

Basis for decision-making in relation to the establishment of a national archiving solution based on LOCKSS PLN

Prepared by the DEFF project Permanent Access and Post-Cancellation Access – National Archiving Based on Private LOCKSS Network (LOCKSS PLN) May 2018.

Recommendations

The project recommends:

1. To establish a Danish archiving solution based on the international infrastructure LOCKSS PLN¹ with an additional registry that keeps track of the content which the individual library has permanent access to and that guarantees the existence of a copy of the content in Denmark – the same solution that the UK, Italy and Switzerland are working on.
2. To base a Danish solution on international best-practice solutions and collaboration, thereby strengthening the collaboration with the EDINA, LOCKSS Stanford and KeepSafe networks during the implementation project.
3. To focus on journals - e-books should be included sometime in the future.
4. That university libraries decide whether they wish to move forward with a national LOCKSS PLN solution as recommended by the project.
5. That university colleges and upper secondary schools decide whether they too wish to become part of a national archiving solution.
6. That DEFF supports the implementation of a national LOCKSS PLN solution, including an entitlement registry, and that they will fund a 4-year implementation project.
7. That the DEFF steering committee discusses the financing options that must be in place in an operational phase, including whether a national LOCKSS PLN solution can get a government grant or otherwise receive national funding.
8. That resources will be allocated both to the DEFF negotiations and to the implementation project. DEFF will handle the LOCKSS PLN negotiations related to the DEFF agreements - in close collaboration with the implementation project. This will ensure that conditions and agreements match actual needs and that DEFF can utilize their knowledge of and relationships with the publishers with respect to archive access, too. The experiences from Edina have demonstrated that publisher negotiations form a significant part of the implementation project and that a strong commitment and support from DEFF is necessary. The negotiations related to local agreements should be handled by the individual library - also in close collaboration with the implementation project.
9. That the collaboration with KeepSafe and EDINA will be maintained and strengthened.
10. That individual libraries consider subscribing to Portico for a period of transition in order to get access to the long list of journals that it will take time to safeguard in a national archival solution.

The background for these 10 recommendations will be presented in the basis for decision-making.

¹LOCKSS PLN: Lots Of Copies Keep Stuff Safe Private LOCKSS Network

Content

Recommendations	1
Background	2
The need for an archiving solution	3
Introduction to the Basis for decision-making	5
Introductory considerations on infrastructure	5
Danish LOCKSS PLN with an entitlement registry	7
Advantages of the chosen infrastructure	8
Content in a national LOCKS PLN	8
Portico and LOCKSS	9
Organization	10
The implementation project	11
Finances	11
Distribution model	12
Appendix 1 Infrastructure	13
Appendix 2 Budget for the implementation project	17
Appendix 3 Distribution models	19

Background

The need to safeguard the access to digital journals was addressed in the original report which set the stage for the establishment of DEF back in 1997: *"A current problem is how to safeguard access to digital journals – also more than three years back. Together with the publishers, the libraries can play a role in the future proofing of material. The problem is particularly severe as regards foreign journals. A national/shared solution model for digital archives or repositories is needed due to resource scarcity and license agreements²."*

The issue has been under continuous discussion and DEFF has supported several investigations into the matter. In 2015, Neil Beagrie issued the report *"Permanent Access to e-Journals in Denmark"*³ commissioned by the DEFF steering committee, who wanted knowledge about and input for a strategy on the issue. The report was based on interviews with seven universities and one university college and also included a review of international initiatives, projects and best practice solutions. The report was followed up by the DEFF project *"Permanent adgang til e-journals i Danmark"* [Permanent Access to e-Journals in Denmark] and the current project *"Basis for decision-making in relation to the establishment of a national archiving solution based on LOCKSS PLN."*

² Forskningsbibliotekernes IT-udvikling – Hovedrapport Danmarks elektroniske forskningsbibliotek. 20. februar 1997. UNI-C. p. 68

³ Permanent Access to e-Journals in Denmark/Neil Beagrie. Charles Beagrie Consultancy report, 12. November 2014.

The need for an archiving solution

The need for a permanent solution as regards access to digital content occurs when libraries purchase electronic material at the expense of printed materials. The gravity of the situation - and the potential consequences of not having an archiving solution - increases because the libraries increasingly purchase electronic content and discard or deselect physical material. There is simply no longer a physical collection in the libraries to fall back on when the material is electronic. In practice, the libraries have become 100% dependent the publishers' ability to handle long-term preservation, access in case of breakdowns, post-cancellation access and so on. In an ever-changing publisher world where revenues and publishing take center stage, few publishers are on top of things when it comes to archiving solutions. Consequently, there is no control over the digital collections for which Danish libraries have spent hundreds of millions of Danish kroner every year for many years.

Libraries have worked with archiving and information distribution for centuries and have professional interests, competences and experiences related to archiving that commercial publishers cannot compete with. The archiving solution will enable libraries to maintain their role as content managers in collaboration with the publishing houses. Without a national archiving solution, a small number of commercial publishers has total control of the libraries' access to content.

A national archiving solution is an archive for the electronic journals – and eventually e-books too - that Denmark has bought access to. Thus, an archiving solution is a way of securing the Danish investment in electronic material – a form of insurance that will not come into effect unless the need arises.

Over the years, the libraries have experienced losing online access to journals. As there is no Kardex, as is the case with printed journals, not all libraries have an overview of what they actually loose access to. Below, we have provided a few examples of journals that we have lost access to during this particular DEFF project and to which we would still have access, had a national archiving solution existed.

