, each Canadian museum would be responsible for updating, maintaining, and publishing its own LOD about its collections. Each museum would be free to use its preferred metadata standard (ontologies and/or element sets) and vocabulary standards. CHIN would become a point of entry in a network of interconnected collections of cultural heritage data, assisting museums in data mapping, converting their data to LOD, reconciling, normalizing, linking, and enriching data, and providing support through training and workshops. For those museums not capable of hosting and/or providing APIs to their own LOD (most museums), CHIN would also act as host. Canadian museum data would be linkable and open to the world. This is the ideal, but there are many obstacles – as we discovered in our pilot project.

The “150 Years of Canadian Art” Pilot Project using CIDOC-CRM

Because CHIN and the Canadian museum community had no experience with Linked Open Data concepts or technologies, and Artefacts Canada is large and contains data from a wide range of museum disciplines, we started with a pilot project using a small sub-set of the Artefacts Canada records.

The pilot project, “150 Years of Canadian Art” used approximately 86,000 records of artworks (a relatively standardized dataset from a single discipline) from eight art museums that already had records and images in Artefacts Canada. We defined a resource description model using CIDOC CRM as the core (to represent the works of art), because CRM is the most comprehensive and expressive reference model for museum documentation, and the only ontology specifically developed for expressing museum data in RDF. We also used other ontologies to handle data for artists (FOAF, xfoaf, DC, and BIO), and for organizations and artist groups (W3C VCard and ORG). For our Artefacts Canada Taxonomy (a faceted vocabulary tool built for facilitating

La collectivité canadienne du patrimoine considère Artefacts Canada comme la ressource la plus précieuse du RCIP. Il s'agit d'une vitrine pour les collections locales, provinciales, nationales, une présence en ligne pour les musées qui n'ont pas de site Internet, et c'est la seule façon de chercher à travers toutes les collections. Mais Artefacts Canada connaît de sérieux problèmes tels qu'il existe aujourd'hui. Le RCIP n'a pas la capacité d'effectuer beaucoup de contrôle de qualité ou de validation (pour les données ou le mappage), et il y a des problèmes avec le modèle des données, l'architecture et le processus d'approvisionnement de données. L'interface d'Artefacts Canada est dépassée et la fonctionnalité est faible. Les musées ont peu d'incitation à y fournir des données.

La solution idéale: Données ouvertes liées

Idéalement, chaque musée canadien serait chargé de mettre à jour, d'entretenir et de publier ses propres données ouvertes et liées sur ses collections. Chaque musée serait libre d'utiliser son standard de métadonnées préféré (ontologies et / ou ensembles d'éléments) et ses vocabulaires. Le RCIP deviendrait un point d'entrée dans un réseau de collections interconnectées de données sur le patrimoine culturel, aiderait les musées à mapper leurs données, convertirait leurs données en données ouvertes et liées, concilierait, normaliserait, lierait et enrichirait les données et fournirait un appui par le biais de formations et d'ateliers. Pour les musées qui ne sont pas capables d'accueillir et / ou de fournir des API à leur propre données ouvertes et liées (la plupart des musées), le RCIP jouera également le rôle d'hôte. Cela rendrait les données des musées canadiens accessibles par le monde entier. C'est le scénario l'idéal, mais il y a beaucoup d'obstacles – comme nous l'avons découvert dans notre projet pilote.

Le projet pilote «150 ans d'art canadien» utilisant CIDOC-CRM

Étant donné que le RCIP et la communauté muséale canadienne n'avaient aucune expérience avec les concepts ou technologies liés aux données ouvertes et liées, et parce que Artefacts Canada est vaste et contient des données provenant de diverses disciplines muséales, nous avons commencé par un projet pilote utilisant un petit sous-ensemble des notices d'Artefacts Canada.

Le projet pilote «150 ans d'art canadien» a utilisé environ 86000 notices d'œuvres d'art (un ensemble de données d'une seule discipline relativement normalisé) provenant de 8 musées d'art qui possédaient déjà des documents et des images dans Artefacts Canada. Nous

public access) we used SKOS.

