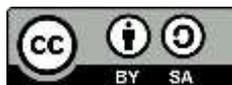


Bern University of Applied Sciences
E-Government Institute

Swiss Heritage Institutions in the Internet Era **Results of a pilot survey on open data and crowdsourcing**

Beat Estermann

Version 1.0
Bern, 26 July 2013



The text of the report and our own graphics are published under a
“Creative Commons Attribution-ShareAlike 3.0 Unported” license.

Picture on cover page:

Swiss National Museum Zurich. Wikimedia Commons, Roland Fischer, Zurich, CC-BY-SA-3.0 Unported.



Management Summary

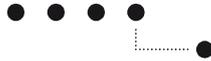
The rise of the internet in the 1990s and the innovations that have resulted from it have not left cultural heritage institutions unaffected. Libraries, archives and museums have since undergone several consecutive trends. This study focuses on two of them: open data and crowdsourcing.

Both trends are a result of the increasing digitisation of cultural heritage-related material and the corresponding metadata. From the beginning of the new millennium, Europe's cultural heritage institutions have coordinated their efforts to digitise cultural heritage-related material as comprehensively as possible. As a result, 'single points of access' in the form of inter-organisational catalogue systems and virtual libraries have developed alongside the increase in networking between institutions for the purpose of exchanging know-how and avoiding duplication. An additional trend towards greater interaction options during the second half of the first decade went hand in hand with the rise of the social web. This trend found expression in new forms of personalisation and the increased inclusion of users and visitors as 'co-producers'. At the same time, the rise of Wikipedia, Flickr and similar community-based services has focussed attention on the potential for collaborative creation and compilation of information on the internet. Cultural heritage institutions and online communities have started to collaborate, whilst some institutions have initiated independent crowdsourcing projects. The Wikipedia community recognises that cultural heritage institutions are important as partners and has actively sought to collaborate with them for a number of years. From 2009 onwards, the Open Government Data movement began to spread from the USA and UK and soon became rooted in around 50 nations around the world. The movement's demand: public sector data should be made freely available in a machine-readable format for use by third parties. From the viewpoint of state governance, this can be seen as the consequent implementation of the principle of public access, whereby administrative data is made available in a machine-readable format as standard practice. Making available 'open government data' is intended to lead to greater transparency, participation and cooperation. There is, however, another driving force behind the 'open government data' initiative, namely an economic-technological objective that involves the creation of a 'semantic web' by linking together as many 'open' datasets as possible. The semantic web converts the data held by public authorities and possibly private enterprises into a freely accessible infrastructure resource that serves third parties as a basis for value-added services such as data refinement, the production and visualisation of innovative links between datasets, and other data-based services.

At the end of 2012, we carried out a pilot survey to establish where Switzerland's cultural heritage institutions stand with regard to open data and crowdsourcing, how they see the opportunities and risks associated with these trends and what potential benefits they consider these have to offer. Around 200 Swiss cultural heritage institutions of national significance were invited to participate, of which approximately one third completed the questionnaire. The results of the survey paint the following picture of the current situation with regard to open data and crowdsourcing:

The digitisation of cultural heritage material and the associated metadata is an important prerequisite for open data and crowdsourcing. The majority of surveyed institutions (60%) make both metadata and images of at least some of their holdings available on the internet. A good two-fifths of the institutions (43%) are members of organisations in which the exchange of metadata plays an important role. A similarly high proportion considers the exchange of metadata to be relevant to the achievement of their core objectives. However, one half of the surveyed institutions indicated a need for improvement to the quality and interoperability of metadata in the medium term, if not sooner.

Hardly any of the surveyed institutions currently fully embrace open data. There are, however, indications that this innovation could rapidly become accepted. A majority of the surveyed institutions consider that open data is of importance and the opportunities it offers outweigh the risks. Concerns about copyright infringements and violations of confidentiality regulations are major obstacles. Better visibility and accessibility to cultural heritage institutions and the establishment of a more comprehensive network between them are perceived as the main benefits of open data. Additionally, many institutions have reservations concerning the 'free' licensing of holdings.



Their concerns relate specifically to the commercial exploitation of holdings and modifications to them by third parties. Concerns relating to loss of control over holdings are shared by 68% of the surveyed institutions and this could therefore become a major obstacle to the adoption of open data by cultural heritage institutions.

In general, use of crowdsourcing is developing at a slower rate than the trend towards acceptance of open data. Although around 10% of the surveyed institutions have experimented with crowdsourcing, a major breakthrough has yet to happen. Whilst nearly half of the institutions consider the subject to be of importance, the majority is nonetheless sceptical with regard to the idea. In general, crowdsourcing is seen as being associated with a lot of risk and little benefit. Opportunities to use crowdsourcing are mostly seen for the improvement of metadata and for tasks relating to correction and transcription work. However, Switzerland's cultural heritage institutions still need to be convinced that this would improve efficiency. The majority of Swiss cultural heritage institutions that collaborate with online communities to create material do so mainly through Wikipedia: 11% of surveyed institutions have employees who contribute to Wikipedia within the scope of their work. Employees at another 14% do so in their free time.

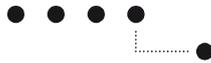
The survey results show that cultural heritage institutions perceive the benefits of open data to be in the fields of education and research and in making material available to the general public. Furthermore, it is assumed that the adoption of an open data policy would enhance the interconnectedness of cultural heritage institutions, increase the visibility of their holdings and generally improve how the institutions are perceived by the general public.

However, the anticipated benefits of open data and crowdsourcing are considered to be offset by additional effort and expense: the surveyed institutions name these as the greatest challenge facing the implementation of open data and crowdsourcing. Loss of revenue, on the other hand, is not a major concern. Public and private funding organisations that provide a large proportion of the funds required by cultural heritage institutions could play a key role in promoting open data. In view of the fact that the majority of cultural heritage institutions recognises the benefits of open data and is not concerned about a loss of revenue, it would seem logical for funding organisations to promote and encourage the 'free' licensing of holdings.



Table of Contents

1	Introduction	6
2	Cultural heritage institutions in the internet era – current developments and trends	8
2.1	Digitisation and increased cooperation and coordination	8
2.2	More interaction options and customised services	9
2.3	User co-production (crowdsourcing)	10
2.4	‘Free’ licensing and open data	13
2.5	Linked data / semantic web	16
3	Results of the pilot survey	20
3.1	Research questions	20
3.2	Methodology / research design	20
3.3	Representativeness and validity of results	21
3.4	Characteristics of the surveyed cultural heritage institutions (sample description)	23
3.5	Where do Swiss heritage institutions stand today with regard to open data and crowdsourcing?	30
3.6	What opportunities and risks do the institutions see in relation to open data and crowdsourcing?	44
4	Conclusions and outlook	49
4.1	Main insights gained from the pilot survey	49
4.2	Aspects omitted or inadequately addressed by the pilot survey	50
4.3	Outlook	51
5	List of references	52
6	Links for further information	54
7	About the author of this study	55
	Appendix: Questionnaire	56



1 Introduction

Cultural heritage institutions are currently undergoing significant changes, especially in the field of online activities. The 'free' availability of data on the internet and new forms of online collaboration represent not only numerous opportunities for archives, libraries and museums but also many new challenges.

Some cultural heritage institutions have already managed to use the new culture of online collaboration to their advantage and are able to involve online communities in the development of their programmes thanks to innovative approaches. In the majority of cases, however, such an approach is still in the planning and testing phase. They first want to establish more clearly how they can best adapt to the new circumstances.

The Wikipedia community, one of the largest communities based on online cooperation, has continuously increased its collaborations with cultural heritage institutions over recent years. Its objective is to acquire new material for the online encyclopaedia and to adopt new approaches together with cultural heritage institutions to promoting and conveying free knowledge.

In order to gain an initial overview of the extent to which cultural heritage institutions in German-speaking Switzerland are adapting to the new trends and also to obtain the information required to be able to develop tailor-made support programmes as and where needed, Bern University of Applied Sciences decided to conduct a pilot survey among these institutions. The results of the survey are being made available to interested parties and the general public in this report.

Content structure:

Chapter 2 provides a brief synopsis of current developments and trends in relation to cultural heritage institutions and presents a number of concrete examples to enable those unfamiliar with the subject to better understand the areas addressed by the survey.

Chapter 3 offers background information on the pilot survey, its results and a discussion of the insights gained from the survey.

Chapter 4 contains a summary of the main results of the survey and indicates the potential next steps that could be taken on the basis of these insights.

At the end of the report are a list of references and useful internet links, a brief description of the E-Government Institute at Bern University of Applied Sciences in whose name the survey was carried out, and a profile of the study's author.

The questionnaire itself is appended. The text of this report and the questionnaire are provided under a 'free' copyright licence. The copyright holder not only permits free reproduction of the material but actually desires its reproduction and dissemination. If you plan to use the questionnaire or parts of it for your own research projects, we would be grateful if you could contact us beforehand to coordinate your approach if necessary and to avoid duplication. Thank you in advance for your cooperation and understanding.



Acknowledgements:

My thanks go first and foremost to Daniel Felder, David Studer and Markus Vogler who made a valuable contribution towards literature research, helped develop the questionnaire, carried out the survey and an initial statistical evaluation of the results as part of their Master of Business Administration course at Bern University of Applied Sciences. In this connection, I would also like to thank Christoph Urwyler of the Social Work Division for his help with the use of the 'Unipark' online questionnaire tool.

My thanks go to the following people for their valuable input to the development of the questionnaire and helpful feedback on the draft of this report: Doris Amacher (Swiss National Library), Barbara Fischer (Wikimedia Germany), André Golliez (opendata.ch), Frank von Hagel (Institute for Museum Research, State Museums of Berlin), Alessia Neuron (Bern University of Applied Sciences), Hartwig Thomas (Digitale Allmend Association), and David Vuillaume (Association of Swiss Museums).

I'm also grateful to Sharon Willingham and her team for the translation of the report into English.

And finally, I would like to thank all the pre-test participants whose feedback enabled us to tweak the questionnaire, and to every participant in the survey who took the time to respond to the questionnaire in spite of their many other duties.

Feedback/contact

We would be delighted to receive feedback and enquiries from you in relation to this survey; they will provide us with valuable input for our future work.

Beat Estermann
Research Associate
Bern University of Applied Sciences
E-Government Institute

beat.estermann@bfh.ch
+41 31 848 34 38



2 Cultural heritage institutions in the internet era – current developments and trends

2.1 Digitisation and increased cooperation and coordination

As the internet has developed, cultural heritage institutions have begun to digitise their metadata and holdings and make them available online. The Lund Action Plan¹, for instance, was launched in 2001 in order to coordinate digitisation efforts at the European level.

Various new forms of cooperation among cultural heritage institutions followed, including joint projects that offered users a 'single point of access', such as cross-organisational catalogues and 'virtual libraries' and 'virtual museums'. The "DigiCULT Report" of the European Union (European Commission & Salzburg Research 2002) called for, among other things, more cooperation and coordination between cultural heritage institutions and was characteristic for this phase of development. It proposed that archives, libraries and museums should explore new relationships with their environments by networking with other institutions across segment boundaries and forming contacts with private enterprises, intermediaries and new user groups. Partnerships of this nature would improve the efficiency of providing new services, coordinating digitisation projects and developing standards and structures to enable seamless access and shared resources. It was postulated that participation in cross-segment networks would become an important aspect of every organisation (European Commission & Salzburg Research 2002, p. 10).

The European Commission reiterated its recommendations four years later and suggested creating a European library – a standardised, multi-lingual online portal to Europe's cultural heritage that is currently being developed under the name of 'Europeana'. In its statement the Commission also named various challenges that would surface along the way (European Commission 2006, p. 3):

- Economic challenges: Who will pay for digitisation?
- Organisational challenges: How can synergies be exploited and duplication avoided? How can fruitful partnerships be established between public entities and private players?
- Technical challenges: How can the cost of digitisation be reduced but good data quality be maintained at the same time?
- Legal challenges: What approach should be adopted towards third-party copyright claims to ensure works protected by copyright are covered?

The Commission also anticipated similar challenges with regard to long-term archiving and promoted the creation of explicit strategies in member states and more cooperation at a European level (European Commission 2006, p. 3).

The report about the Swiss cultural heritage policy – "Memopolitik" - of the Swiss Federal Office of Culture published in 2008 also addressed these developments and made out a certain deficit in comparison to neighbouring countries: "Switzerland is in danger of falling behind, at least with regard to digital libraries; this will eventually lead to dependencies, at the latest if German or French portals become fee-based services." (Bundesamt für Kultur 2008, p.8). Whilst various projects are tackling the challenges associated with the digitisation of cultural heritage items held by libraries and archives and making these available online, the report identified a deficit with regard to the links between the services offered by libraries and archives and those offered by museums (Bundesamt für Kultur 2008, p. 58). Moreover, there were hardly any meta-projects between

¹ <http://cordis.europa.eu/ist/digicult/lund-principles.htm>



individual museums. The report also pointed up the differences in professional culture between archives, libraries, and museums; they would impede harmonisation efforts while there are already tensions in many cases due to the competition for public resources (Bundesamt für Kultur 2008, p. 27).

2.2 More interaction options and customised services

The emergence of the social web and its interactive methods of exchange enhanced the trend among cultural heritage institutions to develop more interaction options and customised services. The increasing popularity of mobile devices has had a reinforcing effect. Some institutions make use of devices (e.g. iPads) that they lend to users. Others exploit the availability of internet-enabled devices owned by visitors to the institutions.

While some of the new interaction options are located within the institution's internet presence – e.g. in the form of Facebook pages or Twitter accounts, where interested parties can engage in dialogue with curators – other forms of interaction require physical visits to collections. The following four factors are relevant to how museum collections are presented (Luyten et al. 2011):

- *Mobility*: How freely should the visitor be allowed to wander through the collection? - The possibilities range from the provision of stationary information to mobile information that follows a set path to mobile, context-dependent information that adapts to the movement patterns of visitors.
- *Interaction with the environment*: What interaction takes place between visitors and the museum environment? – The spectrum ranges from an inactive environment that provides static information to museum landscapes in which every artefact encourages the visitor to engage in dialogue.
- *Personalisation*: To what extent can the presentation of a collection be tailored to individual visitors? – The range stretches from a standard programme for all visitors to a differentiation by user group to a customised programme designed around an individual's user profile.
- *Social interaction*: To what extent does the presentation of the collection promote communication between visitors? – While some solutions make every visit a personal experience (e.g. through the use of an audio guide with headphones that runs permanently and that is more likely to impede interaction with other visitors than encourage it), some museums at the opposite end of the spectrum make use of ludic approaches to encourage visitors to interact with each other.

The example of Derby Museum shows how the information provided by an institution can merge with content made available by third parties on the internet to which hyperlinks are provided (see Case example 1).

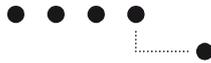
Case example 1: QRpedia at Derby Museum

Derby Museum in the UK was the first to use so-called Quick-Response Codes (QR codes) with its exhibits to draw visitors' attention to the fact that additional information is available through Wikipedia. QR codes are similar to barcodes in that they represent information in a graphic form that can be read by electronic devices. The QR codes in use at Derby Museum can be scanned using a mobile end device (e.g. a smartphone) and refer the user to the corresponding article in Wikipedia. The Wikipedia article is displayed in the language the user has set for his mobile device.

In this example the user is explicitly encouraged to gather additional information online using their own mobile device. Using the user's language settings to display information is a form of personalisation. Visitors following links to Wikipedia articles can also edit the information online or leave a comment on the talk page.



Photo: [Wikimedia Commons](#), User: Victuallers, CC-by-sa.



2.3 User co-production (crowdsourcing)

The term ‘crowdsourcing’ was first used in 2006. Jeff Howe published the article “The Rise of Crowdsourcing“ in *Wired Magazine* and described in detail the new phenomenon and how it works. Howe defined the neologism ‘crowdsourcing’, an amalgamation of the words ‘crowd’ and ‘outsourcing’, as follows:

“Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential labourers.” (Howe, 2006)

The term ‘crowdsourcing’ is not without controversy. Trevor Owens, responsible for digital archiving at the Library of Congress, points out that many typical crowdsourcing projects do not involve a large, anonymous crowd; in fact the majority of contributions are made by a small, manageable community of volunteers, so that the phenomenon is not all that different to the traditional practice of cultural heritage institutions of inviting people to contribute voluntarily to the achievement of the institution’s objective. Moreover, the ‘sourcing’ part of the term stems from the world of business and could imply that, under certain circumstances, those who contribute are being unfairly exploited. This is definitely *not* how cultural heritage institutions should understand the concept; in fact, crowdsourcing should be seen as an invitation extended to amateurs (in the sense of enthusiasts) to participate in the creation, development and enhancement of cultural heritage (Owens 2012, p.1-2).

Moreover, this aspect of ‘sourcing’ is not really applicable to collaborative projects initiated by volunteers. This is the case with Wikipedia, for instance, or with online museums created from bottom-up initiatives such as the “Museum of Menstruation and Women’s Health”². Again, the difference between the internet-based phenomena of today and the customary practice of cultural heritage institutions is not so great – after all, many of them originated in the initiative of a small group of people who initially spent a lot of painstaking time and effort creating a new offer in a niche previously unoccupied by established institutions.

Typology of crowdsourcing approaches

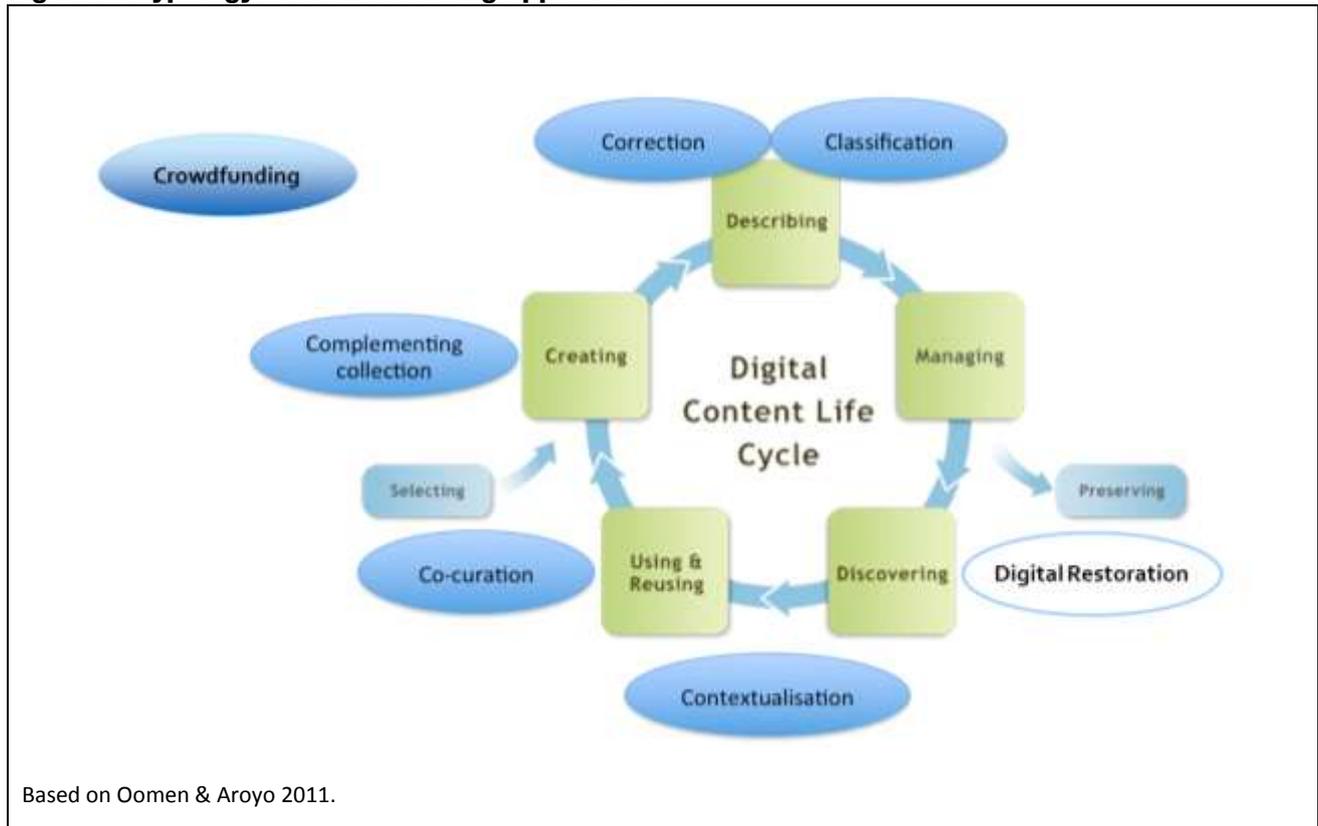
There are meanwhile many examples of crowdsourcing being used by cultural heritage institutions³ as well as approaches to defining its typology. The typology suggested by Oomen & Aroyo (2011), for instance, is derived from the life cycle model developed for digital material by the National Library of New Zealand (see Figure 1). According to this model, crowdsourcing can be utilised in the following areas: Correction and transcription tasks; classification or supplementation of metadata; contextualisation of cultural heritage items; co-curation; supplementation of collections by contributing additional items or by facilitating their acquisition. In addition, crowd-funding solutions can be used to finance activities during various life cycle stages (see case examples 2-7). Another area of utilisation that is at least partially documented in Wikipedia is the restoration of digital items by volunteers.

² <http://www.mum.org/>

³ For examples see Holley (2009) and Oomen/Aroyo (2011).



Figure 1: Typology of crowdsourcing approaches



Case example 2: Project Trove, National Library of Australia (correction and transcription tasks)⁴

'Trove' is the name given to the digitisation project at the National Library of Australia. In 2012, the online archive comprised more than 300 million digital sources such as images, historic magazines and newspapers, maps and music. Thousands of Australians are participating in this crowdsourcing project by correcting texts, reworking photographs or improving other digital material for public use. Trove is intended to be an access point for all types of digital content relating to Australia. The voluntary group consists of professional and amateur researchers, teachers, students and other interested parties and is making sure that 200 years of Australian history are preserved for posterity. The project was started in 2008; by 2009, 7 million lines of text in 318,000 articles had already been edited (National Library of Australia, 2012, D-Lib, 2012).

Case example 3: German Federal Archives, image donation to Wikimedia Commons (classification / supplementation of metadata)

The German Federal Archives hold around 10 million photographs (status as per 2008). The archives' core objective lies in collecting, archiving and publishing photographs and other documents relating to the history of the Federal Republic of Germany and its predecessors, such as the Holy Roman Empire (1495-1806) (Federal Archives, 2011). The internet was identified as an additional publication platform as it can embrace a broad audience and therefore make the task of data publishing a more efficient process. As a consequence, around 100,000 photographs have been published through Wikimedia Commons since 2008 as a result of collabo-

⁴ Cases 2, 3 and 4 are derived from an unpublished students' paper: Felder, Daniel / Studer, David / Vogler, Markus (2013) Gedächtnisinstitutionen im Web2.0-Zeitalter. Crowdsourcing und Open Data. Practical academic project supervised by Beat Estermann, Bern University of Applied Sciences.



ration with Wikimedia. Wikipedia authors then began to link the historic images to articles. The images were licensed under the Creative Commons license “Attribution-ShareAlike 3.0 Germany”. This means that the originator must be mentioned even in modified forms of the pictures and that all derivative works can only be published under the same licence conditions.

The collaborative project also involved the supplementation of metadata in addition to the publication of photographs. For decades, the Federal Archives had used lists of names to identify persons depicted in photographs. The problem was, however, that these lists did not make it possible to clearly identify each individual; furthermore, the data format was not standardised which made it impossible to share the data with other institutions. A Wikipedia volunteer provided a solution by developing a software programme that enabled Wikipedia staff and archive personnel to link the lists of names held by the Federal Archives to Wikipedia articles and the German name authority file (Personennamendatei, PND). In 2008, 59,000 datasets were linked in just four months (Wikimedia Outreach, 2011).

This example illustrates the potential of open data and crowdsourcing. The German Federal Archives would hardly have been able to provide such global availability and links to its photographs with the resources available to it. Publishing the photographs on Wikimedia Commons enabled the archive to better meet one of its core objectives, namely the publication of data. Crowdsourcing the metadata made it possible to provide better links to photos, making them easier to find and enhancing the value of the material. Allowing their reproduction under a free licence meant that others were actually able to make use of the material.

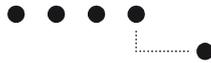
Case example 4: Tropenmuseum Amsterdam (a tropical museum), cooperation with Wikipedia (contextualisation)

The Tropenmuseum’s objective is to inform the people of the Netherlands about the former Dutch colonies overseas. The collection comprises 155,000 photographic documents in addition to 175,000 cultural artefacts and 10,000 other images (status as per 2011). The tropical institute made thousands of high-resolution photographs available in the scope of a project that Wikimedia described as extremely efficient. A number of them were digitally reworked by members of the Wikimedia community and then published. Images of artefacts in the museum and exhibitions as well as films containing background information were also published alongside digitised photographs. The images are located either in the public domain or made available under a ‘free’ licence (Wikimedia Commons, 2012).