- Partridge Dictionary of Slang Online (ceased publication)
- International journal of Learning & Media
- Philosophy Documentation Center (access ceased due to cancellation)
- PsycCritiques
- Siam

For a number of years, DTU Library has had a deal with Elsevier about the delivery of metadata and full texts. A comparison of Elsevier's Freedom Collection 2017 with DTU Library's archive showed that 1.129 journals were not available in Freedom Collection. This does not mean that the consortium has lost access to all these titles. Many of the titles are former titles which the consortium has access to but they appear under new titles in the title index. This is the case for 352 titles. 209 titles are current titles but do not appear in the Freedom Collection, e.g. Cell Press titles. 43 titles are no longer current (i.e. active) subscriptions and consequently do not appear in Freedom Collection but there is nonetheless still access to them. Add to this an unknown number of titles that are available Open Access with Elsevier but do not form part of Freedom Collection.

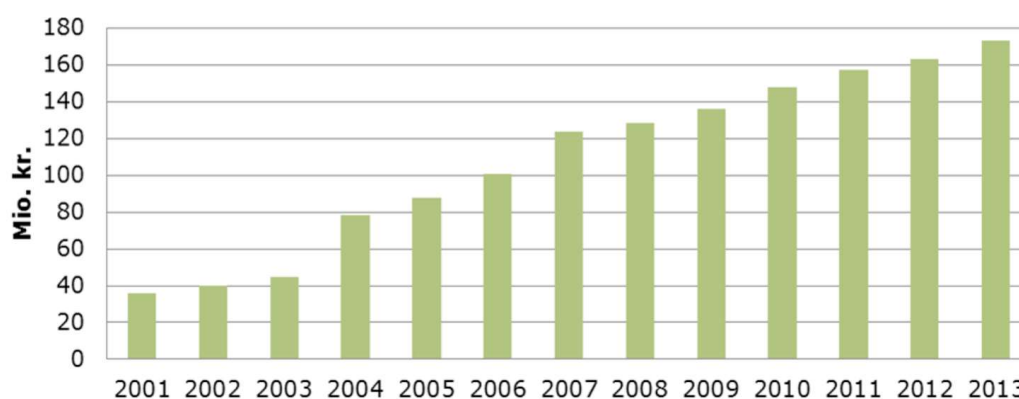
A comparison between Freedom Collection 2014 and 2017 shows 103 titles that are no longer available in 2017. 5 of these are title changes and 7 have ceased publication, but 41 titles have been

transferred to a different publishing house. Of the 50 titles that are still with Elsevier, quite a few have become Open Access journals.

In other words, it is difficult to put a number on the titles that actually disappear but judging from the comparison between 2014/2017, quite a few titles are transferred to other publishers.

Over the years, the Danish DFFU libraries [i.e. libraries that are members of the Danish Research Library Association] have spent large sums of money on subscriptions. These investments must be protected and it is in this context that the costs of a national archiving solution must be seen.

In the figure below, you can see the development of the subscription fees of DEFF for the period 2001-2013⁴



Figur 2: Udviklingen i DEFF's licensomsætning gennem årene

Figure 1: The development of DEFF's license budget over the years

In 2014, 174 million Danish kroner were spent on DEFF license subscriptions⁵ and in 2015, this amount had risen to 196 million kroner⁶. Figure 2 shows that over a 5-6 year period, DFFU libraries have invested one billion Danish kroner on DEFF subscriptions alone and that this expenditure has been rising steadily since 2001.

Advantages of a national archiving solution:

- We protect the billion kroner investment in electronic publications
- The publications will be protected/cannot be changed/removed
- We (Denmark) gain control of what we have bought access to
- We secure continued access in the case of cancellations. Cancellations due to budget cuts is an increasing problem for all institutions.

⁴ DEFFs årsberetning 2013.

https://www.deff.dk/fileadmin/user_upload/DEFF/publikationer/Aarsberetninger/DEFF_a__rsberetning_2013_180215.pdf

⁵ Permanent access to e-journals in Denmark. 8 oktober 2014. Charles Beagrie. s. 3

https://www.deff.dk/fileadmin/user_upload/DEFF/publikationer/Oevrige/Permanent_Access_to_e-Journals_in_Denmark.pdf

⁶ DEFF årsberetning 2015.

https://www.deff.dk/fileadmin/user_upload/DEFF/publikationer/Aarsberetninger/Deff_A__rsrapport2015_WEB.pdf

- Documentation of our access rights
- Archiving solution and permanent access rights to the electronic publications that we can negotiate access to in LOCKSS PLN, e.g. control over content.
- Possibility of offering Text and Datamining (TDM) services to researchers

Introduction to the Basis for decision-making

The frame of reference for this Basis for decision-making is that a Danish solution should be based on the experiences of JISC and EDINA in the UK and that an implementation project should be undertaken in collaboration with EDINA and other European countries that participate in the KeepSafe network. In addition, the project is based on and incorporates the recommendations of the first DEFF project on permanent access – including the recommendation that DEFF should support and prioritize the development of a sustainable and robust national archiving solution based on Private LOCKSS Network that safeguards copies of journals in Denmark.

The following sections describe the project's basis for decision-making for the implementation of a national archiving solution based on LOCKSS PLN.

Introductory considerations on infrastructure

In general, the project recommends that a digital copy of e-resources in Denmark is archived via a national non-commercial solution. This control of content means that the technical solution can be replaced if a different solution than LOCKSS PLN is needed or preferred in 10 or 20 years.

