LoCloud

Local content in a Europeana cloud

D1.1: Report on the State-of-the Art Monitoring and Situation Analysis

Authors:
Kristine Hoff Meyer (KUAS)
Mikkel Christoffersen (KUAS)
Henk Alkemade (RCE)
Maria Luisa Martinez-Conde (MECD)
Rimvydas Lauzikas (VUFK)
Ingrida Vosyliute (VUFK)
Martin Krajnak (PriUF KAEG)
Tatiana Shamarina-Heidenreich (UDE)
Adam Dudczak (PSNC)

Revision: 2

LoCloud is funded by the European Commission's
ICT Policy Support Programme

cip competitiveness and innovation framework programme 2007–2013

Revision History

Revision	Date	Author	Organisation	Description
1	13-08-2013	Kristine Hoff Meyer Mikkel Christoffersen Henk Alkemade Maria Luisa Martinez-Conde Rimvydas Lauzikas Ingrida Vosyliute Martin Krajnak Tatiana Shamarina-Heidenreich	KUAS KUAS RCE MECD VUFK VUFK PriUF KAEG UDE	1st version of the report
2	05-11-2013	Kristine Hoff Meyer Adam Dudczak	KUAS PSNC	The report was updated after review by Adam Dudczak and with information from the Europeana Cloud deliverable 2.2 "Europeana Cloud Architectural Design".

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Contents

Executive Summary

Cloud computing has become ubiquitous, but the concept has no strict definition. Ideally, cloud computing is meant to turn computing into a utility like water or power. Elasticity, availability, improved resource utilisation and support for multiple tenants are key features of the concept. There are three main models of service: Infrastructure as a service (Iaas), Platform as a service (Paas) and Software as a service (Saas).

Cloud computing may aid heritage institutions with its oft-cited benefits such as cost effectiveness, quick deployment and access to resources beyond the abilities of individual small institutions. Developers of cloud-based services in the heritage sector should distinguish between the three user groups: content providers & aggregators, the general public, and scholars.

Though cloud computing is still emerging, a stamp of approval is that The European Commission has adopted a cloud computing strategy based on the reports from expert working groups and open consultations. It was adopted in September 2012 and is part of the "Digital Agenda for Europe."

There is high awareness and willingness to participate in cloud-based development from the heritage institutions and agencies voicing their opinion in this report. The barriers to participating cited are mainly lack of knowledge and skills, trust and legal issues. The main legal obstacle is the fact that many institutions are charged with the governance of their data and there will often be restrictions as to where that data may be placed and to whom it may be given. It lies at the heart of cloud computing that the customer may not know exactly where the data resides.

There are a number of Saas providers providing services for the cultural sector. Some of the commercial vendors of collections management systems offer cloud based versions of their software, and in the library domain the OCLC offers a number of relevant services. However, none of these come with plug-in aggregation tools for Europeana.

There is probably still a need for online tools with a very low barrier to entry which are suited to the needs (and budgets) of smaller local and community museums. This is the window of opportunity for the LoCloud project. The LoCloud project builds on past successful projects such as Europeana Local and CARARE and aims to bring the benefits of cloud computing to especially small- and medium-sized cultural institutions to aid them in aggregating their data to Europeana. The LoCloud project will also cooperate with other Europeana related projects especially Europeana Cloud.

Introduction to this document

The purpose of this report is to monitor the state-of-the art of cloud computing and make an assessment of aspects of the cloud relevant to the needs of the project and to small and medium sized institutions. The report is supposed to inform content providers in their further action planning.

The situation analysis has been conducted by The Danish Agency for Culture, Rijksdienst voor het Cultureel Erfgoed in the Netherlands, The Spanish Ministerio de Educación, Cultura y Deporte, Vilniaus Universitas in Lithuania, Universitaet Duisburg-Essen in Germany and Univerzita Komenskeho v Bratislave in Slovakia.

The first section of the report offers a general description of cloud computing, the different kinds of infrastructure and models of service available, and the advantages and potential risks associated with the technology, in particular legal issues of data storage.

The second section offers a background for the LoCloud project. It presents an introduction to the uptake of cloud computing by small and medium-sized enterprises in the EU and the barriers that exist. It also presents a brief overview of European policy regarding cloud computing, and an analysis of the potential for cloud computing in the cultural heritage sector.

In *the third and final section*, special attention is paid to the needs of the LoCloud project and to small and medium sized cultural heritage institutions. It presents three situation reports from Spain, Holland and Denmark. It looks into existing cloud based collection management and digital library services and offers a description of the relation of LoCloud to the Europeana Cloud project.

The methodology for writing the report is primarily desktop research and analysis of the available literature.

Section 1: The concept of Cloud Computing

Characteristics

The term *Cloud computing* is used to describe a variety of computing services or resources handled by off-site computers, often virtual, connected by a network, and owned and operated by a third party.

The term 'cloud computing' has no strict definition. An early but visionary description of cloud computing was formulated by a group of scientists from the University of Melbourne and Vienna University of Technology in 2009. They defined cloud computing as a model of virtual services

> *"...that are commoditized and delivered in a manner similar to traditional utilities such as water, electricity, gas, and telephony. In such a model, users access services based on their requirements without regard to where the services are hosted or how they are delivered...".*[1]

Such a service-oriented definition determines that ordinary computing as the whole of user-owned hardware and software is transformed into a different type of computing model that can be described as a unity of user-owned hardware and hard-service and soft-service acquired by a user via the Internet - even though hardware may also to a large degree become a utility.

The report "Advances in clouds" from 2012 spends no less than eleven pages discussing what distinguishes cloud computing from other internet based services, and concludes, that elasticity, availability, improved resource utilisation and support for multiple tenant are *intrinsic* capabilities of cloud systems, as opposed to *extrinsic* characteristics like cost-effectiveness, virtualisation and use of internet connectivity. [2]

---

1 Rajkumar Buyya, Chee Shin Yeo, Srikumar Venugopal, James Broberg and Ivona Brandic. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. In Future Generation Computer Systems Volume 25, Issue 6, June 2009, pages 599–616.

[2] Advances in Clouds. Research in Future Cloud Computing . Expert Group Report . Public version 1.0 . Editors: Lutz Schubert [USTUTT-HLRS], Keith Jeffery [STFC]. 2012 [interactive]. Internet access: <http://cordis.europa.eu/fp7/ict/ssai/docs/future-cc-2may-finalreport-experts.pdf >.

Table 1 provides an overview over cloud computing characteristics and their classification.[3] The intrinsic capabilities are highlighted in bold.

Technical	**Business / Economic**	**Social / Legal**	**Other**
Elasticity / Scalability Virtualisation Agility & Adaptability **Availability** Data Management Reliability Programmability	Outsourcing Pay per use **Resource utilisation** Energy efficiency Cost efficiency Metering	Security Provenance Privacy	**Multi-Tenancy** Ease of Use

Table 1. Overview over Cloud Computing characteristics and their classification.

Elasticity (or scalability)

Automatic or manual resource allocation based on user needs. A suddenly very popular service can satisfy demand by deploying thousands of servers running concurrently. When demand subsides the superfluous servers will automatically be decommissioned. This characteristic is most important for services experiencing infrequent usage peaks where a traditional on-site infrastructure would demand expensive over-capacity for the long periods of time where usage is low. They can "turn on the tap", so to speak.

Availability

Cloud computing providers typically offer to deploy services running in data centres located in different parts of the world. Even if an entire data centre is down, customers can continue

[3] Table by: Advances in Clouds . Research in Future Cloud Computing . Expert Group Report . Public version 1.0 . Editors: Lutz Schubert [USTUTT-HLRS], Keith Jeffery [STFC]. 2012 [interactive]. Internet access: <http://cordis.europa.eu/fp7/ict/ssai/docs/future-cc-2may-finalreport-experts.pdf >.

operating their services in a different data centre. This characteristic is important for services requiring high availability and it makes cloud computing very cost effective compared to owning and operating multiple data centres.

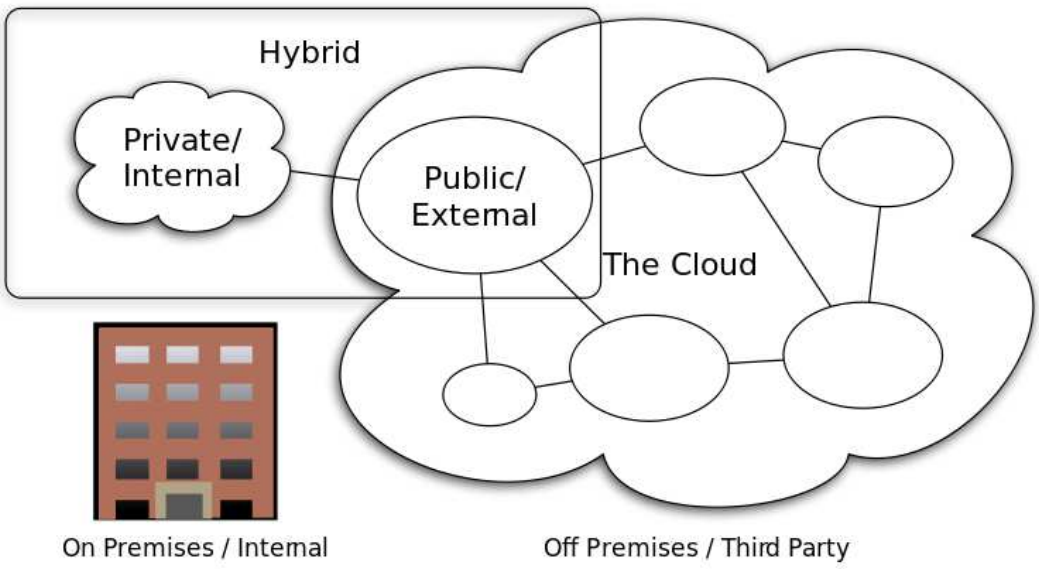
Improved resource utilization and support of multiple tenants

Servers and storage can be shared to increase utilization thus making the hardware more cost-effective for cloud computing providers. This also provides for easy scaling of services.