Susanne Ton, responsible for multimedia projects at the tropical institute, spoke of the benefits of collaborating with Wikipedia in an interview at the GLAM conference held in Australia in 2009. The Tropenmuseum considers the cost and effort involved in publication as negligible. Much of the data was already available and only had to be transferred. Better accessibility to the publications is named as the main benefit. Instead of looking for a specific institution, which they may not even know, many interested parties carry out context-based searches, for instance for images of representatives of a specific ethnic group. Publishing images through Wikimedia Commons makes the chance that they will be found higher than if the material was published exclusively on the Tropenmuseum’s homepage. She also observes that the online material does not pose a threat to ‘offline’ exhibitions. In fact, she expects an increase in demand as online publication has enabled people around the world to become acquainted with the Tropenmuseum and will visit it should they travel to Amsterdam (Ton, 2009).

Case example 5: Brooklyn Museum – “Click! A Crowd-Curated Exhibition” (co-curation)

“Click! A Crowd-Curated Exhibition” is a good example of an approach to co-curation. Co-curation is aimed at allowing the general public to participate in the selection process for an exhibition or (online) publication. The Brooklyn Museum encouraged artists to submit their photographic work relating to an exhibition on “The Changing Faces Of Brooklyn” together with a brief statement via an online platform. A total of 389 photographs were submitted and then evaluated by visitors to the website who used an online evaluation tool. In all, 3344 evaluators participated in the evaluation process. The 78 best photos were then exhibited by the museum. Interestingly, the opinion of the ‘crowd’ largely matched that of experts (Oomen/Aroyo 2011, p. 144-145).



Case example 6: The Great War Archive, Oxford University (supplementing a collection)

The Great War Archive holds more than 6,500 artefacts that were donated by the general public through a crowdsourcing project. The archive, which initiated the project in 2008, used a special website to reach out to the general public. There was also an opportunity to donate artefacts on ‘open days’ held at museums across the country. It took just under three months to gather the collection. Every artefact hails from the time of the First World War or belonged to somebody who used it during the conflict. Once the three-month period had expired, the archive contacted the donators and offered to publish images of the artefacts on a Flickr page set up specifically for the purpose (The Great War Archive, 2012).

Europeana 1914-1918 is a similar crowdsourcing project. Visitors to the website are encouraged to submit photos, letters, postcards, souvenirs and other artefacts dating back to the time of the First World War so they can be made available to the general public. It is also possible to submit anecdotes and stories in addition to artefacts (Europeana, 2012).

Case example 7: E-Books on demand (crowdfunding)

“Photocopying an old publication is arduous, expensive and uses up a lot of paper. If it is necessary to reproduce it, it is better to scan it. This makes it possible to read the book on a computer and to produce electronic copies. This is the principle followed by the European project ‘E-Books on demand’ (EOD) in which the Central Library in Zurich is participating. It does not lend out books that are more than 100 years old. The EOD service, however, makes it possible to order electronic copies of copyright-free works printed between 1500 and 1900. The service makes the holdings visible to a broad public and available to users located far away from the library. [...]

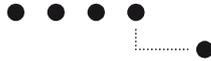
The user simply has to click on the EOD button in the library’s catalogue to purchase a digitised copy. The fee consists of a basic charge of 10 Swiss francs per book and 20 centimes per page. The copy is delivered by email 7 - 14 days later as a high-quality scan. Other interested parties and the library itself benefit from the investment made by an individual. Digitised works are entered into the central library’s catalogue and made available worldwide free of charge.

For this reason, upon a request museum staff first check to see if a digitised copy of a book is already available from a library participating in the project. Thirty libraries in twelve European countries are currently participating, including the National Library of Switzerland and the Library Am Guisanplatz in Bern. [...] Character recognition and payment by credit card are processed by the University and Regional Library of Tyrol in Innsbruck.” (NZZ, 4 Jan. 2013, *translated from German*)

2.4 ‘Free’ licensing and open data

Some of the cases described above reveal an additional trend alongside the use of crowdsourcing: that of making material available under ‘free’ licences that permit the content to be edited and disseminated at no cost. It is only thanks to these licences that the collective authorship approach of large online communities, such as Wikipedia, is actually possible. Often, so-called ‘copyleft’ licences (as opposed to copyright) are used; their provisions stipulate that derivatives of the original published work must also be made available under the same ‘free’ licence. This ensures that the derivatives can be used by the original authors and remain in the pool of ‘free’ works.

Whilst ‘free’ licences have already been a reality for many decades now – initially in connection with the free/open source software movement and later with the open access movement in the field of academic publications, in the field of music (where free licences make it easier to sample music) and projects such as Wikipedia that are based on a large collective authorship – the open data movement is much younger. Initial discussions on ‘open’ data outside academic circles, where the first appearance of the concept can be traced back to the 1950ies, began around 2006. It was, however, most likely the Open Government initiative of the Obama administration in the year 2009 that is responsible for making the topic a matter of international interest. US president Obama launched



various projects in the scope of an ambitious directive to push through a new strategy of transparency, participation and cooperation between politicians, authorities, private enterprises and citizens. Every public authority was required to set up a website within 60 days to provide information to the country's citizens. This open government strategy has inspired similar efforts around the globe: more than 50 countries have meanwhile joined the Open Government Partnership⁵ and have committed to uphold the principle of free accessibility to public sector data.

The concept of "open" data is relatively simple: data is considered open if it can be accessed, processed and disseminated free of charge by anyone for any purpose (see the definition of open data in the text box). The term 'data' covers various types of data: reports, maps, satellite images, photos and paintings, geographic and environmental data, data from scientific surveys and research results such as genomes, medical data and scientific formulae (Bundesamt für politische Bildung 2011). Machine readability is an important prerequisite for enabling data to be used efficiently. Data needs to be structured for this purpose. Structured data is characterised by a sequence of predefined elements with an identical design. A simple address database in which every entry consists of the same repeated elements serves as an example: surname, first name, address, telephone number, etc.

Whilst the concept of "open" data is not limited to data in the public sector, it is this type of information that is the initial concern of the Open Government Data movement. Its demand that information produced by public authorities or financed by taxpayers' money be made available for general use is hardly likely to meet with political resistance as long as compliance with applicable data protection and confidentiality regulations is ensured and financing is available for producing the data. The demand for Open Government Data can be seen as a logical extension of the public access policies that many countries adopted during the first decade of this century.

Definition of Open Government Data

The Sunlight Foundation has formulated 10 formal principles⁶ that have become established internationally as a guideline for the publication of public sector data in the sense of a "best effort"⁷ (cited from Golliez et al. 2012, Open Government Data Studie Schweiz):

Completeness: Official data should be published as comprehensively as possible. This provides for transparency to the extent permissible by data protection regulations with regard to data generated and stored by a public authority in a specific area.

Primary sources: Open public sector data should be published directly from its original sources and supplemented by information on how the data is generated and maintained. This enables third parties to understand where the data comes from and how it was created.

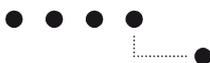
Temporal proximity: Official data should be made publicly accessible as soon as possible after its creation. Data with a usefulness that is time-dependent should be published as a priority. Updates in real time enhance the utility of the data for the general public.

Easy physical and electronic access: Dataset accessibility should be kept as simple as possible with regard to infrastructure and electronic access. It should not be necessary to visit specific premises or fill in special forms or use specific browser-based technology (e.g. Flash, Javascript, cookies, etc.) to gain access to the data. Data collections should be easy to find and downloadable as a unit (so-called 'bulk' access). The use of application programming interfaces (APIs) further increases the technical usefulness of data.

⁵ See <http://www.opengovpartnership.org/>.

⁶ See <http://sunlightfoundation.com/policy/documents/ten-open-data-principles/>

⁷ These principles represent a guideline for the publication of Open Government Data. It is unlikely that all Open Government Data will be able to comply with these right from the start.



Machine readability: Making data machine-readable enables its inclusion in software applications. The use of freely available and widely used data formats and provision of support information on the structure and use of the formats will increase the usefulness of the data.

Freedom from discrimination: Access to data should not be subject to person-related restrictions (e.g. membership of a specific organisation) or time restrictions and should not require the provision of identity details (via registration) or any form of justification for access. The compulsory use of a specific software application to access data is also a form of discrimination.

Use of common standards: The use of standard formats that can be read and processed by applications other than proprietary software makes it easier for a large number of users to access the data. The use of common standards makes it possible to access using various programmes without having to pay licence fees to individual software manufacturers.

Licensing: Restrictive licences represent an obstacle to the usage of data. Public sector data is a product of public authorities and should therefore be marked as public property that can be used without restriction.

Durability: Open public sector data should be permanently locatable and available. Changes, updates and deletions should be comprehensibly documented. Data archiving must be assured.

Cost of usage: The provision of public sector data against payment is an impediment to the use of this data. A large proportion of public sector data is created independently from any subsequent fee-based usage. Charging fees restricts the group of users and impedes the use of data for commercial purposes.

Open data in Europe

The European Commission has made the subject of Open Government Data a priority in its Digital Agenda. In her speech to the OpenForum Summit Europe in September 2011, Neelie Kroes, Vice President of the European Commission and responsible for the implementation of the Digital Agenda, announced the launch of a data portal for the data of the EU Commission and, for 2013, the possible launch of a European data portal⁸.

Open data in Switzerland

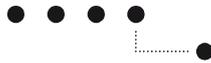
The provision of Open Government Data has been a prioritised aim of the Swiss e-government strategy since the autumn of 2012:

“Data created as a product of administrative activity can be of benefit to the general public and the economy. It represents significant potential for innovation and generates additional added value through its further use and processing by private business and can create greater transparency with regard to governmental and administrative activities. Public authorities shall ensure that all person-unrelated data created as a product of administrative activity shall be made accessible and usable in accordance with the principles of the Open Knowledge Foundation (<http://okfn.org>) as far as is possible.”

(E-Government Schweiz 2012, p. 10, translated from German)

The Federal Council's Strategy for an Information Society in Switzerland has identified the promotion of digital culture and the making of the country's cultural heritage accessible via the internet as targets within the cultural sector:

⁸ Neelie Kroes: “Opening up Europe: from Common Standards to Open Data”. http://europa.eu/rapid/press-release_SPEECH-11-596_en.htm
A beta version of the pan-European open data portal is located at: <http://publicdata.eu/>



“ICT make possible new forms of artistic expression and distribution. They contribute towards promoting cultural and linguistic diversity and facilitate the access to culture. Digital culture is as much a part of our national heritage as any other cultural material. The use of ICT to collect and publish our cultural heritage should be promoted to ensure that it is made readily and comprehensively available to the general public.”

(Schweizerische Eidgenossenschaft 2012, p. 13, translated from German)

In the strategy, the following focal points are defined:

“The federal government supports the development and production of digital culture, its general availability and educational measures to enable its use and application. The main focus will be on networking and the creation of partnerships among the producers of cultural artefacts, business, educational institutions and research.

The federal government will promote the use of ICT in connection with the collection and presentation of the cultural heritage of various societal groups.

The preparation of electronic indexes of the holdings of libraries, archives, museums and collections will be continued with the objective of making them accessible to all via the internet over the long term.”

(Schweizerische Eidgenossenschaft 2012, p. 13, translated from German)

2.5 Linked data / semantic web

‘Linked data’ often relies on data that is made available under a ‘free’ licence – as do many crowdsourcing approaches – as it guarantees legal certainty with regard to the secondary usage of data without any need for extensive negotiations of usage rights. As the term ‘linked data’ implies, the concept refers to creating links between different datasets containing structured data that relate to one another in some way. They can then be used as a basis for establishing new relationships between the objects of various datasets (the box below contains a simple example).

Linked open data – a simple example

Dataset 1 provides information on actors and the productions they have been involved in: Name of the actor, name of the play, place of production, year of production.

Dataset 2 provides information about stage plays and the time of their creation: Name of the play, playwright, date of creation.

Connecting datasets 1 and 2 makes it possible to establish whether a specific actor has been mainly involved in plays that are of more modern or older date. It could also be used to show whether his preferred involvement in older or more modern productions changed over the course of his life, etc. This simple example illustrates how easy it is to imagine other datasets and possible links between them – there are virtually no limitations to what is conceivable in this regard in a world of ‘open’ data. This is because open data is not bound to a specific institution or specific applications that determine the data analysis and recombination options as was previously the case, but is made ‘freely’ available in the form of ‘open’ datasets on the internet and can be linked to other datasets, analysed and visualised by third parties as they see fit.

The so-called ‘semantic web’ is the result of the creation of links between different datasets. In contrast to the conventional internet (Web 1.0), it is not documents that are linked by means of hyperlinks but links are created between individual datasets by defining their relationship with each other. The basis of the semantic web is therefore a large pool of data available online in a standardised format together with a definition of the



relationships between the datasets and their relationship to actual artefacts. In the case of cultural heritage items, the idea is to provide metadata alongside a digital image of the item and to link it to other datasets. Ideally, these links will be based on standardised data, i.e. generally accepted unique identifiers such as those contained in the Integrated Authority File (GND) used by the German National Library. Tim Berners-Lee developed a five-star model to illustrate the route from open data to linked open data; in the following example it has been extended to include the perspectives of users and data holders (see table below).

The route from open data to linked open data – the five-star model

Stage	Description	The user can...	The data holder ...
★	Web-based data (regardless of its format) available under a free licence	view the data print the data save the data to disk manually convert the data to other formats	can readily publish data
★★	Data in a structured format (e.g. Excel)	do everything one can do with one-star data process the data with proprietary software convert the data to other formats	can readily publish data
★★★	Data in a structured, non-proprietary format (e.g. CSV instead of Excel)	do everything one can do with 2-star data process the data without the need to use proprietary software	may require converters or connectors to transform and provide data from its proprietary format
★★★★	Use of unique URLs to enable data to be linked	do everything one can do with 3-star data create links to other resources apply bookmarks to datasets	enhances data and datasets assigns URIs to data receives data via access points and can optimise the availability of data accordingly
★★★★★	Creating links between one's own data and other data to create context	do everything one can do with 4-star data use data as hypertext by linking one dataset to another discover and use the data schema	requires resources to create links makes data effectively locatable for search engines increases the use value of the data

Source: Kaltenböck & Thurner 2011, chap. 6.3 (*translated from German*)

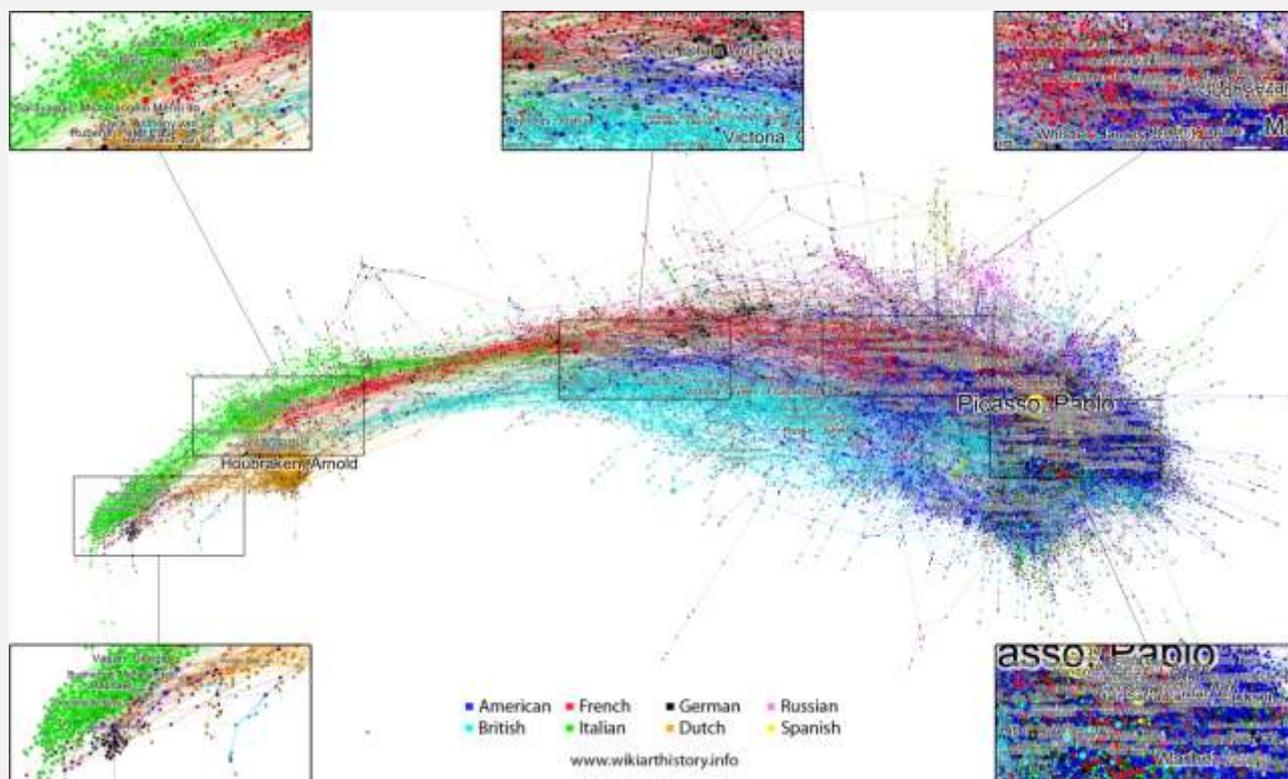


The real value of linked data and the semantic web becomes apparent when a new service, a new form of visualisation or new information is made available by creating new links between existing datasets (see case examples 8 and 9). As shown by the two examples, linked data can also be used to integrate and visualise data derived from Web 2.0 applications or crowdsourcing projects, such as Wikipedia or folksonomies.

Case example 8: wikiarthistory.info (visualisation of art history)

Wikiarthistory is an experimental form of visualisation of art history developed by computer scientists at the Technical University of Vienna. It uses data from the authority file “Getty Union List of Artist Names (ULAN)” and DBpedia, which contains structured data extracted from Wikipedia. The two datasets are linked using another authority file: the Virtual Internet Authority File (VIAF).

The visualisation focuses on the relationships between people of relevance to art history such as artists, patrons, collectors, clients, politicians, and monarchs. It employs an analysis of the links between individual Wikipedia articles and those between individual entries in the ULAN. The result of the visualisation process shows the chronological progression of art history from the Renaissance (left) to today (right; see image below). By displaying the dots in different colours according to the person’s nationality it became possible to visualize patterns of intercultural interaction as seen from the viewpoint of the English-speaking Wikipedia community. The left side is dominated by the Italians (green), whose cluster stretches from the Renaissance to the Baroque period and fizzles out towards the end of the 17th century. A small cluster at the bottom left represents the German Renaissance centred around Albrecht Dürer (black). To the right of the German cluster is a cluster of Dutch artists from the Renaissance and Baroque periods (orange). This is followed by two elongated clusters of British (light blue) and French (red) artists that stretch over the same time period but have few connections between them until they are merged by American artists (blue) in a large melting pot from the 19th century onwards (Goldfarb et al. 2012).



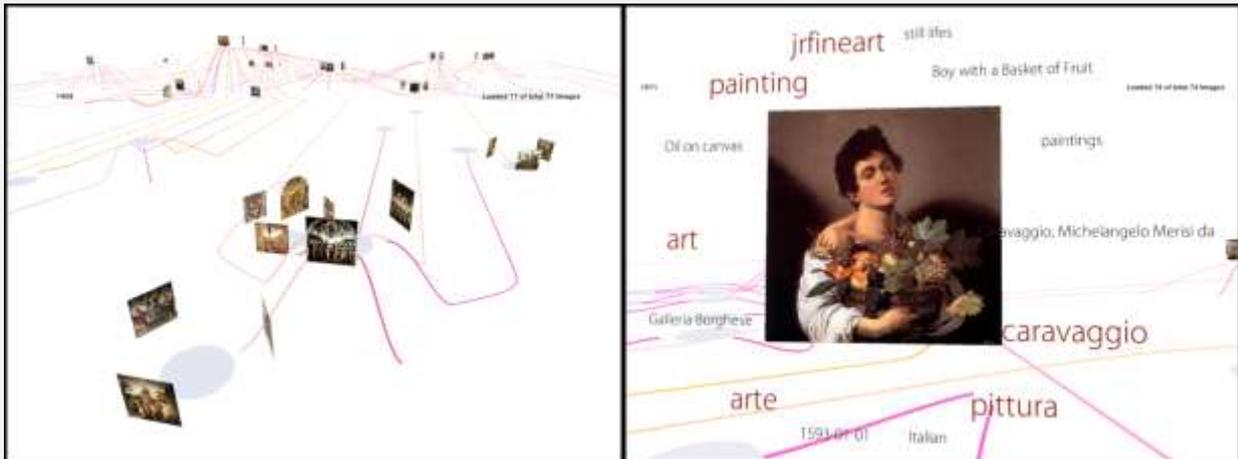
Visualisation of Wikilinks between persons of relevance to the history of art with a maximum age gap of 75 years.

Source: <http://www.wikiarthistory.info>.



Case example 9: The virtual 3D Social Experience Museum

The virtual 3D Museum, which was developed by the same group of researchers at the Technical University of Vienna, links data from DBpedia, images of paintings and metadata from the Web Gallery of Art (WGA) as well as culture-related vocabularies from the Getty Foundation and map them in a three-dimensional landscape. Artists and their works are displayed in the resulting network as dots, while relationships such as parent-child or teacher-student are visualised as lines. Additionally, the network has a chronological structure that makes it suitable for researching art-historic correlations in a new manner (Goldfarb et al. 2011).



"The Virtual 3D Social Experience Museum". Source: http://vsem.ec.tuwien.ac.at/?page_id=23



3 Results of the pilot survey

3.1 Research questions

The following three questions were at the heart of the pilot survey:

1. Where do Swiss cultural heritage institutions stand today with regard to open data and crowdsourcing?

In addition to concrete activities in the fields of crowdsourcing and open data, we were also interested in the factors that could facilitate their participation in these sectors: How good is the quality and interoperability of metadata? Is there any collaboration across organisations? To what extent is material already being made available online? Do institutions have experience of working with volunteers? Do any employees participate in collaborative projects on the internet in their spare time? We were also interested in finding out whether institutions were more likely to perceive the opportunities or the risks associated with open data and crowdsourcing and whether there were any reservations relating to making content ‘freely’ available. Finally, we wanted to know whether linked data is of interest to the institutions.

2. What opportunities and risks do the institutions see in relation to open data and crowdsourcing?

We asked the institutions to assess various risks and opportunities so we could better understand which enabling and inhibiting factors are the most important from their viewpoint.

3. To what extent would society benefit from open data and crowdsourcing in their opinion?

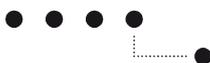
In addition to the benefits to society, we were also interested in determining whom the institutions see as the potential main beneficiaries of crowdsourcing and open data.

3.2 Methodology / research design

The pilot survey was designed as a quantitative investigation using an online questionnaire. The questionnaire itself was designed in a number of stages: The first stage comprised an initial draft drawn up on the basis of literature research, which was presented to various subject experts who contributed various additions and suggestions for improvement. A revised version was then pretested by representatives from ten randomly selected cultural heritage institutions of which six were willing to complete the questionnaire in full and provide feedback by telephone. The questionnaire was revised once again on the basis of the results from the preliminary test.

Invitations to participate in the survey were sent by email to cultural heritage institutions figuring on the lists of A-objects of the national inventory of cultural property (KGS inventory). The list names every institution with holdings of national significance, whereby it should be noted that relevance is determined on the basis of the criteria applied in the context of the protection of cultural property and that other relevance criteria may well apply in other contexts. We limited the survey to institutions located in German-speaking Switzerland for reasons of time and cost. Obvious duplications in the KGS list were eliminated before the invitations were dispatched⁹. Invitations

⁹ Engelberg Abbey (Kloster Engelberg) is a good example of this as it is listed by the KGS as main institution in charge of the Benedictine monastery's collection, the monastery's archives, its library and its collection of sheet music.



to participate in the survey were sent to various addressees at large organisations such as universities, where organisational units could be expected to have a certain degree of independence.

The survey was carried out between 12 November and 14 December 2012. Two reminder emails were sent at intervals of 10 days to help achieve a better return rate. The first reminder also referenced a PDF version of the questionnaire that could be requested if needs be and completed by hand. However, only one institution made use of this option.

3.3 Representativeness and validity of results

A total of 233 invitations to participate in the survey were sent to addresses belonging to 197 organisations. Four addressees could not be contacted three others chose not to participate and provided reasons for their decision. The questionnaire was begun by 132 participants, of whom 99 answered at least two questions and 72 completed the questionnaire. The completion rate is therefore around one third (31% of addressees or 34% of institutions contacted on the basis of the adjusted sample) and represents a respectable result under the given circumstances¹⁰.

The question as to what extent the results of the survey can be applied to Switzerland's cultural heritage institutions as a whole depends mainly on two factors: To what extent is the sample of contacted institutions representative of Switzerland's cultural heritage institutions as a whole? And to what extent does failure to complete the survey bring systematic bias into play?

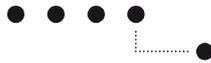
3.3.1 Switzerland's cultural heritage institutions as a whole

Finding an answer to the first question is not straightforward as it is difficult to define Switzerland's cultural heritage institutions as a whole. According to the figures published by umbrella associations, Switzerland has around 750 museums that are members of the Association of Swiss Museums (VMS), around 350 institutions organised in the association Library Information Switzerland (BIS) and 289 archives registered with the Association of Swiss Archivists (VSA). These figures equate to just under 1,400 institutions. There is, however, probably a fair amount of overlapping as some institutions belong to more than one association. A total of around 1,000 independent cultural heritage institutions organised in the three associations as a whole is probably closer to reality. It should however be considered that not every cultural heritage institution is a member of one of the associations: 1101 museums were included in the survey of visitor numbers carried out by the association of Swiss museums (VMS 2012, p.2), which suggests that only around two-thirds of Switzerland's museums are members of the umbrella organisation. The number of archives that are not organised in an umbrella association is probably much higher if one includes smaller municipal archives and the like.