With the help of the galleries, sample records were selected and mapped to the resource description model. In addition to the data from galleries on artworks, we also included artist records from the "Artists in Canada" database (creating biographies where they were not available) and institution records. We identified and linked external and internal resources to enrich the resource descriptions: places reference Geonames.org; organizations and artist groups reference Getty ULAN and OCLC VIAF; individual artist names reference OCLC VIAF, Getty ULAN, and Wikidata; and concepts from the Artefacts Canada taxonomy reference AAT, OCLC FAST, LCSH, ICONCLASS, and BnF Rameau.

The pilot project included an investigation of the tools, technologies, and approaches to building applications in the LOD environment. Various triple store products for storing the RDF datasets and enabling SPARQL queries were evaluated, but we found that implementing all the complex search functionalities we wanted using only SPARQL queries in the triple store was a challenge due in part to our requirements for facet lists/counts, pagination, and sorting. So in parallel with the SPARQL query, we used Apache SOLR search platform to do the indexing and searching. Custom tools are used to populate a SOLR instance with data from the triple store. A SPARQL query retrieves the values for each search field from the triple store and a document is generated that is sent to SOLR for indexing. Mapping formulas were developed and applied to the data using tools such as Excel (although we plan to use mapping tools in the future). The pilot project also involved identifying visualizations (data dimensions presented as entry points allowing users to explore) and search functionality ("faceted" searching or filtering using the AAT and our own Artefacts Canada Taxonomy) to be supported in an online application.

The pilot project effectively demonstrated the use of CIDOC-CRM as the core ontology linking a subset of the data in our Canadian national database of museum collections, Artefacts Canada, to other related information resources. It showed the technical feasibility of using LOD tools and methods and the business value of linking cultural and heritage collections across the country and around the world. And it has allowed CHIN to engage the Canadian museum community in exploring ways to improve the value of cultural and heritage collections using LOD. Based on the success of the pilot project, we believe that Linked Open Data is the best approach for modernizing Artefacts Canada. But we have many challenges ahead.

avons défini un modèle de description des ressources en utilisant le CIDOC CRM comme noyau (pour représenter les œuvres d'art), parce que le CRM est le modèle de référence le plus complet et le plus expressif pour la documentation muséographique et la seule ontologie spécifiquement développée pour exprimer des données de musée en RDF. Nous avons également utilisé d'autres ontologies pour gérer les données pour les artistes (FOAF, xfoaf, DC et BIO) et pour les organisations et groupes d'artistes (W3C VCard et ORG). Pour notre taxonomie d'Artefacts Canada (un outil de vocabulaire à facettes conçu pour faciliter l'accès du public), nous avons utilisé SKOS.

Avec l'aide des galeries, des exemples de notices d'objets ont été sélectionnés et mappés au modèle de description des ressources. En plus des données provenant des galeries d'art, nous avons également inclus les dossiers d'artistes de la base de données «Artistes au Canada» (créant des biographies où elles n'étaient pas disponibles) et les dossiers des institutions. Nous avons identifié et relié des ressources externes et internes pour enrichir les descriptions de ressources: les localités référencent Geonames.org; Les organisations et groupes d'artistes référencent Getty ULAN et OCLC VIAF; Les noms d'artistes individuels renvoient à OCLC VIAF, Getty ULAN et Wikidata; Et les concepts de la taxonomie d'Artefacts Canada référencent AAT, OCLC FAST, LCSH, ICONCLASS et BnF Rameau.

Le projet pilote a inclut une revue des outils, technologies et approches de construction d'applications dans l'environnement LOD. Nous avons évalué plusieurs triplestores pour stocker les données RDF et les interroger via SPARQL, mais nous avons constaté que la mise en œuvre de toutes les fonctionnalités de recherches complexes que nous voulions faire en utilisant seulement des requêtes SPARQL constituait un problème en partie au niveau de générer des listes/comptes, la pagination et le tri. Ainsi, parallèlement aux interrogations SPARQL, nous avons utilisé la plate-forme de recherche Apache SOLR pour effectuer l'indexation et la recherche. Des outils personnalisés sont utilisés pour renseigner une instance de SOLR avec les données de la triplestore. Une requête SPARQL extrait les valeurs de chaque champ de recherche à partir de la triplestore et un document est généré et envoyé à SOLR pour l'indexation. Des formules de mappage ont été développées et appliquées aux données à l'aide d'outils tels que Excel (bien que nous prévoyions d'utiliser des outils de mapping dans l'avenir). Le projet pilote comprenait également l'identification de visualisations (dimensions des données présentées en tant que points d'entrée

Limitations and Challenges

The pilot project focused primarily on identifying and testing technology, creating and selecting ontologies and authorities, and creating a basic online prototype for access – but some other challenges remain to be tackled in future phases.