Cloud infrastructure

There are three different types of cloud infrastructure, distinguished by who hosts the cloud and the level of security offered.

- **Public Cloud**. The cloud infrastructure is made available to the general public or multiple organisations on a shared basis. It is typically owned by a provider selling cloud services.
- **Private Cloud**. The cloud infrastructure is operated solely for an organisation. It is managed by the organisation or a third party and may exist on or off premises.
- **Hybrid Cloud**. A cloud infrastructure composed of two or more clouds, private or public, that remain unique entities but are bound together by standardized technology that enables data and application portability. Hybrid cloud architecture requires both on-premises resources and off-site (remote) server-based cloud infrastructure.



Figur 1. Cloud infrastructure. CC-BY-SA 3.0 by Sam Johnston. Source Wikipedia.

It is important to note that there is not much difference between a private cloud and a public cloud from the perspective of the cloud characteristics, other than a private cloud can theoretically be more secure than a public cloud. Private clouds on the other hand, have attracted criticism because users still have to buy, build, and manage them, and thus do not benefit from less hands-on management.

According to the deliverable 2.2 “Europeana Cloud Architectural Design” The Europeana Cloud project has settled on a Hybrid Cloud solution consisting of a private, community-based part where the necessary hardware resources will be provided by voluntary, technically advanced institutional users of the eCloud system and a public part, based on resources leased from commercial providers. Thus the eCloud system will benefit from the security of a private cloud, and the rapid scalability of a commercial provider.

Models of Service

Cloud computing provides different models of service that the customer can choose from. As illustrated below, the different models of service have different target groups, and require different levels of IT expertise and knowledge.

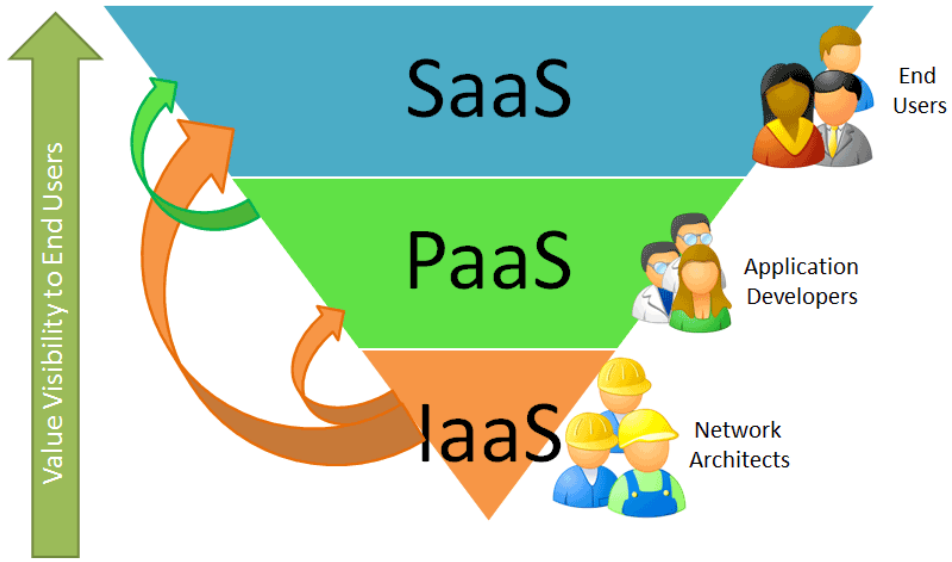


Figure 2. The user types connected with different models of service. [4]

Infrastructure as a Service (IaaS)

The basic form of cloud computing. This is essentially bare servers where the customer has to install operating systems and applications. This model fits IT departments wanting to either scale applications cheaply or test new applications before committing to an infrastructure investment. IaaS is the most flexible model for the customer but also the model requiring the most expertise. Typical examples are Amazon EC2, Google Compute Engine and Rackspace. The Europeana Cloud system will also offer IaaS services for storage of cultural data records.

[4] Source: http://www.saasblogs.com/saas/demystifying-the-cloud-where-do-saas-paas-and-other-acronyms-fit-in/

Platform as a Service (PaaS)

Paas is the next step up from IaaS. This model gives the customer a running server with base applications installed, typically a web server and a database. The customer can deploy their own applications on top of the PaaS. It is suited for small to medium software development businesses with little or no internal IT operations. Typical examples are Amazon Elastic Beanstalk, Google App Engine, Redhat Open Shift and Windows Azure.

Software as a Service (SaaS)

Cloud computing companies provide access to specific applications and charge users for use. Customers need no IT skills and usually only a browser to access and operate the SaaS, simplifying support to the point where no IT department is needed.

Examples of services available in the Cloud are listed below[5]:

- Financial Management and Accounting packages
- Customer Relationship Management (CRM) solutions
- Project Management
- Document creation and management tools
- Email and Calendar tools
- Collaboration and virtual meeting Tools
- Cloud Storage Tools

All of these tools might be as relevant for small and medium cultural heritage institution as for small businesses. Below is given an example of a museum, MOMA, which has taken up cloud computing in order to replace legacy business systems.

However, more relevant to the LoCloud project is to evaluate the offerings out there when it comes to cloud based collections management systems catering for the GLAM sector, as the LoCloud project aims to establish its own SaaS service to small and medium sized cultural heritage institutions. In sector three of this report, we will return to this point and offer a closer look at systems like Duracloud, Omeka.net, Vernon systems, Zetcom, vesica and OCLC.

Last but not least the Europeana Cloud will also offer SaaS solutions for the GLAM sector, and we will also have a closer look at the eCloud System in section three.

---

5 Australian Government Initiative, Enterprise Connect. Using the "Cloud" to support your business. Quick guide http://www.enterpriseconnect.gov.au/ecservices/tkc/Documents/TKC-ITGuides/Using-the-Cloud.pdf , www.enterpriseconnect.gov.au

Example of a museum using cloud SaaS services

Museums and other cultural institution may get great benefits from using standard SaaS solutions to replace expensive commercial software or antiquated home grown legacy systems. An example of a museum which has taken up cloud computing is the Museum of Modern Art in New York (MOMA).

Juan Montes and Diana Pan have been the driving force behind an IT department that is impacting business change, all while keeping the museum's primary mission in mind. MOMA has switched from using Microsoft Exchange to using the cloud based gmail, and they have replaced their antiquated, home grown Customer Relationship Management System with a cloud based alternative. The reasoning behind the switch is that the museum's goal is not to build solutions and host servers, but to bring modern art to as broad an audience as possible.

Juan Montes explains the benefits:

> *We've been able to redeploy our team, previously dedicated to maintaining our internal systems, to take care of things that are more relevant to the museum, like our digital asset and collection management systems and to focus on a strategy for keeping video assets in such a way that they can be deployed easily.*

Diana Pan goes on to argue, that the timing and the need for change in the museums IT infrastructure came along at a time when the technology was mature enough for MoMA to reliably take advantage of it:

> *We probably wouldn't have made the same decision 3-5 years ago. But the technology has evolved, so we aren't really pioneers at all. We were just looking for best solution and that happened to be cloud.* [6]

6 Customer Spotlight - Q&A with Juan Montes and Diana Pan, MoMA, Tuesday, November 13, 2012 http://blog.appirio.com/2012/11/customer-spotlight-q-with-juan-montes.html

Advantages of Cloud Computing

If used properly and to the extent necessary, working with data in the cloud can vastly benefit all types of businesses. Mentioned below are some of the advantages of this technology. [7]

Cost Efficient

Cloud computing is probably the most cost efficient method to use, maintain and upgrade. The up-front costs of establishing an IT infrastructure are greatly reduced. There is no need to procure, install and configure hardware or software.

Traditional desktop software costs companies a lot in terms of finance. Adding up the licensing fees for multiple users can prove to be very expensive for the establishment concerned. The cloud, on the other hand, is available at much cheaper rates and hence, can significantly lower the company's IT expenses. Besides, there are many one-time-payments, pay-as-you-go and other scalable options available.

Almost Unlimited Storage

Storing information in the cloud gives you almost unlimited storage capacity. Hence, you no longer need to worry about running out of storage space or increasing your current storage space availability.

Backup and Recovery

Since all your data is stored in the cloud, backing it up and restoring the same is relatively much easier than storing the same on a physical device. Furthermore, most cloud service providers are usually competent enough to handle recovery of information. Hence, this makes the entire process of backup and recovery much simpler than other traditional methods of data storage.

Downtime and delays are decreased, which results in improved business continuity. Cloud services typically guarantee available network connectivity.

---

[7] Cloud Computing – Is it Really All That Beneficial? Advantages and Disadvantages of Cloud Computing http://mobiledevices.about.com/od/additionalresources/a/Cloud-Computing-Is-It-Really-All-That-Beneficial.htm

Automatic Software Integration

In the cloud, software integration is usually something that occurs automatically. This means less additional efforts to customize and integrate your applications as to your preferences. This aspect usually takes care of itself. Not only that, cloud computing allows you to customize your options with great ease. Hence, you can handpick just those services and software applications that you think will best suit your particular enterprise.

Easy Access to Information

Once you register yourself in the cloud, you can access the information from anywhere, where there is an Internet connection. This convenient feature lets you move beyond time zone and geographic location issues.

Quick Deployment

Lastly and most importantly, cloud computing gives you the advantage of quick deployment. Once you opt for this method of functioning, your entire system can be fully functional in a matter of a few minutes. Of course, the amount of time taken here will depend on the exact kind of technology that you need for your business.

Computing capability can also be rapidly expanded on demand. This is important for businesses with variable workload or where technology capability needs to support business growth.

