

Digital bevaring med Archivematica

Arne-Kristian Groven,
Riksarkivet

arngro@arkivverket.no

KDRS-samling, Trondheim, 8.juni 2016

Litt om hva Archivematica er...

BAKGRUNN

Archivematica...

- Er et digitalt bevaringssystem, basert på dagens beste praksis og «state of art» standarder
- DC (Dublin Core), METS, PREMIS, Bagit og OAIS
- Tilbyr en (veldokumentert) OAIS-arbeidsflyt med manuelle eller automatiserte sjekkpunkter, via et web-basert grensesnitt («dashboard»)
- Tar imot mapper med filer som input, uten krav til filformatene. Konvertering av filformater skjer automatisert og regelstyrt ved generering av visningspakker (DIPer) og arkivpakker (AIPer)

Archivematica

- Tilbyr avanserte søk via verktøyets web-basert grensesnitt («dashboard»), hvor hele pakker (AIPer) kan lastes ned, eller (AIC) samlinger av AIPer, eller digitale objekter inneholdt i AIPer.
- Tilbyr integrasjon mot ulike depot- og aksess-systemer: AtoM, DSpace, CONTENTdm, Islandora, LOCKSS, DuraCloud, Arkivum, OpenStack and Archivists' Toolkit
- Eller, i sin enkleste form, så lagres pakkene i dedikerte fil- og mappestrukturer.

Fri programvare

- Fri programvare / åpen kildekode, under AGPLv3 (GNU Affero General Public License)
- Programkoden (kildekoden) er fritt tilgjengelig for enhver
- Dokumentasjonen likeså, under en Creative Commons Share-alike (del-på-samme vilkår) lisens
- Enhver kan studere, kvalitetssikre, tilpasse og distribuere ut fra spillereglene disse lisensene gir.

Leverandør og forretningsmodell

- Artifactual Systems, fra Canada, har utviklet og er hovedansvarlige for vedlikehold og videreutvikling av programvaren.
- De har kontroll over kodebasen, men andre (organisasjoner, fortrinnsvis) kan spille inn endringsønsker og bidra med finansiering for å få realisert sine ønsker (dersom disse aksepteres).

Eksempel fra versjon 1.4

Archivematica 1.4

Released May 27, 2015

This release improves and facilitates integration with a number of access/deposit systems, including CONTENTdm, Islandora and DSpace, as well as bug fixes and other small usability improvements.

New features

- **Sponsored** (University of Saskatchewan Library) Fedora/Islandora Integration
- **Sponsored** (GE Aviation) Bag ingest improvements ([#8309](#), PR191)
- **Sponsored** (University of British Columbia Library) DSpace Ingest improvements ([#6273](#), [#5889](#), PR154)
- **Sponsored** (University of British Columbia) CONTENTdm dip upload enhancements ([#8039](#), PR148)
- Transfer/SIP Creation feedback ([#7853](#), PR171)
- SIP Arrange enhancements (improved transfer METS) ([#7714](#), PR145)
- Review/Download DIP objects (See CONTENTdm DIP upload enhancements, [#8039](#))
- Processing location cleanup via dashboard ([#7921](#), PR199)
- Improved logging (backend only, no user-facing functionality) ([#6647](#), PR34, PR201)
- **Sponsored** (Columbia University) Extract Packages Recursively ([#8438](#), PR164)

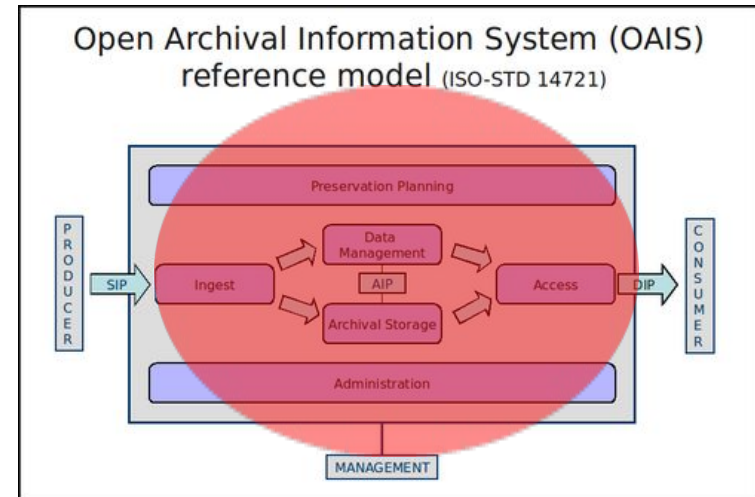
Bug fixes

- Manual normalization fixes ([#6870](#), PR97, PR186)
- Improvements to metadata.csv handling ([#7901](#), [#8007](#), PR210)
- Accession ID's now working ([#7442](#), PR175, PR134)
- Sanitize Transfer Names ([#7937](#), PR170)
- Fix Remove Transfer Components ([#7482](#), PR172)
- Timestamps not updating fix ([#7859](#), PR162)
- Remove events for sanitize directory names ([#8033](#), PR177)
- AIC fixes ([#8258](#), [#7155](#), PR194, PR168)
- Verify checksum fail causes SIP failure ([#8225](#), PR190)
- Various Archival Storage search fixes (PR213, PR215, PR216, PR217, PR221, PR222, PR223)
- DIP Store location preconfiguration ([#7321](#), PR209)

New and upgraded tools

- Elasticsearch Updated to version 1.3
- FITS Updated to version 0.8.4
- FFmpeg Updated to version 2.6
- Automation tools: DSpace fixes (AT PR 2)
- **New** Support for Siegfried 1.0 for file identification (PR202)

Archivematica realiserer arbeidsprosessene i OAIS standarden



archivematica. Transfer ¹ Ingest Archival storage Preservation planning Access Administration elark ▾

Standard ▾ /home ▾ Browse Start transfer

Type Transfer name Accession no.

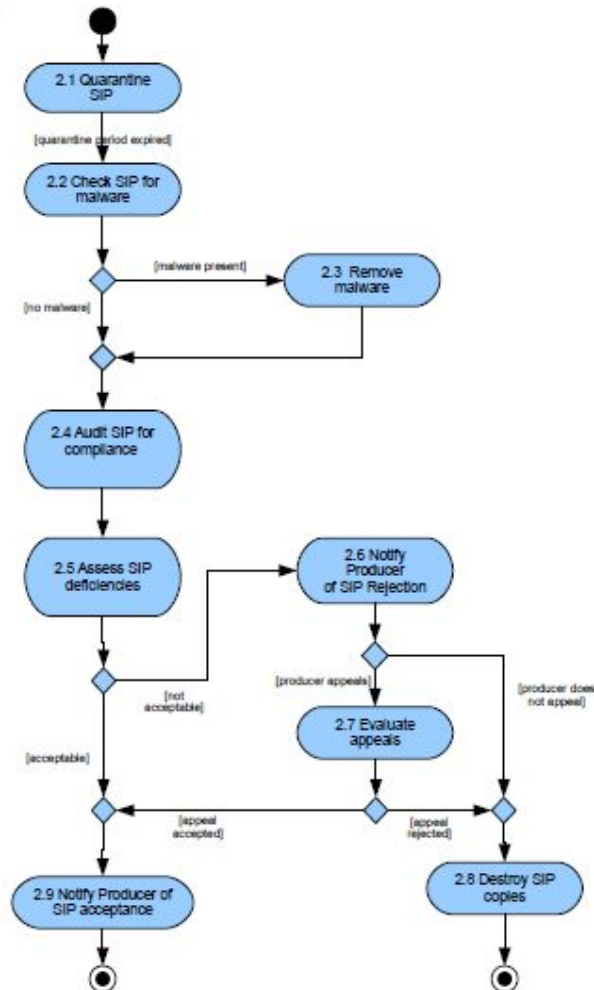
Transfer	UUID	Transfer start time	
test 1 Sofieprisen	1b156366-3638-4ef8-8d88-96eccf0cb049	2016-06-03 14:21	
▶ Micro-service: Approve transfer			
Job: Approve standard transfer		Awaiting decision	Actions ▾

Arbeidsflyt

- Et sentralt begrep i Archivematica
- Grad av automatisering kan forhåndsdefineres (via konfigurasjonsfiler), for ulike use case
- Eller valg kan gjøres underveis, via brukerens interaksjon med GUI-et
- Er hvert steg dokumenteres, logges, samtidig som METS og PREMIS filer bygges opp underveis
- Unike identifikatorer genereres også, noe som muliggjør større søkeflater
- Feilsituasjoner blir synliggjort umiddelbart, via fargeforandringer i GUI-et og feillogger.