The data in this project so far is not fully open. We are using LOD to supplement, connect and enhance the data within our own repository, and we have converted our data to RDF for users to access in an online application – but the information is not yet open for use by others. We have a number of challenges yet to address in order to make the data truly open for reuse (either as a download or as a SPARQL endpoint / other form of data triple querying capability). For example, we have not yet obtained permission from all (any) of the museums to provide open access to the data.

Canadian museums are not publishing their own LOD. They are generally not even able to export data from their systems in a CRM-compatible form, like LIDO XML – let alone publish CIDOC CRM RDF. So for now, Artefacts Canada must accept data in any format (Excel, tab-delimited, XML). Museums will hopefully someday have LOD support from their software vendors. But the largest group of museums (mostly small community museums) are unlikely to have LOD publishing capabilities for years to come – maybe never. Relying on CHIN to map, convert, host, and publish their data as RDF may be the only means they have of making their data available to the world as LOD.

We cannot expect a data refresh from most museums in the short term. Even though the quality and context may have been compromised during previous mapping/contribution to the old Artefacts Canada, some museums will not be able or willing to refresh their data. We will need to migrate some of the old data by mapping it to the new Artefacts Canada ontology, and converting to RDF.

Creating a sustainable data provisioning process across a wide range of museum capabilities is proving to be difficult. Because of the level of effort required, many museums will not be able or willing to switch from the old contribution mechanism to a new one – so for now we will maintain the old database and contribution mechanism. In the longer term, CHIN will receive new data exports from museums (some in an LOD-compatible format, but most not), and work with them to map their data directly to the new Artefacts Canada ontology.

We have not yet decided on an approach for data cleaning, linking, and URI generation. CHIN is currently

permettant aux utilisateurs d'explorer) et la fonctionnalité de recherche (recherche ou filtrage «facetté» à l'aide de l'AAT et à notre propre taxonomie d'Artefacts Canada) pour être pris en charge dans une application en ligne.

Le projet pilote a démontré de façon persuasive l'utilisation du CIDOC-CRM comme étant l'ontologie centrale reliant un sous-ensemble des données de notre base de données nationale canadienne des collections de musées, Artefacts Canada, à d'autres ressources d'information. Il a démontré la faisabilité technique de l'utilisation d'outils et de méthodes LOD et la valeur commerciale de lier les collections culturelles et patrimoniales partout au pays et dans le monde. Il a également permis au RCIP d'engager la communauté muséale canadienne à explorer les moyens d'améliorer la valeur des collections culturelles et patrimoniales en utilisant LOD. Compte tenu du succès du projet pilote, nous croyons que les données ouvertes et liées (Linked Open Data) est la meilleure approche pour moderniser Artefacts Canada. Mais nous avons de nombreux défis à relever.

Limites et défis

Le projet pilote a principalement porté sur l'identification et la mise à l'essai de technologies, la création et la sélection d'ontologies et d'autorités, et la création d'un prototype en ligne de base pour l'accès – mais d'autres défis restent à relever dans les phases futures.

Les données de ce projet ne sont pas encore entièrement ouvertes. Nous utilisons LOD pour compléter, connecter et améliorer les données dans notre propre référentiel, et nous avons converti nos données en RDF pour que les utilisateurs puissent y accéder avec une application en ligne – mais les informations ne sont pas encore ouvertes pour être utilisées par d'autres utilisateurs. Nous avons encore un certain nombre de défis à relever pour rendre les données vraiment ouvertes à la réutilisation (soit comme téléchargement, soit comme point d'accès SPARQL/ autre forme d'interrogation de triplets). Par exemple, nous n'avons pas encore obtenu l'autorisation d'aucun des musées pour fournir un accès ouvert aux données.

Les musées canadiens ne publient pas leur LOD. En général, ils ne sont même pas capables d'exporter des données de leurs systèmes sous une forme compatible CRM, comme LIDO XML – et encore moins publier CIDOC CRM RDF. Pour l'instant, Artefacts Canada doit accepter des données sous n'importe quel format

experimenting with cleaning/enriching the centralized data, using rule-based scripts that will standardize the data as it is contributed, and linking to authorities. But in the longer term, we would like to provide additional tools and training that would allow museums to do their own data cleaning, mapping to authorities, and URI creation.