Potential risks of Cloud Computing

The many benefits of cloud computing, especially the cost, can seem tempting but some issues need to be taken into account before committing to using the various services cloud computing offers:

Privacy

Data is stored off site on the storage systems of a third party entity. That third party can always gain access to this data unless specific care is taken to protect it using e.g. encryption. Communication between the cloud systems and the customer's users and/or local systems is handled by the cloud provider's network which means that transmitted data is available to the cloud provider unless it is encrypted. Sensitive data might not belong in the cloud.

Businesses must have confidence that their data is secure and that their privacy and that of their customers is properly managed. This means having assurance that the Cloud provider has both the infrastructure (encryption, firewalls, authentication and authorisation mechanisms) and policies and procedures in place to guarantee that a business’s data is only accessible by authorised personnel within the business. Businesses also need to be aware of their own privacy obligations and ensure the protections offered by a Cloud vendor enable them to meet these.

Compliance

Cloud computing providers can operate in a different country than the customer resides in. Data can be distributed to and stored in different parts of the world. Laws governing the data can be quite different from what the customer is expecting.

Commitment

When a customer commits all or most of their services to a specific cloud provider, they essentially base their company on that cloud provider and become dependent upon its future existence. Should it go out of business, customers can potentially lose everything running and stored in the cloud providers systems. Commitment also makes migration to another cloud provider difficult or impossible, thus making pricing an important factor to monitor.

Reliability

A cloud computing provider can experience technical difficulties like any other company running an IT infrastructure. Companies committing to a cloud provider must assess the consequences of down time and plan accordingly.

Security

Cloud computing providers are often targets of attacks because of their size and the amount of data they store. Should a cloud company somehow be compromised, attackers could gain access to customer data.

If the security of the cloud is compromised potentially all services and associated data within the cloud are at risk. That means that the data centre must be architected with security in mind and it must be considered a priority for every application, service, and network infrastructure solution that is deployed.

The application delivery solution, as the control node in the data centre, is necessarily one of the first entry points into the cloud data centre and itself must be secure. It should also provide full application security in order to contravene potential attacks at the edge. Network security, protocol security, transport layer security and application security should be prime candidates for implementation at the edge of the cloud, in the control node. While there certainly will be, and should be, additional security measures deployed within the data centre, stopping as many potential threats as possible at the edge of the cloud will relieve much of the risk to the internal service infrastructure.

Quality of Service

Businesses need to ensure that the service they receive from a Cloud vendor enables them to meet the service quality expectations of their customers, i.e. service reliability and availability, access to data when required, privacy and security.

Technical Issues: Data access, interoperability and portability

Businesses must have assurance that they can access their data, retain and move it if required (e.g., if they wish to move to a different vendor). If this is not possible, businesses risk being 'locked in' to vendors that no longer meet their needs, who fail to remain competitively priced or who have a poor performance record.

Alhough it is true that information and data on the cloud can be accessed anytime and from anywhere at all, there are times when this system can have some serious dysfunction. You should be aware of the fact that this technology is always prone to outages and other technical issues. Even the best cloud service providers run into this kind of trouble, in spite of keeping up high standards of maintenance. Besides, you will need a very good Internet connection to be logged onto the server at all times. You will invariably be stuck in case of network and connectivity problems.

Special funding and legal issues for cultural heritage institutions

Small and medium sized cultural institutions are no different from other small businesses in some respects. They could greatly benefit from using cloud based services (like email, document management, and collections management), thus reducing the need for in-house IT expertise.

However, funding may be a special issue for cultural institutions, as the funding mechanisms of the cultural sector may prevent institutions from taking advantage of the cost-efficient cloud services. Hardware used in public institutions is often bought from project related grants. It is a completely different funding model than the one commonly used in cloud services, where you pay as you go. If organisations should decide to use cloud based solutions, the funding mechanisms must be changed accordingly.

They would also face the same risk issues as any other business. However, legal issues regarding where the data is stored may be a special issue for cultural institutions.

There is a need to dwell for a moment on the legal issues associated with cloud computing that arise from the very nature of the cloud. When using cloud computing the data owner may not really know exactly where the data resides and where the software that performs operations on the data resides. This may need to be known, because it may, for example, be illegal for the data owner to send the data outside the European Union.

Some public institutions are prohibited from keeping any data on external servers. For example, this is the case for some of the public institutions in UK. Some institutions have special regulations related to geo distribution of their data. For example, the National Library of Israel, which has a policy that restricts the geographical location where data are stored, indicated in an Interview conducted by the Europeana Cloud Project, that if data are replicated across different systems, so that there is never a full copy of an object at a single location, this might be a solution to overcome this issue.

However, if the data contains personal data about individuals rules are particularly restrictive. If the data includes personal data, European Union member states are bound by Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Being a directive, member states have implemented national versions of the directive which must adhere to the directive's wording and meaning and principles. Personal data are defined in the directive as:

> *[…] any information relating to an identified or identifiable natural person ("data subject"); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.*

The directive governs the processing of the data regardless of whether the data processing agent (the data controller) resides within or outside the EU itself. The directive was written before the breakthrough of the Internet and is in any case being superseded by the new European Union Data Protection Regulation, a draft of which was presented in January 2012, and which makes it clear that the regulations cover all processing of EU data no matter where or how it may take place.

Responsibility for compliance with the directive lies with the controller.

Personal data may be processed if and only if:

- the data subject has given his consent
- the processing is necessary for the performance of or the entering into a contract
- the processing is necessary for compliance with a legal obligation
- the processing is necessary in order to protect the vital interests of the data subject
- it is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed
- it is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject.

The data subject has the right to access all data processed about him. The data subject even has the right to demand the rectification, deletion or blocking of data that is incomplete, inaccurate or is not being processed in compliance with the data protection rules.

The regulation of transferring data to third parties outside of the EU is such that the data can only be transferred if the receiving country employs similar protection as the European directive. Dealings between the United States of America and the European Union countries have been streamlined with the International Safe Harbor Privacy Principles negotiated by the US and the EU. US companies may be certified to handle the data in question if they meet the seven principles below:

- Notice: Individuals must be informed that their data is being collected and about how it will be used.
- Choice: Individuals must have the ability to opt out of the collection and forward transfer of the data to third parties.
- Onward Transfer: Transfers of data to third parties may only occur to other organizations that follow adequate data protection principles.
- Security: Reasonable efforts must be made to prevent loss of collected information.
- Data Integrity - Data must be relevant and reliable for the purpose it was collected for.
- Access: Individuals must be able to access information held about them, and correct or delete it if it is inaccurate.
- Enforcement: There must be effective means of enforcing these rules.

If no personal data is involved, it is much easier for the institutions to comply with relevant legislation, but the overall picture may become more muddled, because national legislations may and do differ as do the types of relevant institutions.

Generally, the location where processing of any kind takes place is also the location whose legislation regulates the processing. This may cause problems for two major reasons.

First, it may in itself be illegal for the institution to let go of its data and hand it over to third parties and third party hardware and software by the nature of the governance entrusted to the institution. Second, the jurisdiction of the destination of the data may allow for technical or other operations not allowed for by the data owner's jurisdiction. For example, data residing in the US is subject to the provisions of the Patriot Act.

The big players in the cloud market have tended to respond to these obstacles by erecting server parks in several major jurisdictions. Amazon, for instance, offers both the US and the EU (Ireland) for its EC2 service.

Section 2: Policy and Potential – background for the LoCloud project

Potential for up-take of cloud computing by SMEs in the EU

According to the report "Quantitative Estimates of the demand for Cloud Computing in Europe and the Likely Barriers to uptake", the market for public cloud services is growing strongly in Europe. However, overall adoption and usage of cloud services remain somewhat behind the United States. Small and medium sized enterprises (SMEs) have been slower to take up cloud computing than large enterprises:

> *Companies in all vertical markets and company sizes will increasingly rely on public cloud services. Some differences will however apply. Large enterprises (over 250 employees), who already represent more than 80% of current cloud spending, will continue investing more than SMEs. Among SMEs, larger ones (with 100-249 employees) are expected to increase spending faster than smaller ones. Many SMEs adopt free cloud services.* [8]

As might be expected in Europe, where nation states have different laws governing the area, one clear cluster of correlated barriers is unclear legal jurisdiction and data location issues, complex security and data protection regulations, uncertain trust in suppliers, and lack of guaranteed data access and portability between cloud systems.

Other barriers particularly relevant to SMEs identified in the report were a tendency to undervalue the usefulness of cloud computing and something as simple as slow connectivity. The report sums up the main differences between large corporations and SMEs here:

> *As SMEs lag behind large companies in the cloud journey, it is interesting to analyse data by company size to understand what is hampering adoption at the low-end of the market. IDC went through the analysis of the indicators presented for the total market, for both SMEs (<250 employees) and large companies (>250 employees).*
>
> *The following considerations emerge:*
>
> - *Evaluation of usefulness is a much higher barrier for SMEs than for other enterprises.*

[8] http://ec.europa.eu/information_society/activities/cloudcomputing/docs/quantitative_estimates.pdf, page 10

- *Trust is one of the most important barriers for SMEs (with almost the same "impact" score of evaluation of usefulness), because of their need to rely on their suppliers, lacking specialised resources to deal with them.*
- *Lack of solutions in local language has a medium relevance for SMEs.*
- *Ownership of customisation is less relevant for SMEs than for other companies.* [9]

European policy regarding cloud computing

The so-called CLOUD Computing Expert Working Group came out in September 2012 with a report called "Advances in Clouds" following up on the January 2011 report called "The Future of CLOUD Computing." The "Advances in Clouds" report analyses the progress in the two years prior, discusses various barriers and makes the following eight recommendations:

1. Ensure progress in CLOUD research
2. Focus on concerns of long term relevance
3. Enable the fast transition to the CLOUD
4. Encourage large scale European providers
5. Encourage SME providers
6. Promote open source solutions
7. Encourage the development and adoption of standards
8. Think ahead

The European Commission revealed a cloud computing strategy in the same month under the flagship "Digital Agenda for Europe" – one of the "Europe 2020" initiatives. The strategy is called "Unleashing the potential of cloud computing in Europe." It is not clear whether it is as ambitious as the recommendations of "Advances in CLOUDS." The European Commission presents the purpose of the strategy in this manner:

> *The strategy outlines actions to deliver a net gain of 2.5 million new European jobs, and an annual boost of EUR 160 billion to EU GDP (around 1%), by 2020. The strategy is designed to speed up and increase the use of cloud computing across the economy. This strategy was the result of an analysis of the overall policy, regulatory and technology landscapes and of a wide consultation of stakeholders, undertaken to identify what needs to be done to make the most of the potential that the cloud has to offer us. [The associated documents] set out the most important and urgent additional actions. It represents a political commitment of the Commission and serves as a call on all*

[9] http://ec.europa.eu/information_society/activities/cloudcomputing/docs/quantitative_estimates.pdf, page 39

stakeholders to participate in implementing these actions. In fact, select working groups are already working on this.[10]

The Commission's stated aim is to introduce new, pan-European certification schemes for cloud computing, including data protection, by 2014. The European Network and Information Security Agency (ENISA) and other relevant parties will be asked to assist in this process. These certification schemes will address data protection, especially data portability, and focus on increased transparency of cloud service providers' security practices. Participation is voluntary. A new model contract terms for cloud computing will be drafted by the end of 2013 to ensure consistency and fairness in contracts for cloud computing services across Europe. The Commission places particular emphasis on how data is handled.