The design of an archiving solution based on the LOCKSS system depends on how you relate to these four questions:

- Do you want a "trusted" or an "untrusted" model?
- Do all participants have the same content?
- A solution with or without rights management?
- A solution with or without dedicated ingest boxes?

Trusted or untrusted model?

A trusted model presupposes 1) that all participants have the same content or, 2) that an entitlement registry has been added to the architecture so that you can control the access rights to the archive.

If all participants have the same content, we will be able to make a simple trusted solution without rights management – a solution which only contains the shared content. An example of such a model exists in Germany where they have bought national back files and base their LOCKSS PLN on the shared content.

In an untrusted model, the archive boxes do not contain identical content; as such, there is no rights management involved and so each box must harvest data from the publisher. Only the boxes that have a license can have a copy and so the idea of having a national common body of copies is lost. Each box must be able to harvest data, i.e. this part of the process must be repeated each time by 6 boxes. There are many disadvantages associated with an untrusted solution and it cannot be recommended.

Does everyone have the same content?

The content, and the extent to which the participants have access to identical content, is of vital importance in deciding on a solution. For this reason, the project has analyzed the extent of shared content of e-journals in Denmark.

License agreements with shared content:

- Elsevier Journals Freedom Collection 2014-18
- T & F DEFF database SSH Collection 2018-20
- Springer eJournals 2017-19*
- ACM Digital Library - Association For Computing Machinery 2015-2018*
- Wiley Journals 2016-18*
- JSTOR. Arts & Sciences – 1*
- CUP journals 2018*
- OUP journals 2016 collection*
- Emerald 60 collection*
- Sage 2016 collection*

License agreements marked with an * have the same permanent access rights for all if anyone withdraws from the agreement. As regards the others, the university libraries have access to different journals and volumes and the content of an archiving solution will therefore not be identical. The intersection of shared content amongst the Danish university libraries is minimal and if we include the agreements of the University Colleges, there is no shared content left. Thus, a trusted model based on 100% shared content is not an option in Denmark.

With or without rights management

If a trusted model based on 100% shared content is not an option, some sort of rights management must be implemented to make sure that the institutions only gain access to the content they have the rights to. This means that an entitlement registry must be set up for all boxes in the network, containing information on the rights of the participating institutions.

With or without dedicated ingest boxes

Ingest boxes harvest data from publishers and can be used in a trusted model. A few ingest boxes harvest data from publishers on behalf of all participants in the network. This way, the traffic to the publishers as well as the strain on the individual LOCKSS boxes is minimized – in contrast to an untrusted model.

Conclusion

The conclusion is that since everybody has access to different content, we need a solution where all boxes can archive all content, interdependent of licenses. The condition for this to work is the set-up of an effective access management - i.e. a shared entitlement registry must be added to all boxes. Dedicated ingest boxes are an advantage and will also contribute towards securing an effective and cheap operations set-up.

Danish LOCKSS PLN with an entitlement registry

A LOCKSS PLN based national solution with an entitlement registry will be able to manage various content; ingest boxes can be used to optimize the daily tasks and you can make copies of the content on all boxes, thereby improving archive safety. This is the model that is recommended by Stanford's LOCKSS team and that has been chosen in the UK, Italy and Switzerland.

The project recommends to base the architecture for a Danish LOCKSS PLN with an entitlement registry on EDINA's model, which is illustrated in Figure 1 below.