Next Steps

We hope to launch the new Artefacts Canada in 2018, but we still have many questions and obstacles. We are now working to incorporate an RDF dataset provided by a large Canadian museum, and this will be a real demonstration of the power of the LOD approach – but the bigger challenge is the data from the vast majority of museums that do not have LOD capabilities. Perhaps CHIN will be able to act as a centralized service to help these museums with data cleaning and linking, assigning URIs, converting to RDF, and hosting within the Artefacts Canada triplestore. We have formed an advisory committee to help guide us and provide input on some of the work that remains to be done:

- adjusting the ontologies and extensions used and expanding them to address non-art data
- mapping data from museums that contribute fresh data; migration of old data for those that don't
- incorporating the Nomenclature classification system as part of the Artefacts Canada taxonomy, evaluating the use of the taxonomy for data access and usability, and publishing parts of the taxonomy created by CHIN as LOD
- expanding and improving the interface, and adding reporting, metrics, dashboard
- scaling up to handle more data
- opening the data where possible (to be negotiated with museums) and clarification of copyright statements
- developing tools and processes to allow CHIN to map, clean, link, convert the data centrally, or to allow museums to do this
- creating sustainable data provisioning processes, given the wide range of museum LOD capabilities, while maintaining the original Artefacts Canada application on an interim basis

These are huge challenges for a small team that is new to the LOD approach, but we can see a way forward. Where in the past we relied on mapping to a centralized data structure, semantic frameworks like the CIDOC CRM have given us a new way ahead for cultural heritage data integration and harmonization that will be much more flexible, powerful, and sustainable over the long term.

(Excel, délimité par des tabulations, XML). Les musées espèrent qu'un jour ils bénéficieront du support LOD de leurs éditeurs de logiciels. Mais le plus grand groupe de musées (principalement de petits musées communautaires) ne sera probablement pas doté de capacités de publication LOD pour les années à venir – peut-être jamais. S'appuyant sur le RCIP pour cartographier, convertir, héberger et publier leurs données comme RDF peut être le seul moyen qu'ils ont de mettre leurs données à la disposition du monde comme LOD.

Nous ne pouvons pas nous attendre à un rafraîchissement des données de la plupart des musées à court terme. Même si la qualité et le contexte ont peut-être été compromis au cours de la précédente cartographie/ contribution à l'ancien Artefacts Canada, certains musées ne seront pas en mesure de rafraîchir leurs données. Nous aurons besoin de migrer certaines des données anciennes en les cartographiant vers la nouvelle ontologie d'Artefacts Canada et en les convertissant en RDF.

La mise en place d'un processus d'approvisionnement durable de données à travers un large éventail de capacités de musée s'avère difficile. En raison du niveau d'effort requis, de nombreux musées ne seront pas capables ou désireux de passer de l'ancien mécanisme de contribution à un nouveau. À plus long terme, le RCIP recevra de nouvelles données provenant de musées (certains dans un format compatible avec la LOD, mais la plupart ne le seront pas), et travaillera avec eux pour cartographier leurs données directement à l'ontologie d'Artefacts Canada.

Nous n'avons pas encore décidé d'une approche pour le nettoyage des données, la liaison et la génération d'URI. Le RCIP expérimente actuellement le nettoyage/ enrichissement des données centralisées, en utilisant des scripts basés sur des règles qui normaliseront les données au fur et à mesure qu'elles sont apportées, et en reliant les autorités. Mais à plus long terme, nous aimerions offrir des outils et une formation supplémentaires qui permettraient aux musées de faire leur propre nettoyage des données, la cartographie auprès des autorités et la création de l'URI.