Et lite utsnitt av arbeidsflyten som gjennomføres i Archivematica

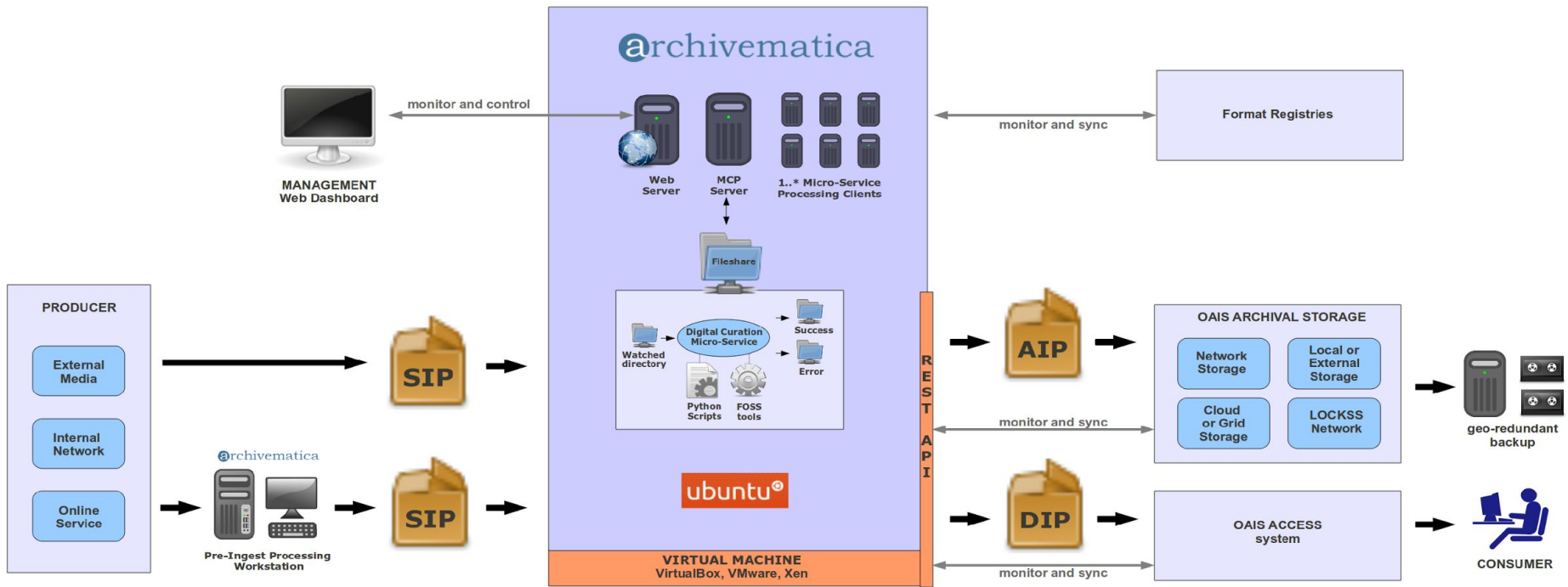
Archivematica– Activity Diagram 2
OAIS: Audit SIP



Litt om...

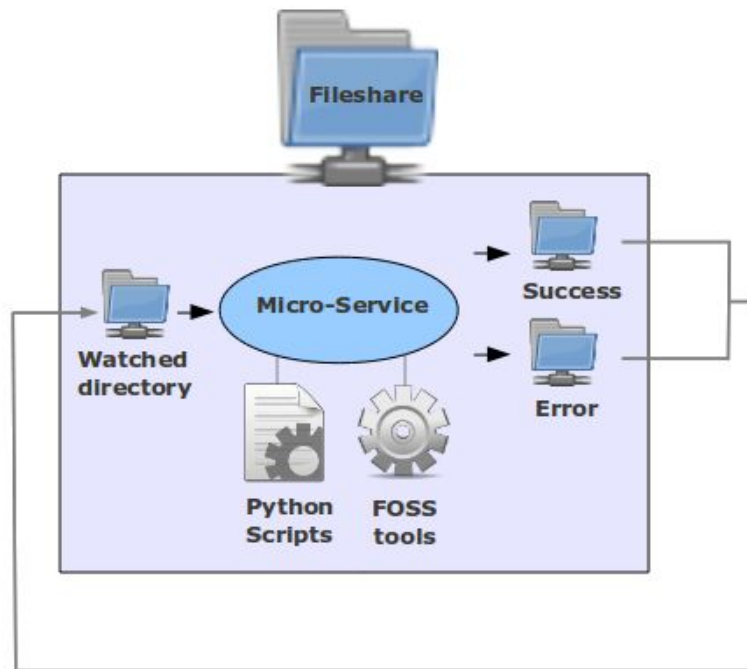
TEKNOLOGI

Det store bildet...



Kilde: <https://www.archivematica.org/wiki/File:Archivematica-0.8-beta-architecture.png>

Mikrotjeneste (MCP) -arkitekturen


























- En mikrotjeneste utfører et steg i OAIS arbeidsflyten
- Mikrotjenester bearbejder informasjonspakker (IP-er), SIP-er/AIP-er/DIP-er
- IP-er inneholder filer, sjekksummer, logger, XML-metadata
- En mikrotjeneste er enten et Archivematica Python skript eller (ekstern) fri programvare

Fri programvare som benyttes

Release 1.4.1

[Main Page](#) > [External software tools](#) > [Release 1.4.1](#)

Archivematica integrates a suite of free and open-source tools that allows users to process digital objects from ingest to access in compliance with the ISO-OAIS functional model. In addition to the core Archivematica which is released under AGPL v3 license, the following tools are bundled with Archiver

Tool	Description	License
BagIt 	Standard and script to package digital objects and metadata for archival storage	BSD License
bulk_extractor 	Disk image and file contents analysis tool	Public domain
[http://www.clamav.net/]  Clam AV (anti-virus) 	Anti-virus toolkit for UNIX	GNU General Public License (GPL)
ElasticSearch 	Indexing and search	Apache License 2.0
ExifTool 	Multimedia metadata extraction	GNU General Public License and Artistic License
FFmpeg 	Converts a wide variety of audio and video formats	GNU Lesser General Public License (LGPL)
File Information Tool Set (FITS) 	File format identification and validation software integration tool	GNU Lesser General Public License (LGPL)
fido 	File format identification tool	Licensed under the Apache License, Version 2.0 (the "License")
Seigfried 	File format identification tool	Licensed under the Apache License, Version 2.0 (the "License")
JHOVE 	File format validation tool	GNU Lesser General Public License (LGPL)
MediaInfo 	Multimedia metadata extraction	BSD (2-clause), Zlib
Tesseract 	Optical Character Recognition -read image files and convert to text	Apache License Version 2
AtOM 	Web-based archival description and access tool.	GNU General Public License (GPL)
Imagemagick 	Converts a wide variety of bitmap images	GPL compatible Imagemagick license 
Inkscape 	Converts vector images to Scalable Vector Graphic (SVG) format	GNU General Public License (GPL) version 2
NFS-common 	Network File System Access - allows access to files on network storage devices.	GNU General Public License (GPL)
Python-lxml 	Python binding for libxml2 and libxslt	GNU General Public License (GPL)
The Sleuthkit 	Disk image management and extraction toolkit	Common Public License / IBM Public License
md5deep 	Checksum generation and verification scripts	GNU General Public License (GPL)
UUID 	command line interface (CLI) for the generation of DCE 1.1, ISO/IEC 11578:1996 and IETF RFC-4122 compliant Universally Unique Identifier (UUID).	GNU General Public License (GPL)
Ubuntu Linux 	Interface with computing hardware. Ubuntu Linux server edition.	GNU General Public License (GPL)