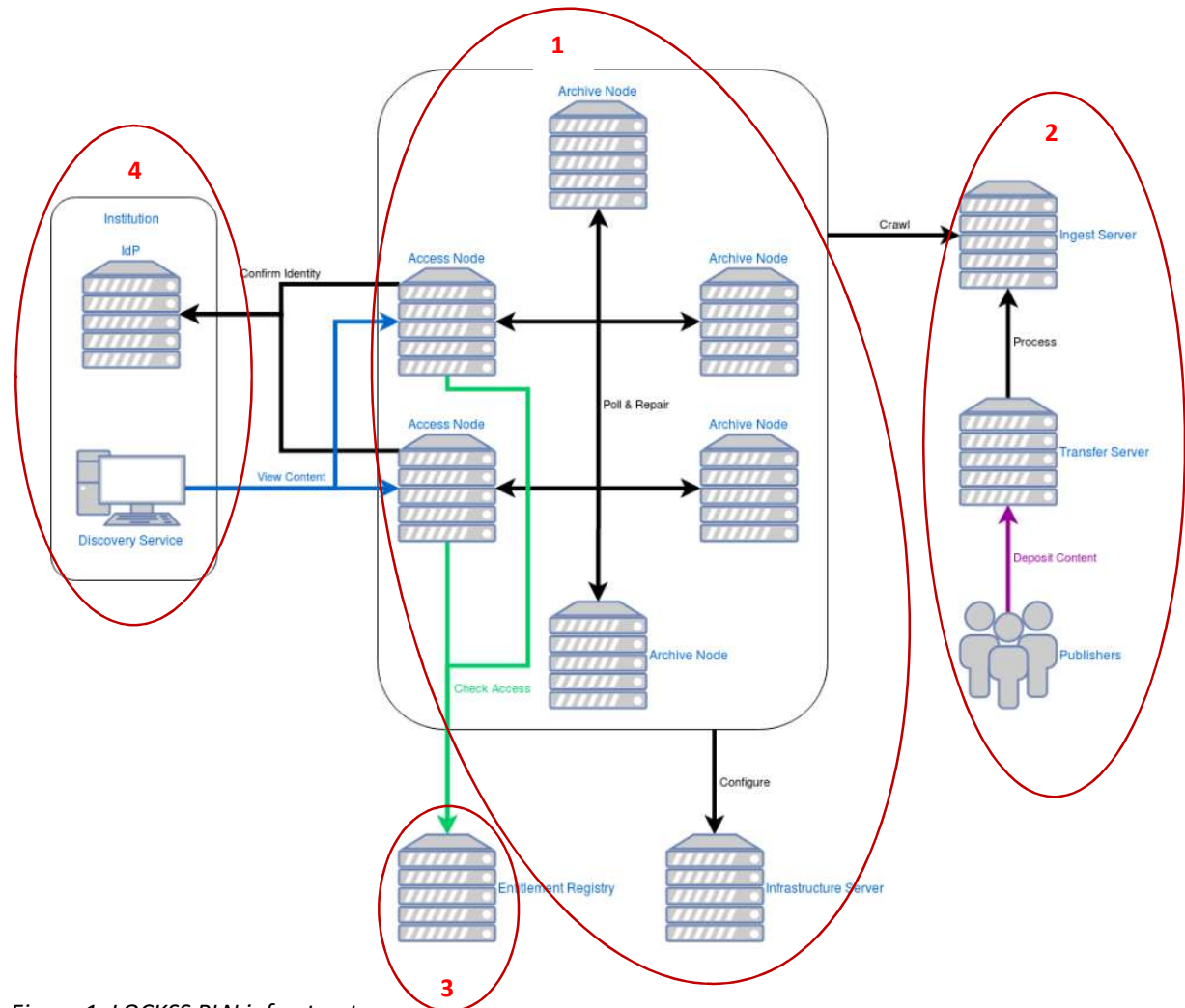


Figure 1: LOCKSS PLN infrastructure

The infrastructure has 4 main functions:

1. LOCKSS PLN
2. Import function
3. Entitlement registry
4. Access control

The infrastructure is described in detail in Appendix 1.

Advantages of the chosen infrastructure

The project finds that the recommended infrastructure has the following advantages:

- Metadata as well as full texts will be archived on Danish soil which will safeguard both access and ownership on national as well as local level.
- The solution works independently of publishers and is not sensitive to changes in publisher ownership
- The solution will secure access during temporary breakdowns
- The import function will be built on existing architecture and know-how
- The entitlement registry will be built on existing architecture and know-how
- LOCKSS PLN builds on the national experience gained during the PLN test in the previous DEFF project as well as on international best-practice and network collaboration
- Access is based on international best-practice and experiences (EDINA/WAYF)
- The infrastructure will safeguard Lots of Copies and thus a high level of security
- We can manage different content, which is necessary in Denmark
- Less traffic to the publisher due to the ingest boxes
- In an audit repair process between boxes, the boxes do not have to document their rights to data amongst themselves – the rights are handled in the entitlement registry
- Since the box will not reflect the rights of individual libraries, it can be hosted by others. Thus, there will be fewer boxes, lower costs and a more effective infrastructure
- Easier for libraries to participate as the institutions do not need to operate or manage a system themselves
- An entitlement registry is being developed by EDINA and we can reuse their solution and knowhow
- Many institutions are contributing with plugins and it is an international solution
- Under implementation in the UK, Italy and Switzerland
- Recommended by Stanford's LOCKSS team

Content in a national LOCKS PLN

The experiences from EDINA shows that the rights negotiations in relation to permanent access based on a national LOCKSS PLN are comprehensive, both negotiation- and contract-wise. In the UK, license and archive access negotiations are handled separately - i.e. EDINA negotiates agreements in relation to their LOCKSS PLN. EDINA has made a standard agreement for permanent access specifically which e.g. contains information about publishers and about EDINA's responsibility as an archive host. The advantages of a separate agreement for archiving is that:

- 1) You can keep the two negotiations apart contract-wise and thereby obtain an agreement that is as clear and concise as possible
- 2) The archive agreement will be valid even if the license agreement is cancelled
- 3) Products that have been purchased (e.g. books and archives) can be negotiated independently of renewal
- 4) The license agreement will not be "polluted" by technical details
- 5) You can draw up an almost finalized contract template to DEFF so that they do not need to be familiar with the technical details

EDINA is currently in dialogue and negotiation with five publishers: Oxford University Press, Edinburgh University Press, Taylor & Francis, Brill and Cambridge University Press.

In order to concretize and estimate an implementation project, it is necessary to be realistic about the content. Obviously, the solution is scalable and more content can be added in time, but the starting point should be as concrete as possible. In addition, EDINA has paved the way for the implementation project due to their dialogue with the five above-mentioned publishers, who will therefore be familiar with the project's needs and solution models. For this reason, the project's estimation of the implementation project is based on these four publishing houses:

- Oxford University Press
- Taylor & Francis
- Brill
- Cambridge University Press

Edinburg University Press has been left out as this publisher is not part of a consortium agreement with the Danish university libraries. In addition to the four publishers, the budget for the implementation project includes resources for the addition of at least 3 more publishers. Taking into consideration the EDINA experiences, the amount of common content as well as the negotiation experiences, the final choice of both negotiation strategy and the publishing houses that should be negotiated with must be considered in the implementation project.

Portico and LOCKSS

Portico is a digital preservation service offered by the non-profit organization ITHAKA. Portico was launched in 2002 and is a digital archive for academic literature that libraries can subscribe to. I.e. there is no local technology and no copy of the content in the libraries.