Prochaines étapes

Nous espérons lancer le nouveau Artefacts Canada en 2018, mais nous avons encore beaucoup de questions et d'obstacles. Nous travaillons maintenant à intégrer un ensemble de données RDF fourni par un grand musée canadien, ce qui sera une véritable démonstration de la puissance de l'approche LOD –

mais le plus grand défi est les données de la grande majorité des musées qui n'ont pas de capacités LOD. Peut-être le RCIP sera-t-il en mesure d'agir comme un service centralisé pour aider ces musées à nettoyer et à relier des données, à attribuer des URI, à convertir en RDF et à héberger le triplestore d'Artefacts Canada. Nous avons formé un comité consultatif qui nous aidera à nous guider et à fournir des commentaires sur certains des travaux qui restent à faire:

- ajuster les ontologies et les extensions utilisées et les étendre aux données non-art
- cartographier les données des musées qui fournissent de nouvelles données; Migration de données anciennes pour ceux qui ne
- incorporer le système de classification de la nomenclature dans la taxonomie d'Artefacts Canada, évaluer l'utilisation de la taxonomie pour l'accès aux données et la convivialité et publier des parties de la taxinomie créée par le RCIP en tant que LOD
- l'expansion et l'amélioration de l'interface et l'ajout de rapports, de métriques, de tableau de bord
- mise à l'échelle pour traiter plus de données
- l'ouverture des données lorsque cela est possible (à négocier avec les musées) et la clarification des déclarations de droits d'auteur
- élaborer des outils et des processus pour permettre au RCIP de cartographier, de nettoyer, de lier, de convertir les données de façon centralisée ou de permettre aux musées de le faire
- la création de processus de provisionnement de données durables, compte tenu de la vaste gamme de capacités LOD du musée, tout en maintenant provisoirement l'application originale d'Artefacts Canada

Ce sont des défis énormes pour une petite équipe qui est nouvelle à l'approche LOD, mais nous pouvons voir une voie à suivre. Dans le passé, nous nous sommes appuyés sur la cartographie d'une structure de données centralisée, des cadres sémantiques comme le CIDOC CRM nous ont donné une nouvelle voie pour l'intégration et l'harmonisation des données du patrimoine culturel qui sera beaucoup plus flexible, puissant et durable à long terme.

PERSISTANT IDENTIFIER SERVICES FOR THE HUMANITIES WORKSHOP

GORDON MCKENNA

Members of CIDOC were invited to take part in a workshop on *Persistent Identifier Services for the Humanities* which took place at the British Library in December 2016. CIDOC was invited to give a museum perspective for the discussions. Three, UK-based, members were able to attend.

The workshop was run by the THOR¹¹ project, and began with a presentation on the project.

SERVICES D'IDENTIFICATION PERSISTANTE POUR L'ATELIER SUR LES SCIENCES HUMAINES

GORDON MCKENNA

Les membres du CIDOC ont été invités à participer à un atelier sur les services d'identification persistante pour les humanités qui a eu lieu à la British Library en décembre 2016. Le CIDOC a été invité à donner une perspective muséale aux discussions. Trois membres du Royaume-Uni ont pu y assister.

L'atelier était géré par le projet THOR11 et a commencé par une présentation sur le projet.



THOR (Technical and Human infrastructure for Open Research)

by Adam Farquhar, British Library.

THOR is funded by the European Commission, for 30 months (2015-2017), as part of its major research programme *Horizon 2020*. The project's aim is to:

"To ensure every researcher, at any phase of their career, or at any institution, will have seamless access to persistent identifiers (PIDs) for their research artefacts and their work will be uniquely attributed to them"

[Presentation slide]

It has goals to:

- Place PIDs at the fingertips of researchers;
- Integrate PIDs into services researchers already use;
- Ensure PIDs are embedded in research outputs;
- Make persistent identifier the default;
- Seamless integration among articles, data, and researchers.

THOR (Infrastructure technique et humaine pour la recherche ouverte)

Par Adam Farquhar, British Library.

THOR est financé par la Commission européenne pour une période de 30 mois (2015-2017) dans le cadre de son programme majeur de recherche Horizon 2020. Le projet vise à:

"Assurer à chaque chercheur, à n'importe quelle phase de sa carrière ou dans n'importe quelle institution, un accès transparent à des identificateurs persistants (PID) pour leurs artefacts de recherche et leur travail leur sera attribué de manière unique"

[slides]

Il a pour objectif de:

- Placer les PID à la portée des chercheurs;
- Intégrer les PID dans les services que les chercheurs utilisent déjà;
- Veiller à ce que les PID soient intégrés dans les résultats de la recherche;
- Rendre l'identifiant persistant par défaut;
- Intégration transparente entre les articles, les données et les chercheurs.