The European Cloud Partnership (ECP) will consist of high level procurement officers from European public bodies and key players from IT and telecom industry. The ECP will, under the guidance of a Steering Board, bring together public procurement authorities and industry consortia to implement pre-commercial procurement actions. The ECP does not aim at creating a physical cloud computing infrastructure. Rather, via procurement requirements that will be promoted by participating Member States and public authorities for use throughout the EU, its aim is to ensure that the commercial offer of cloud computing in Europe, both of the public and of the private sector, is adapted to European needs.

The Commission's new strategy also aims to undertake a review, by the end of 2013, of the current standard contractual clauses for international data transfers to make them more cloud-friendly; and to encourage national data protection authorities to approve Binding Corporate Rules tailored for cloud services; as well as draft a new industry code of conduct for the unified application of data protection provisions that would be developed in collaboration with the cloud computing industry and endorsed by all national data protection authorities in the EU. Further, it is intended to increase coordination with the United States, India and other countries concerning issues such as access to data by law enforcement agencies as well as data and cyber security at the global level.

In summary, it may be said that the strategy aims at facilitating Europe‘s participation in the global growth of cloud computing by setting out a clear plan on how to address the issues related to cloud computing while at the same time considering the global context of cloud services that will be offered in the European markets.[11]

---

[10] https://ec.europa.eu/digital-agenda/en/european-cloud-computing-strategy

[11] See also "Unleashing the Potential of Cloud Computing in Europe" http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2012:0529:FIN:EN:PDF

This strategy has been criticised for not going far enough by the European Economic and Social Committee (EESC). The EESC claims a European cloud strategy should go beyond mere awareness raising and offer actual incentives to taking up cloud computing by European institutions and organisations on one hand and foster a cloud computing infrastructure development within Europe's borders on the other. Otherwise, the benefits of the technological development may end up elsewhere.

In short, the EESC proposes:

- increasing use of the cloud
- developing software based on cloud computing
- implementing a cloud computing infrastructure in Europe.

Potential of cloud computing in cultural heritage

The public sector has a strong role to play in shaping the cloud computing market. As the EU's largest buyer of IT services, it can set stringent requirements for features, performance, security, interoperability and data portability and compliance with technical requirements.

The usage of cloud computing in the space of digitization of heritage may be based on the action programme "Promoting Common Public Sector Leadership through a European Cloud Partnership", Key Action 312 stated in the Communiqué of the European Commission "Unleashing the Potential of Cloud Computing in Europe":

- identify public sector cloud requirements; develop specifications for IT procurement and procure reference implementations to demonstrate conformance and performance;
- advance towards joint procurement of cloud computing services by public bodies based on the emerging common user requirements;
- set up and execute other actions requiring coordination with stakeholders as described in this document.

12 Unleashing the Potential of Cloud Computing in Europe. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS (Brussels, 27.9.2012 COM(2012) 529).

From this viewpoint, the cloud computing definition, characteristics and possibilities described in section 1 allows us to distinguish three basic groups of digital heritage users, for which the usage of cloud computing would be a question of present interest.

1. Digital content providers and aggregators
2. The general public
3. Scholars

The usage of cloud computing in each group would allow creating specific infrastructures and service packages that would be relevant in enhancing digital heritage usability and creating surplus values in European creative industries. [13]

Digital content providers and aggregators

The largest EU portal of digital heritage, Europeana, currently has 134 content providers (Europeana content providers are mainly aggregators and a small number of major national institutions), and the Europeana Network has 694 members.[14] Practically, institutional content providers constitute a less complex case than aggregators because one aggregator usually embodies many individual content providers.

The most important challenges that content providers and aggregators are faced with are high-quality digitization of heritage and interaction between the local systems metadata and the Europeana metadata model (EDM). In dealing with these challenges, the following cloud computing services could be offered for content providers and aggregators:[15]

---

[13] Correspond the provider perspective by Advances in Clouds. Research in Future Cloud Computing. Expert Group Report. Public version 1.0. Editors: Lutz Schubert [USTUTT-HLRS] , Keith Jeffery [STFC] . 2012. [interactive]. Internet access: <http://cordis.europa.eu/fp7/ict/ssai/docs/future-cc-2may-finalreport-experts.pdf >.

14 Europeana partners. 2013. [interactive]. Internet access: <http://pro.europeana.eu/web/guest/about/partners>.

15 Services given in this part of the text are just the examples and it was not aimed to encompass the entire available spectrum of cloud computing services. The New Services Priorities for content providers are discussed in more detail in separate document “Service Priorities and Best Practices for Digital Cultural Heritage” (Edited by the DC-NET Working Group 3: New Services Priorities , 2012). Challenges emerging for creative industries and the ways of dealing with them (by defining the service priorities) are discussed in WORK PROGRAMME 2013 . COOPERATION . THEME 3 ICT – INFORMATION AND COMMUNICATIONS TECHNOLOGIES (European Commission C(2012)4536 of 09 July 2012). A part of services in the Europeana environment are already created and offered for content providers, aggregators and the

- controlled vocabularies – controlled vocabularies created and supported in the Europeana environment would help content providers to provide higher quality metadata for digitized heritage and would enhance the local information systems interoperability with Europeana and provide services such as cross-content provider vocabulary linkage as well as cross-language and cross-domain creating a semantic heritage web;

- local metadata to EDM cross-walking tools – software that helps to perform XML based mapping from local information systems to EDM model (e.g. MINT tool);

- GIS based web visualization tools – intended for attractive presentation of local data on the Internet (e.g. CARARE visualization tool);

- tools for digitization of geodata, 3D and virtual reality data where still needed – services intended for digitization of specific data and presentation on the Internet;

- virtual exhibitions tool – service that allows creating virtual exhibitions using the data available on Europeana;

- context interoperability tool – service that allows linking the content stored in digital heritage systems with context data (e.g. Wikipedia) or data stored in the information system of scientific data (e.g. IS operating in DARIAH or CLARIN environment);

- EDM based information system as informational infrastructure service – for those content providers who have no local information systems;

- long term preservation repository – for those content providers who have no servers for long term preservation of digital heritage;

- networking, online consultation, good practice and content providers and aggregators professional discussions space;

- Cloud storage space

The surplus value created by these services arises when taking into account that most of content providers and aggregators are institutions of the public sector that are limited-budget organizations (especially during the economic crisis). Furthermore, a part of them are small institutions that do not have financial possibilities and intellectual resources to develop software and information products similar to the aforementioned services. Controlled vocabularies, metadata crosswalking

---

general public, another part is being implemented via “Europeana Cloud: Unlocking Europe’s Research via The Cloud ” and in “Local Content for Europeana Cloud” projects.

tools, etc. are very specific software and information products, the development of which requires collaboration by professionals from different countries.

On the other hand, the provision of such products as free (or available for a token fee) cloud computing services for content providers and aggregators would be a great support (especially for small institutions) that would give an impetus to the activities of digitization of heritage of higher quality and greater volume. From the viewpoint of creation of general European cultural space, controlled vocabularies would play an important role as they would help to bring down historical and cultural boundaries formed by national narratives at the end of the 19th century – beginning of the 20th century that could be treated as the reasons of potential social conflicts when moved to a digital space.[16]

The general public

One of the greatest contemporary challenges is a perception of social context. Although much public funding has been invested in the development of digital heritage, the number of its users is relatively small. For instance, the Europeana's portal currently has 500-600,000 of unique visitors every month - which corresponds to about 1/1000 of the population of the European Union (and EU residents are not the only ones who visit the portal). [17]

Insufficient perception of social context turns the infrastructures of digital heritage into some kind of 'frozen capital' that does not generate any surplus value. We could look for the reasons of this phenomenon in the changing society. Although the information and communication technologies that have widely spread during the recent decades have caused great changes in society, most of the internet service developers follow the perception of the needs of the old world.

The developed heritage internet services operate mostly like scientific archives; they are unfriendly for the general public, and the objects of digital heritage presented in them often require context data or professional interpretation.

---

[16] The prime example of such boundaries could be presentation of historical personal names and place names in Europeana metadata. Usually these data are presented in such a form as they are recorded in information systems of national content providers and aggregators, and when they are moved to Europeana portal they become non-interoperable. That means different names of the same person or place are not linked with each other.

[17] Correspond the (non-technical) user perspective by Advances in Clouds. Research in Future Cloud Computing. Expert Group Report. Public version 1.0 . Editors: Lutz Schubert [USTUTT-HLRS] , Keith Jeffery [STFC]. 2012 [interactive]. Internet access: <http://cordis.europa.eu/fp7/ict/ssai/docs/future-cc-2may-finalreport-experts.pdf>.