The first project on permanent access took as its starting point the requirements that the libraries had asked Neil Beagrie to take into consideration and the project compared three systems in relation to the libraries' requirements of an archiving solution. The result of the comparison can be seen in the figure below:

DANISH REQUIREMENTS UPDATED JANUARY 2016			
MANDATORY NOW	• GLOCKSS	PLOCKSS	Portico
Supports post-cancellation access	• YES	YES	YES
Supports long-term access	• YES	YES	YES
Improves on current position	• YES	YES	YES
International solution	• YES	YES/Partial	YES
Coverage of DEFF national framework licences	• 13%-43,4%	Require negotiation	27,3%-82,7%
Cost is affordable	• ?	?	?
DESIRABLE NOW			
Out-sourced	• NO	NO	YES
Subscription model	• YES/Partial	YES/Partial	YES
Response time instant for customer	• YES) (+temporary breakdown)	YES (+temporary breakdown)	NO (30 days delay. Not temporary breakdown)
Based in or mirrored in Europe	• YES	YES	NO
100% coverage of DEFF national framework licences	• NO	NO	NO
HIGHLY DESIRABLE OR MANDATORY NOW			
Based in or mirrored in Europe	• YES	YES	NO
100% coverage of DEFF national framework licences	• NO	NO	NO

The coverage of Portico varies from library to library but the advantage of Portico is that the solution is already up and running and that it can be used in a transition period until an adequate national archiving solution has been established.

Apart from storing local copies of the content in Denmark, one of the advantages of a Danish LOCKSS PLN is that we can control when to open for access and that we can open for access in the case of temporary publisher breakdowns. In comparison, it takes on average 30 days to accomplish a request to open for access at Portico and there is no option of opening in the event of a temporary publisher breakdown.

Organization

The project's considerations on organization are:

1. That an archive agreement, in sync with the EDINA experience, should be considered an agreement in itself in addition to a license agreement
2. That there is a need for focus on archiving and that such a focus is time consuming
3. That the organization must be able to accommodate all research libraries

The project's recommendation is that national LOCKSS consortium agreements are set up via DEFF. In addition, it is recommended that we establish a project organization in connection with the

implementation of a national archiving solution. The project organization must have a steering committee, a working group as well as a coordinating function that handles administration, coordination of development wishes, development, dialogue with DEFF etc.

The implementation project

The implementation of a national archiving solution based on LOCKSS PLN is fraught with many unknown factors that are difficult to foresee and estimate beforehand. Due to the collaboration in the KeepSafe network, with EDINA as well as Stanford's LOCKSS team, we already know a lot. *In other words, we cannot prepare more at this point – next step must be to set in motion the implementation of a national archiving solution based on LOCKSS PLN.* We can be pragmatic about timeframes, though, and ensure that we learn as much as we can before the implementation project develops into a regular operational task. The project's recommendation is that a future implementation project must run for at least for 4 years, two of which must be set aside for a designated preparatory period for the operational stage when actual experiences on how to operate the overall set-up are collected: infrastructure, addition of new import sources, administration of entitlement registry and negotiation of agreements. In addition, the implementation project will provide an opportunity of gradually adding more libraries.

Finances

There are a number of conditions that must be met if the 4 year budget is to be realistic

- 1) Purchase of EDINA in order to develop further an entitlement registry for Denmark. The advantage of this is that we will be able to utilize the skills and competences inherent in EDINA to the advantage for all who use the entitlement registry.
- 2) Purchase of EDINA to manage the entitlement registry for Denmark.
- 3) KB and DTU Bibliotek will manage the national LOCKSS PLN during the 4-year project period. KB København, KB Aarhus and DTU Bibliotek have the necessary technical competences and have tested LOCKSS. When the necessary experiences, skills and knowhow have been obtained, it must be decided whether this will be the management set-up in the future or if another organization should take over the operation after the 4 years have passed.
- 4) DTU Bibliotek will be responsible for the import function during the 4-year project period. This is due to the library's 20-year experience with data import and normalization of data - DTU's current infrastructure for this is necessary for getting the project implemented. What it would cost for another organization to establish an equivalent set-up and knowhow, is difficult to ascertain. It must be decided whether this will be the management set-up in the future or if another organization should take over the operation after the four years have passed.
- 5) The project has not budgeted for the time that DEFF needs for project participation or for negotiating archiving agreements. Experiences from Edina show that negotiations with publishers are a major part of the implementation project and that a strong commitment and support from DEFF is necessary.

The estimations in the budget have been made in collaboration with EDINA and Stanford's LOCKSS team. The budget for the implementation project is available in Appendix 1.

Distribution model

Before developing the implementation project, the project group designed a distribution model - the idea being that the distribution model could and should be agreed upon before the initiation of the actual implementation as this would contribute to ensuring the sustainability of the entire project. In the distribution model, the financial starting point is the mutual costs that the project foresees will be current in the 4th year of the project, which is closest to the operational phase. The figures are mere estimates and involve a high degree of uncertainty, i.e. the figures only give an idea of the expected costs but are not precise at present. The idea is to give the libraries an idea of what costs to expect in the operational phase before the project is actually initiated.

The project group recommends that we decide on one of the two following distribution models:

Model 1 operates with a tier-division of the participating institutions (see Appendix 3 for an elaboration of the tier classifications as well as a financial estimate of costs per institution)

The advantages are that the tier-model is known already. If the price is graduated, it will enable other types of institutions than university libraries to participate.

The disadvantages are that an agreement must be reached about the principles behind the tier-model; it takes resources to administer the model and allowances must be made for changes in the organization and size of the individual institutions.