Another source that can be used for the purposes of determining numbers of institutions is the ISIL directory maintained by the Swiss National Library that lists around 1,500 libraries, archives, documentation offices and similar entities. The problem here is that many larger institutions (e.g. universities) have multiple entries, and we may assume that decisions on open data and crowdsourcing are in many cases made not at the level of individual libraries or archives but rather at a higher organisational level.

Furthermore, with regard to open data and crowdsourcing, the question to what extent holdings are unique plays a certain role as it is usually sufficient if a single instance of a specific work is digitised and made available online. The question of the uniqueness of holdings mainly concerns libraries whose primary function is that of a lending library, although it also applies to museums to a certain extent. In contrast, the situation differs with regard to

¹⁰ By way of comparison, a 2011 survey undertaken by the umbrella association of Swiss museums, for instance, achieved a return rate of 52%; the survey addressed every museum in Switzerland and focused on visitor statistics (VMS 2012).



archives, as the vast majority of their holdings are unique items that document the actions of organisational units from various viewpoints (Bundesamt für Kultur 2008, p. 26). The KGS inventory provides a good basis for estimating the quantity of cultural heritage institutions from the viewpoint of the uniqueness of holdings. The A-lists contain 465 collections of national significance, whilst the cantonal B-lists reference 417 collections of regional significance (status as per June 2013). We know from our analyses within the context of the pilot survey that the 287 collections of national significance located in German-speaking Switzerland correspond to just under 200 independent organisations. The number of duplicates in the B-lists is probably lower as the corresponding organisations are usually smaller. If we extrapolate these figures we can assume to be dealing with 600-700 independent organisations that hold collections of national or regional significance.

3.3.2 Bias of the sample with regard to the basis population

Given our selection criteria, the sample includes those institutions that hold collections of national significance from the perspective of cultural heritage protection. It can also be assumed that the majority of large institutions were included in the survey whilst many of the smaller heritage institutions were not taken into consideration. The survey's focus on institutions in German-speaking Switzerland is another source of potential bias. Although we are unaware of any basic differences between language regions with regard to attitudes towards open data and crowdsourcing, one must consider when interpreting the results that many federal and national institutions (often located in Bern) were surveyed, while the international organisations in the Geneva area that also hold collections of national significance were not included.

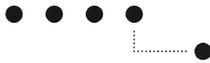
Further potential sources of sample bias can be identified using the surveyed data. Firstly, one can draw a comparison between the institutions that responded to the survey and those that did not, based on the type of institution. Before carrying out the survey, we assigned each institution to one of the following institution types: 'Archive', 'Museum', 'Library', 'Other'¹¹. An analysis of the return rate by category reveals significant differences; 43% of the contacted archives and 34% of libraries completed the survey in comparison to a mere 25% of museums and 20% of the 'others' (mean value 31%). This bias is significant¹².

The surveyed data also allows a comparison between the institutions that completed the survey and those that did not. It is apparent that those institutions that cited 'objects of art' as their characteristic cultural heritage items scored a much lower completion rate of 54% compared to other cultural heritage institutions (79%). Moreover, cultural heritage institutions that cited 'collecting' as one of their core activities had a much higher completion rate of 80% compared to those that do not consider 'collecting' one of their core activities (54%). Organisations that count public authorities among their most important users had a much lower completion rate of 63% in comparison to the others (82%). These three forms of bias are significant. There are no significant differences between the two groups with regard to institution type, number of employees, and legal form.

It is therefore important when interpreting the results of the survey to bear in mind that archives and institutions with 'collecting' as a core activity are over-represented in the survey sample, whilst institutions characterised by works of art or public authorities as main users are under-represented. When extrapolating the results for all cultural heritage institutions in Switzerland, it should be borne in mind that small and very small institutions are probably under-represented in the sample, even though the size of an institution had no influence on whether or not the questionnaire was completed in full.

¹¹ Categorising the institutions in this manner is not unproblematic as various cultural heritage organisations belong to several categories – for instance, the names and email addresses of some institutions refer to different institution types, e.g.: archiv@library.ethz.ch, ETH library, archives and bequests (in this particular case we assigned the institution to the 'archive' category as the word 'archive' occupies the most specific place in the email address). The category of 'other' comprises collections whose names did not allow an allocation to one of the other types (e.g. botanical gardens, Swiss museum of transport, the Federal Art Collection, collections held by foundations, etc.).

¹² In our survey, we used a confidence level of 95% for significance testing; this means that an ascertained correlation is considered to be significant if it can be assumed with a probability of at least 95% that the correlation will apply to the basis population.



3.3.3 Significance of results

Whether a survey returns meaningful results or not depends on the sample size, the variance of observations and the markedness of observed characteristics and correlations. The sample in this survey is rather small. As a consequence, the confidence intervals that indicate the range of values that comprises the true value for the basis population at a certain degree of statistical probability are relatively wide. To illustrate this, we have explicitly stated the confidence intervals for some of the survey's results (see chapters 3.4.9 and 3.5.7); confidence intervals for other results can readily be calculated using the information provided in this report.¹³

3.4 Characteristics of the surveyed cultural heritage institutions (sample description)

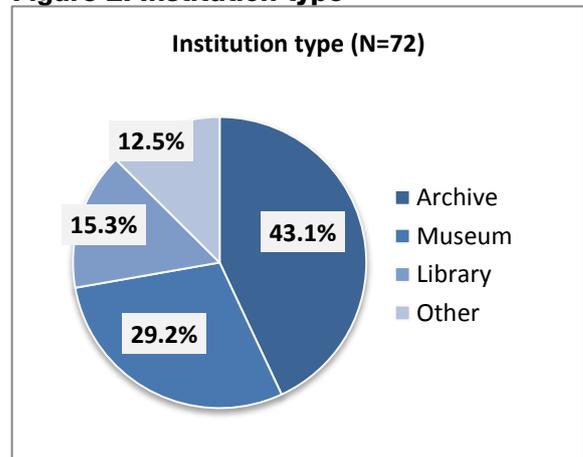
This chapter provides an overview of the characteristics of the surveyed institutions.

3.4.1 Institution type

We used the entries in the KGS inventory and the names contained in the contact details (email addresses) to allocate each cultural heritage institution to one of the following categories: 'Archive', 'Museum', 'Library', 'Other'. Categorising the institutions in this manner is not unproblematic as various cultural heritage organisations can comprise a number of institution types and the names are not always unambiguous¹⁴.

Figure 2 shows that archives form the largest group in the sample with 43%, followed by museums (29%) and libraries (15%). The 'other' institutions form the smallest group (13%). It comprises collections whose names did not allow an allocation to one of the other types (e.g. botanical gardens, collections held by foundations and institutes, etc.).

Figure 2: Institution type



3.4.2 Characteristic types of artefacts

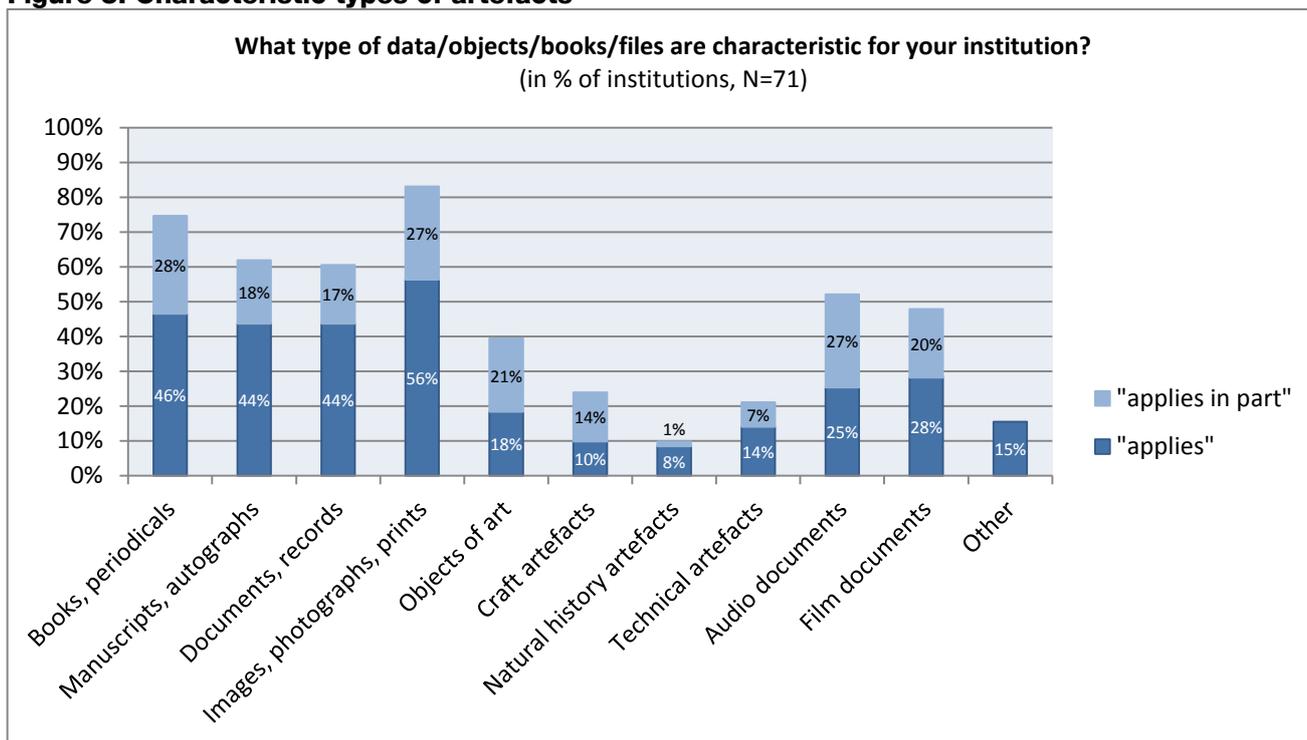
To achieve a more accurate categorisation of the institutions, we asked the respondents what type of cultural heritage artefact is particularly characteristic of their institution. Figure 3 provides an overview of how often individual types of cultural heritage artefact were named. The category 'illustrations, photographs, prints' leads the field; more than half of the surveyed institutions stated that these types of artefact define the special nature of their institution ('applies'). This is followed by 'books, periodicals', 'manuscripts, autographs' and 'documents, records'. Just under half of the surveyed institutions selected these categories. Films and audio documents were cited by just over a quarter of the institutions; less than one fifth cited objects of art and technical artefacts, and less than one tenth cited craft artefacts and natural history artefacts. Some 15% of the institutions cited at least one other type of characteristic cultural heritage item. The most commonly cited (by approximately 5% of institutions in each case) were articles of daily life (fashion, toys, etc.) and digital artefacts (electronic data, media, documents, databases).

¹³ The following online tool can be used to calculate confidence intervals for percentages: <http://www.mccallum-layton.co.uk/stats/ConfidenceIntervalCalcProportions.aspx>

¹⁴ See footnote 11.

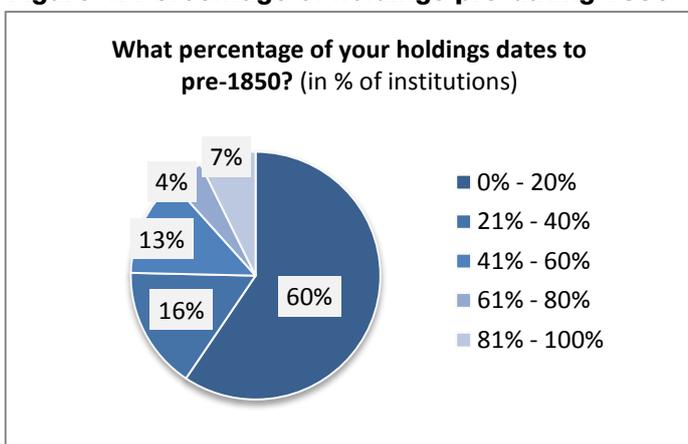


Figure 3: Characteristic types of artefacts



With regard to copyright issues we also wanted to know from the surveyed institutions what percentage of their cultural heritage holdings dates to pre-1850. Figure 4 shows that 60% of institutions hold cultural heritage items that mainly date to a more recent period. Only around 10% of the institutions have holdings where a large majority of artefacts is more than 160 years old and therefore quite certainly out of copyright.

Figure 4: Percentage of holdings pre-dating 1850

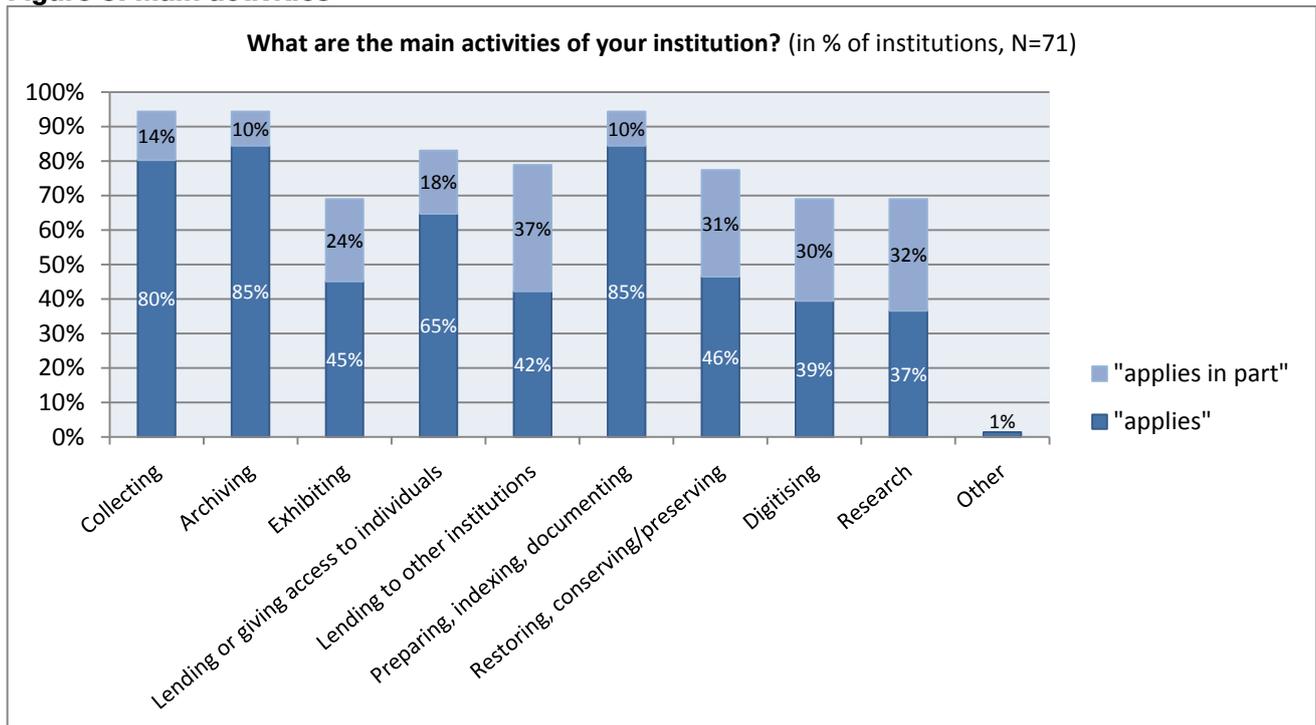


3.4.3 Main activities

We also asked the institutions about their main activities. It is apparent that the surveyed institutions are characterised by a certain degree of homogeneity with regard to their main activities. At least 69% of institutions stated that all of the possible activities defined in the survey counted at least in part among their main activities. Activities such as 'collecting', 'archiving' or 'preparing, indexing, documenting' were identified by more than 90% as counting at least in part among their main activities (see Figure 5). Activities such as 'doing research', 'digitising', 'lending to other institutions', 'exhibiting' and 'restoring, conserving/preserving' are particularly distinguishing features. Each was clearly named as a main activity by just less than one half of the surveyed institutions.



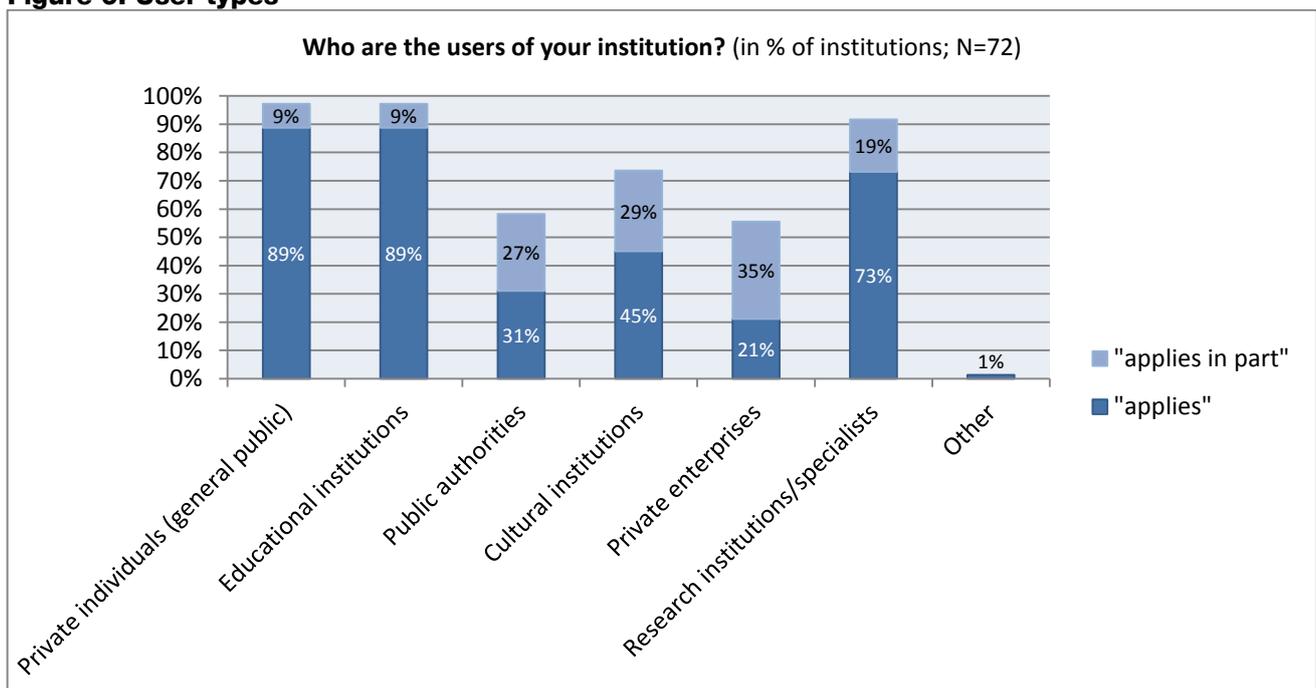
Figure 5: Main activities



3.4.4 Users

We wanted to know who the institutions count among their users. The most frequently named user groups are private individuals (general public) and educational institutions; around 90% of institutions count them explicitly among their users (see Figure 6). Around three-quarters of the surveyed institutions also consider research institutions and specialists as their users. The remaining three user groups were named significantly less frequently: cultural institutions (45%), public authorities (31%) and private enterprises (21%).

Figure 6: User types

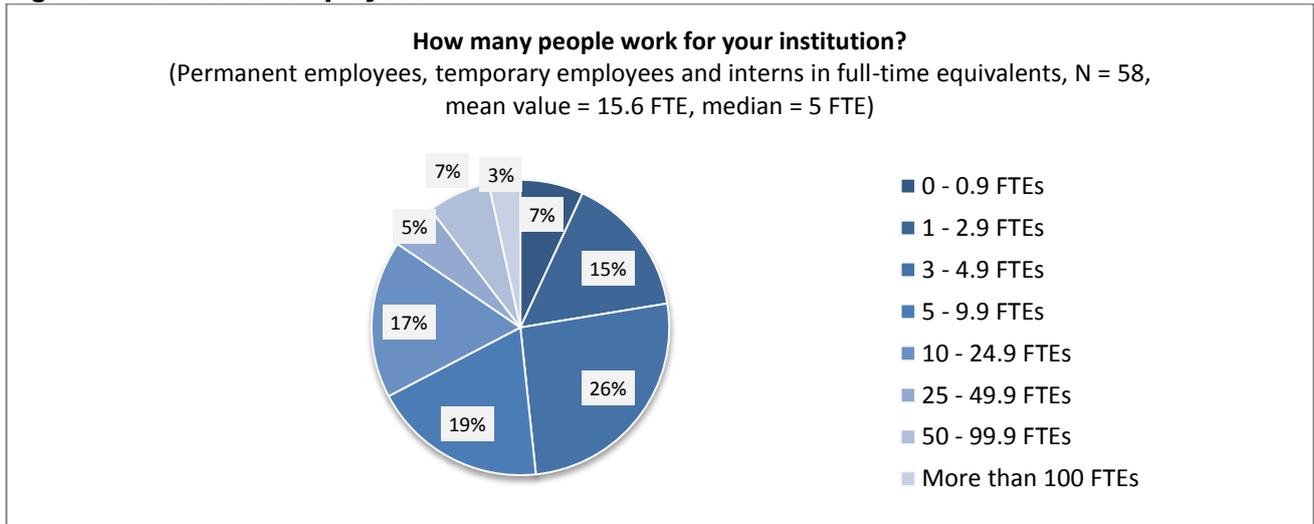




3.4.5 Size of the institution:

We asked for details relating to the number of employees and full-time equivalents (FTEs) to get an idea of the size of the surveyed institutions. The result is a relatively good mix: around one fifth of the institutions are small or tiny institutions with less than 3 FTEs. One quarter have 3 - 5 FTEs. The medium-sized institutions (5 - 25 FTEs) make up a good third of the sample, whilst the larger institutions with more than 25 FTEs account for 15% of the sample. The mean value is 15.6 FTEs, the median is 5 FTEs.

Figure 7: Number of employees

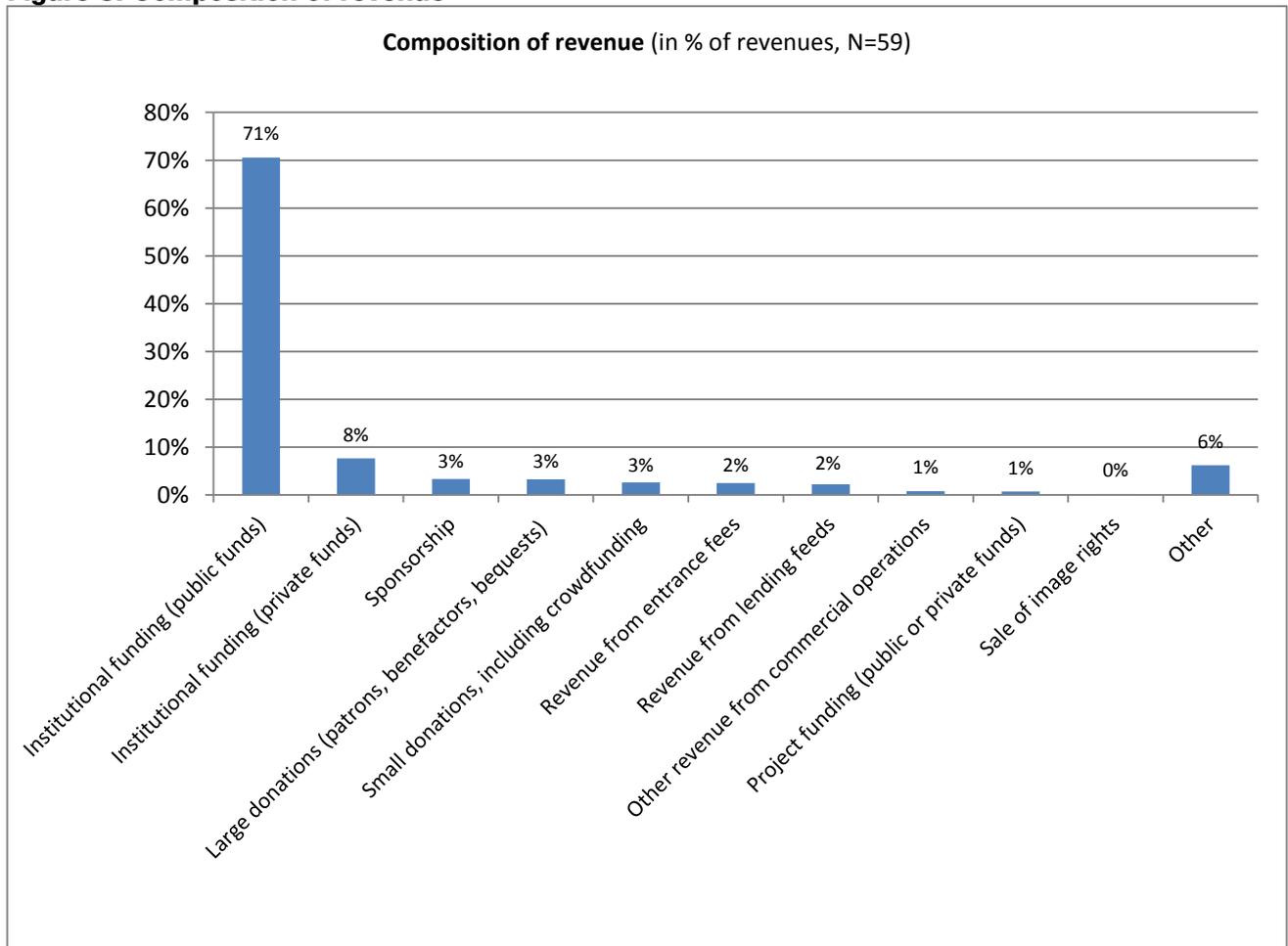


3.4.6 Sources of revenue

Figure 8 shows the composition of revenue obtained by the surveyed institutions. It is apparent that government grants in the form of institutional funding constitute two-thirds of the overall revenues of the surveyed cultural heritage institutions. The second main source of revenue is private institutional funding and accounts for 8%. Donations (large and small) and sponsorship together account for around 7% of revenue, whilst income from commercial activities (entrance fees, lending fees, sale of image rights, etc.) makes up around 6% of total revenue. Finally, 2% of the revenue of surveyed institutions is in the form of project funding. Revenue from commercial operations is of particular interest in view of a potential loss of income as a consequence of making high-quality images and data available free of charge on the internet. This is, however, rather low at 6%, whereby the 'revenue from entrance fees' and 'other revenue from operations' amount to 5% of total revenue and account for the lion's share within this category. 'Revenue from lending fees' (1%) and revenue from the sale of image rights (less than 0.5%) seem negligible by comparison. These figures should, however, be interpreted with caution in view of the fact that the category 'other' accounts for the third-largest share (6%) of overall revenue. It is possible that respondents simply grouped the more insignificant sources of revenue under 'other' due to a lack of knowledge of details.

