Consultancy services for a new infrastructure for digital archiving

Case

The Danish National Archives (DNA) seeks to replace its current infrastructure for digital archiving within a few years. Preferably with an off-the-shelf solution using open standards and open source and providing automation, independence of individuals and easy maintenance. For that purpose, we request a Consultant (or Consultants) who can assist DNA in getting knowledge of relevant infrastructures and specifying the requirements. The number of services required by the Consultant are defined in more detail below. The final product is expected to augment DNAs existing domain knowledge and contribute to future advertisements of call for tenders in relation to potential IT acquisitions.

Background

The DNA is an authority and research institution under the Ministry of Culture with 250 employees that resides in Copenhagen, Odense, Viborg and Aabenraa. For more than four decades, the DNA has performed digital archiving by collecting, preserving and giving access to data from the public and private sector.

Data are submitted from the state, regions and municipalities of Denmark as well as from private organisations, individuals and research institutions. Notification to and approval of new IT systems by the DNA is mandatory for state authorities. Submission of data to the DNA is mandatory for all state authorities.

Access to data is dependent on application, apart from those that are now publicly available. Data are submitted to the DNA in a system independent data format structured as relational databases with or without documents. The DNA also preserves huge amounts of scanned data, *digitized record*s - i.e. photographic scans of traditional paper records.[1] [2]

Current infrastructure and capacities

DNA's current infrastructure and its components is mostly developed in-house. It is based on the OAIS standard and consist of both software and hardware. On a yearly basis, the current infrastructure ingests 400-500 submissions with sizes ranging from 0.1 GB to 15 TB with a yearly average of 50 TB in total. Data is validated in the ingest process, using a validation tool developed by

[1] See *Strategy for archiving digital records at the Danish National Archives:* https://www.sa.dk/wp-content/uploads/2014/12/Strategy-for-archiving-digital-records-2013.pdf
[2] See *Executive Order on Information Packages:* https://www.sa.dk/wp-content/uploads/2020/05/Executive-Order-on-Information-Packages128-1.pdf

the DNA. This assures that the submission information packages comply with the specification as defined in Executive Order 128.

Archival Storage is based on distributed digital preservation using three independent sets, two on tape and one on optical disc. The amount of born-digital data is currently approximately 300 TB per set for submitted data – 900 TB in total. The amount of digitized data, currently stored outside the archival infrastructure, is around 600 TB. Regarding Preservation Planning, a migration from the previous generation of optical media to the current generation is in progress. Further, a migration of data from the previous AIP specification to current is prepared. Access to data is given frequently, with up to one third of the data being migrated from AIPs to DIPs.

Future expectations
In the coming years the DNA expects single submissions in the area of 30 TB with a yearly average above 100 TB and we expect to implement IP specifications based on the E-ARK3 Common Specification for Information Packages (CSIP). Although the DNA requires digital archival materials ingested in certain formats, we also expect the introduction of new submission and preservation formats such as SIARD 2.2, new formats for geodata and other data structures. Furthermore, we expect the around 600 TB of digitized data, increasing with 75 TB annually, to be managed by the future digital archiving infrastructure.

Required services

- Creating detailed system requirements for a Digital Archiving infrastructure (in coordination with the DNA project team). These are described in a system requirements specification.
- Examining the market for potential Digital Archiving solutions.
- Creating a list of solution providers.
- Setting up an evaluation scheme for evaluating the fulfilment of the solutions to the requirements.
- Evaluating the solutions towards the requirements and presenting these in form of an in-depth evaluation report.

Product
The aforementioned services are expected to be presented for the DNA project team in the form of numerous analysis, papers, reports etc. on a regular basis throughout the autumn/winter of 2020.

The final evaluation report must include an in-depth analysis of the possible solutions and scenarios with emphasis on price, modularity and integration. We expect all services to be delivered in close collaboration with the DNA project team.

Delivery of the final product is only considered to have taken place once the Danish National Archives has approved the final report, which allow the project to finalise its analytical phase.

Time schedule

The products must be delivered before the end of 2020 (December 18th). Hence the consultant(s) is/are expected to spend a maximum of 4 calendar months from the end of August until the end of December.

Place of work
No physical presence in Denmark is required. Ongoing communications as well as potential presentations are to take place on platform(s) chosen by DNA.

Requirements for the Consultant(s)

Education:
The Consultant(s) must have a university degree in the areas of *Computer Science*, *Engineering* or similar areas relevant to Digital Archiving.

Positions:
The Consultant(s) must be able to document relevant experience from positions in one or more of the following organisations: research institutes, archive institutions, software development companies, management consulting companies.

Experience:
The Consultant *must* have experience with:

- System Requirements Specification
- System Design and System Integration

The Consultant *should* have knowledge about:

- Digital Archiving
- The OAIS reference model
- The suppliers of systems to the area of Digital Archiving
- Open standards and open source

Language
All communication and writing must be in English or Danish.