¹¹ See: <https://project-thor.eu>

And has 'focus areas' of:

- Biological and Medical Sciences;
- Environmental and Earth Sciences;
- Physical Sciences;
- Humanities and Social Sciences.

It also has a view on PIDs as:

- A name not an address
- Globally unique;
- Globally resolvable;
- Bound to core metadata;
- Interlinkable;
- Interoperable;
- Professionally managed;
- Backed with organisational commitment;
- Designed to last beyond the lifetime of any system or (most) organisation.

The major areas it considered for PIDs were:

- Researchers & Contributors;
- Organisations;
- Funding Agencies;
- Datasets;
- Publications;
- Projects.

In this context it is significant that two partners of the project are responsible for issuing and managing PIDs:

- DataCite – Provides persistent identifiers (DOIs) for research data and other research outputs;
- ORCID – An open registry of unique, persistent identifiers for researchers and contributors.

The second part of the day was three presentations from the humanities area:

Transkribus and the READ project

by Dr Louise Seaward, Bentham Project, Faculty of Laws, University College London.

READ (Recognition and Enrichment of Archival Documents)¹² is a project aimed at making archival material more accessible through the use of handwritten text recognition (HTR) technology. Transkribus¹³ is a research infrastructure for working with historical documents, including a downloadable expert client and web-based tools.

¹² See: <https://read.transkribus.eu>

¹³ See: <http://transkribus.eu>

«Domaines d'intérêt» du projet:

- Sciences biologiques et médicales;
- Sciences de l'environnement et de la terre;
- Sciences physiques;
- Sciences humaines et sociales.

Le projet estime qu'un PID doit être:

- Un nom et non une adresse;
- Globalement unique;
- Avoir une résolution globale;
- Lié aux métadonnées de base;
- Interconnectable;
- Interopérable;
- Gestion professionnelle;
- Soutenu par l'engagement institutionnel;
- Conçu pour durer au-delà de la durée de vie de tout système ou (la plupart) de l'organisation.

Les principaux domaines qu'il a pris en considération pour les PID étaient les suivants:

- Chercheurs et collaborateurs;
- Organisations;
- les organismes de financement;
- Ensembles de données;
- Publications;
- Projets.

Dans ce contexte, il est significatif que deux partenaires du projet soient responsables de l'émission et de la gestion des PID:

- DataCite – Fournit des identificateurs persistants (DOI) pour les données de recherche et autres résultats de recherche;
- ORCID – Un registre ouvert d'identifiants uniques et persistants pour les chercheurs et les contributeurs.

La deuxième partie de la journée inclut té trois présentations d'intervenants des sciences humaines:

Transkribus et le projet READ

Par Dr Louise Seaward, projet Bentham, Faculté des lois, University College London.

READ (Reconnaissance et enrichissement des documents d'archives)¹² est un projet visant à rendre les archives plus accessibles grâce à l'utilisation de la technologie de reconnaissance de texte manuscrit (HTR). Transkribus¹³ est une infrastructure de recherche pour travailler avec des documents historiques, y

- Transcribing manuscripts (for a digital edition);
- Creating training data for HTR technology;
- Participating in crowdsourcing projects.

An example of the use of the infrastructure presented was the Transcribe Bentham¹⁴ project which was part of producing a scholarly edition of the papers of the British philosopher Jeremy Bentham (1748-1832). Here 17 thousand pages were transcribed by volunteers, however the project was dependent on small group of 'super-transcribers'. One issue that was address was: How to improve the user experience and encourage more people to participate? Two methods were used:

- Use of the HTR technology to help transcribers;
- Use of ORCHIDs for those taking part.

This meant:

- Researchers could get credit for their transcription and editing work;
- Volunteers could record their contributions to crowdsourcing projects;
- Work was better recognised, cited and used, leading to improved access to historical documents.

This was a significant for the workshop.

Citation culture – British History Online as an example

by Jonathan Blaney, Institute of Historical Research.

This presentation looked at issues around the citation of historical sources, especially when the source that researchers used was a digital version available on a website such as *British History Online*¹⁵, which is a digital library of printed primary and secondary sources for the history of Britain and Ireland. There is a tendency to only reference the original even if the digital repository provides a ready-made reference as British History Online does.