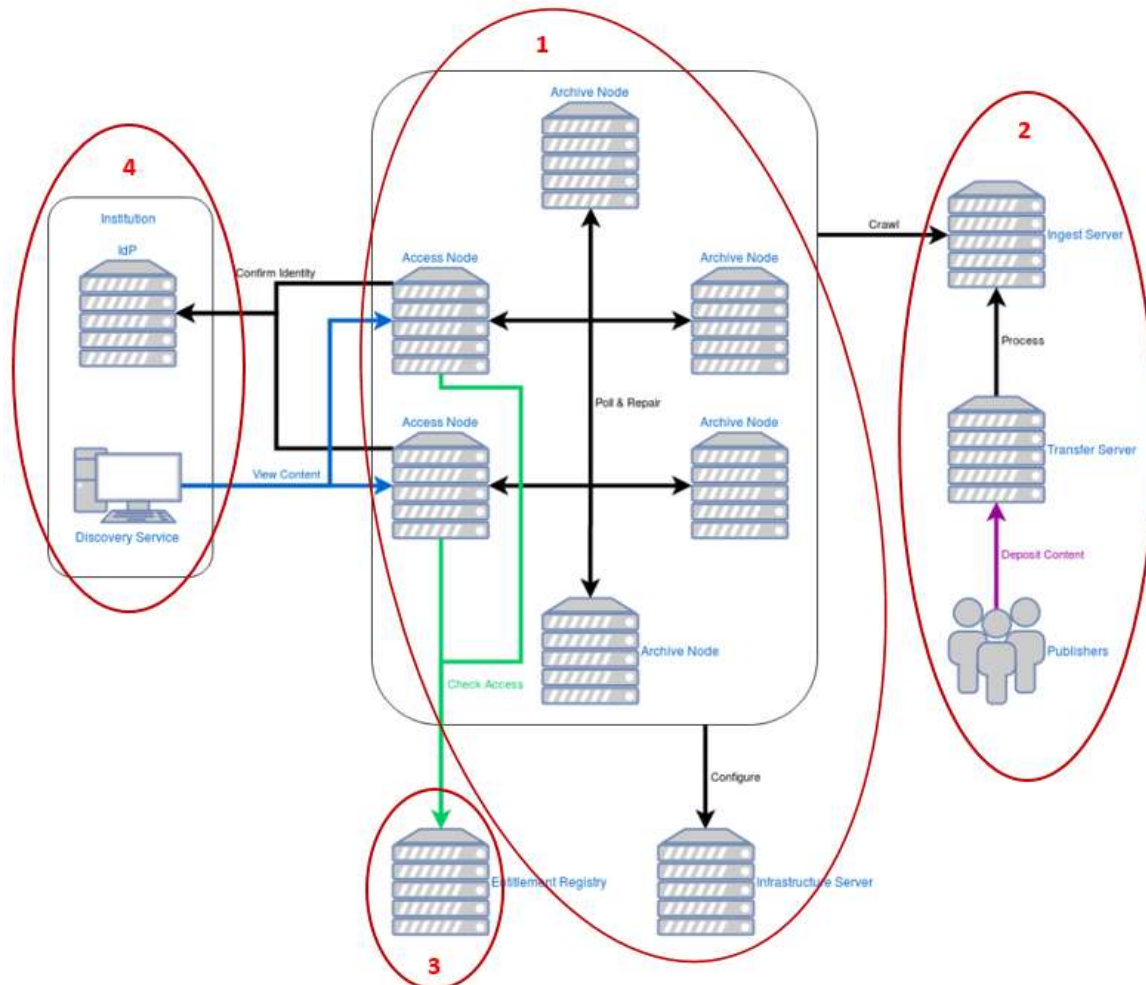
Model 2 is a pledging model in which all institutions pay the same amount no matter their size (see Appendix 3 for an elaboration). The amount that must be paid/pledged will be calculated as a maximum amount based on a given number of participating institutions.

Advantages: the maximum price is known beforehand and it will be easy to make a decision on whether the individual institution can afford to participate in a PLN. The pledging price may also fall if more institutions sign up than had been anticipated in the calculation of the maximum price. The more participants, the better we can guarantee purchasing content at a lower price.

As regards both models, the participating institutions must consent to the following basic principles:

- If an institution participates in the PLOCKSS-collaboration, it participates in an agreement if the institution has or has had a subscription with a publisher – i.e. payment is made on the basis of participation in the publisher agreement
- The number of agreements with the publisher depends on how much must be programmed (journals, books, archive). Payment is per plugin.
- Institutions that subsequently take out a subscription on an agreement that is in, will only need to pay operating costs.
- The pledging model is also based on the fact that an institution will participate if it has or has had a subscription with a publisher with whom an agreement is being made
- FTE figures are disclosed once a year; in the event of a change in institutions with resulting changes in FTE of more or less than 10 %, this will give cause for a recalculation and possibly a replacement in the tier system.

Appendix 1 Infrastructure



Figur 1: Dansk LOCKSS PLN infrastruktur

Figure 1: LOCKSS PLN Infrastructure

LOCKSS PLN

LOCKSS PLN are the actual archive parts of the infrastructure where metadata and full texts are archived on a number of LOCKSS archive nodes, whereby Lots of Copies are guaranteed. This is only possible with an entitlement registry model – i.e. all archive nodes can archive all content and the access is administered via rights management. In addition, the infrastructure makes sure that damaged full text can be repaired between the individual archive nodes without contacting the publisher/supplier.

An archive node is a server that stores content. Archive nodes are connected og perform "poll and repair" with the intention of safeguarding the integrity of the content.