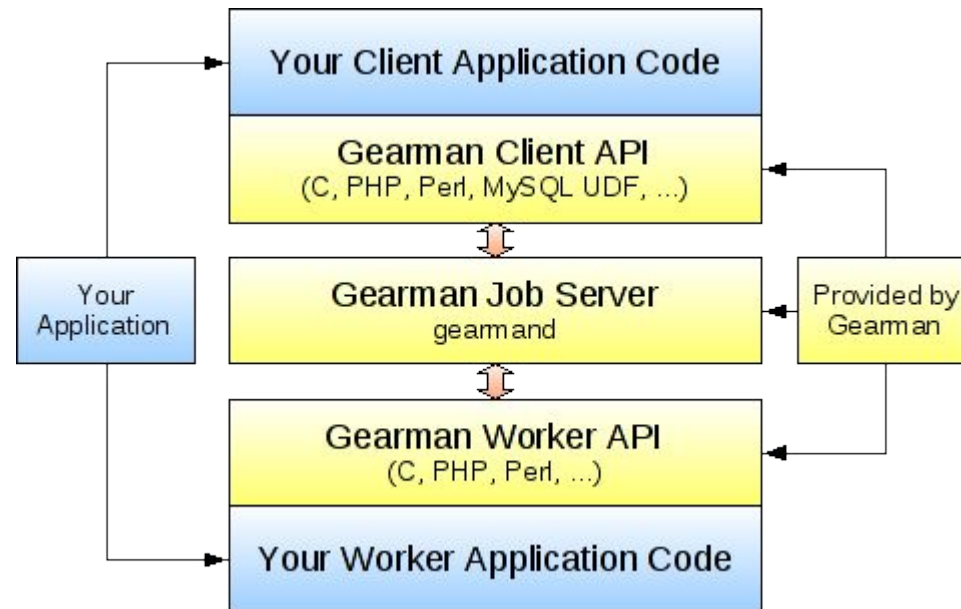
-
-
-
-

Mikrotjenester

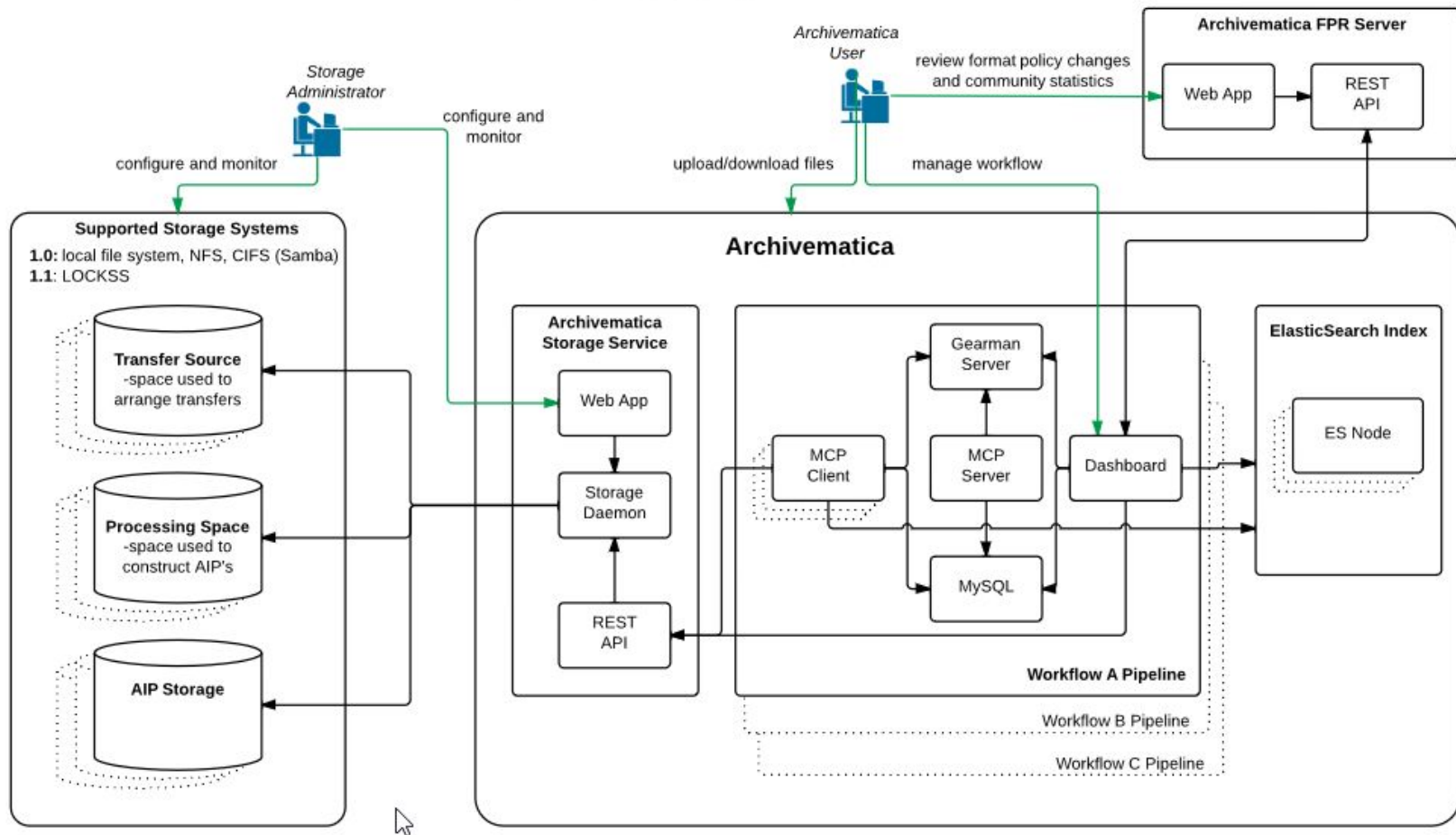
Ingest

Micro-service	Description
Approve SIP creation [Collapse] Approve SIP creation Create removal from backlog PREMIS events	Applies to SIPs that have been created from backlog. Approves creation of SIP, and creates PREMIS event for removal from backlog.
Verify transfer compliance [Expand]	Verifies the METS from the transfer.
Verify SIP compliance [Collapse] Set file permissions Move to processing directory Verify SIP compliance	Verifies that the SIP conforms to the folder structure required for processing in Archivematica. The structure is as follows: <i>/logs/, /metadata/, /metadata /submissionDocumentation/, /objects/.</i>
Rename SIP directory with SIP UUID [Collapse] Rename SIP directory with SIP UUID Check if SIP is from Maildir Transfer	Directly associates the SIP with its metadata by appending the SIP UUID to the SIP directory name and checks if SIP is from Maildir transfer type to determine workflow.
Include default SIP processingMCP.xml [Expand]	Copies the processing configuration file added to the transfer in Include default Transfer processingMCP.xml, above, to the SIP.
Remove cache files [Expand]	Removes any thumbs.db files.
Clean up names [Expand]	Some file systems do not support unicode or other special characters in filenames. This micro-service removes prohibited characters and replaces them with dashes. Original filenames are preserved in the PREMIS metadata.
Normalize [Collapse] Identify manually normalized files Check for Service directory Check for Access directory Set remove preservation and access normalized files to renormalize link. Get normalization options for an existing SIP	<ul style="list-style-type: none"> ▪ ▪ ▪ ▪ ▪

I bunnen av mikrotjeneste-arkitekturen ligger Gearman



Et mer detaljert bilde av en (tilfeldig valgt) Archivemata arkitektur



Kilde: <https://wiki.its.sfu.ca/departments/archives/images/3/31/ArchivemataArchitecture.pdf>

Så til...

ERFARINGER FRA PROSJEKTET VI KJØRER I SAMDOK-PROSJEKTET

Prosjektet (SAMDOK, Privat)

- Oppgave PRIV 10_Metodikk for høsting og langtidslagring av privatarkiv
- **Ellen Røsjø (leder)** (*Riksarkivet*) – ellros@arkivverket.no
- Arne Kristian Groven, RA (*Riksarkivet / fagressurs*) – arngro@arkivverket.no
- Martin Ellingsrud (*Arbark*) – Martin.Ellingsrud@arbark.no
- Pål Mjørland (*Fylkesarkivet i Oppland i samarbeid med OPAM*) – pal.mjorland@oppland.org
- Stig Brunstad (*Bergen byarkiv*) – Stig.Brunstad@bergen.kommune.no
- Torkel Thime (*Statsarkivet i Stavanger/NOGA Norsk olje- og gassarkiv*) – toth@arkivverket.no
- Per Kristian Ottersland (*Riksarkivet*) – peot@arkivverket.no
- Rolf Petter Waage (*Stokke AS / referanseressurs*) – waage@stokke.com