Moreover, the systems are unsuitable for the users' individual needs; quite often they offer hardly any interactivity. In dealing with these problems, cloud computing services such as the following could be considered for offering to the general public:

- service package that enables users to access the digital content on GIS mobile and smart devices;
- service package that allows indexing and thus singling out the digital objects that are the most interesting for the general public by applying the methods of professional selection and crowdsourcing;
- service package that allows to easily link the digital content with QR codes in heritage objects (service of generating automated QR codes for specific objects of digital heritage);
- service package that allows the users to create content and exchanging user generated content with other users;
- virtual exhibitions tool – service that allows creating private virtual exhibitions and presenting them to other users;
- service of creating a personal profile with an opportunity for the users to create a collection of selected link references of digital heritage objects, private gallery, discussion forum;
- service package that allows the users creating thematic presentations;
- service package that allows executing genealogical and local historical pilot surveys in virtual space;
- long term preservation repository, for user generated content;
- global-to-local interoperability tool – a service that links from national digital heritage systems to the related content presented by Europeana providers from other countries;
- applications which foster social engagement of end-users and allow to monitor social media activities.

Scholars

One of the most important problems of digital heritage is the separation of heritage objects and the scientific data related to them. Heritage objects are usually stored in memory institutions (museums, archives, libraries) whereas scientific data are accumulated and analysed by research institutions. In the EU digital space, heritage is accessible via the Europeana portal whereas

infrastructures for the arts and humanities are administered by research infrastructures such as the CLARIN and DARIAH networks, or the Ariadne network for Archaeology. [18]

In the activities of memory institutions, scientific data are important for the description of collections and objects as they interpret heritage, highlight its value, define context information in communicating heritage to the society. From the point of view of science (usually the arts and humanities), heritage objects stored and digitized in memory institutions are the sources of scientific research. Therefore interoperability of the internet services of digital heritage and digital scientific data is important to scholars.

The scholars' needs differ from the needs of the general public. Europeana's digital virtual collection may be interpreted as a collection of scientific sources, and, being the largest online collection of such nature, it already now serves for the arts and humanities scholars and is becoming an important catalyst of such research. Seeking to improve the usage of digital heritage for scientific purposes, the following cloud computing services could be offered:

- scientist's personal profile service that enables a scientist to save and analyse the sources of interest;
- service of thematic source subscriptions that enables a scientist to receive real time information on the new heritage objects of interest uploaded into Europeana according to the pre-indexed list of key words;
- service package that allows the digital content to automatically be analysed in virtual space and the data of digital heritage to automatically be exported to file formats acceptable for analytical programs;
- context interoperability tool – service that allows linking the content stored in the systems of digital heritage with the context data (e.g. Wikipedia) or with the data stored in the information system of scientific data (e.g. IS operating in DARIAH or CLARIN environment).

[18] Correspond the (non-technical) user perspective by Advances in Clouds. Research in Future Cloud Computing. Expert Group Report. Public version 1.0 . Editors: Lutz Schubert [USTUTT-HLRS] , Keith Jeffery [STFC] . 2012. [interactive]. Internet access: <http://cordis.europa.eu/fp7/ict/ssai/docs/future-cc-2may-finalreport-experts.pdf>.

Section 3: The LoCloud Project and cloud computing

Situation reports from LoCloud partners

Three LoCloud Partners were asked to evaluate the situation in their respective countries with respect to the expertise and experience with cloud computing among the small and medium sized cultural institutions, which will be delivering content in to LoCloud. The evaluation resulted in the three situation reports cited below. All three organizations cited below have extensive experience in working with small and medium cultural heritage institutions.

Ministerio de Educación, Cultura y Deporte , Spain

"In our view, small and medium sized institutions should not consider the migration of the three basic sets that make the cloud computing services, i.e. SaaS, IaaS and PaaS. It is much more reasonable in the short and medium term, and even long term, to outsource all these services.

Nevertheless, what is necessary is to carry out the migration of the application in order to move it to SaaS. Probably this service will have to be outsourced since it's not likely that SMIs have the suitable staff for it.

It's essential that institutions have a complete control over their data and on the development of the process as well, particularly through the establishment of a benchmarking and the ability of interacting with the project and having a copy of the resources that will be uploaded to the cloud both in massive bulks and batch as well as the subsequent updates.

We consider it's virtually impossible for a SMI to address these projects individually with guaranteed success. It should be considered the possibility of establishing a working party or a tender board in order to look for a regional or national infrastructure capable of interacting in more than one province or region with the aim of reducing costs".

The Netherlands Institute for Heritage

"We have asked the RCE liaison-officers for museum and other (small) heritage organizations in the Netherlands to give an assessment of their target group.

The results of this survey are:

All heritage organizations have some knowledge of working in the cloud/using cloud services and what that means. However this knowledge is limited and is based on personal/private experience (Google, Flickr, YouTube etc.). We have not found an organization in the target group that is actually (partly) working in the cloud: Neither collection data, nor collection management processes.

Most of the organizations do not have a vision on using cloud services or even understand concepts as SAAS. The general attitude (exceptions excluded) is "waiting and following". Best practices and illuminating examples are means to persuade these organizations to use cloud services.

To use cloud services costs, necessary effort to manage collection data, and necessary IT-knowledge/knowledge to understand and manage applications and contracts for the services are very important factors. These should all be (very) limited. Heritage organizations in the Netherlands generally distrust commercial IT-solutions and interference with how they manage their collections and data. They want to see what's in it for them!

Financing a transition to cloud services and financing the use and maintenance of cloud services is difficult. These heritage organizations have very limited financial means. Transition to the cloud and use of cloud services should be made easy and cheap.

There is a general awareness of the importance of standardization. The heritage organizations also recognize the benefits of reduced effort and costs when using existing standards. However, these standards should be flexible enough to allow for the diversity of the organizations and the way they work. Mapping services for conversion to the standards like EDM, LIDO should be available, simple and easy to use.

Interoperability should be made easy by making thesauri available and making it simple to connect to these thesauri.

Security, confidentiality, legal rights are important issues that must be addressed and solved to satisfaction. There is a general distrust concerning these aspects among the heritage organizations in the Netherlands.

Finally, RCE has some tools in place and in use already that may become part of the LoCloud toolset: the heritage suite to bring data online and connect them to the heritage thesaurus, enabling cross-collection searches. DiMCoN is a national

aggregating platform enabling the uploading of metadata to Europeana".

The Danish Agency for Culture

"In Denmark, the Agency for Culture offers a web-based Collections Management System to all museums officially recognized and state-subsidised. The museums do not have to host any servers or software, but use the system via a browser. The software and database runs on a dedicated server financed by the Agency of Culture, and hosted by a commercial ISP at a facility in Denmark. In effect it is a private cloud SaaS solution.

By choosing hosting at a Danish facility, on a dedicated server, the Agency has avoided many of the security issues of cloud hosting and storage. However, it is indeed possible that the hosting cost could be lower and scalability more flexible if a commercial cloud provider was used.

The solution as such has been a great success, and is used by around 100 small and medium sized institutions. It has saved a lot of money for the museums, who do not need to worry about software licenses, server hosting, data storage or backup. All they need is a browser and an internet connection.

The main disadvantage of the solution is that the museums have limited access to individual adaptions of the system. They all get the same functionality. This has led to some museums building small plug-ins or private solutions, which communicate with the central system. The two national museums The State Museum of Art and the National Museum have chosen not to use the central system at all, as they had too many special requirements.

The Agency is at present working at a new version of the system which will include a Digital Asset Management solution, a placement solution and support export of data to Europeana. The new version of the SaaS solution is expected to be taken up by all state-subsidised museums, including the large national museums.

The centralized web-based collections management system will be part of the LoCloud project, and deliver data directly to MINT. The museums will not be delivering content individually, and therefore have less interest in a future lightweight digital library solution. However, it may server as an example of a successful SaaS solution.

Important issues for the Danish museums when using the system are security, back up, performance and interoperability with other systems such as placement solutions and websites".

Conclusions from the situation reports

One common trait in all three situation reports is that the small and medium sized institutions need and value a centralized solution. They are not able to do it themselves for lack of money and skills, and would greatly benefit from a SaaS solution developed during the LoCloud project.

Another conclusion is that security, confidentiality, legal rights are important issues that must be addressed and solved. There is a general distrust concerning these aspects among the heritage organizations in the three countries, which is likely to be mirrored in other countries.

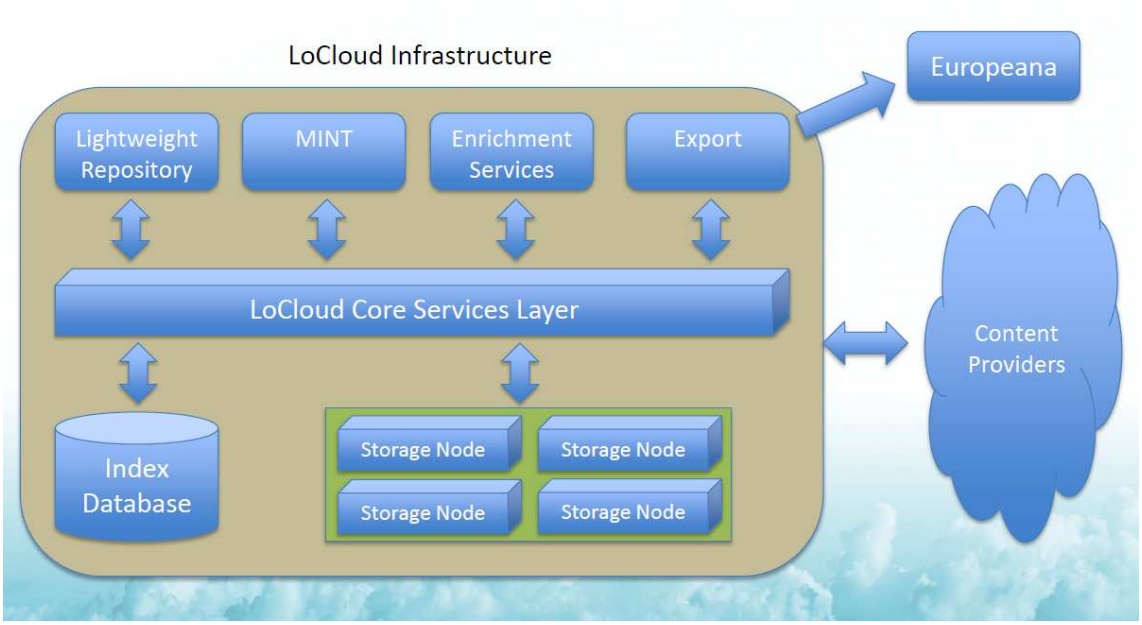
The use of cloud computing in the LoCloud project

The LoCloud project will create a Best Practice Network to support small and medium-sized institutions in making their content and metadata available to Europeana by exploring the potential of cloud computing.