The issue was described as one of the lack of education for academics, and also, perhaps, prejudice against online citation by publishers even if they have a view. Therefore some guidance for researchers is needed.

¹⁴ See: <http://blogs.ucl.ac.uk/transcribe-bentham>

¹⁵ See: <http://www.british-history.ac.uk>

compris un client expert téléchargeable et des outils Web.

Dans Transkribus il y a trois activités:

- Transcription de manuscrits (pour une édition numérique);
- Création de données d'entraînement pour la technologie HTR;
- Participer à des projets de crowdsourcing.

Un exemple de l'utilisation de l'infrastructure présentée fut le projet de Transcribe Bentham¹⁴ qui a permis de produire une édition érudite des documents du philosophe britannique Jeremy Bentham (1748-1832). Ici, 17 000 pages ont été transcrites par des bénévoles, mais le projet dépendait d'un petit groupe de «super transcribers». Une des questions abordées était la suivante: Comment améliorer l'expérience utilisateur et encourager davantage de personnes à participer? Deux méthodes ont été utilisées:

- Utilisation de la technologie HTR pour aider les transcribers;
- Utilisation des ORCHID pour les participants.

Cela signifiait que:

- Les chercheurs pouvaient obtenir un crédit pour leur travail de transcription et d'édition;
- Les bénévoles pouvaient enregistrer leurs contributions aux projets de crowdsourcing;
- Le travail était mieux reconnu, cité et utilisé, ce qui a permis d'améliorer l'accès aux documents historiques.

Cela a été important pour l'atelier.

Citation culture – British History Online en tant qu'exemple

Par Jonathan Blaney, Institut de recherche historique.

Cette présentation a examiné les questions entourant la citation de sources historiques, en particulier lorsque la source utilisée par les chercheurs est une version numérique disponible sur un site Web tel que British History Online¹⁵, une bibliothèque numérique de sources imprimées primaires et secondaires pour l'histoire de l'Angleterre et l'Irlande. On a tendance à ne faire référence qu'à l'original même si l'entrepôt numérique fournit une référence prête comme fait British History Online.

On useful resource that was mentioned in the presentation was TIDSR (Toolkit for the Impact of Digitised Scholarly Resources)¹⁶ published by JISC (Joint Information Systems Committee) in the UK. This might have a wider application in the cultural heritage sector in general.

Networking The (Face)Books of the Dead

by Dr K Faith Lawrence, King's College London. This presentation focused on the PID issues that were found during the Sharing Ancient Wisdoms¹⁷ project. The focus was:

"... collections of ideas and opinions – ranging from pithy sayings to short passages from longer philosophical texts – which make up the ancient genre of Wisdom Literature".

One important aspect of this was being able to define relationships between entities:

- Text;
- Places;
- People.

It was relatively simple to do this for the texts themselves and places. For ancient places there is the Pleiades¹⁸ gazetteer. However for people it proved more difficult, with significant disagreement between scholars in identification. It is possible that the same person is described differently in two authority source, or one person in one list to be mapped to two or more in another source. This needs to be addressed if we want to use PIDs to manage the relationships. This aspect of uncertainty of identification was one of the issues identified by the group discussions in the afternoon.

Group discussions

Over lunch the group was asked to indicate (with stickers) which of:

- Persons (including researchers and contributors);
- Organisations;
- Datasets;
- Publications;
- Funding Agencies;
- Projects.

La question a été décrite comme une lacune de la part des universitaires, et aussi, peut-être, comme un préjugé contre la citation en ligne par les éditeurs, même si elles ont une vue. Il est donc nécessaire de fournir des conseils aux chercheurs.

Parmi les ressources utiles mentionnées dans la présentation est le TIDSR (Toolkit for the Impact of Digitized Scholarly Resources)¹⁶ publié par le JISC (Joint Information Systems Committee) au Royaume-Uni. Cela pourrait avoir une application plus large dans le secteur du patrimoine culturel en général.

Networking Le (Face)Books des morts

Par le Docteur K Faith Lawrence, du King's College de Londres. Cette présentation a porté sur les questions PID qui ont été apparues lors du projet Sharing Ancient Wisdoms¹⁷. L'accent était mis sur:

«[...] des collections d'idées et d'opinions – allant de proverbes brefs à des passages courts de textes philosophiques plus longs – qui constituent le genre ancien de la littérature de la Sagesse».