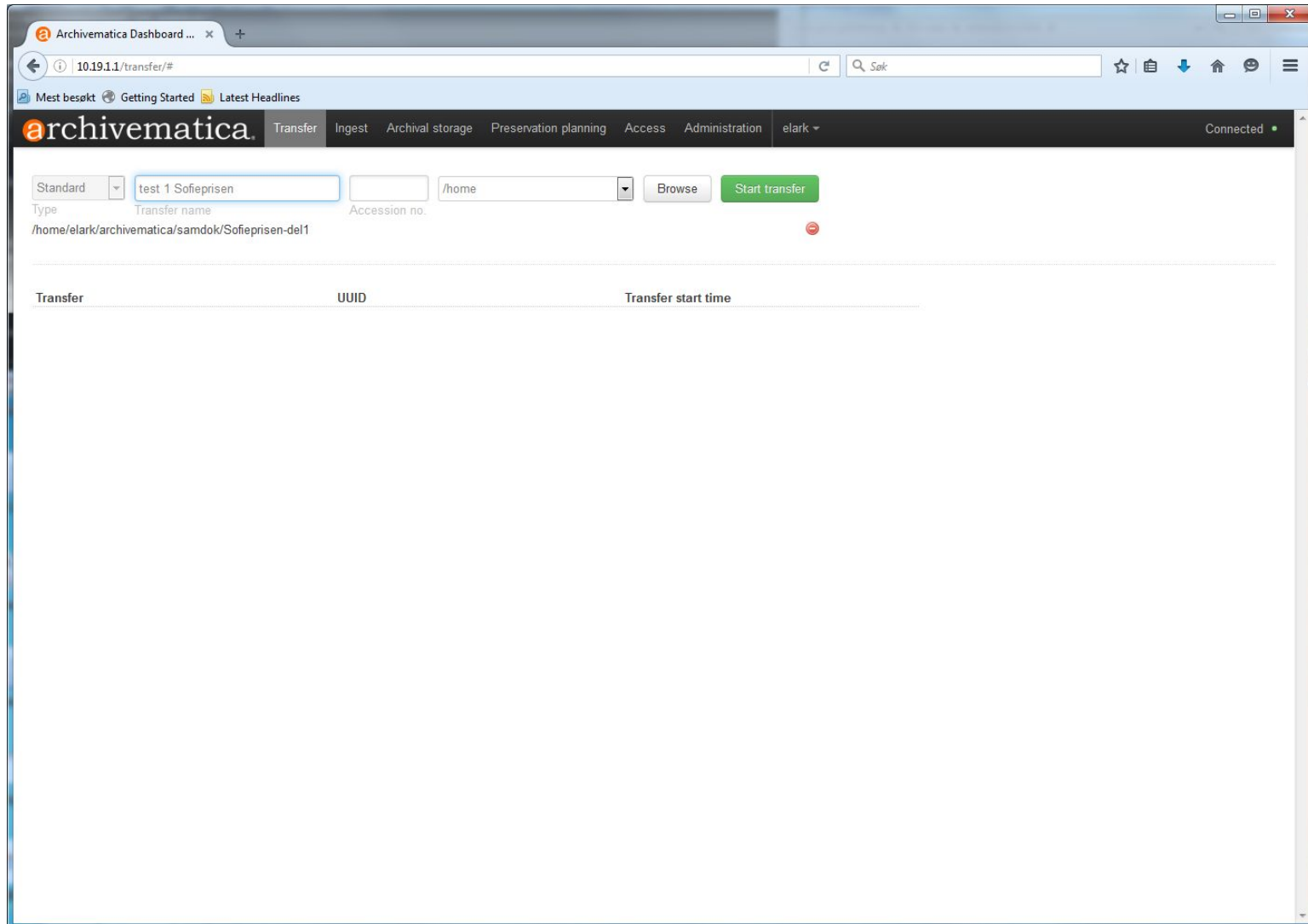
Erfaringer knyttet til bruk av Archivematica

- Verktøyet har lav brukerterskel
- Det er relativt enkelt å bruke, og grad av automatisering kan lett justeres
- Noen eksempler/skjermdump følger...

Digitalt innhold som skal bearbeides hentes fra spesifiserte mapper...

The screenshot shows the Archivematica web interface in a browser window. The address bar displays "10.19.1.1/transfer/". The navigation menu includes "Transfer", "Ingest", "Archival storage", "Preservation planning", "Access", "Administration", and "elark". The main content area features a form with fields for "Type" (set to "Standard"), "Transfer name", and "Accession no.", along with a "Browse" button and a green "Start transfer" button. Below the form is a table with columns for "Transfer", "UUID", and "Transfer start time". A "Select a directory" dialog box is open, showing a tree view of directories: "home", "elark (230 objects)", "archivematica (224 objects)", "samdok (224 objects)", "Sofieprisen-del1 (69 objects)", "Sofieprisen-del2 (110 objects)", and "Sofieprisen-del3 (45 objects)". Each directory entry has an "Add" button next to it. A "Cancel" button is located at the bottom right of the dialog.

Materialet som skal overføres gis et navn



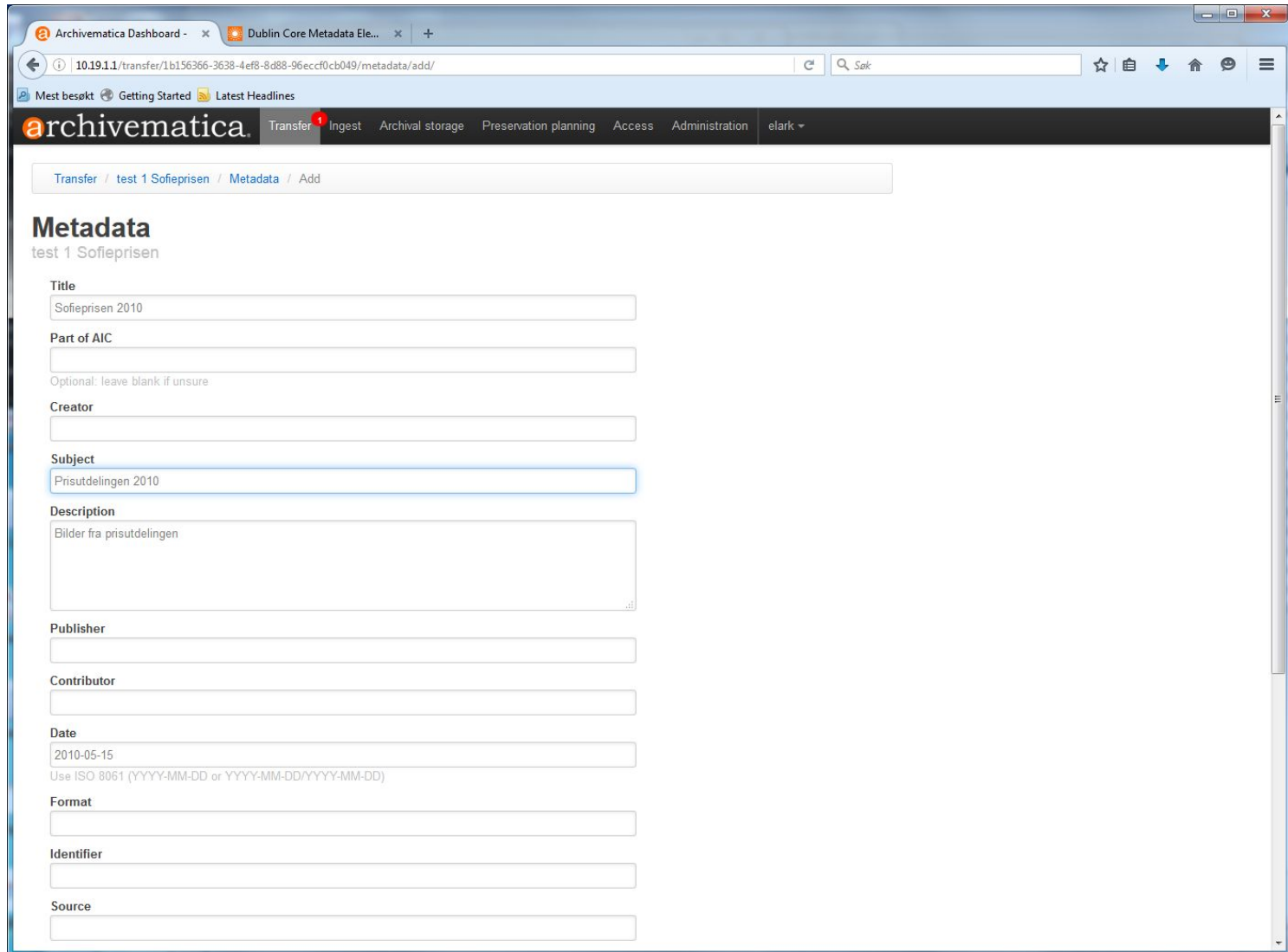
The screenshot shows the Archivematica web interface in a browser window. The address bar displays "10.19.1.1/transfer/#". The navigation menu includes "Transfer", "Ingest", "Archival storage", "Preservation planning", "Access", "Administration", and "elark". The "Transfer" section is active, showing a form with the following fields:

- Type: Standard
- Transfer name: test 1 Sofieprisen
- Accession no.: (empty)
- Path: /home

Buttons for "Browse" and "Start transfer" are visible. Below the form, the path "/home/elark/archivematica/samdok/Sofieprisen-del1" is shown. A table below the path has the following headers:

Transfer	UUID	Transfer start time
----------	------	---------------------

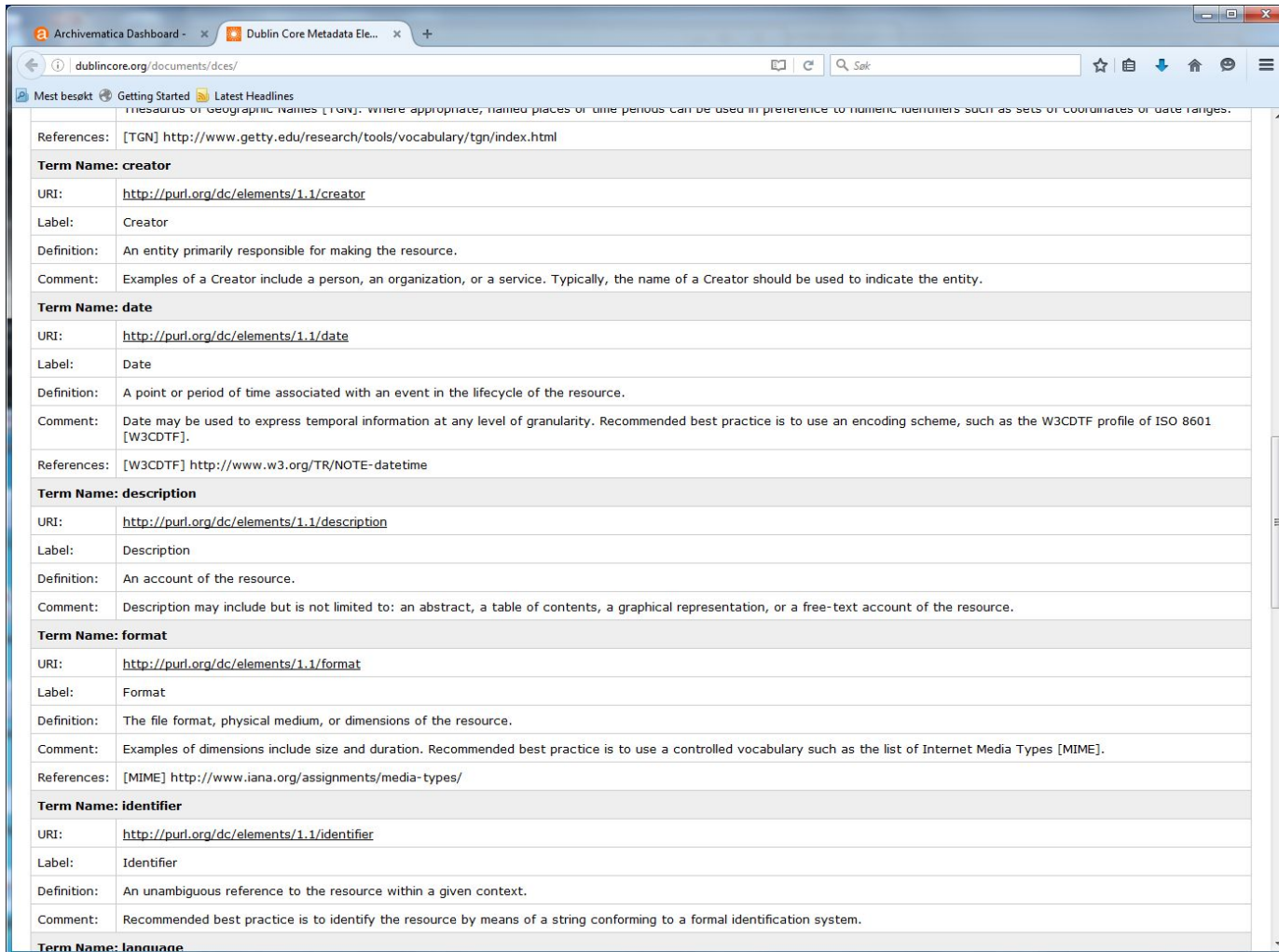
Dublin Core (DC) metadata fylles ut



The screenshot shows a web browser window with the Archivematica dashboard. The address bar shows the URL: `10.19.1.1/transfer/1b156366-3638-4ef8-8d88-96eccf0cb049/metadata/add/`. The browser tabs are 'Archivematica Dashboard' and 'Dublin Core Metadata Ele...'. The browser's search bar contains the text 'Søk'. The dashboard header includes the Archivematica logo and navigation links: 'Transfer', 'Ingest', 'Archival storage', 'Preservation planning', 'Access', 'Administration', and 'elark'. A red notification badge is visible on the 'Transfer' link. Below the header, a breadcrumb trail reads 'Transfer / test 1 Sofieprisen / Metadata / Add'. The main content area is titled 'Metadata' and 'test 1 Sofieprisen'. The form contains the following fields:

- Title:** Sofieprisen 2010
- Part of AIC:** (empty)
- Optional:** leave blank if unsure
- Creator:** (empty)
- Subject:** Prisutdelingen 2010
- Description:** Bilder fra prisutdelingen
- Publisher:** (empty)
- Contributor:** (empty)
- Date:** 2010-05-15
Use ISO 8061 (YYYY-MM-DD or YYYY-MM-DD/YYYY-MM-DD)
- Format:** (empty)
- Identifier:** (empty)
- Source:** (empty)

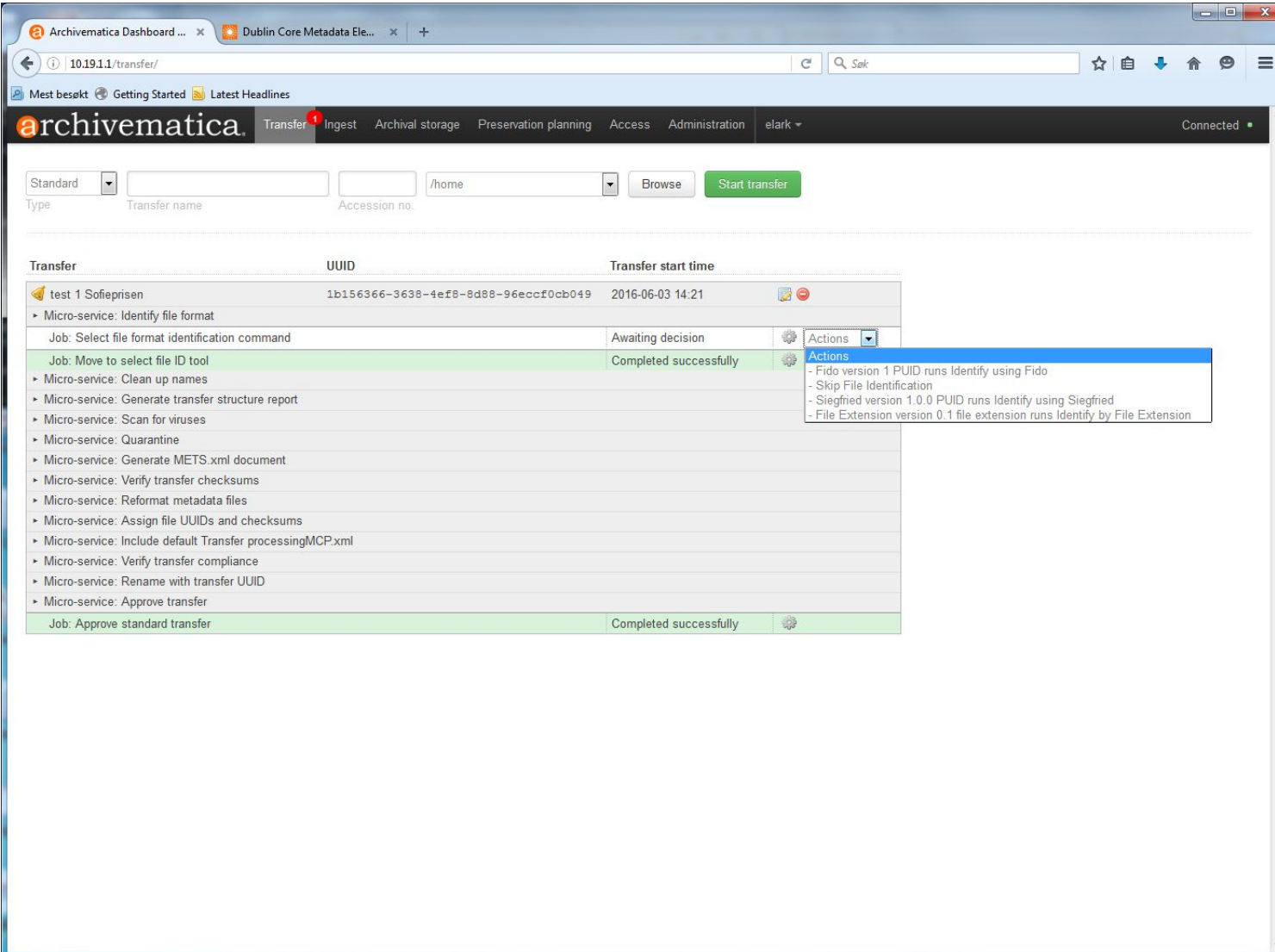
Dokumentasjon av DC er lett tilgjengelig



The screenshot shows a web browser window displaying the Dublin Core Metadata Elements (DC) documentation page. The browser's address bar shows the URL dublincore.org/documents/dces/. The page content is organized into sections for each DC element, including:

- References:** [TGN] <http://www.getty.edu/research/tools/vocabulary/tgn/index.html>
- Term Name: creator**
 - URI: <http://purl.org/dc/elements/1.1/creator>
 - Label: Creator
 - Definition: An entity primarily responsible for making the resource.
 - Comment: Examples of a Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.
- Term Name: date**
 - URI: <http://purl.org/dc/elements/1.1/date>
 - Label: Date
 - Definition: A point or period of time associated with an event in the lifecycle of the resource.
 - Comment: Date may be used to express temporal information at any level of granularity. Recommended best practice is to use an encoding scheme, such as the W3CDTF profile of ISO 8601 [W3CDTF].
 - References: [W3CDTF] <http://www.w3.org/TR/NOTE-datetime>
- Term Name: description**
 - URI: <http://purl.org/dc/elements/1.1/description>
 - Label: Description
 - Definition: An account of the resource.
 - Comment: Description may include but is not limited to: an abstract, a table of contents, a graphical representation, or a free-text account of the resource.
- Term Name: format**
 - URI: <http://purl.org/dc/elements/1.1/format>
 - Label: Format
 - Definition: The file format, physical medium, or dimensions of the resource.
 - Comment: Examples of dimensions include size and duration. Recommended best practice is to use a controlled vocabulary such as the list of Internet Media Types [MIME].
 - References: [MIME] <http://www.iana.org/assignments/media-types/>
- Term Name: identifier**
 - URI: <http://purl.org/dc/elements/1.1/identifier>
 - Label: Identifier
 - Definition: An unambiguous reference to the resource within a given context.
 - Comment: Recommended best practice is to identify the resource by means of a string conforming to a formal identification system.
- Term Name: language**

Så utføres en lang rekke MCP-er og beslutningspunkter krever interaksjon



The screenshot shows the Archivematica web interface. The browser address bar displays '10.19.1.1/transfer/'. The navigation menu includes 'Transfer', 'Ingest', 'Archival storage', 'Preservation planning', 'Access', 'Administration', and 'elark'. The main content area shows a transfer form with fields for 'Type' (Standard), 'Transfer name', and 'Accession no.' (with a '/home' dropdown). A 'Start transfer' button is visible.

Below the form is a table of transfer jobs. The table has columns for 'Transfer', 'UUID', and 'Transfer start time'. The first job is 'test 1 Sofieprisen' with UUID '1b156366-3638-4ef8-8d88-96eccf0cb049' and start time '2016-06-03 14:21'. The job details are expanded, showing a list of micro-services and their status. The 'Job: Move to select file ID tool' is highlighted in green and marked as 'Completed successfully'. An 'Actions' dropdown menu is open, showing options: 'Fido version 1 PUID runs Identify using Fido', 'Skip File Identification', 'Siegfried version 1.0.0 PUID runs Identify using Siegfried', and 'File Extension version 0.1 file extension runs Identify by File Extension'.

Transfer	UUID	Transfer start time
test 1 Sofieprisen	1b156366-3638-4ef8-8d88-96eccf0cb049	2016-06-03 14:21
Micro-service: Identify file format		
Job: Select file format identification command		Awaiting decision
Job: Move to select file ID tool		Completed successfully
Micro-service: Clean up names		
Micro-service: Generate transfer structure report		
Micro-service: Scan for viruses		
Micro-service: Quarantine		
Micro-service: Generate METS.xml document		
Micro-service: Verify transfer checksums		
Micro-service: Reformat metadata files		
Micro-service: Assign file UUIDs and checksums		
Micro-service: Include default Transfer processingMCP.xml		
Micro-service: Verify transfer compliance		
Micro-service: Rename with transfer UUID		
Micro-service: Approve transfer		
Job: Approve standard transfer		Completed successfully

SIP-er genereres..

The screenshot shows the Archivematica web interface. The browser address bar displays '10.19.1.1/transfer/'. The navigation menu includes 'Transfer', 'Ingest', 'Archival storage', 'Preservation planning', 'Access', and 'Administration'. The user is logged in as 'elark' and is connected to the system.

The main interface features a form for starting a transfer with fields for 'Type' (set to 'Standard'), 'Transfer name', 'Accession no.', and a directory path ('/home'). A 'Start transfer' button is visible.

The transfer process is detailed in a table below:

Transfer	UUID	Transfer start time	
test 1 Sofieprisen	1b156366-3638-4ef8-8d88-96eccf0cb049	2016-06-03 14:21	
▶ Micro-service: Create SIP from Transfer			
Job: Check transfer directory for objects		Completed successfully	
Job: Move to SIP creation directory for completed transfers		Completed successfully	
Job: Create SIP from transfer objects		Completed successfully	
Job: Serialize Dublin Core metadata to disk		Completed successfully	
Job: Move to processing directory		Completed successfully	
Job: Create SIP(s) [?]		Completed successfully	
Job: Load options to create SIPs		Completed successfully	
Job: Check transfer directory for objects		Completed successfully	
▶ Micro-service: Complete transfer			
▶ Micro-service: Examine contents			
▶ Micro-service: Validation			
▶ Micro-service: Characterize and extract metadata			
▶ Micro-service: Update METS.xml document			
▶ Micro-service: Extract packages			
▶ Micro-service: Identify file format			
Job: Identify file format		Completed successfully	
Job: Determine which files to identify		Completed successfully	
Job: Select file format identification command		Completed successfully	
Job: Move to select file ID tool		Completed successfully	
▶ Micro-service: Clean up names			
▶ Micro-service: Generate transfer structure report			
▶ Micro-service: Scan for viruses			
▶ Micro-service: Quarantine			
▶ Micro-service: Generate METS.xml document			
▶ Micro-service: Verify transfer checksums			
▶ Micro-service: Reformat metadata files			
▶ Micro-service: Assign file UUIDs and checksums			
▶ Micro-service: Include default Transfer processingMCP.xml			
▶ Micro-service: Verify transfer compliance			
▶ Micro-service: Rename with transfer UUID			
▶ Micro-service: Approve transfer			

AIP/DIP-er genereres, automatisert konvertering til lagrings-/visningsformater

The screenshot displays the Archivematica dashboard interface. The browser address bar shows '10.19.1.1/ingest/'. The dashboard header includes the Archivematica logo and navigation tabs: Transfer, Ingest, Archival storage, Preservation planning, Access, Administration, and elark. The main content area is a table of Submission Information Packages (SIPs) with columns for Submission Information Package, UUID, and Ingest start time. Each SIP entry is expanded to show a list of jobs and their status. A dropdown menu is open for the first SIP, showing actions like 'Normalize for preservation and access'.

Submission Information Package	UUID	Ingest start time
test_3_Sofieprisen	01f32c1e-3f8b-4faa-9f79-d4df1fbc8888	2016-06-03 15:01
Micro-service: Normalize		
Job: Normalize [?]		Awaiting decision
Job: Resume after normalization file identification tool selected.		Completed successfully
Job: Identify file format		Completed successfully
Job: Select pre-normalize file format identification command		Completed successfully
Job: Move to select file ID tool		Completed successfully
Job: Grant normalization options for no pre-existing DIP		Completed successfully
Job: Set remove preservation and access normalized files to renormalize link.		Completed successfully
Job: Check for Access directory		Completed successfully
Job: Identify manually normalized files		Completed successfully
Job: Check for Service directory		Completed successfully
Micro-service: Clean up names		
Micro-service: Remove cache files		
Micro-service: Include default SIP processingMCP.xml		
Micro-service: Rename SIP directory with SIP UUID		
Micro-service: Verify transfer compliance		
Micro-service: Verify SIP compliance		
test2-Sofieprisen	b84c475f-51ce-405c-a86a-e856f99d4fee	2016-06-03 14:52
Micro-service: Normalize		
Micro-service: Clean up names		
Micro-service: Remove cache files		
Micro-service: Include default SIP processingMCP.xml		
Micro-service: Rename SIP directory with SIP UUID		
Micro-service: Verify transfer compliance		
Micro-service: Verify SIP compliance		
test_1_Sofieprisen	4b8121c2-2e32-4323-a029-452001ba1d02	2016-06-03 14:41
Micro-service: Store AIP		
Job: Store AIP (review) [?]		Awaiting decision
Job: Move to the store AIP approval directory		Completed successfully
Micro-service: Prepare AIP		
Micro-service: Upload DIP		
Job: Upload DIP		Awaiting decision
Micro-service: Prepare DIP		
Micro-service: Generate AIP METS		
Micro-service: Verify checksums		

Review av resultatene før endelig godkjenning...