A cloud-based technology infrastructure will enable the aggregation of local content and a number of micro-services will help to reduce technical, semantic and skills barriers and to render the content more discoverable and interoperable.

It is important to note, that digital heritage is a sector where specialized and expensive information services (e.g. controlled vocabularies) are necessary, and that the LoCloud community consists of small and medium-sized institutions which lack intellectual or economical resources to purchase or create these software tools and information services themselves.

The LoCloud project will establish a storage and aggregation infrastructure as described in the illustration below:



A Cloud-based Aggregation Infrastructure, Dimitris Gavrilis, Costis Dallas, presentation at the LoCloud kick-off meeting

The infrastructure will reside in the cloud, and will host a number of Cloud Computing based micro-services and Cloud computing tools:

LoCloud (Cloud shared) metadata schemas

A working team will analyse the different metadata schemas and will define the metadata schemas to be used in LoCloud as intermediaries to EDM, on the basis of the schemas used in prior projects, such as Europeana Local, Athena and CARARE and will define work needed to support their implementation in local metadata to EDM cross-walking tool MINT and their ingestion into the MoRe repository.

LoCloud local metadata to EDM cross-walking tools and EDM based information system and repository as informational infrastructure services

The LoCloud team will design technical specifications of the core infrastructure components, taking into account business process models for small and medium-sized institutions in the context of their interactions in Europeana's ecosystem.

The scope of LoCloud is also to specify and carry out necessary developments in the existing MoRe aggregation service sufficient to enable ingestion, data management and continued delivery to Europeana, not only from the MINT mapping and ingestion service but also by a variety of other methods identified as appropriate for small and medium sized institutions participating in LoCloud. Consulting national and regional aggregators, it will specify and make available a lightweight digital library system for small and medium sized institutions, delivered in the cloud and compatible with the standards established by MINT, MoRe and Europeana, based on their existing work in this field and suitable for deployment where local infrastructure is lacking.

For other institutions with access to IT infrastructure, tools will be specified and delivered as cloud services to prepare content and metadata online in repositories and using provided tools to support metadata harvesting. For small institutions with limited technical capacity tools will be specified and provided in order to prepare content online with embedded EDM meta-tags ready for capture by crawler services and to be incorporated in the LoCloud repository.

Specialized information services tools (controlled vocabularies, GIS tools, metadata enrichment)

In this context, controlled vocabularies, which enable semantic interoperability of local providers aggregated content must be distinguished as one of the most important of information services tools for local small and medium-sized institutions. When analysing the problems of interoperability in terms of communication, two groups of reasons should be mentioned: relationships between terminology and reality (meaning), as well as multilingual terminology.

Relationships between terminology and reality can be interpreted in accordance with semiotic communication theory (theory of meaning), [19] where a term is a conventional sign, which is developed by the interpreter in his mind perceiving the object of reality. Therefore in terms of communication of meanings, terminology is a piece of work of different human groups intended to name the same object of reality. Whereas miscommunication and non-interoperability occurs in the level of signs (words) rather than objects.

Such treatment of terminology raises an obvious multilinguistic problem. Different nations, like different socio-natural groups of people, have developed different terminology to namel the same existing objects. A typical illustration of this could be mentioned is the problem of presentation of historical place-names in information systems.

In many European countries and regions, administrative division, administrative subordination of territories had often changed in the past, place-names were written down in different languages, in different name forms and different writing systems; there were in the past (and still are today) many identical place-names and many extinct place-names (that are mentioned in written sources only but not mapped except for historical maps); in different languages the same geographical objects may be called differently. This creates multilingual confusion, which disrupts the interoperability of information systems. LoCloud team, working together with interested content provider partners, are planning to specify, develop and test:

---

[19] Peirce, Charles Sanders. Logic as Semiotic: The Theory of Signs. In *Philosophical Writings of Peirce*, ed. by Justus Buchler. [interactive]. New York: Dover Publications. INC. Internet access: <http://theory.theasintheas.org/wp-content/uploads/2013/02/Peirce-C-S-Logic-Semiotic.pdf>.

a) multilingual controlled vocabularies for local history and archaeology;

b) historical place name gazetteer;

c) a suite of geolocation enrichment tools;

d) the Natural Language Processing (NLP) tools necessary to analyse and enrich the metadata being provided to Europeana as a test-bed for two languages, English and Spanish, and to include the enriched content into the repository;

e) application, which allows content to be uploaded to Wikimedia during 'Wiki Loves Monuments' campaigns and content uploaded by Pol Mayer (Archive of photographs of historic buildings).

Each micro-services and Cloud computing tools would be developed using a designed schema of information system and information services, also by creating a good practice from investigations that would be made, as well as constructing and testing the quality and impact of evaluation and assessment.

Saas offerings relevant to the LoCloud project

As one of the main tasks of the LoCloud project is to implement a lightweight digital library, this chapter is dedicated to an exploration of existing cloud based, collections management and digital library software. These are relevant to the LoCloud project because they competitors of a sort – they offer solutions that small and medium cultural institutions might take up instead of the services offered by LoCloud. On the other hand, they may be sources of inspiration for the LoCloud project, when building a lightweight digital library service.

The museum perspective

In 2011 Nick Poole, the chief Executive at Collections Trust, wrote an article called "Is now the time for Collections in the Cloud?" and presenting the following scenario for the future:

> *In reality, we are unlikely to see a total transition to Collections Management in the Cloud anytime soon. This is a tremendously diverse sector, with a huge range of different types and scales of institution. There will always be those who need an installer and a locally-based application, for whom this is the simplest and most effective option. What I do think we'll see, particularly over the next 4-5 year cycle, is two significant trends:*
>
> *1. A significant swing in the medium-to-larger end of the museum community towards full-service online Collections Management Software and;*
>
> *2. The emergence of more online tools with a very low barrier to entry (both in terms of cost and complexity) which are suited to the needs (and budgets) of smaller local and community museums.*

The majority of the leading Collections Management Systems has either already developed fully-hosted versions of their applications or is in the process of developing them. Most vendors are reporting significant increases in the uptake of their Software-as-a-Service offers, with some confidently predicting a full transition to Cloud-based, browser-based Collections Management within the next decade.

Well known vendors like The Museum System and AdlibSoft do not yet offer SaaS versions of their software. The following companies are examples of vendors who do provide commercial fully cloud-based collection management systems for museums. The list is probably not comprehensive; there may be other solutions out there.

Vernon systems: http://www.vernonsystems.com

Based in New Zealand, Vernon systems a product called eHive, which is a web-based collection management tool. eHive accounts are available on an annual subscription basis covering use of the software, storage and ongoing development. The amount of storage you use determines the level you need to purchase. Prices range from 76 - 616 Euros for one year's license.

Zetcom: http://www.zetcom.com

Zetcom is an international company with offices in Europe, North America and Asia. They offer a SaaS version of their MuseumPlus collection management product, and a parallel product especially suited for Art Collections called ArtPlus. Prices are not available on the website.

Vesica: http://vesica.ws

Vesica is a UK Company which offers a web based art collection management application that can be used in lieu of traditional desktop software. Vesica pricing plans are based on a flat fee of £0.05 per object per month. If your collection holds 1000 objects, it will cost about 730 Euros per year.

To the best of our knowledge, none of these offer out-of-the-box integration with a metadata aggregator or with Europeana. This functionality will probably only be developed, if the vendors see a possibility of a new source of revenue. They do however, offer export of the data in an xml format, which may be map-able to the metadata schema chosen for LoCloud.

The switch from a legacy database to one of these comprehensive collection management systems mentioned above may be quite a large task and demand quite a lot effort and learning. As Nick Poole suggests, there is probably a need for online tools with a very low barrier to entry which are suited to the needs (and budgets) of smaller local and community museums.

One aim of LoCloud is to establish at least one such solution, called a Lightweight Digital Library, in the framework of the project. The solution must be affordable, very easy for basic use, have a

customizable data model and support creation of high quality metadata. The requirements will be defined during the project. [20]

In the light of the concerns expressed by in the three situation reports above, the solution should also be secure, reliable and interoperable with other IT systems at the museum.

The library perspective

Libraries have taken to the cloud in several different ways. The international library community has arguably had a head start compared to museums, archives and galleries with their tradition for international collaboration on standards and technology as well as the benefit of homogenous main materials; books. The international library committee prepared for the era of the database in the 1960s by developing cataloguing rules and standards that spoke of entries and entity relationships rather than books and authors.

Library materials embody a challenge that does not plague other GLAM institutions' materials in the same way, namely that container and content are two vastly different things. It is the role of library classification and cataloguing to bring together like things together at the content level. Most library online public access catalogues (OPACs) both try to guide the user in the bibliographic universe by allowing for browsing through relationships between content – e.g. 'this author has also written', or' there are also these books on that subject' – while still describing the object in and of itself and tracking its whereabouts.