An access node has the same function as an archive node but in addition makes the content available to the users. All in all, Stanford LOCKSS recommends 6 access nodes (access/archive) in a LOCKSS PLN.

The infrastructure server contains figuration and code components and makes sure that the individual nodes are configured the same way and that maintenance is minimized. Stanford's LOCKSS team and EDINA recommends that this server is operated and maintained at a national level.

The import function

In the import function, metadata and full texts are harvested from publishers to the transfer server via API/FTP. The publishers deliver metadata in various formats, i.e. the data that is received from the publishers must be normalized and standardized before they are made available to the LOCKSS PLN boxes from the ingest server. According to EDINA, the fact that the data from the publishers come in many different formats is the main resource challenge associated with establishing and maintaining an archive. Experiences show that a regular operations and management task must be expected e.g. monitoring that data is received and monitoring ongoing changes in data processing and data collection e.g. when changes occur at the publishers' end.

The ingest server collects content from the transfer server and makes the content available to the individual archive nodes in a suitable format. One of the advantages of the national solution which the project recommends is that only one ingest server is needed - not six.

This happens via API or FTP, depending on which solution the individual publishers support. Stanford LOCKSS has developed a number of plugins for different publishers, plugins that inform archive nodes on how data is to be collected. These plugins can be used by everybody, i.e. for some publishers, it is not necessary to develop plugins – you simply adjust the existing ones. The publishers deliver metadata in various formats, i.e. you need to be able to normalize and standardize the metadata that is received from the publishers before they are made available via the ingest server to the LOCKSS PL boxes.

According to EDINA, the primary resource challenge in relation to establishing and maintaining an archive is the fact that data is received in many different formats from the publishers. Experiences indicate that a regular operational task must be anticipated in relation to monitoring that data is received and monitoring ongoing changes in data management and data harvesting when the publishers change something.

The ingest server harvests content from the transfer server and makes the content available for the individual archive nodes in a suitable format. One of the advantages of the national solution which the project recommends is that only one ingest server is needed (not six)

Entitlement registry

As an integrated part of the LOCKSS PLN network, an entitlement registry is set up which is developed by EDINA. The entitlement registry maintains the access to the content in the archive in compliance with the licenses. In addition, the entitlement registry provides an overview of

permanent access rights that both publishers and libraries can agree on. An entitlement is a record that contains information about access rights in relation to:

- A resource
- An institution
- A publisher
- A specific period of time

The entitlement registry consists of a number of components:

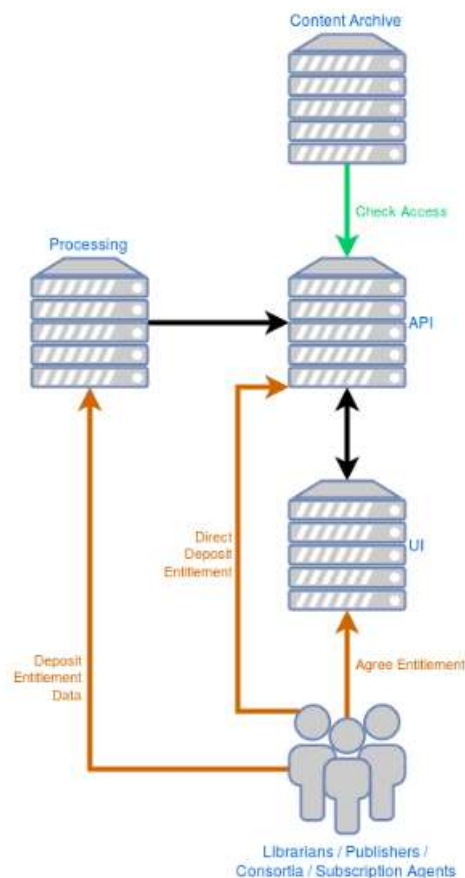


Figure 2: Entitlement registry

The API accepts and preserves information about rights. The API answers calls from the LOCKSS box about whether or not to provide access to content.

The user interface will enable publishers and libraries to see and agree on rights.

The processing of data about rights is the most advanced part of operating the rights registry because data that documents rights are often difficult to find – both for publishers and libraries.


Information about rights can be found in different formats and via different systems.


The figure below shows the demo model of KeepSafe's entitlement registry:

entitlement.ac.uk Home Your Entitlements Summary

KEEPSAFE EUROPE DEMO ENTITLEMENT REGISTRY


Clarifying library entitlement history






DISCOVER PUBLISHER ENTITLEMENT RECORDS

We work with publishers to record institutional entitlements directly from publisher systems.




UPLOAD LIBRARY RECORDS AND COMPARE DIFFERENCES

Upload your own library entitlement records. We then provide a simple interface to compare between library and publisher supplied data.



AGREE A SHARED VIEW WITH PUBLISHERS

We work with publishers to record institutional entitlements directly from publisher systems.



BUILD AN AUTHORITATIVE RECORD IN CASE ACCESS IS LOST

Libraries are making a significant investment in their perpetual access rights. We give assurance that you can add those rights when needed.

Access control

A user finds an article in a library discovery system and has problems accessing the article on the publisher's website, discovers that the journal subscription has been cancelled or that the journal no longer exists. The user's access to the article is verified in the entitlement Registry via the institution's access system (Federated Identity Management/WAYF, e.g. EZproxy).