Archivematica Dashboard ... x Archivematica Dashboard ... x Dublin Core Metadata Ele... x Dublin Core Metadata Ele... x +

10.19.1.1/ingest/preview/normalization/6dd503c7-28cc-41c4-96fd-e87b3a8c8bdf/

Mest besøkt Getting Started Latest Headlines

archivematica Transfer Ingest ³ Archival storage Preservation planning Access Administration elark ▾

Review normalization

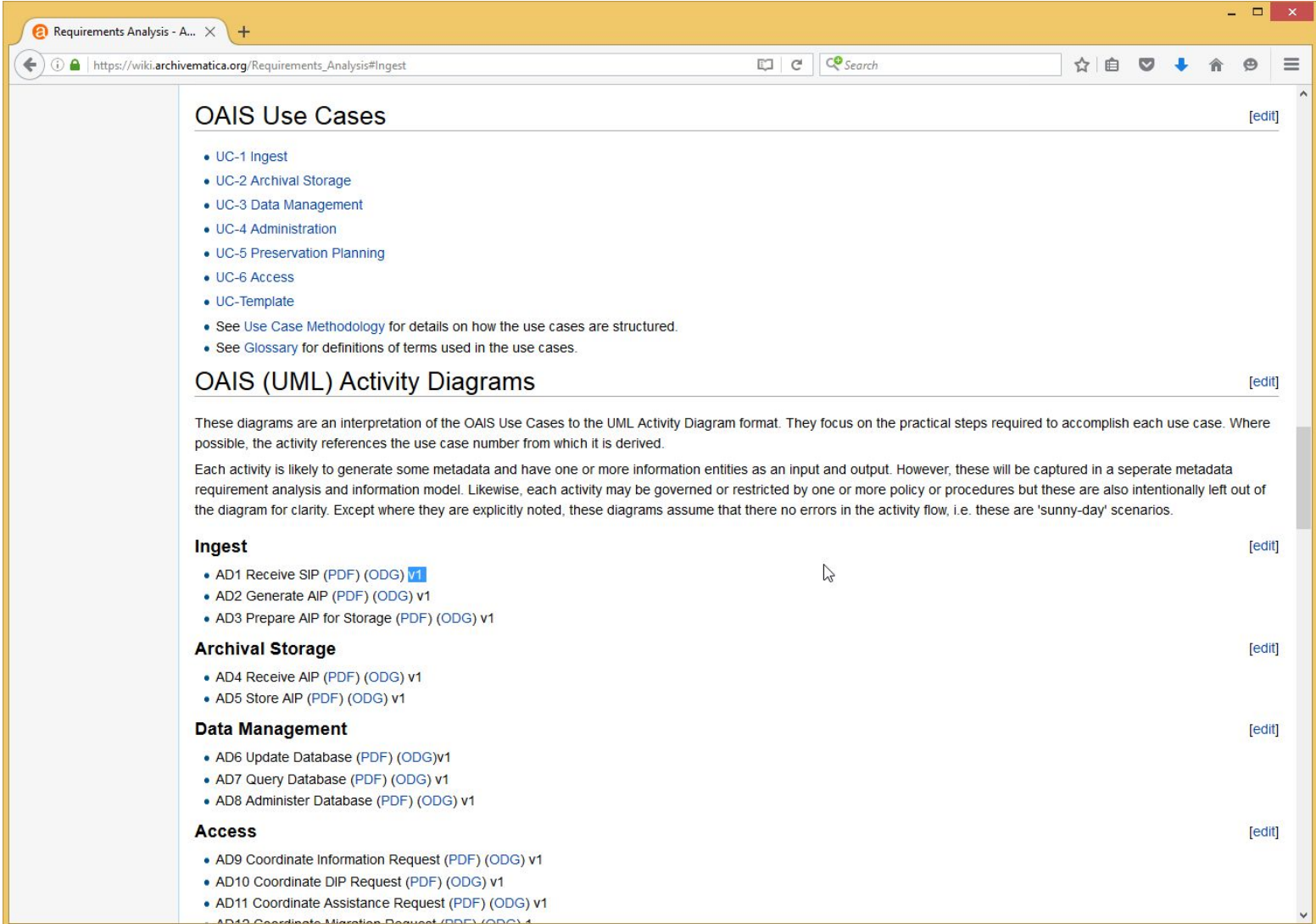
Ingest / test_1_Sofieprisen

- approveNormalization
 - test_1_Sofieprisen-4b8121c2-2e32-4323-a029-452001ba1d02
 - DIP
 - logs
 - fileFormatIdentification.log
 - fileMeta
 - metadata
 - dc.json
 - submissionDocumentation
 - objects
 - 2010
 - Bilder
 - DSC_0005-3823069f-0fc1-4c53-b041-e55ed6876ecd.tif
 - DSC_0005.jpg
 - DSC_0022-46080114-5071-4f67-b7f5-dbbfcf3e3e8d.tif
 - DSC_0022.jpg
 - DSC_0032-2a347cd3-18dd-4cc7-b2f0-b0fbc5861c14.tif

Erfaringer knyttet til bruk av Archivematica

- Det meste kan skje uten særlig bruker-kompetanse
- Men ettersom så mye skjer under panseret så bør brukeren sette seg inn i hva som skjer, for å kunne utnytte verktøyet optimalt
- Dette er veldokumentert
- I tillegg bør brukeren ha forkunnskap om DC, OAIS, METS og PREMIS

Dokumentasjon om arbeidflyt (og mye annet) kan finnes på web-/wiki-sidene



Requirements Analysis - A... X +

https://wiki.archivematica.org/Requirements_Analysis#Ingest

OAIS Use Cases [edit]

- UC-1 Ingest
- UC-2 Archival Storage
- UC-3 Data Management
- UC-4 Administration
- UC-5 Preservation Planning
- UC-6 Access
- UC-Template
- See [Use Case Methodology](#) for details on how the use cases are structured.
- See [Glossary](#) for definitions of terms used in the use cases.

OAIS (UML) Activity Diagrams [edit]

These diagrams are an interpretation of the OAIS Use Cases to the UML Activity Diagram format. They focus on the practical steps required to accomplish each use case. Where possible, the activity references the use case number from which it is derived.

Each activity is likely to generate some metadata and have one or more information entities as an input and output. However, these will be captured in a separate metadata requirement analysis and information model. Likewise, each activity may be governed or restricted by one or more policy or procedures but these are also intentionally left out of the diagram for clarity. Except where they are explicitly noted, these diagrams assume that there no errors in the activity flow, i.e. these are 'sunny-day' scenarios.

Ingest [edit]

- AD1 Receive SIP (PDF) (ODG) [v1](#)
- AD2 Generate AIP (PDF) (ODG) v1
- AD3 Prepare AIP for Storage (PDF) (ODG) v1

Archival Storage [edit]

- AD4 Receive AIP (PDF) (ODG) v1
- AD5 Store AIP (PDF) (ODG) v1

Data Management [edit]

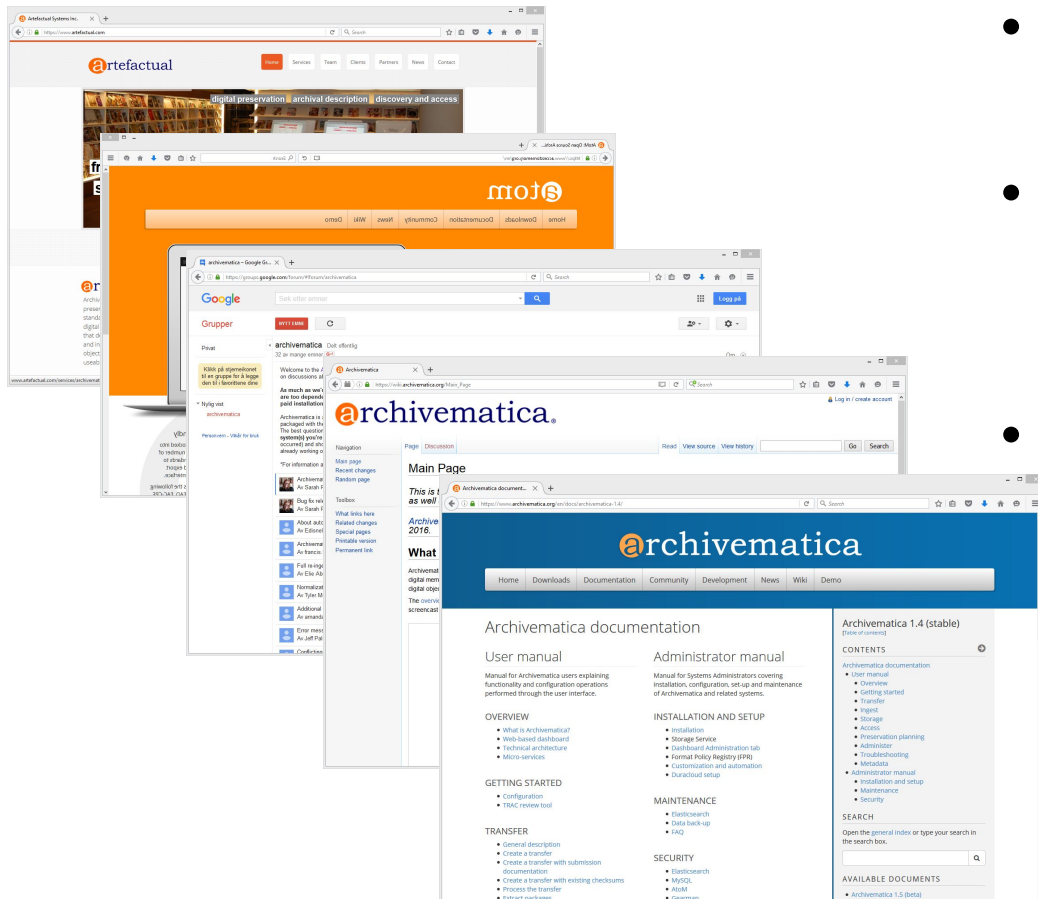
- AD6 Update Database (PDF) (ODG)v1
- AD7 Query Database (PDF) (ODG) v1
- AD8 Administer Database (PDF) (ODG) v1