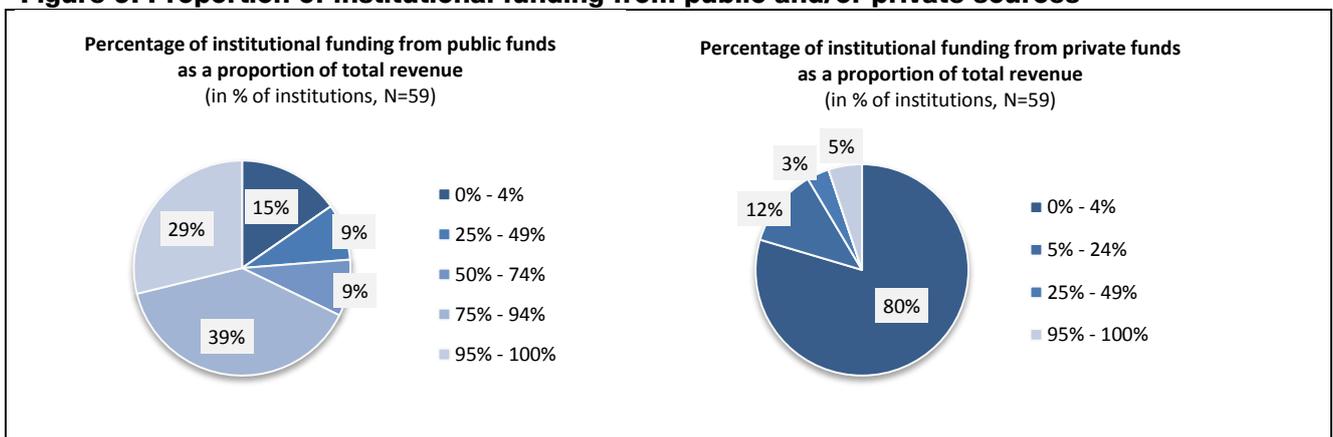


Figure 8: Composition of revenue



This general overview should not obscure the fact that the financing models of cultural heritage institutions are very heterogeneous. Hence, 15% of the institutions receive hardly any public funding, while just under 30% are financed virtually exclusively from this source. On the other hand, institutional funding from private sources hardly plays a role for 80% of cultural heritage institutions, while 5% of those surveyed are financed virtually exclusively through private institutional funding (see Figure 9). Around 4% of the institutions specified that large donations (patrons, benefactors, bequests) are the source of more than three-quarters of their revenue, while this type of financing plays virtually no role at all for 90% of institutions.

Figure 9: Proportion of institutional funding from public and/or private sources

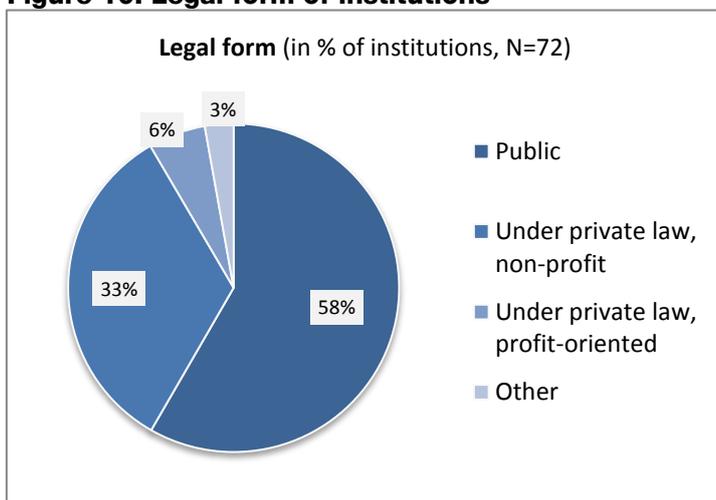




3.4.7 Legal form

In addition to sources of income we were also interested in identifying the legal form of the surveyed institutions. This aspect correlates with the results of the questions on financing: just under 60% of the surveyed institutions are public organisations, whilst one-third consists of private, non-profit organisations. Only 6% of the institutions are profit-oriented or belong to a profit-oriented organisation (e.g. corporate archives) (see Figure 10).

Figure 10: Legal form of institutions



3.4.8 Summary

In summary it can be said that the majority of the cultural heritage institutions we surveyed are public, non-profit organisations financed mainly through public funds and counting private individuals and representatives from the fields of education and research among their most important users. Most of them will therefore be concerned by the demands of civic organisations to make the results of activities carried out by public authorities or financed from taxpayers' money available to the general public (e.g. via Open Data, Open Access, etc.). It should, however, be noted that the financing models of the surveyed institutions differ significantly and that around one-third of the institutions are not predominantly financed from public funds.

The sample is representative with regard to the size of the surveyed institutions. In addition to numerous small and tiny institutions (around one-half of those surveyed have fewer than 5 full-time posts) it includes a number of larger institutions (10% of those surveyed have more than 50 full-time posts).

Archives account for more than 40% of the sample and are, as already mentioned in chapter 2.3, over-represented in comparison to Switzerland's cultural heritage institutions as a whole. Two-dimensional artefacts are the most frequently named characteristic cultural heritage items, whilst audio-visual artefacts take second place and three-dimensional artefacts are more rarely named. Three-quarters of the surveyed institutions hold artefacts that are less than 150 years old. It is likely that these institutions will encounter copyright, data protection and confidentiality issues in connection with open data and crowdsourcing. Conversely, one-quarter of institutions hold significant proportions of cultural heritage items that will hardly be affected by these problems. The surveyed institutions are relatively homogeneous with regard to their main activities.



3.4.9 General applicability of results

As stated in chapter 2.3, caution must be exercised when extrapolating these results for Switzerland's cultural heritage institutions as a whole due to the relatively small sample size. The following statements are correct with a probability of 95% (if one disregards a certain bias in the sample):

- **Legal form:** Between 47% and 69% of Swiss cultural heritage institutions are organisations under public law; 22% to 44% are non-profit organisations under private law, while up to 11% are profit-oriented organisations under private law.
- **Financing:** Switzerland's cultural heritage institutions receive 59% to 83% of their revenue from public funds (project funding excluded). The financing models are, however, very heterogeneous: 19% to 39% of the institutions are financed virtually exclusively from public funds, while 7% to 23% receive hardly any public funding.
- **Users:** Between 82% and 96% of Switzerland's cultural heritage institutions clearly count private individuals (the general public) and educational institutions among their users, 63% to 83% count representatives from the field of research among their users, whilst the other user groups are less significant: cultural institutions (34 - 56% of the surveyed institutions), public authorities (20 - 42%), businesses (12 - 30%).
- **Size:** A Swiss cultural heritage institution has on average (median) between 3.7 and 7.6 full-time employees, 36% to 60% of the institutions have fewer than 5 full-time posts; 30% to 52% have 5 to 49 full-time posts, while the proportion of large institutions with at least 50 full-time employees lies between 3% and 17%.
- **Characteristic cultural heritage items:** Cultural heritage institutions most frequently cite 'images, photographs, prints' as their most characteristic cultural heritage holding; this is the case with 45% to 69% of the institutions. At the opposite end of the scale, natural history artefacts are characteristic for 2% to 16% of the institutions.
- **Age of holdings:** Between 48% and 72% of Swiss cultural heritage institutions hold hardly any artefacts that date to before 1850. Conversely, the majority of artefacts held by 4% to 18% of the institutions pre-date 1850.



3.5 Where do Swiss heritage institutions stand today with regard to open data and crowdsourcing?

We decided to approach this aspect with reference to the innovation-diffusion process described in 1957 by George Beal and Joe Bohlen in their seminal study and subsequently adopted by Everett Rogers (1962). According to this study, there is a five-stage process that leads to the adoption of an innovative technique by an individual entrepreneur (Beal/Bohlen 1957, p. 2):

Awareness stage: In this phase the subject becomes aware of a new idea or innovative practice but lacks detailed knowledge about it.

Interest stage: In this phase, subjects are interested in obtaining further information on the new idea or practice. They want to know what it is about, how it works and what potential it offers.

Evaluation stage: In this phase, subjects make a mental trial of the idea and use the information acquired in the preceding phases to determine how the new idea or practice could be implemented in their own specific situation. They examine if and how they could use the new idea or practice and what effect it would have on their situation.

Trial stage: Subjects that are positively inclined towards the idea following the evaluation phase now start to apply the new idea or technique on a small scale to gain initial concrete experience.

Adoption stage: The new idea or practice can be adopted if the positive assessment is confirmed during the experimentation phase. It then becomes a normal component of daily activities.

We have based our analysis of the survey's results on the premise that we are dealing with a number of innovations (see chapter 2) that will be adopted by a growing number of cultural heritage institutions over time, initially by the so-called 'innovators', then by 'early adopters' and finally by the majority of institutions. Whether a specific innovation will prevail or not must remain an open issue; our pilot survey is unable to provide an answer to this question.

The pilot survey focuses on the following five innovations: Digitisation; cooperation in networks in combination with the electronic exchange of metadata; open data; crowdsourcing; and linked data / semantic web. As discussed in chapter 2, these innovations are in part complementary (e.g. open data and crowdsourcing) or are interdependent (linked data, for example, is only possible if metadata is available in electronic form).

We have identified indicators for most of the five innovations that provide evidence of whether an institution is in the interest stage or trial stage with regard to a specific innovation, as well as indicators that point towards implementation (trial and adoption stages). The questionnaire also contains a number of questions that enabled us to determine the extent to which prerequisites or factors that would facilitate the implementation of an innovation are in place at each institution.

3.5.1 Digitisation

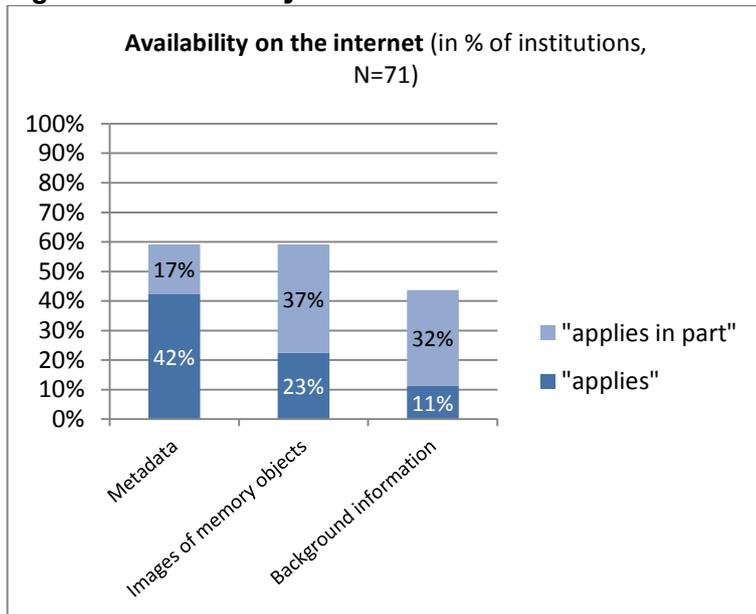
The digitisation of metadata and images of cultural heritage items is an important prerequisite for internet-based cooperation, open data and many forms of crowdsourcing.

Implementation: To establish the status of implementation, we asked the institutions whether they make data relating to their cultural heritage items available on the internet. Around 60% of the surveyed institutions make both metadata and images of at least some of their holdings available on the internet. Another 43% also stated that they make at least some background information available on the internet. A distinct graduation becomes apparent if one considers only those institutions who responded with 'applies'. Most commonly made available on the internet is metadata (42%), followed by digitised copies or photographs of the cultural heritage items (23%)



and finally background information (11%). Around 40% of the institutions seem as yet oblivious to the digitisation wave or they have not arrived at the point where it would have an impact on their online activities (see Figure 11).

Figure 11: Availability on the internet

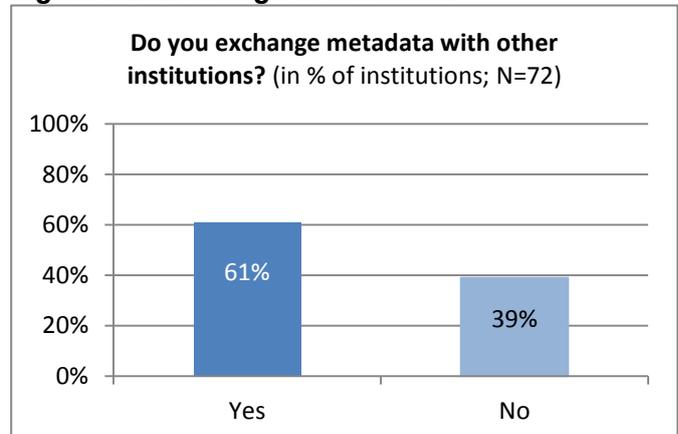


3.5.2 Cooperation in networks

From a historical viewpoint, the widespread trend towards digitisation was followed by increased cooperation in networks aimed at coordinating digitisation and providing users with a single point of access.

Implementation: We wanted to get an idea of the current level of implementation and asked the institutions to tell us whether they exchange metadata and participate in collaborative projects that necessitate an exchange of metadata. Figure 12 shows that around 60% of the institutions exchange metadata with other institutions, whilst 40% do not. This corresponds roughly with the number of institutions that make metadata at least partially available on the internet. Another 30% of the institutions exchange data within the scope of bilateral cooperation projects and 43% do so as part of multi-lateral cooperation projects (see Figure 13; multiple answers permissible).

Figure 12: Exchange of metadata



Perception: Figure 13 also provides information on the relevance of metadata exchange. Just under 30% of institutions consider it relevant to achieving their main objectives. This is partially the case for another 17%. It is also interesting to see that the surveyed institutions consider the exchange of metadata as unsuitable for generating revenue.

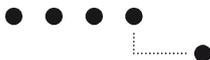
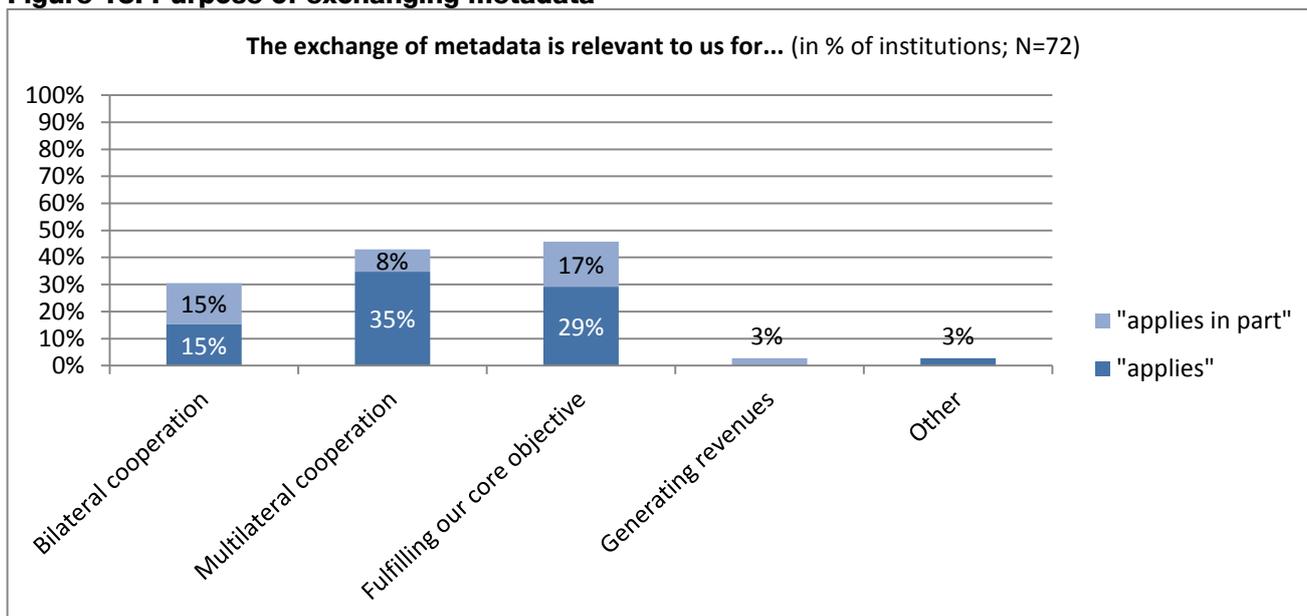
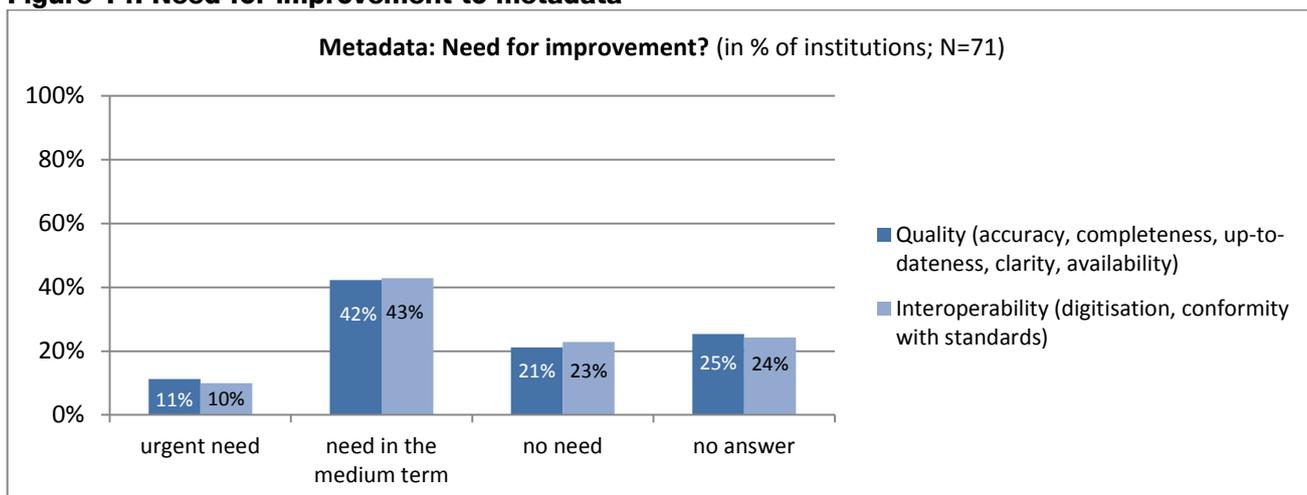


Figure 13: Purpose of exchanging metadata¹⁵



Prerequisites: Metadata must comply with certain quality standards if they are to be suitable for publication and exchange with other institutions. We therefore asked the cultural heritage institutions if and where they perceive a need to improve their metadata. Around 10% of the surveyed institutions perceive an urgent need for improvement, whilst a good 40% identified a need in the medium term. Only one-fifth of the surveyed institutions did not perceive any need for improvement. The proportion of participants who felt unable to answer this question is relatively high at 25%. It is unclear whether these responses reflect the fact that institutions had not addressed the quality of their metadata in recent times or simply because respondents were unable to provide any information. Interestingly, the rates of responses with regard to the improvement of quality and those relating to improving interoperability are roughly the same (see Figure 14). This could indicate that the digitisation and exchange of metadata are often linked to increased demands in terms of data quality.

Figure 14: Need for improvement to metadata

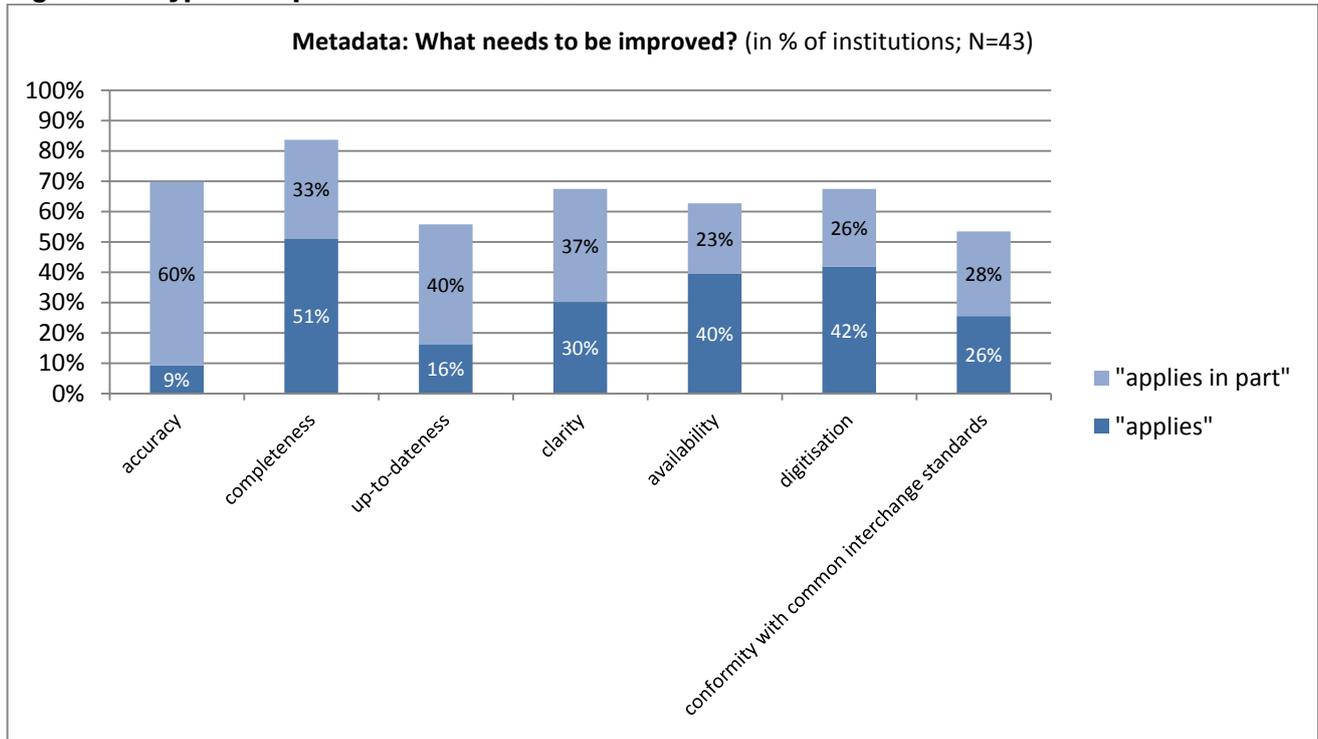


¹⁵ Percentages relate to the basis population of surveyed institutions. However, only those institutions who stated that they exchange metadata with other institutions were asked to respond to this question.



We wanted to know from those institutions that indicated a need for improvement exactly where they see deficits. The answers show that the greatest deficit is perceived in relation to the completeness of metadata. At least partial deficiencies were reported in more than 80% of cases. At the same time, all the other aspects listed in the survey, such as accuracy, up-to-dateness, clarity, availability, digitisation and conformity with common exchange standards, were also cited by more than half of the respondents as problem areas – albeit with relatively large differences as to their perception as general problem areas ('applies') or only partial problem areas (see Figure 15).

Figure 15: Type of improvement to metadata



We also asked the institutions about the type of metadata formats they use (see Figure 16). Interpreting the results is not so straightforward due to the different purpose of each format. Furthermore, a relatively large number of respondents were unable to answer the question. Of the standards most commonly used, in particular by archives, ISAD(G) was named more frequently by the surveyed institutions than EAD (25% versus 7%). The fact that ISAD(G) as an archive metadata format was most frequently named could have to do with the over-representation of archives in our sample in comparison to other types of institution. The metadata formats most frequently used by libraries are MARC21 in first place (17%) followed by AACR (13%) and Dublin Core (11%). Only 3% of the institutions used MAB; according to the national library of Germany,¹⁶ this is a format that was replaced by MARC21 after its development was frozen in 2006.