OCLC

Most of the Western world's library cataloguing, entity relationship modelling between documents and coding standards are national expressions of international standards with Library of Congress and Online Computer Library Center, inc. (OCLC – formerly Ohio College Library Center), a not-for-profit research and development organisation, historically functioning as the *primus inter pares* in the international development. Library cataloguing has changed in recent years with new standards emerging, but the traditional picture of national cataloguing rules congruent with the Anglo American Cataloging Rules (AACR), national adherence to the Functional Requirements of Bibliographic Records (FRBR) E/R-model and national Machine Readable Cataloging standards such as DanMARC, FinMarc etc. being interoperable with the international UNIMARC, is still predominant .

[20] Adam Dudczak and Marcin Werla, PSNC, digital libraries team. Slides from LoCloud kick-off.

With OCLC as project managers, libraries have experimented with interdependent collection development and management in the cloud. Specifically, institutions may share both analogue as well as digital resources in a cloud-based digital repository and shared physical storage, vastly improving efficiency, value-for-money for the participating libraries, as well as freeing up physical space. However, the print-on-demand collections are still troublesome from intellectual rights and licensing perspectives.

OCLC WorldShare Management Services is a more fully-fledged cloud-based collection management system, presented in this manner on the OCLC web site:

> *OCLC WorldShare Management Services provide a unified, Web-based environment that streamlines acquisitions, circulation, license management and metadata management workflows and offers a powerful discovery and delivery tool for library users. Both replacing and standing apart from traditional ILS systems, OCLC's cloud-based library management services enable libraries to share infrastructure costs and resources as well as to collaborate in new way.*
>
> *OCLC WorldShare Management is essentially PaaS to facilitate discovery services in the form of free-for-all app development. This makes OCLC the Google- / Amazon- / Facebook-like platform for the library community.*

The system works by coupling standard library workflows to the WorldCat, which makes the system in and of itself less useful for other GLAM institutions, but makes it interesting as a complimentary system and even more so if and when the WorldCat data becomes linked open data.

DuraCloud

DuraCloud is a very interesting service in the domain of digital libraries designed for storage and preservation with the aim to achieve maximum availability and durability through true redundancy. This means that data are not only distributed into multiple geographical locations managed by one vendor, but also to infrastructures of different vendors. DuraCloud uses Amazon S3, Amazon Glacier and Rackspace as their underlying storage services.

DuraCloud offers cloud storage, and a set of automated services:

- **Backup and Syncing.** DuraCloud automatically copies your content onto several different cloud storage providers.
- **Health Checkup.** DuraCloud´s health checking service lets you verify that the content you have stored in DuraCloud is safe and secure.
- **Upload Tools.** DuraCloud is flexible; there are several stand-alone tools and command-line utilities that give you more direct ways to upload content to DuraCloud.
- **Video and Audio Streaming.** DuraCloud can stream video and audio files directly from DuraCloud to any internet-linked device.

DuraCloud offers several different subscription plans, starting from $1,500 for 1st terabyte per year.

Omeka.net

Omeka.net is web-publishing platform that allows anyone with an account to create or collaborate on a website to display collections and build digital exhibitions. No technical skills or special server requirements are necessary. Omeka.net may be used by all of the GLAM sector and offers the following features:

- Share Collections
- Gather Stories
- Build Exhibits
- Map Photos
- Create Simple Web Pages
- Customize Web Design Templates
- Share data through feeds
- Collaborate

Omeka.net offers different payment plans:

- Basic (free): 500 MB storage; 1 site
- Plus ($49/year): 1 GB storage; 2 sites
- Silver ($99/year): 2 GB storage; 5 sites
- Gold ($299/year): 5 GB storage; 10 sites
- Platinum ($999/year): 25 GB storage; unlimited sites

LoCloud and the Europeana Cloud project

As LoCloud aims to leverage the cloud for its infrastructure needs, so do other stakeholders within the Europeana network. The Europeana Cloud project aims to build a cloud-based infrastructure to support Europeana and its aggregators.

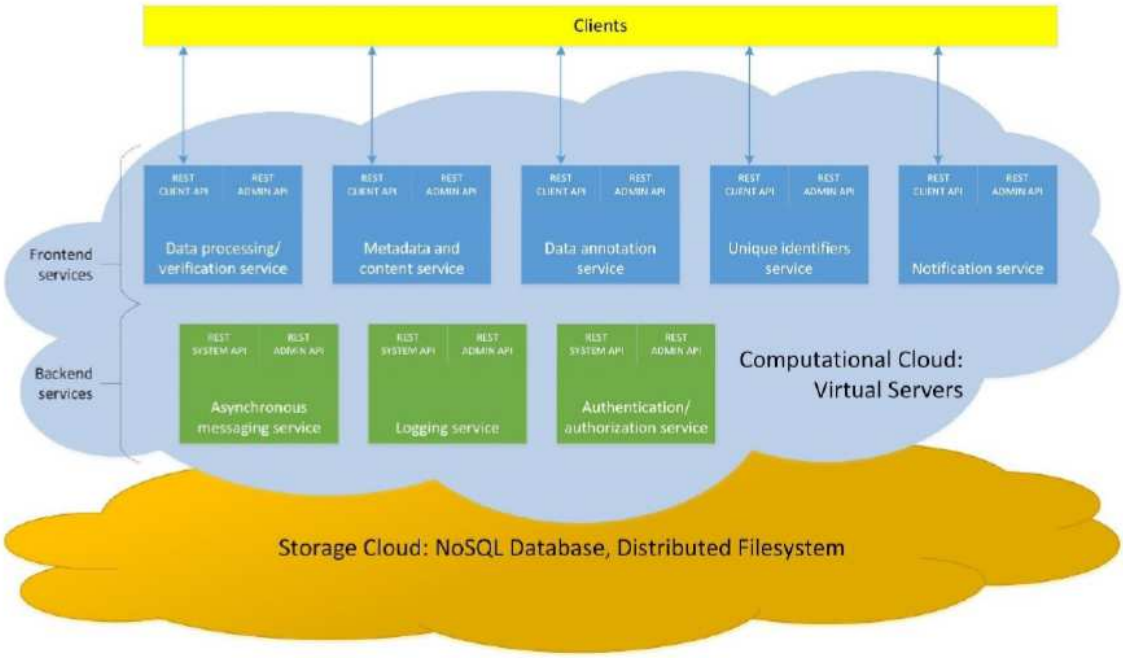
The envisioned Europeana Cloud infrastructure is meant specifically to provide cost-effective storage for both digital content and its metadata. Within this infrastructure, the project aims to deliver tools and services to researchers, supporting discovery, use of and research based on the content and metadata in Europeana by developing the Europeana Research platform.

The Europeana Cloud infrastructure is seen as a shared space for European digital heritage with data able to transition more easily and flexibly between aggregators, content providers and Europeana than is the case today. The current infrastructures supporting delivery of content to Europeana have been built as individual solutions as aggregators emerged. As a result, some duplication of solutions and effort has occurred. The project aims to reduce this through a more unified architecture, where the time spent processing and quality-checking metadata is reduced, technical resources are shared and cost is saved as a result.

In Deliverable 2.2 - “Europeana Cloud Architectural Design” the future architecture of the eCloud system is described. As mentioned earlier, the eCloud project has chosen a Hybrid Cloud solution consisting of a private, community-based part where the necessary hardware resources will be provided by voluntary, technically advanced institutional users of the eCloud system and a public part, based on resources leased from commercial providers. Thus the eCloud system will benefit from the security of a private cloud, and the rapid scalability of a commercial provider.

In relation to LoCloud, it is also worth noting that the Europeana Cloud project will evaluate cloud technologies suitable for the implementation of the project’s storage system and analyse their performance, scalability and suitability as a platform for the delivery of services.[21]

LoCloud may consider a similar strategy to eCloud and build a hybrid cloud. However, it may not be necessary for LoCloud to look any further for an IaaS solution, since eCloud will offer just that. Based on the Deliverable “2.2 - Europeana Cloud Architectural Design” from the Europeana Cloud Project, we expect that this project will offer storage facilities. It seems logical that LoCloud should consider using the Europeana Cloud infrastructure if it is mature enough at some point to deliver the LoCloud SaaS solutions.



Overview of the Europeana Cloud architecture. The eCloud system will look like a SaaS cloud from an end user point of view, or IaaS in case of cultural data records storage.