Appendix 2 Budget for the implementation project

Budget for the 4-year implementation project

Spreadsheet - Major Project Grant to Denmark's Electronic Research Library	
Project title	Permanent Access
Project manager (person who has the legal responsibility)	
E-mail	
Phone	
Revenue	Budget
Self-financing	kr 1.752.000
DEFF Grant Application Amount	kr 6.774.792
Revenue in all (incl. VAT)	kr 8.526.792
Notes to revenue, if any	
Expenditure	Budget
Wages, project participation	kr 3.650.000
Overhead of project participation, if any	kr 730.000
Wages project participation + overhead, in all	kr 4.380.000
Wages, project management	kr 730.000
Overhead of project management, if any	kr 146.000
Wages project management + overhead, in all	kr 876.000
Expenses for meetings and travel	kr 200.000
Expenses for consultancy services	kr 2.075.792
Hardware acquisition (server+storage)	kr 250.000
Expenses for management of infrastructure (for 4 years)	kr 745.000
Expenses for the documentation of effect (in addition to the expenses listed above)	kr -
Expenditure in total (incl. VAT)	kr 8.526.792
Result	kr -
Notes to expenditure, if any	
Consultancy services: Expenses for entitlement register/Edina (year 1: 24.875€, year2:22.749€, year3:21.000€, year4:21.000€ and expenses for LOCKSS fee (year1: 50.000\$, year2:47.500\$, year3:45.000\$, year 4:42.500\$), and in addition 200.000 DKK to further development of entitlement register	
Regards VAT, please note:	

Expenditure in total per year

Expense	Year1	Year2	Year3	Year4	In total year1-year4
Participation in KeepSafe	0,25	0,25	0,25	0,25	1
Hardware maintenance	0,75	0,75	0,75	0,75	3
Import function maintenance	0,5	0,25	0,25	0,25	1,25
Import function new sources	2	1	1	1	5
LOCKSS PLN	0,75	0,75	0,75	0,75	3
Coordination of management of infrastructure (20% PL)	0,85	0,6	0,6	0,6	2,65
Entitlement registry (coordination of development plans, wishes etc.)	6	1,5	1,5	1,5	10,5
Entitlement registry (clarification of entitlements)	6	6	6	6	24
Access	1,5	0	0	0	1,5
Participation in working group incl. assignments/support DEFF	12	12	6	6	36
Expenses for management of infrastructure (in total per year)	4,85	3,35	3,35	3,35	14,9
Manmonths (in total per year)	25,75	19,75	13,75	13,75	73
Project management	5,15	3,95	2,75	2,75	14,6
Wages excl. overhead (MM x price per MM)	1.287.500	987.500	687.500	687.500	3.650.000
Overhead (OH), 20% of wages	257.500	197.500	137.500	137.500	730.000
Wages incl. 20% OH	1.545.000	1.185.000	825.000	825.000	4.380.000
Self-financing (40% of wages)	618.000	474.000	330.000	330.000	1.752.000
Wages, project management excl. overhead	257.500	197.500	137.500	137.500	730.000
Overhead (OH), 20% of project management	51.500	39.500	27.500	27.500	146.000
Wages, project management incl. 20% OH	309.000	237.000	165.000	165.000	876.000
Expenses for management of infrastructure (no OH)	242.500	167.500	167.500	167.500	745.000
DEFF grant (wages + management of infrastructure)	1.478.500	1.115.500	827.500	827.500	4.249.000
Expenses for meetings and travel	50.000	50.000	50.000	50.000	200.000
Hardware acquisition (2x server & 2x10 TB)	62.500	62.500	62.500	62.500	250.000
Expenses for LOCKSS fee (year1: 50.000\$, year2: 47.500\$, year3: 45.000\$, year4: 42.500\$)	303.500	288.325	273.150	257.975	1.122.950
Expenses for entitlement register/Edina (year1: 24.875€, year2: 22.749€, year3: 21.000€, year4: 21.000€)	208.950	191.092	176.400	176.400	752.842
Further development of entitlement register	50.000	50.000	50.000	50.000	200.000
Expenses for the documentation of effect	-	-	-	-	0
Expenses financed by DEFF:	2.153.450	1.757.417	1.439.550	1.424.375	6.774.792
AFTER THE PROJECT, EXTERNAL EXPENSES PER YEAR WILL BE (=EXCES:	817.450	709.417	679.550	664.375	&ltaverageyear> 717.698
[=expenses above for Hardware, LOCKSS and EDINA]					
OBS: Based on exchange rates as mentioned in the partner budget					
To this must be added staff resources, the costs of which the institutions already cover - as well as meeting and travel expenses that occur after the implementation project has ended. Expenses for the development of the entitlement register after the project has ended have not been included.					

Appendix 3 Distribution models

The Tier Model (Model 1)

	Price per institution	In total per tier	Per cent share per institution
Tier 1	149.625	299.250	22,5 %
Tier 2	99.750	199.500	15%
Tier 3	66.500	133.000	10%
Tier 4	33.250	33.250	5%
Tier 5	N/A		
In total		665.000	

The tier model has been made on the basis of the universities' share of the collective FTE figures for the 7 universities (round up/down to the nearest higher/lower figure to make the figures fit)

KU 25,1 %
 AU 23,16 %
 SDU 14,42 %
 AUB 13,8 %
 DTU 8,8 %
 CBS 9,6 %
 RUC 4,8 %

The pledging model (model 2)

The estimated cost to be pledged for is 640.000 Danish kroner. In this scenario, only the 7 universities participate.

Price DKK 665.000

Number of institutions 7

Price per participant

DKK 99.000

If more institutions sign up, the agreement will be cheaper.

In the pledging model, all institutions will pay the same amount no matter their size.