The German name authority file (PND) is an authority file for persons that served mainly to index literature in libraries. It was merged with other authority files in 2012 to create the Integrated Authority File (GND). The GND is also linked to other national authority files in the Virtual International Authority File (VIAF) project by means of a concordance file. The GND is used together with the MARC21 standard. It is unclear how to interpret the fact that only one institution mentioned PND among the formats it uses.

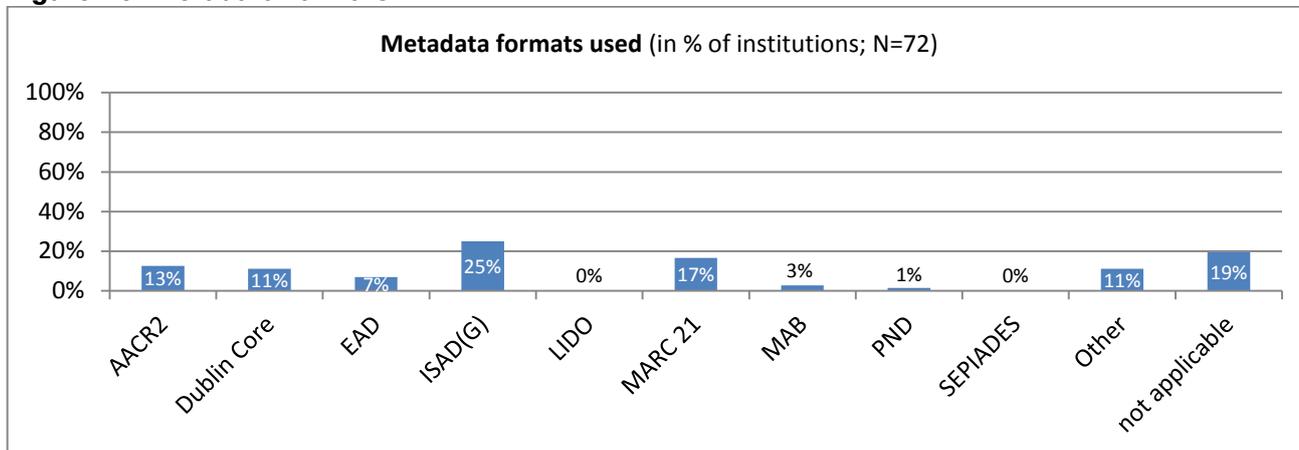
LIDO, a metadata exchange format used mainly by museums, is not commonly used by the surveyed institutions. The same applies to SEPIADES, a metadata standard for cataloguing photographic collections, whereby this

¹⁶ http://www.dnb.de/DE/Standardisierung/Formate/MAB/mab_node.html



format is based on the directory standard ISAD(G). PREMIS, a metadata standard for long-term archiving, and METS were each cited once in the category 'other'. A number of institutions also cited application-specific standards.

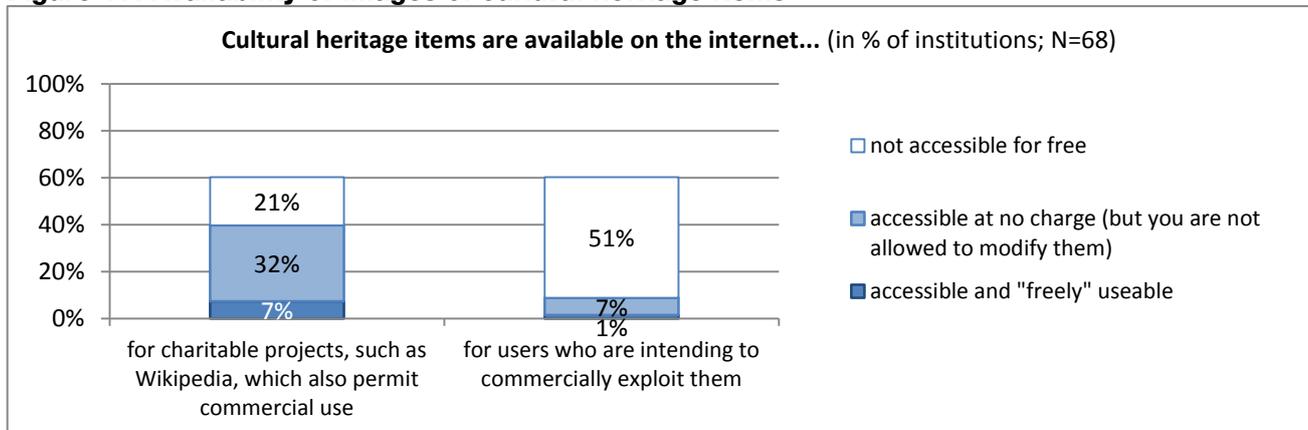
Figure 16: Metadata formats¹⁷



3.5.3 Open data / 'free' licensing

Implementation: To find out how many institutions have already adopted an approach to making images or scans of their heritage objects available that is compatible with the principles of open data, we compared the responses to the question relating to the availability of cultural heritage items on the internet with the responses to the question relating to the conditions of free availability (see Figure 17). The comparison shows that of the 60% of institutions that make at least some of their cultural heritage items available on the internet, only a very small number make them available for 'free' (including commercial) use by third parties. The proportion is probably somewhere around 1% - 7% of the institutions. This 'fuzziness' is due to the fact that respondents apparently differentiated between the provision of data 'for non-profit projects such as Wikipedia that also permit commercial use' and the provision of data 'for users with commercial intentions', even though the former always includes the latter. A basic misunderstanding is also present in the case of 32% of the surveyed cultural heritage institutions that declared their willingness to provide images of their cultural heritage items for use in Wikipedia free of charge whilst simultaneously wanting to protect the images from modification. Only material available under 'free' licences is accepted by Wikipedia and Wikimedia Commons; they *per se* permit the modification of works.

Figure 17: Availability of images of cultural heritage items

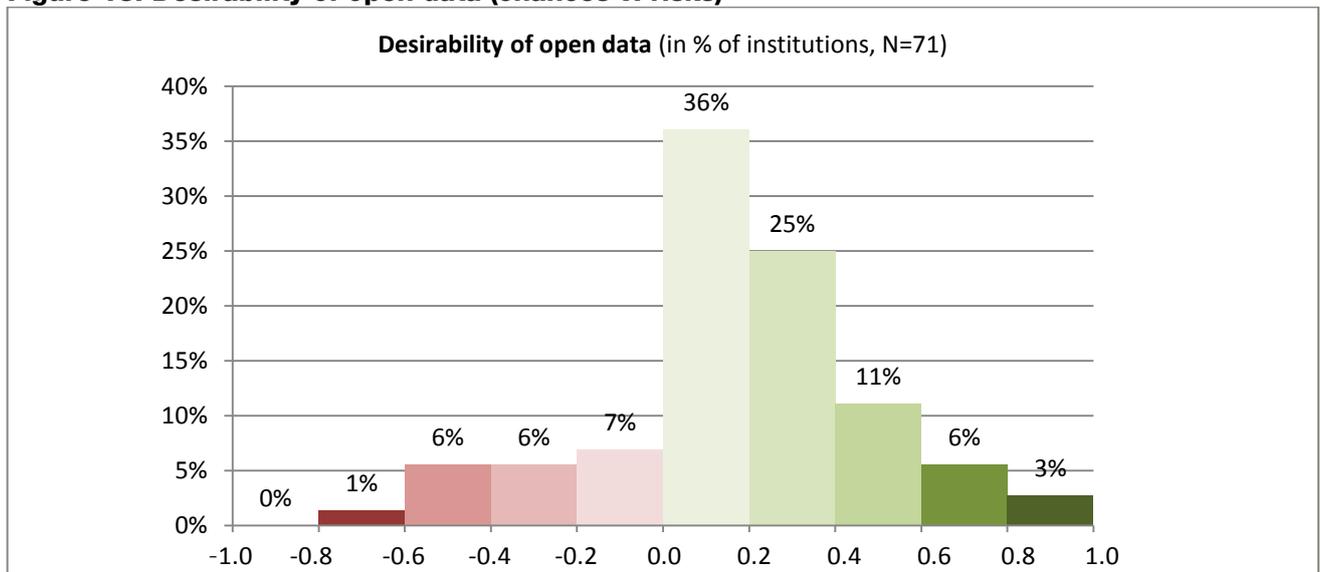


¹⁷ Percentages relate to the basis population of surveyed institutions. However, only those institutions who stated that they exchange metadata with other institutions were asked to respond to this question.



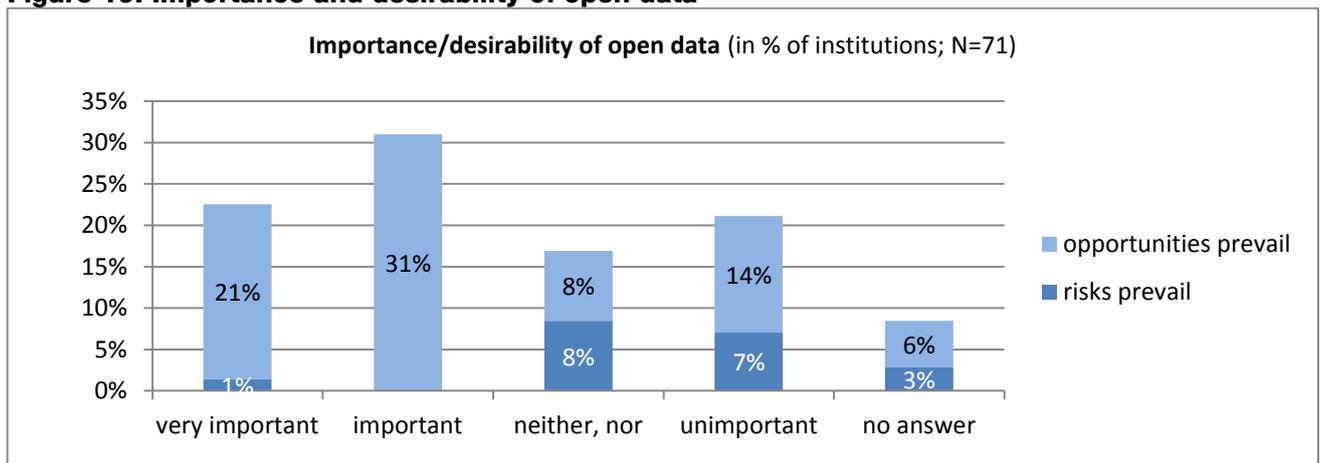
Perception: We have two indicators that relate to the perception of open data: An indicator for the perceived ‘opportunity’ of open data can be obtained by comparing how the opportunities and risks of open data are perceived. We have converted the results into figures on a scale from -10 to 10, whereby 0 represents a neutral attitude. Figure 18 shows that 80% of the surveyed institutions consider the opportunities of open data to outweigh the risks; for more than 40% this is clearly the case.

Figure 18: Desirability of open data (chances v. risks)



The importance attached to open data is a further indicator of how this trend is perceived: more than half of the surveyed institutions consider open data to be important or very important. Figure 19 shows that nearly all of them also think that the opportunities of open data outweigh the risks. Conversely, the institutions that attach less importance to open data are more likely to think that the risks outweigh the opportunities.

Figure 19: Importance and desirability of open data



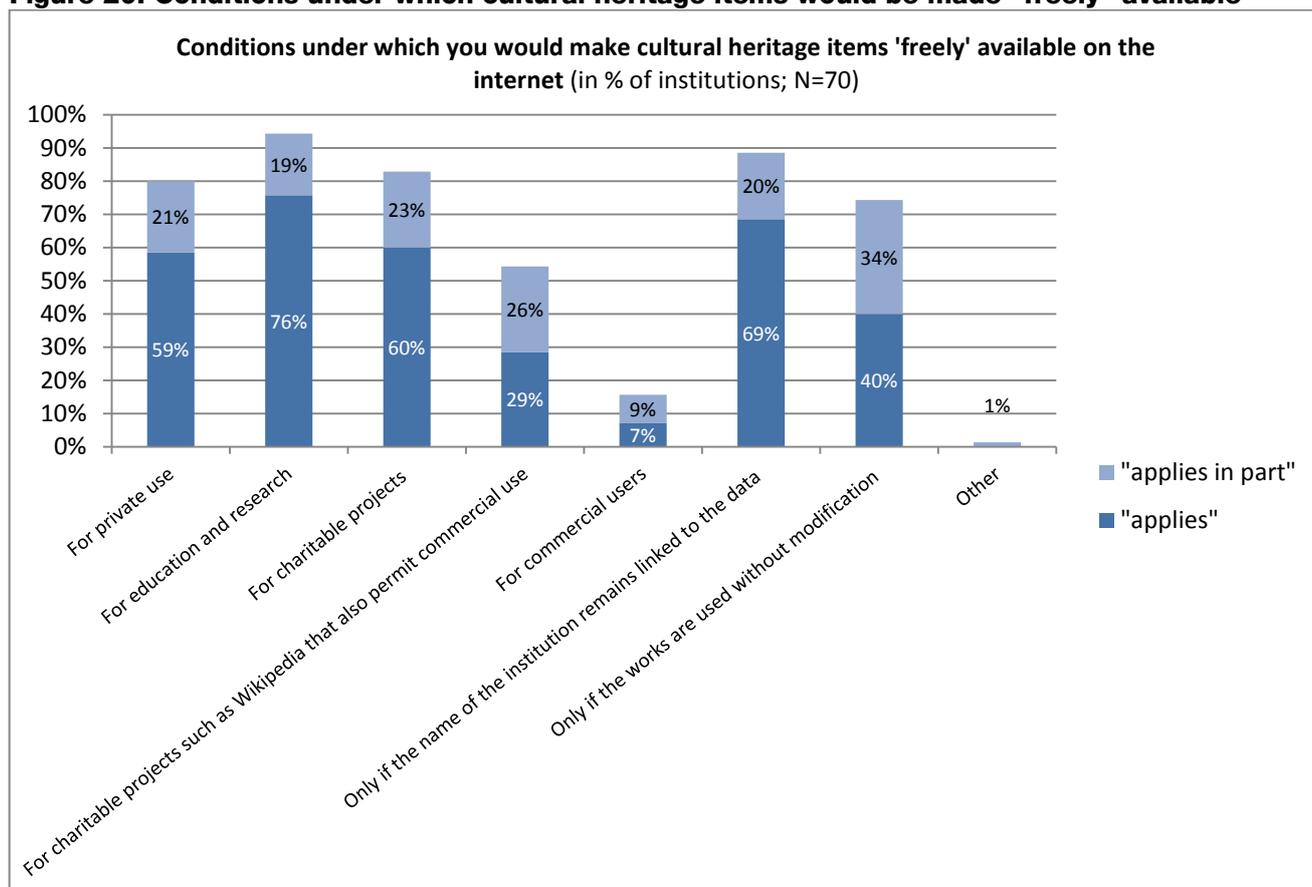
Prerequisites: Various preconditions are necessary to enable a cultural heritage institution to pursue a consistent open data strategy with regard to images of its cultural heritage items: The artefacts must be available in digitised form and must either be in the public domain or the institution must hold the necessary rights of use and the permission to make them available under a ‘free’ licence. Furthermore, the institution must also be in favour of ‘free’ licensing that permits the images to be modified and commercially exploited. But there can also be other legal obstacles, such as restrictions with regard to the commercial use of archive material. The Swiss Federal Archiving Act, for instance, states that the commercial use of archive material requires a permit (Art. 19,



Sec. 1). The Federal Council's message concerning the act states: "The commercial use of archive material shall be subject to approval. The confederation shall be financially compensated for the in part extensive initial financial outlay necessary to maintain and adequately store archive material."¹⁸ Furthermore, it may be helpful to be aware of the terms of the standard licences that are generally used for the purpose of 'free' licensing.

Figure 20 provides an overview of the conditions under which the surveyed institutions could imagine making content or images of their cultural heritage artefacts available on the internet free of charge without acquiring any revenue in return – provided that the material is already available in digital format and the publication is not contrary to either third party copyright claims or confidentiality regulations. It is apparent that the willingness to make images available on the internet free of charge for educational and research purposes is very high: 95% of institutions are at least partly willing to do so. The figures are slightly lower with regard to making images available to non-profit projects or private use, but still reach the 80% mark nonetheless. Willingness drops significantly with regard to the release of images for commercial use. When non-profit projects such as Wikipedia that are committed to promoting 'free' access to knowledge and therefore permit the commercial use of content are specified in this context, 55% of the institutions are still willing to make their content available free of charge, even if commercial use of the content by third parties cannot be ruled out. The value drops to 15% if the focus on non-profit projects is removed. A majority of cultural heritage institutions stipulates additional conditions - for instance, 89% of the surveyed institutions require the institution's name to remain linked to the data in at least some of the cases, and 74% require the material provided to be used in its original form in at least some of the cases.

Figure 20: Conditions under which cultural heritage items would be made "freely" available

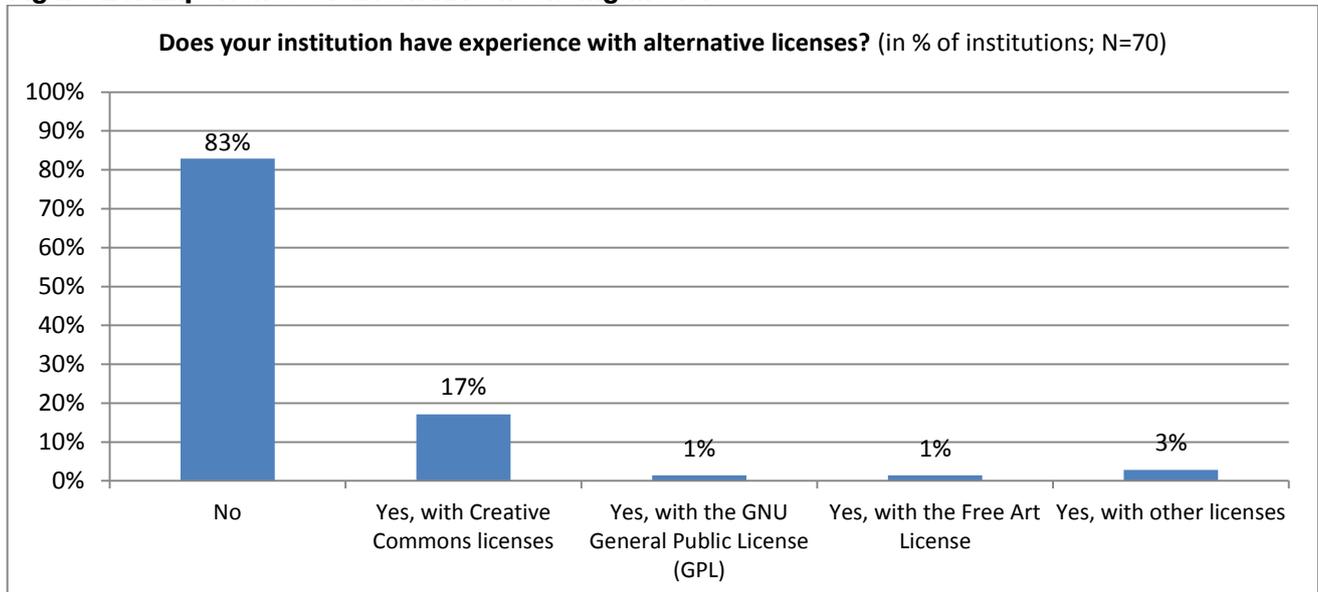


¹⁸ <http://www.amtsdruckschriften.bar.admin.ch/viewOrigDoc.do?id=10054231>



Figure 21 shows that the vast majority (83%) of institutions do not yet have any experience of alternative licensing models; 17% of the surveyed cultural heritage institutions stated that they had already employed Creative Commons licensing. However, it must be taken into consideration that not all Creative Commons licences are 'free' licences – some of them do not allow the commercial use of the material, whilst others prohibit modifications. Both types of 'unfree' licences are incompatible with the principles of open data.

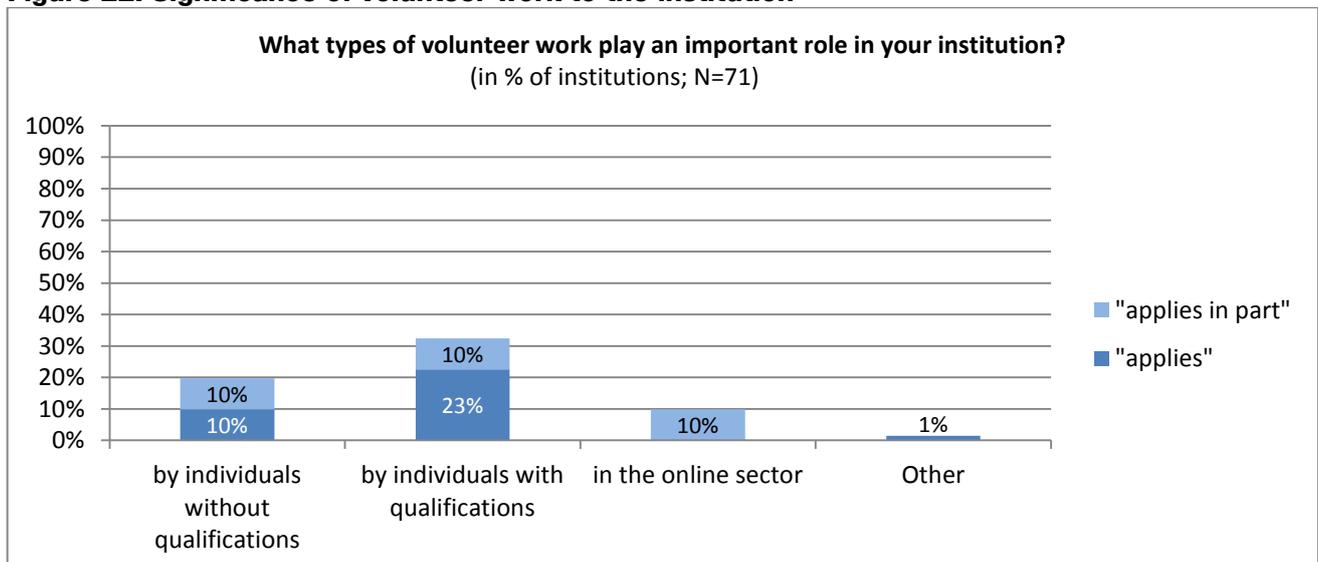
Figure 21: Experience of alternative licensing models



3.5.4 Crowdsourcing

Implementation: The questions relating to the importance of volunteer work in the online sector and to the involvement of employees in collaborative projects on the internet serve as indicators as to how approaches towards crowdsourcing are being implemented. Figure 22 shows that 10% of the surveyed institutions stated that volunteer work in the online sector plays partly an important role. The surveyed data does not allow any conclusions to be drawn as to the type of online work or the community projects involved.