[21] http://pro.europeana.eu/documents/1414567/0/Europeana+Cloud+-+Description+of+Work

It would seem that the services and tools offered by LoCloud and eCloud would actually complement each other very well. The services being developed by the eCloud project are:

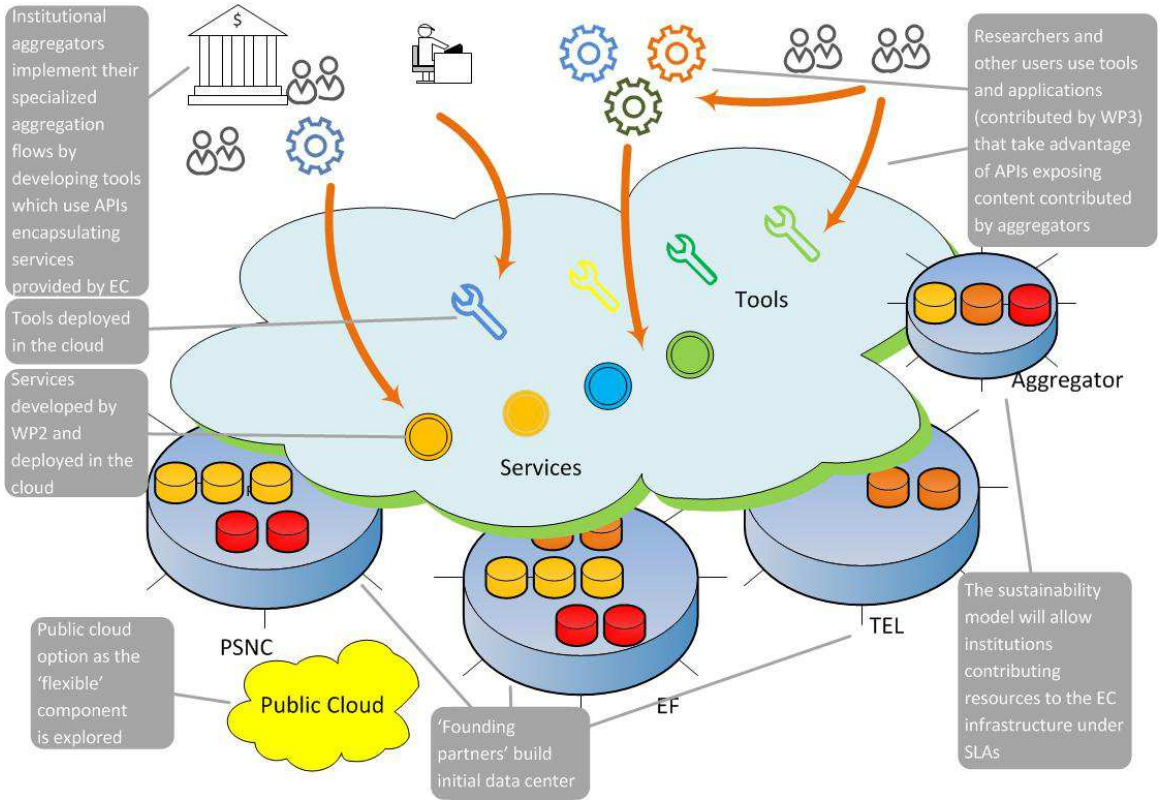
- Unique Identifier Service
- Metadata and Content service
- Notification Service
- Data annotation service
- Data processing service

And three backend system services:

- Authentication/Authorization service
- Logging service
- Asynchronous Messaging Service

The requirements have been collected from notably large organisations, namely Europeana itself, The European Library and the Polish Digital Libraries Federation. The resulting services will be very relevant for the LoCloud aggregator which could build tools based on the APIs of eCloud, but not for the small and medium sized cultural heritage institutions. Therefore, the LoCloud tools could possibly supplement the eCloud offerings.

The diagram below shows that the eCloud team is thinking along the same lines. LoCloud could be one of the “aggregators [who] implement their specialized aggregation flows by developing tools which use APIs encapsulating services provides by eCloud”. Exactly how the two projects could cooperate in the future will be explored in the near future through the Cloud coordination Group which has been established.



Conclusions

There are many advantages to cloud computing which could be taken advantage of by cultural and heritage institutions. Cost-effectiveness and access to resources beyond the abilities of the individual institutions are among the primary advantages.

There is high awareness and willingness to participate in cloud-based development from the heritage institutions and agencies voicing their opinion in this report. However, both a literature review and our own situation reports show a hesitancy to delve into the new service offerings on their own. Hesitancy is mainly based on lack of knowledge and skills. The aim of the LoCloud project - to create a Best Practice Network which will support institutions in making their content and metadata available to Europeana by using a cloud-based technology infrastructure – therefore is very relevant at this point in time.

There are a number of SaaS providers providing services for the cultural sector. Some of the commercial vendors of collections management systems offer cloud based versions of their software, and in the library domain the OCLC offers a number of relevant services. However, none of these come with plug-in aggregation tools for Europeana. There still is a need for online tools with a very low barrier to entry which are suited to the needs (and budgets) of smaller local and community museums. This is the window of opportunity for the LoCloud project.

The LoCloud community consists mainly of small and medium-sized institutions which lack financial and intellectual resources to create, purchase, regularly update and maintain software tools and services for digital heritage use. Therefore the scope of LoCloud project is to create these required Cloud computing based micro-services and Cloud computing tools in the area of metadata interoperability, content aggregation and harvesting, informational infrastructure, multilingual controlled vocabularies, historical place names, geolocation and metadata enrichment, usage of Wikimedia applications and professional networking.

However, there are possible risks, which must be taken into consideration, and the infrastructure must be designed with security in mind and considered a priority for every application, service, and network solution which is provided. The LoCloud partners should thoroughly analyse their present and future needs in view of cloud-computing possibilities and develop a shared vision of an infrastructure. The LoCloud team will be monitoring what is emerging from Europeana Cloud in the way of infrastructural thinking and will take it into account in recommendations to the CH partners in LoCloud.

Literature and online resources

Advances in Clouds. Research in Future Cloud Computing. Expert Group Report. Public version 1.0 . Editors: Lutz Schubert [USTUTT-HLRS] , Keith Jeffery [STFC] . 2012. [interactive]. Internet access: http://cordis.europa.eu/fp7/ict/ssai/docs/future-cc-2may-finalreport-experts.pdf

Bernard, Allen:"Move IT Infrastructure to the Cloud? What SMBs Need to Know" (2012). http://www.smallbusinesscomputing.com/testdrive/move-it-infrastructure-to-the-cloud-what-smbs-need-to-know.html

Budriene, Daiva, and Zalieckaite, Laima: "Cloud computing application in small and Medium-Sized enterprises. In: Issues of Business and Law, Volume 4 (2012), http://www.ibl.ttvam.lt/index.php/ibl/article/view/47/pdf

Buyya, Rajkumar; Chee Shin Yeo, Srikumar Venugopal, James Broberg and Ivona Brandic. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. In Future Generation Computer Systems Volume 25, Issue 6, June 2009.

Cambell, Sara: Customer Spotlight - Q&A with Juan Montes and Diana Pan, MoMA (2012). http://blog.appirio.com/2012/11/customer-spotlight-q-with-juan-montes.html

Dallas, Costis and Gavrilis, Dimitris: A Cloud-based Aggregation Infrastructure, presentation at the LoCloud kick-off meeting 2013. Available on www.locloud.eu

Dudczak, Adam and Werla, Maczin PSNC, digital libraries team: Presentation at the LoCloud kick-off meeting 2013. Available on www.locloud.eu

Enterprise Connect, Australian Government Initiative: "Using the "Cloud" to support your business". Quick guide http://www.enterpriseconnect.gov.au/ecservices/tkc/Documents/TKC-ITGuides/Using-the-Cloud.pdf

Europeana Cloud: Description of Work. http://pro.europeana.eu/documents/1414567/0/Europeana+Cloud+-+Description+of+Work

Europeana Cloud: D2.2: Europeana Cloud Architectural Design

Frank Gens. Defining "Cloud Services" and "Cloud Computing". 2008 [interactive]. Internet access: <http://blogs.idc.com/ie/?p=190>.

Gartner: "Magic Quadrant for Cloud Infrastructure as a service (2012)

Hadidi, Rassule: “What managers of small and medium enterprises (SMEs) need to know about cloud computing, Report.” Published in: European Journal of Management and services, 2/22/2011, ISSN: 1555-4015, http://www.freepatentsonline.com/article/European-Journal-Management/272511391.html

Idea.org: “Collection Management Systems: Museums and the Web 2011” http://www.idea.org/blog/2011/04/07/collection-management-systems-museums-and-the-web-2011/

Information security breaches survey 2013. Technical report. 2013. [interactive]. Internet access: <http://www.pwc.co.uk/assets/pdf/cyber-security-2013-exec-summary.pdf>

Jamsa, Kris: “Cloud computing : SaaS, PaaS, Iaas, virtualization, business models, mobile, security, and more”, 1. ed. Jones & Bartlett Learning, 2013

Peirce, Charles Sanders. Logic as Semiotic: The Theory of Signs. In Philosophical Writings of Peirce, ed. by Justus Buchler. [interactive]. New York: Dover Publications. INC. Internet access: <http://theory.theasintheas.org/wp-content/uploads/2013/02/Peirce-C-S-Logic-Semiotic.pdf>.

Poole, Nick: ”Is now the time for Collections in the Cloud?” (2011) http://www.collectionstrust.org.uk/is-now-the-time-for-collections-in-the-cloud/#more-665

Prakash, Sri: “Risk Management: Cloud Computing Considerations” (2011) http://ecomcanada.wordpress.com/2011/06/09/risk-management-cloud-computing-considerations/

Quantitative Estimates of the Demand for Cloud Computing in Europe and the Likely Barriers to Up-take . SMART 2011/0045 . D4 – Final Report . 2012 . [interactive]. Internet access: <http://ec.europa.eu/information_society/activities/cloudcomputing/docs/quantitative_estimates.pdf >.

Rackspace: Moving your Infrastructure to the Cloud: How to Maximize Benefits and Avoid Pitfalls. http://www.rackspace.com/knowledge_center/whitepaper/moving-your-infrastructure-to-the-cloud-how-to-maximize-benefits-and-avoid-pitfalls

Service Priorities and Best Practices for Digital Cultural Heritage (Edited by the DC-NET Working Group 3: New Services Priorities, 2012).

Unleashing the Potential of Cloud Computing in Europe. COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS (Brussels, 27.9.2012 COM(2012).

Viswanathan, Priya: “Cloud Computing – Is it Really All That Beneficial? Advantages and Disadvantages of Cloud Computing”. http://mobiledevices.about.com/od/additionalresources/a/Cloud-Computing-Is-It-Really-All-That-Beneficial.htm

Wikipedia: “Cloud Computing” http://en.wikipedia.org/wiki/Cloud_computing

WORK PROGRAMME 2013. COOPERATION. THEME 3 ICT – INFORMATION AND COMMUNICATIONS TECHNOLOGIES (European Commission C(2012)4536 of 09 July 2012).

Zeinstra, Maarten and Fallon, Julia: Europeana Cloud: Establishing the High Level Principles, http://pro.europeana.eu/web/europeana-cloud/blog/-/blogs/europeana-cloud%3A-establishing-the-high-level-principles