Access [edit]

- AD9 Coordinate Information Request (PDF) (ODG) v1
- AD10 Coordinate DIP Request (PDF) (ODG) v1
- AD11 Coordinate Assistance Request (PDF) (ODG) v1
- AD12 Coordinate Migration Request (PDF) (ODG) v1

Generelt veldokumentert

- <https://www.artefactual.com/>
- <https://www.accesstomemory.org>
- <https://groups.google.com/forum/#!forum/archivematica>
- <https://wiki.archivematica.org>
- <https://www.archivematica.org>



Tillit til programvare

- Negativ risiko forbundet med (innkjøp og bruk av) programvare, fri eller lukket, er stor dersom:
 - Få har kompetanse til å modifisere programvaren
 - Få bruker programvaren
- Dessuten, dersom dataene må gjennom mange transformasjoner:
 - Så vil lukkede løsninger forhindre kvalitetssikring, og øke negativ risiko
 - Fri og åpne løsninger vil redusere denne risikoen og øke etterprøvbarhet
- Dette er hovedgrunnen til at mange velger fri programvare/åpen kildekode for langtidsbevaring og aksess

Produktleverandøren har publisert følgende liste av kunder

International

- International Council on Archives
- UNESCO Memory of the World
- UNESCO Archives
- United Nations Archives and Records Management Section
- The World Bank Group Library and Archives of Development
- International Monetary Fund Archives
- NATO Archives
- International Records Management Trust
- Dutch Ministry of the Interior and Kingdom Relations
- Dutch Institute for Archival Research and Education
- United Arab Emirates Center for Documentation and Research
- Institut de Cultura de Barcelona
- Al-Dhakira Al-Arabiyya
- Association of Brazilian Archivists
- Botswana National Archives and Records Service
- Caribbean Regional Branch of the International Council on Archives
- University of the Witwatersrand Archives
- Glasgow Caledonian University Archives
- University of Strathclyde Archives
- Stirling Council Archives
- National Library of Wales
- Zuse Institute Berlin

Canada

- Canadian Council of Archives
- Canadiana
- Library and Archives Canada
- City of Vancouver Archives
- City of Montreal Archives
- City of Edmonton Archives
- Archives Association of British Columbia
- British Columbia Museum and Archives
- British Columbia Ministry of Management Services
- Archives Association of Ontario
- Association for Manitoba Archives
- Archives Society of Alberta
- Provincial Archives of Alberta
- Alberta Government Services Ministry
- Association of Newfoundland and Labrador Archives
- Council of Archives New Brunswick
- University of British Columbia Library
- Simon Fraser University Archives
- Simon Fraser University Library
- University of Alberta Library
- University of Saskatchewan Archives
- University of Victoria Archives
- University of Toronto iSchool Institute
- University of Northern British Columbia Archives
- Irving K. Barber Learning Centre
- Trinity Western University Archives
- Anglican Church of Canada Archives
- Catholic Diocese of Calgary Archives
- Sisters of the Child Jesus
- West Vancouver Memorial Library
- City of Toronto Information Management Services
- City of Sudbury Archives
- Squamish Public Library
- Whistler Museum and Archives
- St. Albert Heritage Museum
- Council of Prairie and Pacific University Libraries
- Council of Ontario University Libraries

U.S.A.

- Yale University Library
- Museum of Modern Art (MoMA)
- Harvard Business School Library
- Columbia University Library
- MIT Libraries
- Rockefeller Archive Center
- American Institute of Architects
- Pasadena Art Center College of Design
- Dumbarton Oaks Research Library
- Bentley Historical Library (University of Michigan)
- University of Houston Library

Prosjektets kontakt med leverandør (Artifactual Systems)

- Vi har (generelt) i prosjektet ikke hatt mye kontakt med leverandøren
- Med et unntak, ARBARK- Arbeiderbevegelsens arkiv og bibliotek
- ARBARK leide tjenester i skyen av Artifactual Systems, fra servere som sto i Tyskland
- Erfaringene med dette var stort sett positive, så vidt jeg forstår.
- Ellers har vi diskutert norsk skyproblematikk med leverandøren
- Og de svarte uoppfordret at de sikkert kunne tilby skytjenester fra servere i Norge, dersom det var nødvendig

Vi som ikke leide tjenester i skyen...

- Stolte på vår egen evne til å installere og drifte Archivemata og Atom... (det var jo også et formål med prosjektet, å teste ut teknologi)
- Vi fant ut etter hvert at vi behøvde kursing i Linux server-administrasjon, med fokus på installasjon av programvare, noe vi fikk hos Redpill-Linpro
- Da kunne installasjon av Archivemata og AtoM gjennomføres i løpet av en dag eller to

Noen ord om installasjonen

INSTALLING ARCHIVEMATICA 1.4 PACKAGES (NEW INSTALL)

Archivematica packages are hosted on Launchpad, in an Ubuntu PPA (Personal Package Archive). In order to install software onto your Ubuntu system from a PPA:

1. Add the archivematica/release PPA to your list of trusted repositories (if add-apt-repositories is not available you must install python-software-properties first):

```
sudo apt-get update
sudo apt-get install python-software-properties
sudo add-apt-repository ppa:archivematica/1.4
```

2. Add the ElasticSearch apt repository next:

```
sudo wget -O - http://packages.elasticsearch.org/GPG-KEY-elasticsearch | sudo apt-key add -
sudo sh -c 'echo "deb http://packages.elasticsearch.org/elasticsearch/1.4/debian stable main"
< >
```

3. Update your system to the most recent 12.04 release (12.04.5 at the time of this writing).

This step will also fetch a list of the software from the PPAs you just added to your system.

```
sudo apt-get update
sudo apt-get upgrade
```

4. Install the storage service package

```
sudo apt-get install archivematica-storage-service
```

5. Configure the storage service

```
sudo rm -f /etc/nginx/sites-enabled/default
sudo ln -s /etc/nginx/sites-available/storage /etc/nginx/sites-enabled/storage
sudo ln -s /etc/uwsgi/apps-available/storage.ini /etc/uwsgi/apps-enabled/storage.ini
sudo service uwsgi restart
sudo service nginx restart
```

6. Install the Archivematica packages (each of these packages can be installed separately, if necessary). Say YES or OK to any prompts you get after entering the following into terminal:

```
sudo apt-get install archivematica-mcp-server
sudo apt-get install archivematica-mcp-client
sudo apt-get install archivematica-dashboard
```

- Det er veldokumenterte oppskrifter som beskriver Archivematica server-installasjon, og AtoM server-installasjon
- Vi slet litt med å konfigurere koblingen mellom de. Her var dokumentasjonen litt mangelfull

Noen avsluttende ord...

- Jeg skulle gjerne sagt noe om AtoM (Access to Memory), aksess og «content management» modulen som er «bundlet» med Archivematica, men det får bli en annen gang...
- Samdok-prosjektet vårt skal etter planen levere en avsluttende rapport innen (ca.) 1. september.
- Riksarkivet planlegger også å evaluere Archivematica opp mot et annet tilsvarende verktøy som heter Preservica. Dette vil skje i et prosjekt kalt DIAS 2.

Noen avsluttende ord

- Archivematica har en fleksibilitet som muliggjør mange ulike konfigurasjoner og oppsett langs verdikjeden fra arkivskaper til depot
- Archivematica effektiviserer bearbeiding av digitalt arkivmateriale, slik at en kan regne minutter og timer, framfor dager og uker
- Og tilliten til prosessene ligger dokumentert i systemet på standardiserte formater

arngro@arkivverket.no

TAKK FOR MEG!