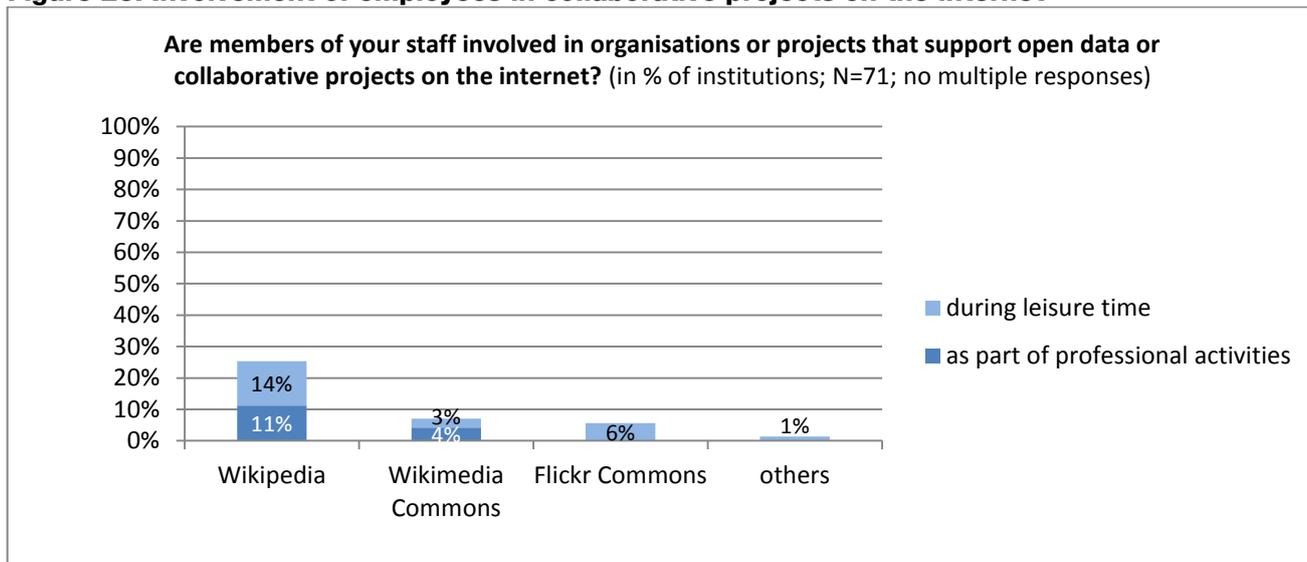
Figure 22: Significance of volunteer work to the institution





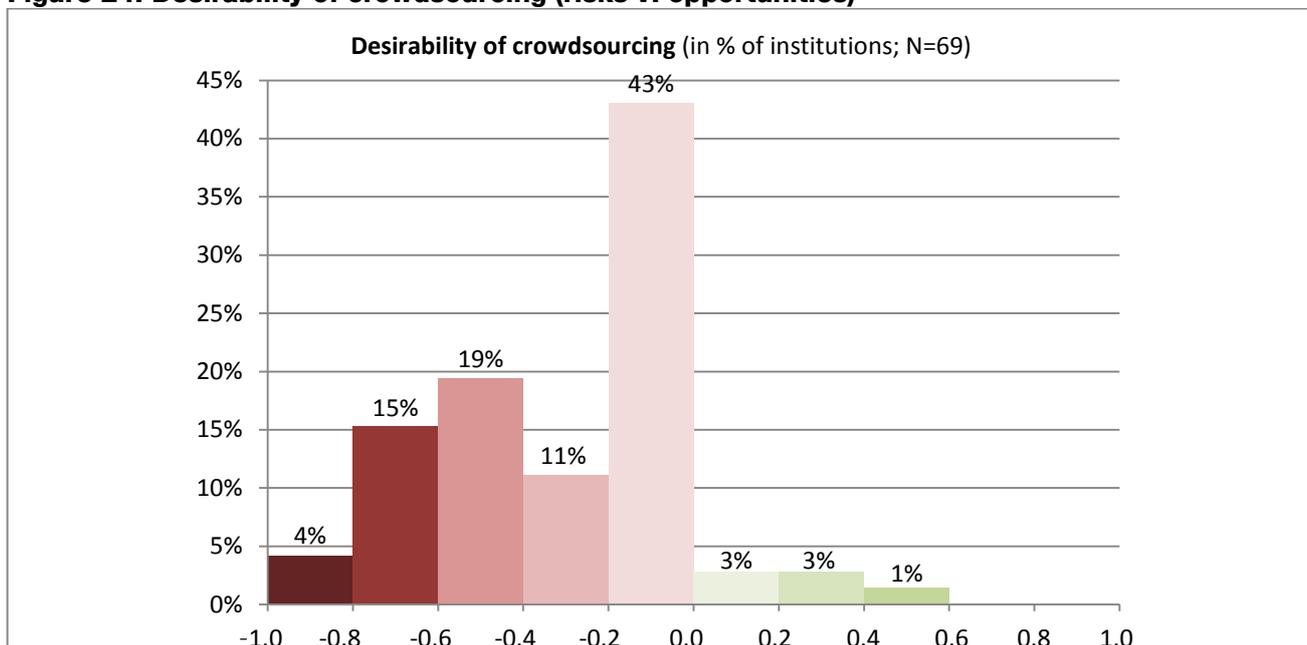
Things are slightly different when it comes to employee participation. The results clearly show that employees are most likely to be involved in Wikipedia: 11% of the surveyed institutions stated that their employees participate in the online project as part of their work, whilst the employees of a further 14% do so in their free time. Proportions of 7% and 6% of institutional employees are involved in Wikimedia Commons and Flickr Commons respectively (see Figure 23). Musicbrainz.org, an open encyclopaedia of music that collects music metadata, and the International Music Score Library Project (IMSLP), the largest online collection of music scores available freely without charge, were also named as online communities in which employees of Swiss cultural heritage institutions participate. Both collections are based on the same Wiki principle as Wikipedia.

Figure 23: Involvement of employees in collaborative projects on the internet



It is interesting to see that there is no clear correlation between the two indicators: it cannot be assumed that institutions whose employees participate in Wikipedia projects as part of their professional activities consider voluntary work in the online sector as at least of partial importance to the institutions (this was in fact the case with only 2 of 8 institutions). The case is similar with regard to the institutions whose employees participate in Wikipedia in their free time (2 of 10 institutions consider voluntary work in the online sector as important to their institution).

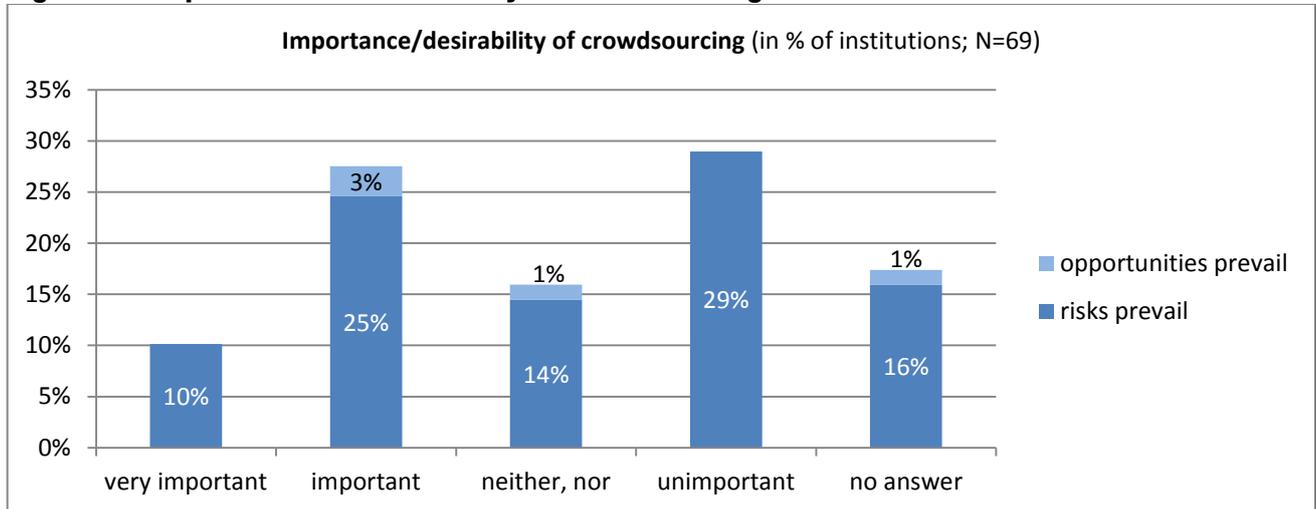
Figure 24: Desirability of crowdsourcing (risks v. opportunities)





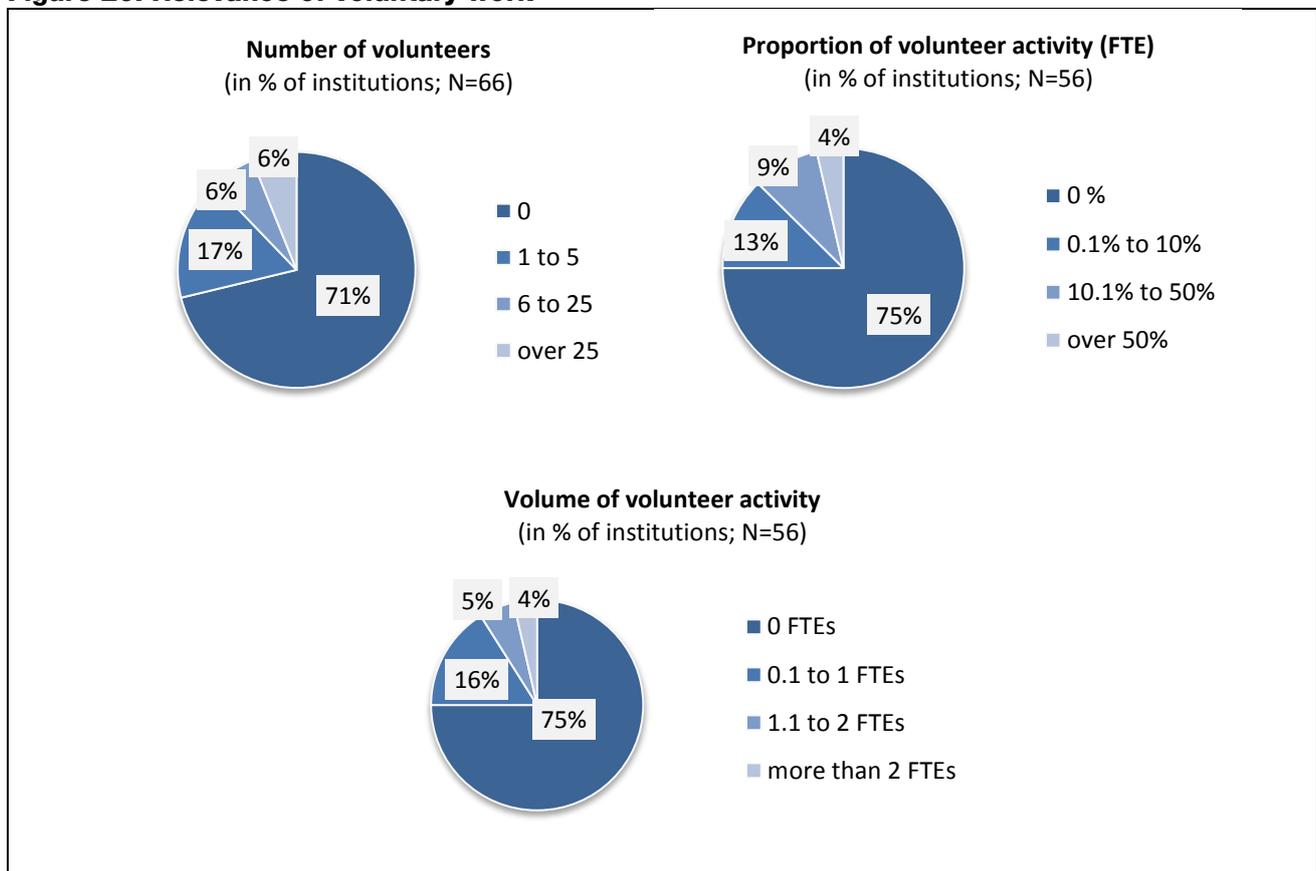
Perception: A review of the weighting scores regarding opportunities and risks shows that the surveyed institutions consider crowdsourcing to be less promising than open data: 92% of the surveyed institutions see at least an equilibrium in terms of risks versus opportunities, whereby for one half the risks clearly outweigh the opportunities (see Figure 24).

Figure 25: Importance and desirability of crowdsourcing



We also asked the institutions how important the subject is to them. Figure 25 shows that the surveyed institutions consider crowdsourcing as slightly less important than open data, although the differences are not significant (38% of the surveyed institutions consider crowdsourcing important, while 53% consider open data important). The fact that the institutions that considered crowdsourcing to be important were also sceptical about its opportunities is somewhat thought-provoking.

Figure 26: Relevance of voluntary work





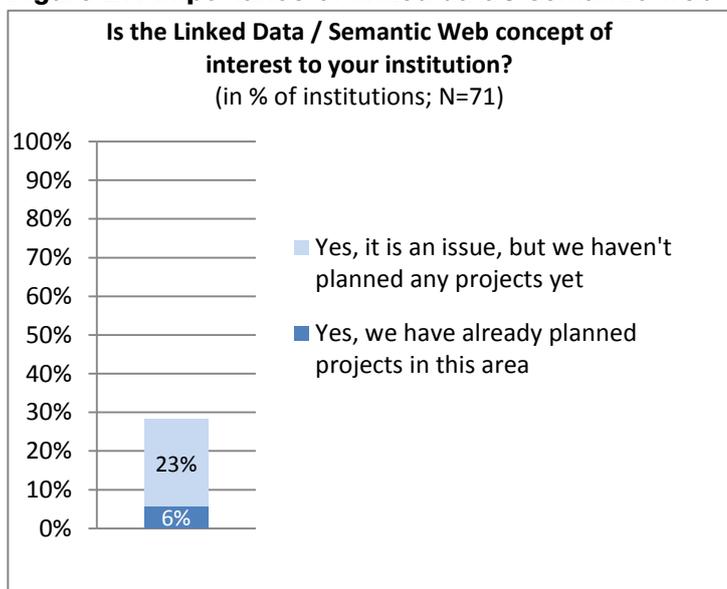
Prerequisites: The factors that would better enable an institution to participate in crowdsourcing activities are the importance it attaches to volunteer work in general and the extent to which its employees are already involved in collaborative projects. The three graphics in Figure 26 together with Figure 22 show the relevance of volunteer work to the participating institutions: Whilst voluntary work hardly plays any role in around three-quarters of the surveyed institutions, 12 - 13% claimed it played an important if not decisive role. For instance, two institutions stated that more than 50% of the workload is handled by volunteers. All in all, around 9% of the institutions specified that the work provided by volunteers equates on average to more than that provided by one full-time employee.

3.5.5 Linked data / semantic web

Perception: As the linked data concept is not yet that well-known among Switzerland's cultural heritage institutions, we asked only one question: "Is the linked data/semantic web concept of interest to your institution?" It would seem that 6% of the surveyed institutions have already planned related projects, while an additional 23% affirmed that they were interested, but have not yet planned any projects (see Figure 27).

Prerequisites: Although the linked data / semantic web concept could also be implemented within a closed framework, the general tendency among cultural heritage institutions indicates that it will mainly be about linking open data – more or less as a logical progression of the other trends; namely cooperation in associations, establishing single points of access, opening up availability of data, avoiding duplication, and outsourcing some of the workload to volunteers on the internet. All of the factors discussed in the preceding sections are thus likely to facilitate the implementation of a linked data strategy.

Figure 27: Importance of linked data / semantic web



3.5.6 Summary

Figure 28 provides a comparative overview of the indicators of how each innovation is perceived and the status of its implementation. The results should nonetheless be interpreted with caution as the indicators are based on somewhat different questions and do not always 'measure' every trend in exactly the same way. However, they do make it possible to obtain a rough overview of the current diffusion status of each innovation. The following points stand out:

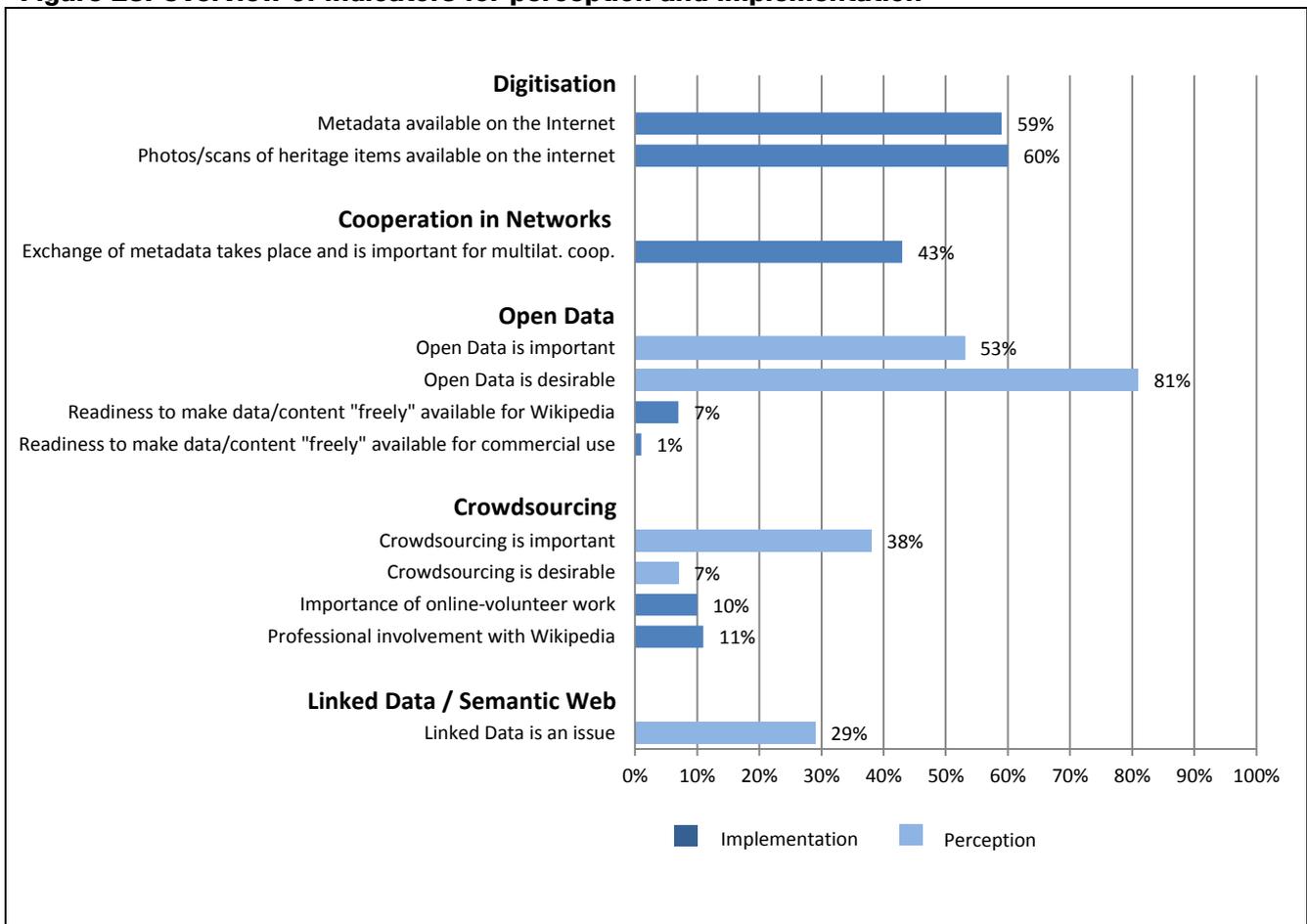
- A distinct graduation becomes apparent when we consider the status of implementation of each innovation. Digitisation achieves the highest value as 60% of the surveyed institutions already make at least some metadata and images of cultural heritage items available on the internet. This is followed by cooperation in associations in second place: 43% of the surveyed institutions exchange metadata and state that they do this in the context of multi-lateral cooperation. The indicators show that crowdsourcing is being implemented in some 10% of cases, while only 1% of the surveyed institutions have fully adopted the practice of open data. There is a certain amount of ambiguity as many institutions have

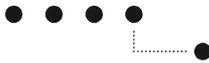


indicated a willingness to release content for non-profit projects such as Wikipedia that allow the commercial use of material, while simultaneously opposing the commercial use of material and also specifying that their material should not be modified.

- If we consider how open data is perceived, we can see that 80% of the surveyed institutions rate the opportunities of open data higher than the risks and that more than half of them consider open data to be important. These are very high values in view of the degree of implementation and indicate that dynamic forces are at work in this area that may well result in speedy progress.
- The surveyed institutions are a lot less positive when it comes to crowdsourcing. Only 7% see more opportunities than risks. However, 38% consider the subject to be of importance, while 11% would appear to be already experimenting with Wikipedia. This constellation could indicate that successful crowdsourcing is perceived as being far more complex than making data available on the internet in a machine-readable format and under a 'free' licence. Due to the considerable importance attached to the trend, it can be assumed that many cultural heritage institutions are inclined to participate in crowdsourcing experiments, even though the basic attitude is significantly more sceptical than towards open data.
- Linked data is important to 29% of the surveyed institutions, although only 6% have concrete plans for projects. The figures seem to indicate a trend that is only now beginning to take root in cultural heritage institutions.

Figure 28: Overview of indicators for perception and implementation



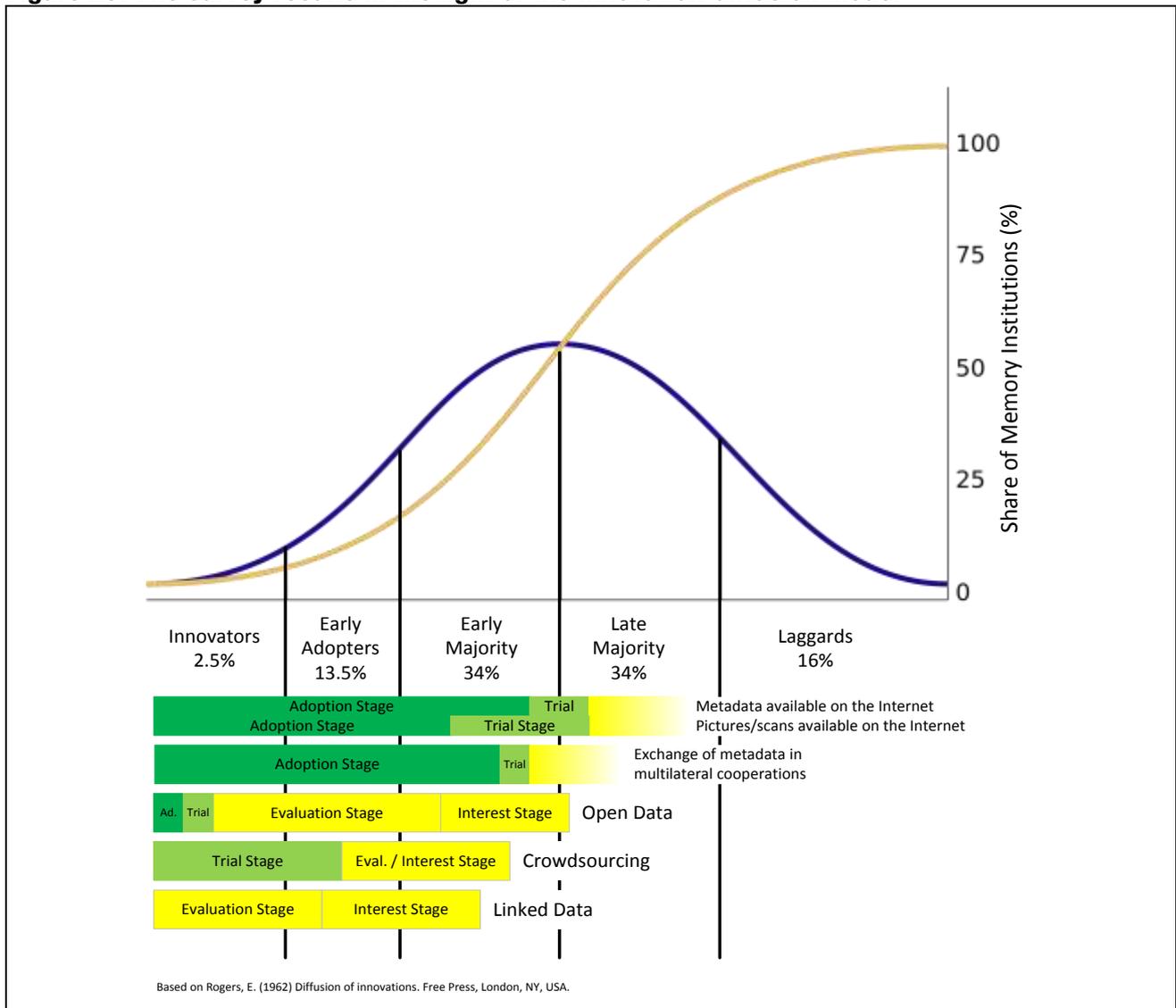


Assessing the results of the survey in the light of the diffusion model of Beal/Bohlen (1957) produces the following conclusions (see Figure 29):

- The early majority is already in the adoption stage or about to leave the trial stage with regard to making metadata and images available on the internet; the late majority is entering the trial stage.
- With regard to the exchange of metadata within the framework of associations, a large proportion of the early majority is already in the adoption stage, whilst the remainder is still in the trial or evaluation stage.
- The first innovators are already in the adoption or trial stage with regard to open data. All early adopters are already in the evaluation stage, whilst the early majority has reached the interest stage.
- With regard to crowdsourcing, the innovators and some early adopters are in the trial stage; the other early adopters and a good number of the early majority are in the interest or evaluation stage.
- With regard to linked data, the innovators and some early adopters are in the evaluation stage; the remaining early adopters and a proportion of the early majority have entered the interest stage.

These results should be viewed with a certain amount of caution. Firstly, the data leaves room for interpretation in many aspects when applied to the model; secondly, the underlying sample size is relatively small.

Figure 29: The survey results in the light of the innovation-diffusion model





3.5.7 General applicability of results

As stated in chapter 2.3, caution must be exercised when extrapolating these results for Switzerland's cultural heritage institutions as a whole due to the relatively small sample size. The following statements are correct with a probability of 95% (if one disregards a certain bias in the sample):

- Of cultural heritage institutions in Switzerland, 49% to 71% make at least some of their metadata and representations of their memory objects available on the internet. However, no more than 3% provide their content under 'free' licences.
- Of cultural heritage institutions in Switzerland, 32% to 54% exchange metadata in the context of multi-lateral cooperation.
- Of cultural heritage institutions in Switzerland, 4% to 18% are already involved in Wikipedia projects, while 3% to 17% consider voluntary work in the online sector as important. There is no significant correlation between the two aspects.
- Of cultural heritage institutions in Switzerland, 72% to 90% see open data as an opportunity, while 41% to 65% consider the subject to be of importance.
- Only 1% to 13% of cultural heritage institutions in Switzerland consider crowdsourcing as an opportunity; however, 27% to 49% at least consider the subject to be of importance.
- Linked data is important to 18% to 40% of cultural heritage institutions in Switzerland; 0.5% to 12% of institutions already have concrete plans for related projects.