Un aspect important est de pouvoir définir les relations entre les entités:

- Texte;
- Lieux;
- Gens.

C'est relativement simple à faire pour les textes eux-mêmes et les lieux. Pour les lieux anciens il y a la nomenclature des Pléiades¹⁸. Toutefois, pour les gens, c'est plus difficile, avec un désaccord important entre les chercheurs en matière d'identification. Il est possible que la même personne soit décrite différemment dans deux sources autoritatives, ou qu'une personne dans une liste soit assignée à deux ou plus dans une autre source. Cela doit être abordé si nous voulons utiliser les PID pour gérer les relations. Cet aspect de l'incertitude de l'identification fut l'un des problèmes identifiés par les discussions de groupe dans l'après-midi.

Discussions de groupe

Au cours du déjeuner, le groupe a été invité à choisir (avec des autocollants):

- les personnes (y compris les chercheurs et les contributeurs);
- Organisations;
- Ensembles de données;
- Publications;
- les organismes de financement;

¹⁶ See: <http://digitisation.jiscinvolve.org/wp/2009/05/18/you-can-now-measure-the-impact-of-your-online-resource>

¹⁷ See: <http://www.ancientwisdoms.ac.uk>

¹⁸ See: <https://pleiades.stoa.org>

They used at present, and which of them they would like to use in future. Those from a humanities, non-academic background indicated the first four but did not deal with funding agencies and projects. Those who were academics did have these.

We were also asked to add any other entities that we dealt with or wanted to have. Museum-based attendees added:

- Objects;
- Places;
- Subjects (including periods);
- Events.

In the afternoon the workshop the attendees, divided into four groups, discussed two questions:

- What do you want to do with persistent identifiers?
- What are the problems that persistent identifiers can help to address?

The group with all the CIDOC members came up with the aim: *“To be able to describe the relationships between everything at all times in history”*. PIDs were identified as one of the ways in which these relationships can be made clear. However it was recognised that there has to be the ability to disagree and for that to be clear also.

The groups also examined: What are the obstacles to adoption? The following issues were identified:

- Financial;
- Technical;
- Cultural (e.g. lack of buy-in from end users);
- Time/effort;
- Lack of added value;
- Lack of functional support (e.g. can get a DOI for a dataset, but can't link it to publications).

None of these were unexpected. However it was foreseen that the THOR project might be able to address some of these in their work.

- Projets.

Ils ont indiqué ceux qu'ils utilisent à l'heure actuelle, et ceux qu'ils voudraient utiliser à l'avenir. Les participants issus d'une formation arts libéraux non-universitaire ont indiqué les quatre premiers, mais n'ont pas choisi les organismes de financement et les projets. Les universitaires ont choisi les 2 derniers.

Nous avons également été invités à ajouter toutes les autres entités avec lesquelles nous travaillons ou voudrions avoir. Les participants de formation muséale ont ajouté:

- des objets;
- Des endroits;
- les sujets (y compris les périodes);
- Événements.

Dans l'après-midi, les participants, divisés en quatre groupes, ont débattu de deux questions:

- Que voulez-vous faire avec les identifiants persistants?
- Quels sont les problèmes que les identificateurs persistants peuvent aider à résoudre?

Le groupe avec tous les membres du CIDOC choisirent l'objectif: *«Pouvoir décrire les relations entre tout ce qui se passe à tout moment de l'histoire»*. Les PID ont été identifiés comme l'une des façons par lesquelles ces relations peuvent être clarifiées. Cependant, il fut reconnu qu'il doit y avoir la capacité de ne pas être d'accord et que ce désaccord soit clair aussi.

Les groupes ont également examiné la question suivante: Quels sont les obstacles à l'adoption des PID? Les problèmes suivants furent identifiés:

- financier;
- Technique;
- culturels (par exemple manque d'acceptation des utilisateurs finaux);
- Temps/effort;
- Manque de valeur ajoutée;
- Manque de support fonctionnel (par exemple, possibilité d'obtenir un DOI pour un ensemble de données, mais impossibilité de le lier à des publications).

Aucun de ces problèmes n'était inattendu. Cependant, il était prévu que le projet THOR puisse être en mesure d'aborder certaines de ces questions dans son travail.