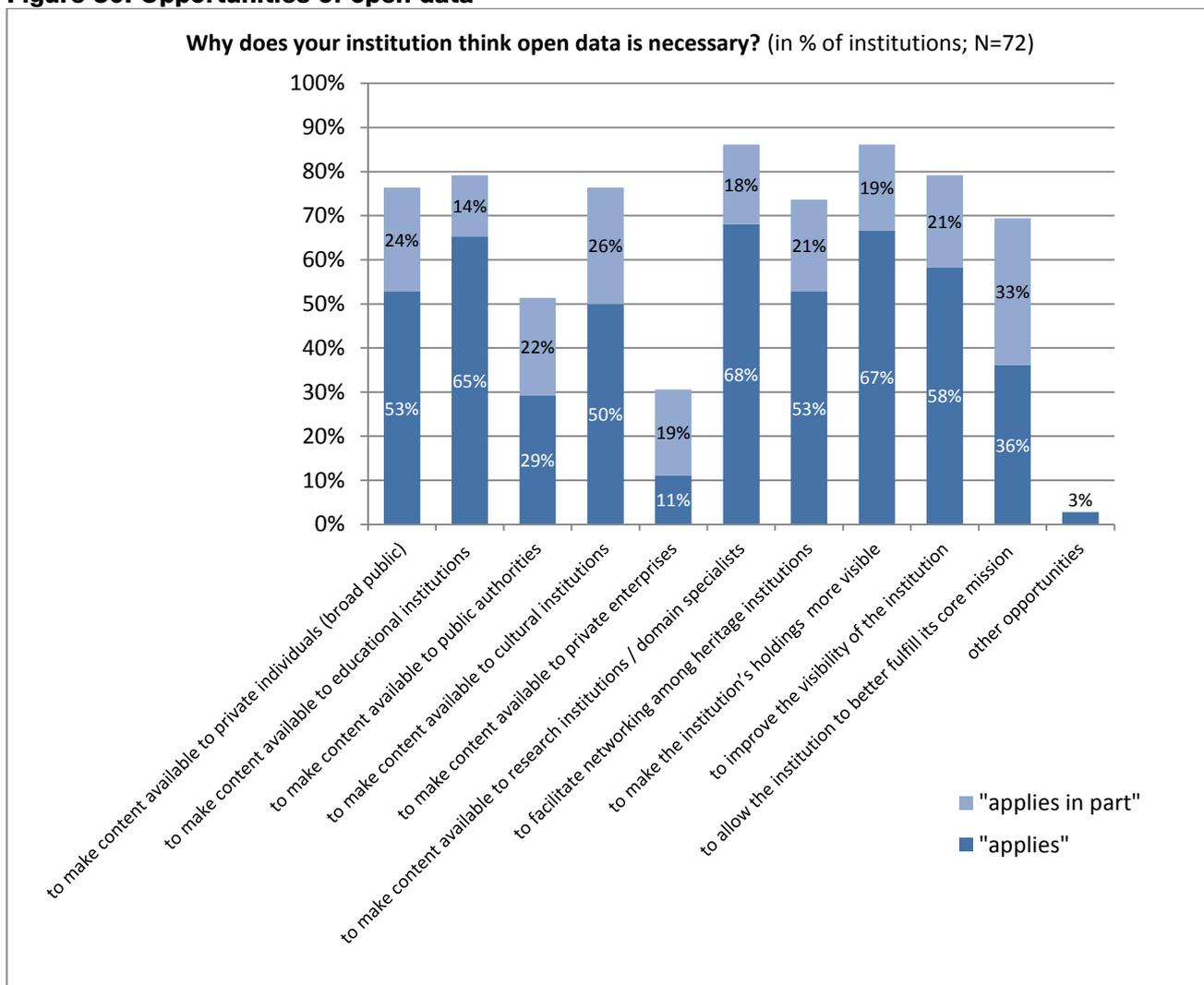
3.6 What opportunities and risks do the institutions see in relation to open data and crowdsourcing?

We asked the cultural heritage institutions to assess various opportunities and risks so we could better understand what the main drivers and inhibitors are from their viewpoint with regard to implementing open data and crowdsourcing. In this chapter we look at the results and attempt to reach some initial conclusions. Furthermore, we consider the question of how open data and crowdsourcing can be of benefit to society.

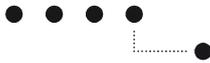
3.6.1 Opportunities and risks of open data

Figure 30 shows that the surveyed institutions see the following target groups as the main beneficiaries of open data: Research institutions/specialists, educational institutions, private individuals and cultural institutions (all scored 'applies' in more than 50% of cases). Making data available to public authorities or private enterprises appears to be less important (selected by 29% and 11% respectively). These results correspond roughly with those for the question of who the users of cultural heritage institutions are – the only difference being that research takes first place with regard to open data.

Figure 30: Opportunities of open data



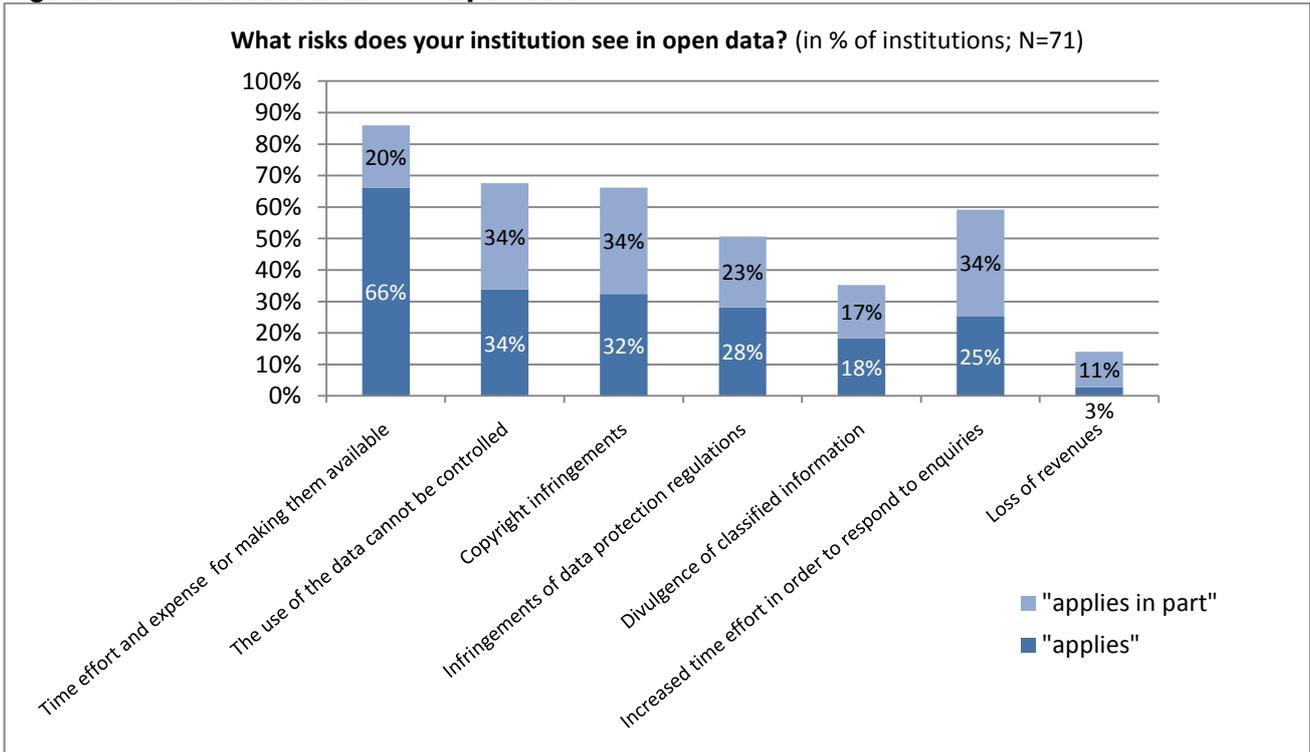
A clear benefit of open data is perceived by a majority of cultural heritage institutions with regard to the visualisation of holdings (67% of the surveyed institutions responded with 'applies'), to greater public awareness of the institution (58%) and to the facilitation of networking between cultural heritage institutions (53%). The perceived benefit with regard to the fulfilment of their core mission is less clear: Only one-third responded with



‘applies’, while another third thinks that open data may be beneficial ‘in part’ to the fulfilment of their core mission. All in all, the benefits of open data are seen as considerable.

In addition to the benefits of open data we also asked the cultural heritage institutions about the risks. The effort and cost of providing the data is clearly the leading issue (66% of the institutions responded with ‘applies’). Other risks such as loss of control (34%), increased time required to deal with enquiries (25%) and problems relating to copyright infringements (32%), data protection violations (28%) and the divulgence of classified information (18%) played a much less important role. Fears of a loss of revenue played virtually no role at all (3%) (see Figure 31).

Figure 31: Risks associated with open data



3.6.2 Opportunities and risks of crowdsourcing

We based our question about the opportunities of crowdsourcing on the typology of crowdsourcing approaches presented in chapter 2.3. The results show that very few of the surveyed institutions are convinced of its opportunities (see Figure 32). The surveyed institutions consider the use of crowdsourcing for classification tasks and the supplementation of metadata as its greatest potential (11% responded with ‘applies’); this is followed by correction and transcription tasks (6%), additions to collections (4%), crowdfunding (3%) and improvements to and supplementing texts (1%). The fewest opportunities were perceived in relation to the tasks of curators. It should be noted that a relatively large proportion of the surveyed institutions (10% - 17%) stated they were unable to answer the question; this suggests that they have not yet addressed the issue of crowdsourcing to any significant extent.

The main risk perceived in relation to crowdsourcing is once again the additional cost: 42% of the institutions clearly identified extensive preparation and follow-up work as a risk. The other risks listed in the survey followed closely behind with values in range 30% - 38%: little influence on results; difficulties in estimating the time investment required; lack of a guarantee with regard to the continuity of data maintenance and a low degree of planning security. Employee anxiety (for instance in relation to job loss, changed roles and tasks, etc.) played virtually no role at all and was cited by a mere 6% of the institutions (see Figure 33). Again, it should be noted that the proportion of institutions unable to answer the question is relatively high at 24% - 28%. On the one hand, this



would suggest that the institutions have not yet addressed the issue in detail. On the other hand, it shows that the proportion of institutions that consider the risks do not or tend not to apply is very low. This ranges from 0% (extensive preparation and follow-up work) to 11%, with the exception of employee anxiety, where more than 50% of the surveyed institutions said this plays virtually no role at all.

Figure 32: Opportunities of crowdsourcing

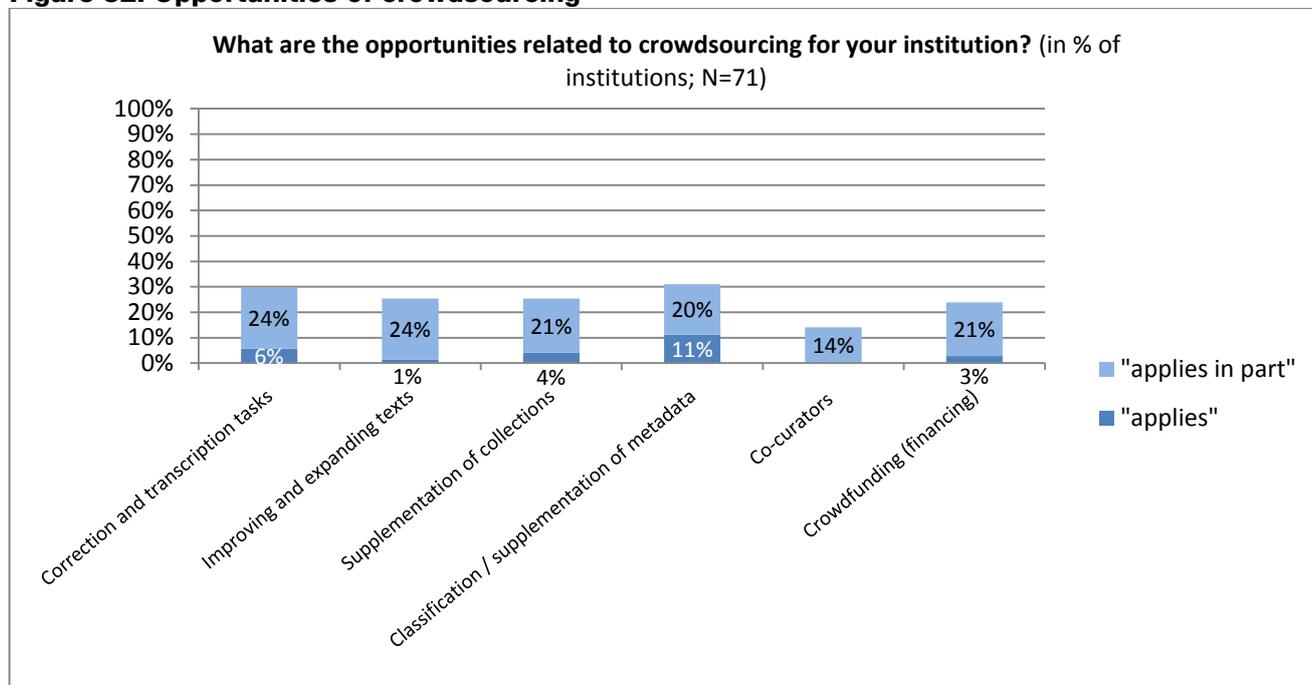
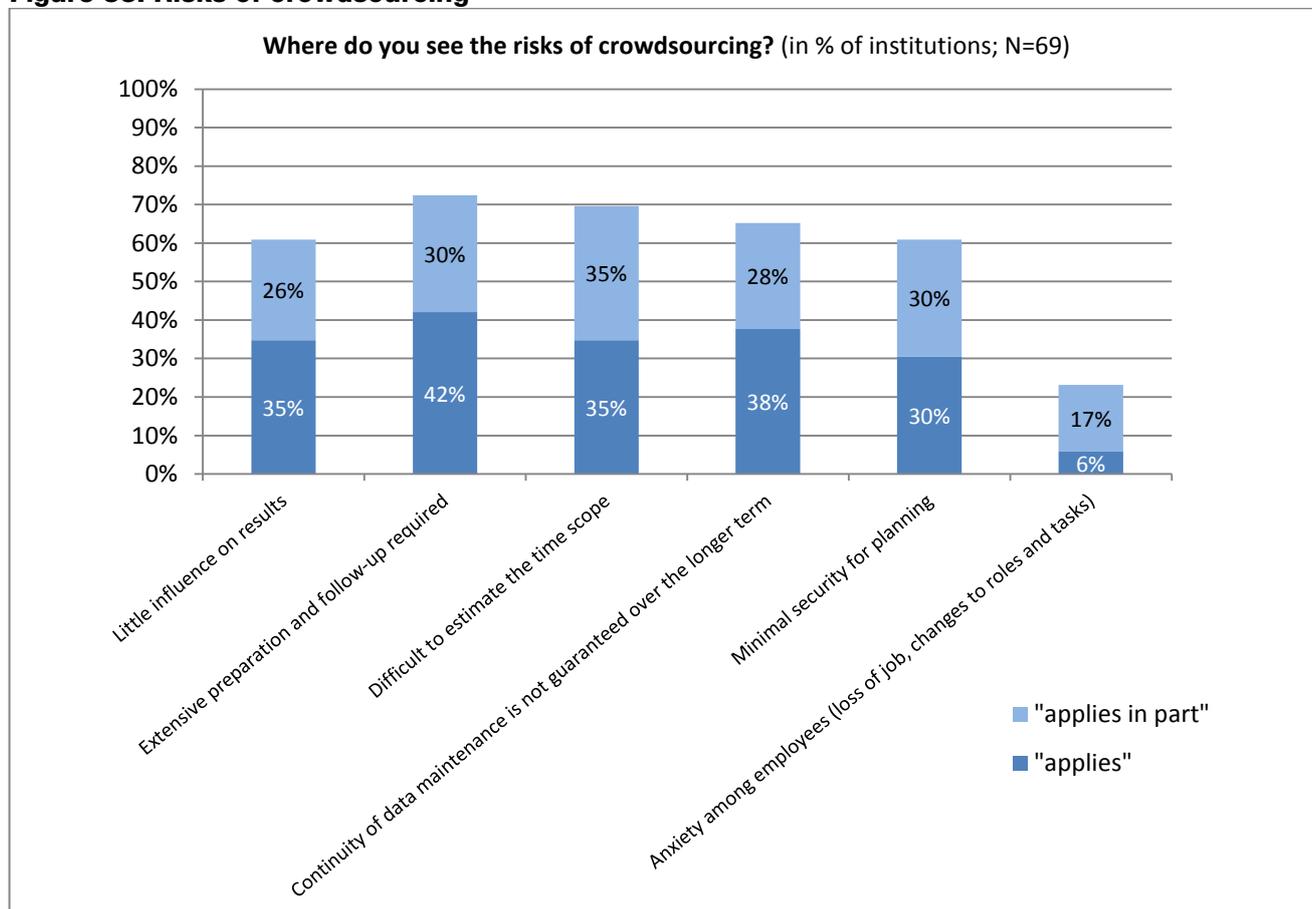
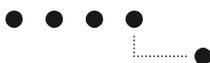


Figure 33: Risks of crowdsourcing





3.6.3 Benefits of open data and crowdsourcing to society when applied in the cultural heritage sector

The survey does allow certain conclusions to be drawn with regard to the benefits of open data and crowdsourcing to society when applied in the cultural heritage sector:

The user groups that in the eyes of the cultural heritage institutions would benefit most from an open data policy are located in the fields of education and research, followed by the general public. Furthermore, adoption of an open data policy would promote the networking of cultural heritage institutions, improve the visibility of their holdings and generally enhance how these institutions are perceived by the general public.

Existing crowdsourcing activities are centred mainly around Wikipedia. Collaborations with specialised online communities is also seen as an option for specific niche subjects (e.g. music metadata, music scores). Furthermore, it can be assumed that individual cultural heritage institutions are experimenting with their own approaches to crowdsourcing. The view of the benefits of crowdsourcing is ambivalent. On the one hand, the benefits of Wikipedia as a first port of call for people seeking information are uncontested. For some institutions, the online encyclopaedia is meanwhile an irreplaceable tool to reach their publics: 500 million people use Wikipedia and its sister projects, accounting for 19 billion page views every month¹⁹; according to Alexa, these figures make it the sixth most visited website in the world.²⁰ Or, to say it in the words of the US federal archivist: “The Archive is involved with Wikipedia, because that’s where the people are!”²¹ On the other hand, Swiss cultural heritage institutions do not currently see crowdsourcing as an opportunity to reduce their workload.

The added benefits don’t come without extra costs. The surveyed institutions cite additional effort and expense as the greatest challenge with regard to both open data and crowdsourcing. A loss of revenue, on the other hand, is not a major concern. There are two ways of covering the additional outlay. One is to provide the institutions with extra funding. Public funds would probably be the most appropriate sources in view of the financial structure of Swiss cultural heritage institutions and the expectation that education and research will reap the greatest benefit. The other option is for the institutions to improve efficiency through the application of an open data policy – for instance through improved and facilitated networking. The institutions would also need to pursue crowdsourcing approaches where increased outlays are compensated by a reduction of workload. In view of tightened public purse strings, it is unlikely that open data and crowdsourcing will be implementable unless they improve efficiency in the shape of added benefit whilst keeping costs at the current level. Especially in the case of crowdsourcing, Switzerland’s cultural heritage institutions first need to see how it would specifically improve efficiency.

3.6.4 Summary

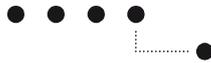
The analysis of risks and opportunities sheds further light on the obstacles that will need to be overcome if an open data policy and/or crowdsourcing approaches are to be adopted by Swiss cultural heritage institutions and points to the corresponding driving factors.

Firstly, regarding open data, internal processes and structures will need to be adapted to reflect the reversal of previously valid principles relating to the publication of data: copyright infringements and violations of data protection and confidentiality regulations will have to be avoided in a system designed specifically to make all

¹⁹ See the Wikimedia Foundation’s annual report for 2011-2012: http://wikimediafoundation.org/wiki/Annual_Report

²⁰ www.alexa.com/topsites

²¹ Wikimania 2012 (Washington D.C.): Closing Plenary by David Ferriero, Archivist of the United States. <http://www.youtube.com/watch?v=47pEcmXjt8E>



information ‘freely’ available in a machine-readable format. All other cases are considered to be exceptions. The survey results give rise to the assumption that these adjustments will make necessary a fundamental rethinking process in many institutions and will probably pose considerable challenges. Moreover, the results of the survey indicate that many institutions are worried that adopting the principles of open data will result in a loss of control to a certain extent. Again, at least some of the institutions will be forced to undergo a change in culture.

A number of uncertainty factors need to be further explored with regard to crowdsourcing: What results can be expected? How much time will the institution have to invest? What activities could be planned? What activities call for a spontaneous reaction to events occurring within the partnering community? What about the continuity of data maintenance? How sustainable are the results of cooperation? – Basically, this uncertainty can only be reduced by engaging in small, manageable experiments and pilot projects. It is also necessary to form a relationship with an online community; either with a community developed by the institution itself or with an existing community. As with any relationship, engaging with a community brings uncertainties and unknowns with it and demands real commitment.

The greatest obstacle with regard to open data and crowdsourcing is the fear of additional costs. The challenge here will be to show individual institutions how open data and crowdsourcing can reduce workload and lead to greater efficiency in certain areas.

The main driving factor behind open data is the positive perception of its benefits in the form of better networking between cultural heritage institutions, greater public awareness of the institutions and improved services for users as a consequence of better accessibility to cultural heritage items. Thus, the surveyed institutions are entirely orientated towards the Federal Council’s strategy for a Swiss information society that will “make the country’s cultural heritage comprehensively and easily accessible” to the general public through the use of IT (Schweizerische Eidgenossenschaft 2012, p. 13).

The driving factors behind crowdsourcing are probably much weaker than those behind open data. The positive expectations of some of the institutions with regard to classification of heritage items and supplementation of metadata could play an important role – especially in view of the fact that more than half of the institutions perceive a need to improve conditions in this area urgently or in the medium term, a need that will become more apparent with the consistent application of an open data policy. Correction and transcription tasks are likely to be another area where cultural heritage institutions will try to reap advantages. The partially positive perception of opportunities by a good quarter of the institutions is likely to have an enabling effect, as they may be ready to engage in pilot projects. However, it should be noted that the dynamics behind the development of crowdsourcing are significantly weaker than those behind open data.



4 Conclusions and outlook

4.1 Main insights gained from the pilot survey

The pilot survey provides valuable information on the state of diffusion of innovative practices among Swiss cultural heritage institutions that have developed as a consequence of the expansion of the internet.

Digitisation: The digitisation of metadata and images of cultural heritage items are important prerequisites for internet-based cooperation, open data and many forms of crowdsourcing. A majority of surveyed institutions (60%) is already active in this area and makes both metadata and images of at least some of their cultural holdings available on the internet.

Cooperation in networks: A good two-fifths of the institutions (43%) participate in networks in which the exchange of metadata plays an important role. A similarly high proportion considers the exchange of metadata to be relevant to the fulfilment of their core mission. However, the need to improve the quality of metadata is relatively high: one half of the surveyed institutions indicated a need for improvement in the medium term, if not more urgently.

Open data: Hardly any of the surveyed institutions currently fully embrace open data. There are, however, indications that this innovation could rapidly become accepted. A majority of the surveyed institutions consider that open data is of importance and the opportunities it offers outweigh the risks. There are, however, a number of obstacles to overcome beforehand: the risks of copyright infringements, breaches of data protection and confidentiality regulations must be eliminated by introducing appropriate measures. Additionally, aspects relating to copyright, such as the question of orphaned works, and legal restrictions concerning the commercial use of cultural heritage items will continue to pose considerable challenges. Better visibility and accessibility of holdings, better visibility of cultural heritage institutions, and the establishment of a more comprehensive network between them are perceived as the main benefits of open data. It is, however, first necessary to overcome reservations concerning the 'free' licensing of material to enable a cultural heritage institution to pursue a consistent open data strategy with regard to images of its holdings. These reservations concern the commercial use of and modifications to holdings. Concerns relating to loss of control over holdings are shared by 68% of the surveyed institutions; this could therefore become a major obstacle to the adoption of open data by cultural heritage institutions.

Crowdsourcing: In general, use of crowdsourcing is developing at a slower rate than the trend towards acceptance of open data. Although around 10% of the surveyed institutions have experimented with crowdsourcing, a major breakthrough has yet to happen. Nearly one half of the institutions considers the subject to be of importance, yet many of those surveyed have not yet addressed it. Moreover, the majority of institutions is sceptical of crowdsourcing. In general, crowdsourcing is seen as being associated with a lot of risk and little benefit. Institutions mostly see opportunities to use crowdsourcing for the improvement of metadata and for tasks relating to correction and transcription work. However, Switzerland's cultural heritage institutions first need to be convinced that this would improve efficiency.

Linked open data: Developments relating to linked open data are still at the very beginning. Linked open data is, however, already of interest to 29% of the surveyed cultural heritage institutions; 6% have planned initial projects.

The survey results show that cultural heritage institutions perceive the **benefits of open data** to be primarily in the fields of education and research and in relation to the general public. Furthermore, it is assumed that the adoption of an open data policy would enhance the interconnectedness of cultural heritage institutions, increase the visibility of their holdings and generally improve how the institutions are perceived by the general public. In addition to facilitating networking among institutions, the benefits of open data will manifest themselves in particular in new approaches to linking and visualising data based on semantic web technologies as well as in the form of crowdsourcing activities. However, the results of the survey indicate that most of the surveyed institutions do not yet have a clear idea how they could use these innovative approaches to their advantage.



The anticipated benefits of open data and crowdsourcing are at least in part offset by **costs**: the surveyed institutions state additional effort and expense as the greatest challenges with regard to both open data and crowdsourcing. Loss of revenue, on the other hand, is not a major concern. Public entities and private foundations, which provide a large proportion of the funds required by cultural heritage institutions in Switzerland, could play a key role in promoting open data. In view of the fact that the majority of cultural heritage institutions recognises the benefits of open data and is not concerned about a loss of revenue, it would seem logical for the funding institutions to promote or even demand the ‘free’ licensing of holdings.

4.2 Aspects omitted or inadequately addressed by the pilot survey

The survey falls short in three aspects. Firstly, a number of limitations arise from the relatively small sample size; secondly, certain questions that would have better addressed specific aspects didn’t find their way into the questionnaire for a variety of reasons; and thirdly, there are of course always important questions that should be addressed by means other than a quantitative survey.

4.2.1 Limitations due to the small sample size

Care should be taken when drawing conclusions from the results of the survey about cultural heritage institutions as a whole due to the relatively small sample size. There is thus marked statistical fuzziness with regard to the results. This led us to largely forego ascertaining the correlations between individual variables of the study as the underlying data is in most cases insufficient to allow clear conclusions to be drawn. However, there is sufficient data to make possible an explorative analysis of these correlations to provide well-reasoned hypotheses that could be used as the basis of future surveys. These could then subsequently be tested at a later date in the scope of a broader survey. This would make it possible to identify factors relating to a positive attitude towards open data or crowdsourcing: Is the size of the institution important? Does experience of working with volunteers make it easier to adopt a crowdsourcing approach? Is there a correlation between the type of cultural heritage items and involvement in Wikipedia? And so on.

4.2.2 ‘Blind’ spots that could have been covered by additional questions

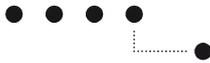
The pilot survey did not ask the institutions about the extent to which they are already using Web 2.0 applications, i.e. in relation to Facebook, YouTube, etc. Neither did it pose the question as to how volunteers contribute towards the institution’s objective via the internet (online volunteers). We would very probably include these questions in the questionnaire if we were to carry out the survey again. One could also shed more light on the question as to why 40% of the surveyed institutions do not provide any metadata or images of their holdings on the internet: What are the needs of these institutions? Why does digitisation not take place (yet) in certain circumstances?

4.2.3 Further research questions

Of course, there are many interesting questions relating to the subject of ‘open data and crowdsourcing in cultural heritage institutions’ that the survey omitted or addressed insufficiently for reasons of methodology:

Questions of a legal nature: What proportion of holdings will not be made ‘freely’ available in the foreseeable future due to copyright restrictions? What consequences will this have? What role do ‘orphaned’ holdings play whose legal status or copyright holder is unknown or ascertainable only with disproportionate effort? In what areas do legal restrictions other than copyright restrictions apply with regard to the commercial use of cultural heritage items that are incompatible with the principles of open data? Can they be overcome?

Questions of a cultural nature: To what extent is the self-concept of cultural heritage institutions and the associated professions compatible with the ‘free’ licensing of content and participation in collaborative projects on the internet? What is the reason why institutions are concerned about loss of control if they make holdings



that are already in the public domain ‘freely’ available? In what areas are cultural heritage institutions innovative? In what areas are they less innovative? Why?

Questions of an economic nature: What business models for financing digitisation processes will prove successful? How can the additional costs relating to open data and crowdsourcing be financed?

These are but a few examples; we do not raise any claim to completeness with regard to relevant research questions.

4.3 Outlook

The results of the pilot survey indicate that the implementation of an open data policy in the field of cultural heritage institutions will develop dynamically. The traffic lights would appear to be on ‘green’ as far as the majority of institutions are concerned. The speed of implementation will probably depend heavily on the political commitment with regard to ‘free’ licensing of data and content. While the Federal Council’s Strategy for an Information Society in Switzerland defines the objective of making cultural heritage comprehensively and easily accessible, it lacks a clear definition of what that means.

Once data is freely available, cultural heritage institutions will face the challenge of cutting costs through networking efficiently and avoiding duplication while at the same time creating added value through the use of crowdsourcing and new approaches to linking and visualising data. With regard to the latter, we are standing only at the beginning of developments the potential of which is difficult to predict at this moment in time.



5 List of references

Beal, George M./Bohlen, Joe M. (1957). "The Diffusion Process", Special Report No. 18 (Agriculture Extension Service, Iowa State College) 1: 56–77. Reprinted in 1981.

Boos, Daniel (2013). „Gemeinfreiheit – Enteignung oder Gewinn?“ In: Fabrikzeitung Nr. 288, January/February 2013.

Bundesamt für Kultur (2008). Memopolitik. Eine Politik des Bundes zu den Gedächtnissen der Schweiz. Bericht des Bundesamtes für Kultur.

Bundesamt für politische Bildung (26. October 2011). Open Data. Retrieved 23 April, 2012 from: <http://www.bpb.de/gesellschaft/medien/opendata/64055/was-sind-offene-daten>

Bundesarchiv (21. Januar 2011). Aufgaben. Retrieved 12 June 2012 from: <http://www.bundesarchiv.de/bundesarchiv/aufgaben/index.html.de>

D-Lib (June 5, 2012). D-Lib Magazine. Retrieved June 5, 2012 from: <http://www.dlib.org/dlib/march10/holley/03holley.html>

E-Government Schweiz (2012). Katalog priorisierter Vorhaben, Stand 24. Oktober 2012. http://www.egovernment.ch/dokumente/katalog/E-Gov-CH_Katalog_2012-24-10_D.pdf

Europeana (2012). Europeana 1914-1918. Retrieved June 4, 2012 from: <http://europeana1914-1918.eu/en>

European Commission, Information Society DG., & Salzburg Research (Firm) (2002). The DigiCULT report: technological landscapes for tomorrow's cultural economy: unlocking the value of cultural heritage: executive summary. Office for official publications of the European Communities.

European Commission (2006). Communication from the Commission on the Digitisation and Online Accessibility of Cultural Material and Digital Preservation. Brussels, 24.08.2006.

Goldfarb, Doron / Arends, Max / Froschauer, Josef (2012). Art History on Wikipedia, a Macroscopic Observation GoldWebSci 2012, June 22–24, 2012, Evanston, Illinois, USA. <http://www.ec.tuwien.ac.at/~dieter/research/publications/websci2012.pdf>

Golliez, André/Aschwanden, Cécile/Bretscher, Claudia/Bernstein, Abraham/Farago, Peter/Krügel, Sybil/Frei, Felix/Laux, Christian/Bucher, Bruno/Neuroni, Alessia/Riedl, Reinhard (2012) Open Government Data Study of Switzerland, Bern University of Applied Sciences, Bern.

Holley, Rose (2009). Crowdsourcing and social engagement: potential, power and freedom for libraries and users. Published 18 November 2009.

Howe, Jeff (June 2, 2006). "Crowdsourcing: A Definition". Crowdsourcing Blog. Retrieved January 24, 2013.

Kaltenböck, Martin / Thurner, Thomas (eds.) (2011). Open Government Data Weissbuch, Universitätsverlag der Donau-Universität Krems. <http://open.semantic-web.at/display/OGDW/Home>. Retrieved January 25, 2013.

Luyten, K., Coninx, K., Flerackers, E., Gabriels, K., Robert, K., Schroyen, J., & Teunkens, D. (2011). iDiscover: Towards the Next Generation of Contextualised Mobile Museum Guides.

National Library of Australia (June 5, 2012). National Library of Australia. Retrieved June 5, 2012 from <http://trove.nla.gov.au/>

Oomen, Johan / Aroyo, Lora (2011). Crowdsourcing in the Cultural Heritage Domain: Opportunities and Challenges. C&T' 11, 20 June - 2 July 2011, QUT, Brisbane, Australia.

Opendata network (December 10, 2009). Retrieved April 20, 2012 from: <http://opendata-network.org/2009/12/obama-online-gegen-die-burokratie/>



Owens, Trevor (2012). The Crowd & the Library. The Agony and the Ecstasy of “Crowdsourcing” our Cultural Heritage. IIPC Crowdsourcing Workshop, May 4, 2012.

Rogers, Everett (1962). Diffusion of Innovations, Free Press, London, NY, USA.

Schweizerische Eidgenossenschaft (2012). Strategie des Bundesrates für eine Informationsgesellschaft in der Schweiz. Eidgenössisches Departement für Umwelt, Energie, Verkehr und Kommunikation UVEK, März 2012.

The Great War Archive (04 June 2012). The Great War Archive. Retrieved June 4, 2012 from:
<http://www.oucs.ox.ac.uk/ww1lit/gwa>

Ton, S. (August 11, 2009). Interview with Susanne Ton from the Tropenmuseum about the Wiki loves art / NL project. (H. K. Rose Klaver, Interviewer). Retrieved 25 March, 2013 from:
http://commons.wikimedia.org/wiki/File:Wikilovesart_tropenmuseum_interview.ogv.

United States Government. (April 18, 2012). Open Government Data Sites. Retrieved April 18, 2012 from:
<http://www.data.gov/opendatasites>

VMS (2012). Museum visits in Switzerland. Statistical Report 2011. Association of Swiss Museums.

Wikimedia Commons. (February 12, 2012). Wikimedia Commons. Retrieved June 2, 2012 from:
<http://commons.wikimedia.org/wiki/Commons:Kooperationen>

Wikimedia Outreach (22 May 2011). GLAM/Case studies/German Federal Archives. Retrieved June 11, 2012 from: http://outreach.wikimedia.org/wiki/GLAM/Case_studies/German_Federal_Archives



6 Links for further information

Below you will find a number of links to further information on the subjects of Open Data, Crowdsourcing, Linked Open Data and Digital Humanities. The list is not exhaustive.

openglam.org

OpenGLAM is an initiative of the Open Knowledge Foundation to promote free public access to digital cultural heritage items held by libraries, archives and museums. The Open Knowledge Foundation is a UK-based non-profit foundation for the promotion of open knowledge.

<http://openglam.org/>

International GLAM-WIKI-portal

Contains case studies and best practice examples of collaborations between cultural heritage institutions and the Wikipedia community.

<http://glamwiki.org>

opendata.ch

Opendata.ch is the Swiss chapter of the Open Knowledge Foundation. The association's website contains information on current events relating to open data in Switzerland as well as the Swiss Open Government Data Manifesto.

<http://opendata.ch/>

Swiss Open Government Data Study

A study by the Bern University of Applied Sciences and its partners on the subject of "Open Government Data" in Switzerland.

http://www.wgs.bfh.ch/en/ueber_uns/news/newsdetails/article/studie-zu-open-government-data-in-der-schweiz.html

Linked Open Data: The Essentials. A Quick Start Guide for Decision Makers

An easy to understand introduction to the subject of Linked Open Data by Florian Bauer (REEEP) and Martin Kaltenböck (Semantic Web Company)

<http://www.semantic-web.at/LOD-TheEssentials.pdf>

Linked Open Data – What is it?

Introductory video on the subject of Linked Open Data by Europeana.

<http://www.youtube.com/watch?v=uju4wT9uBIA>

data.europeana.eu

The Linked Open Data portal of Europeana with an overview of the available datasets and practical examples developed in the framework of various Hackathons.

<http://data.europeana.eu>

LODLAM: Linked Open Data in Libraries, Archives, and Museums

Community of Practice for questions relating to "Linked Open Data" in cultural heritage institutions.

<http://lodlam.net/>

infoclio.ch

The professional portal of the historical sciences in Switzerland with a particular focus on developing the digital infrastructure for historical research in Switzerland.

<http://www.infoclio.ch/>



7 About the author of this study

Beat Estermann is a member of the research team dealing with 'Open and Linked Data' at the E-Government Institute of Bern University of Applied Sciences which has contributed to the Swiss Open Government Data Study (Golliez et al. 2012), advises various Swiss authorities on the subject of 'Open Government Data' and plays a leading role in research and development in the field of 'linked open data' within the European Union's "Fusepool"²² project.

Beat Estermann has studied commons-based peer production and crowdsourcing for a number of years. Since 2012, he has focused on open data and crowdsourcing in relation to cultural heritage institutions. He cooperates closely with Wikimedia CH²³, the national association for the promotion of free knowledge recognised by the Wikimedia Foundation. He and other members of Wikimedia CH cooperate with libraries, archives and museums that wish to develop a collaborative relationship with Wikipedia/Wikimedia. In addition, he is a member of 'opendata.ch'²⁴, the Swiss chapter of the 'Open Knowledge Foundation'²⁵, and a founding member of the 'Digitale Allmend'²⁶ association that acts as the Swiss chapter of 'Creative Commons'²⁷.

²² <http://www.fusepool.eu/>

²³ <http://www.wikimedia.ch>

²⁴ <http://www.opendata.ch>

²⁵ <http://okfn.org/>

²⁶ <http://www.allmend.ch>

²⁷ <http://www.creativecommons.ch/>, <http://creativecommons.org>



Appendix: Questionnaire²⁸

Survey: Cultural heritage institutions in the internet era

Dear Sir/Madam,

Cultural heritage institutions are developing and transforming – especially with regard to their online activities. Making data ‘freely’ available on the internet and new forms of online collaboration represent not only numerous opportunities for archives, libraries and museums but also many new challenges.

Some cultural heritage institutions have already managed to use the new culture of online collaboration to their advantage and are thus able to use online communities in the development of their programmes thanks to innovative approaches. In most cases, however, such concepts are still in the planning and testing phase and institutions want to establish more clearly how they can best adapt to the new conditions.

Conversely, the Wikipedia community, one of the largest communities based on online collaboration, has continuously been extending its range of partnerships with cultural heritage institutions over recent years. Its objective is to acquire new content for its online encyclopaedia and to adopt new approaches together with cultural heritage institutions for promoting and conveying free knowledge.

Bern University of Applied Sciences is undertaking a pilot survey among cultural heritage institutions to gain an initial overview of how they are adapting to new trends and in order to develop tailor-made support programmes as and where needed. The results of the survey will be made available to the general public.

We would be delighted if you would take a few moments to answer the questions. Your opinion is important to us.

Best regards,

Beat Estermann, research associate, Bern University of Applied Sciences

Contact details

Bern University of Applied Sciences
E-Government Institute

Beat Estermann
beat.estermann@bfh.ch

+41 31 848 34 38

²⁸ The survey was based on an online questionnaire in German. Only one institution made use of the option to complete the questionnaire offline.



Notes on completing the questionnaire

Many of the questions relate to **“your institution”**. If you are unsure as to which organisation you should refer to in your answers, please choose the one that seems most appropriate to you. You can always add explanatory notes at the end of the survey if necessary.

Data protection policy

Please note that provision of your contact details at the end of the survey is not mandatory. Bern University of Applied Sciences will treat your answers in confidence. Neither Bern University of Applied Sciences nor third parties will contact you as a consequence of your participation in the survey unless you explicitly invite us to do so.

Thank you for your cooperation!

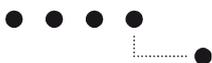


1. What types of data/objects/books/records are characteristic for your institution?

Type/Description	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Books, periodicals	<input type="checkbox"/>				
Manuscripts, autographs	<input type="checkbox"/>				
Documents, records	<input type="checkbox"/>				
Images, photographs, prints	<input type="checkbox"/>				
Objects of art	<input type="checkbox"/>				
Craft artefacts	<input type="checkbox"/>				
Natural history artefacts	<input type="checkbox"/>				
Technical artefacts	<input type="checkbox"/>				
Audio documents	<input type="checkbox"/>				
Film documents	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

2. What are the main activities of your institution?

Main activities with regard to data/objects/books/records...	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Collecting	<input type="checkbox"/>				
Archiving	<input type="checkbox"/>				
Exhibiting	<input type="checkbox"/>				
Lending or giving access to individuals	<input type="checkbox"/>				
Lending to other institutions	<input type="checkbox"/>				
Preparing, indexing, documenting	<input type="checkbox"/>				
Restoring, conserving/preserving	<input type="checkbox"/>				
Digitising	<input type="checkbox"/>				
Research	<input type="checkbox"/>				



Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

3. Who are the users of your institution?

Users	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Private individuals (general public)	<input type="checkbox"/>				
Educational institutions (school classes, students, teachers, lecturers)	<input type="checkbox"/>				
Public authorities	<input type="checkbox"/>				
Cultural institutions	<input type="checkbox"/>				
Private enterprises	<input type="checkbox"/>				
Research institutions/specialists	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

4. How many people work for your institution? (approximate figures are sufficient)

When stating full-time equivalents, please indicate the number of full-time positions made up by part-time positions (e.g. two 60% positions = 1.2 FTEs). Please state the approximate long-term average if figures vary significantly.

Description	Number of people	Approximate number of full-time equivalents
Permanent employees (including part-time)	____ people	_____FTEs
Temporary employees	____ people	_____FTEs
Trainees/interns	____ people	_____FTEs
Volunteers	____ people	_____FTEs



5. What are the sources of your institution's revenue? (approximate figures are sufficient)
 Figures in % of total revenue

Revenue sources	Proportion of total revenue in %
Institutional funding from public funds	Approx. _____%
Institutional funding from private funds (non-profit foundations, etc.)	Approx. _____%
Sponsorship	Approx. _____%
Large donations (patrons, benefactors, bequests)	Approx. _____%
Small donations, including crowdfunding	Approx. _____%
Revenue from entrance fees	Approx. _____%
Revenue from lending fees	Approx. _____%
Other revenue from commercial operations (events, cafeteria, book sales, merchandising, venue rentals, courses, digitisation assignments and similar)	Approx. _____%
Project funding (public or private funds)	Approx. _____%
Sale of image rights	Approx. _____%
Other:	Approx. _____%

6. What is your institution's legal form?

Legal form	Applies
Institution under public law	<input type="checkbox"/>
Non-profit institution under private law	<input type="checkbox"/>
Profit-oriented institution under private law	<input type="checkbox"/>
Other:	<input type="checkbox"/>

7. Are open data (making data 'freely' available on the internet) and collaborative content creation important to your institution?

Sector	Very important	Important	Neutral	Un-important	No response
Open data	<input type="checkbox"/>				
Collaborative content creation	<input type="checkbox"/>				



8. Do you exchange metadata with other institutions? (Metadata provide a description of the data/objects/books/records held by your institution.)

Yes No

If yes:

8a) Which metadata formats do you use? (You may select more than one of the following)

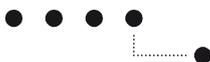
- AACR2
- Dublin Core
- EAD
- ISAD(G)
- LIDO
- MARC 21
- MAB
- PND
- SEPIADES
- Other: _____
- Not applicable

8b) What is the relevance of the exchange of metadata?

The exchange of metadata is important to us for:	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Bilateral cooperation	<input type="checkbox"/>				
Multilateral cooperation (networks)	<input type="checkbox"/>				
Fulfilling our core objective	<input type="checkbox"/>				
Generating revenues	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

9. Do you think there is a need for improvement with regard to the quality or interoperability of metadata in your institution?

Need for improvement with regard to...	Urgent need for improvement	Need for improvement in the medium term	No need for improvement	Not applicable
Quality of metadata (accuracy, completeness, up-to-dateness, clarity, availability)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Interoperability of metadata (digitisation, conformity with standards)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



9a) If there is a need for improvement, where exactly do you see a need for improvement to metadata?

Need to improve the quality or interoperability of metadata	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Accuracy	<input type="checkbox"/>				
Completeness	<input type="checkbox"/>				
Up-to-dateness	<input type="checkbox"/>				
Clarity	<input type="checkbox"/>				
Availability	<input type="checkbox"/>				
Digitisation	<input type="checkbox"/>				
Conformity with common interchange standards	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

10. Do you make the data of your objects/books/records available on the Internet?

Description	Applies	Tends to apply	Tends not to apply	Does not apply	Not applicable
Metadata available on the internet	<input type="checkbox"/>				
Images of the data/objects/books/records are available on the Internet (digitised copies, photos)	<input type="checkbox"/>				
Background information on the data/objects/books/records is available on the Internet (results of research, analyses, etc.)	<input type="checkbox"/>				

11. The term 'open data' is used to refer to data that is made available to the general public free of charge in a machine-readable format for unrestricted use. Why does your institution think open data is necessary?

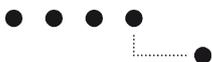
Why does your institution think open data is necessary?	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
to make content available to private individuals (general public)	<input type="checkbox"/>				
to make content available to educational institutions	<input type="checkbox"/>				
to make content available to public authorities	<input type="checkbox"/>				



to make content available cultural institutions	<input type="checkbox"/>				
to make content available to private enterprises	<input type="checkbox"/>				
to make content available to research institutions / domain specialists	<input type="checkbox"/>				
to facilitate networking among heritage institutions	<input type="checkbox"/>				
to make the institution's holdings more visible	<input type="checkbox"/>				
to improve the visibility of the institution	<input type="checkbox"/>				
to allow the institution to better fulfill its core mission	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

12. What risks does your institution see in open data?

Risks associated with open data	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Time effort and expense for making them available (including selection and content checks)	<input type="checkbox"/>				
The use of the data cannot be controlled	<input type="checkbox"/>				
Copyright infringements	<input type="checkbox"/>				
Infringements of data protection regulations	<input type="checkbox"/>				
Divulgence of classified information	<input type="checkbox"/>				
Increased time effort in order to respond to enquiries	<input type="checkbox"/>				
Loss of revenues	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				



13. The trend among cultural heritage institutions is to make images of/information on their holdings freely available on the Internet.

Under what conditions would you make images of/information on your holdings freely available on the Internet without receiving payment in exchange? (This assumes that your holdings are already available in a digitised format and publication would not violate copyright or confidentiality regulations)

Scope of use/conditions	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
For private use	<input type="checkbox"/>				
For education and research	<input type="checkbox"/>				
For non-profit projects	<input type="checkbox"/>				
For non-profit projects such as Wikipedia that also permit commercial use	<input type="checkbox"/>				
For commercial users	<input type="checkbox"/>				
Only if the name of the institution remains linked to the data	<input type="checkbox"/>				
Only if the works are used without modification	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

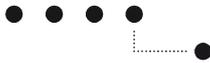
14. Does your institution have experience with alternative licenses?

(several responses possible)

- No
- Yes, with Creative Commons licenses
- Yes, with the GNU General Public License (GPL)
- Yes, with the Free Art License
- Yes, with other licenses

15. What percentage of your holdings dates to pre-1850?

- 81-100 %
- 61-80 %
- 41-60 %
- 21-40 %
- 0-20 %



16. Is the Linked Data/Semantic Web concept of interest to your institution?

Description	Yes, some projects have already started	Yes, we have already planned projects	Yes, it is of interest but we have not yet planned any projects	No, it is not of interest	Not applicable
Is the Linked Data/Semantic Web concept of interest to your institution?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

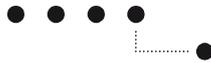
17. What types of volunteer work play an important role in your institution?

The following types of volunteer work play an important role in your institution:	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Voluntary work by individuals without qualifications	<input type="checkbox"/>				
Voluntary work by individuals with qualifications	<input type="checkbox"/>				
Voluntary work in the online sector	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

18. Public and private organisations including cultural heritage institutions are outsourcing various tasks to volunteers via the internet (crowdsourcing).

What are the opportunities related to crowdsourcing for your institution?

Opportunities related to crowdsourcing?	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Correction and transcription tasks	<input type="checkbox"/>				
Improving and expanding texts	<input type="checkbox"/>				
Supplementation of collections (contributing additional objects or facilitating their acquisition)	<input type="checkbox"/>				
Classification/supplementation of metadata	<input type="checkbox"/>				
Co-curators	<input type="checkbox"/>				
Crowdfunding (financing)	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				



19. Where do you see the risks of crowdsourcing?

Risks of crowdsourcing?	Applies	Applies in part	Tends not to apply	Does not apply	Not applicable
Little influence on results	<input type="checkbox"/>				
Extensive preparation and follow-up required	<input type="checkbox"/>				
Difficult to estimate the time scope	<input type="checkbox"/>				
The continuity of data maintenance is not guaranteed over the longer term	<input type="checkbox"/>				
Minimal security for planning	<input type="checkbox"/>				
Anxiety among employees (loss of job, changes to roles and tasks, etc.)	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				
Other:	<input type="checkbox"/>				

20. Are members of your staff involved in organisations or projects that support open data or collaborative projects on the internet?

Organisations/projects	Involvement as part of professional activities	Involvement during leisure time	not known
Wikipedia	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Wikimedia Commons	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Flickr Commons	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other:	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

21. Would you like to receive further information on open data and crowdsourcing or collaborative projects on the internet?

Type of information	Yes	No
Consultation without obligation	<input type="checkbox"/>	<input type="checkbox"/>
Informative event	<input type="checkbox"/>	<input type="checkbox"/>
Information brochure	<input type="checkbox"/>	<input type="checkbox"/>



21a) If you would like to receive an information brochure, request a consultation without obligation or attend an informative event, please give your contact details below and let us know what aspects most interest you.

Contact details	
Institution	
Surname	
First name	
Street	
Postal code	
City	
Country	
Telephone number	
Email	

Description of the aspects that most interest you

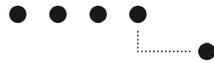
Do you consent to our evaluation of the information you have provided in the survey specifically with regard to your institution to optimise any future information and consultation offers? (Bern University of Applied Sciences will not publish individualised data or make it available to third parties.)

- Yes
- No

22. Feel free to make any remarks or comments on the survey in the space provided below.

Remarks and comments

Thank you for participating.



Bern University of Applied Sciences, Business Division

Morgartenstrasse 2c, Postfach 305

3000 Bern 22, Switzerland

T +41 31 848 34 00 F +41 31 848 34 01

wirtschaft@bfh.ch

www.wirtschaft.bfh.ch