

## ADD Mellanarkiv Journal II (1.4)

Projekt: Mellanarkiv Journal II  
Ansvarig arkitekt: Magnus Lönnroth  
Version: 1.4  
Datum: 2024-11-08

## Versionshistorik

Version	Datum	Beskrivning	Utfärdat av
0.1	2024-02-12	Ny ADD skapad från ADD MAJ I	Magnus Lönnroth
0.7	2024-03-22	Färdigställd inför intern granskning	Magnus Lönnroth
1.0	2024-04-05	Färdigställd inför AL granskning	Magnus Lönnroth
1.1	2024-04-11	Rättade fel i tabell över bilagor + hänvisningar	Magnus Lönnroth
1.2	2024-05-10	Återkoppling från AL + revision inför nästa PI	Magnus Lönnroth
1.3	2024-10-09	Omarbetade kap. 4 och 5. Informationsflöden och integrationsarkitektur kraftigt reviderade. La till första utkast till teknisk arkitektur (endast översikt).	Magnus Lönnroth
1.4	2024-11-08	Omarbetade kap. 6 och 7	Magnus Lönnroth

## Referenser

Detta dokument ska ses som ett nytt inkrement, där referenser som använts i tidigare inkrement bibehåller sitt nuvarande referensnummer och där gamla referensnummer inte återanvänds för nya referenser.

Nr	Källa
1	EA-principer - Version 1.0, VGR IT, 2019-06-24 <a href="https://mellanarkiv-offentlig.vgregion.se/alfresco/s/archive/stream/public/v1/source/available/sofia/rs4336-314204600-317/surrogate/EA%20principer%20v%201.0.pdf">https://mellanarkiv-offentlig.vgregion.se/alfresco/s/archive/stream/public/v1/source/available/sofia/rs4336-314204600-317/surrogate/EA%20principer%20v%201.0.pdf</a>
16	ADD - ARK Strategisk Målarkitektur för Mellanarkivering <a href="https://iserver.vgregion.se/#!/document/viewer/bc64ff99-986e-4647-8848-170ff597dcbf">https://iserver.vgregion.se/#!/document/viewer/bc64ff99-986e-4647-8848-170ff597dcbf</a>
19	Informationsmodell R7 e-arkiv (1.1) (II) <a href="https://iserver.vgregion.se/#!/document/viewer/e0430435-2ff2-436e-9a0b-b62a9d683d71">https://iserver.vgregion.se/#!/document/viewer/e0430435-2ff2-436e-9a0b-b62a9d683d71</a>
20	Plan för digital informationshantering (2.0) (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=a24e91fb-9f0c-4fc9-9d90-f0177b733ef9">https://iserver.vgregion.se/#!/object/details?objectId=a24e91fb-9f0c-4fc9-9d90-f0177b733ef9</a>
21	Preserving Databases - from digital volatility to digital permanence (dec 2003) (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=85ca0125-e8f0-4893-92e1-c38f3fd21f31">https://iserver.vgregion.se/#!/object/details?objectId=85ca0125-e8f0-4893-92e1-c38f3fd21f31</a>
22	E-ARK publiceringar <a href="https://e-ark-project.com/resources/project-deliverables.html">https://e-ark-project.com/resources/project-deliverables.html</a>
23	CCSDS 650.0-M-2 - Reference model for OAIS (jun 2012) (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=f8098f23-474a-41bf-a0ab-7d0e3f812178">https://iserver.vgregion.se/#!/object/details?objectId=f8098f23-474a-41bf-a0ab-7d0e3f812178</a>
24	Kvalificering av arkitekturuppdraget fas 2 (II) <a href="https://iserver.vgregion.se/#!/document/viewer/377643ae-5b44-488d-9bcd-7898b1d3f558">https://iserver.vgregion.se/#!/document/viewer/377643ae-5b44-488d-9bcd-7898b1d3f558</a>

25	Referensinformationsmodell (RIM) för informationshantering (inom hälso- och sjukvård) <a href="https://vgrinformatik.se/vgr-rim/">https://vgrinformatik.se/vgr-rim/</a>
26	Försäkringskassans beslutsunderlag för vägval framtida e-arkiv (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=d0e2db04-37d8-4057-a710-45eb52b255bc">https://iserver.vgregion.se/#!/object/details?objectId=d0e2db04-37d8-4057-a710-45eb52b255bc</a>
27	Slutrapport - Utredning av Skånes digitala arkiv (SDA) (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=022d52b2-db7f-443b-8f40-d1d05f529ac0">https://iserver.vgregion.se/#!/object/details?objectId=022d52b2-db7f-443b-8f40-d1d05f529ac0</a>
28	Tullverket E-arkiv Projektanalysrapport (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=51523168-0303-4603-9994-75784f2c5dfc">https://iserver.vgregion.se/#!/object/details?objectId=51523168-0303-4603-9994-75784f2c5dfc</a>
29	Strategi datamigrering Millennium version 1.5 (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=7dcf6f46-9f2d-4d28-a6f8-8c46913974a7">https://iserver.vgregion.se/#!/object/details?objectId=7dcf6f46-9f2d-4d28-a6f8-8c46913974a7</a>
30	CITS eHealth1 Patient Medical Records 1.0.0 (II) <a href="https://dilcis.eu/content-types/cits-ehealth1-patient-medical-records">https://dilcis.eu/content-types/cits-ehealth1-patient-medical-records</a>
31	E-ARK Common Specification for Information Packages (CSIP) 2.1.0 (II) <a href="https://dilcis.eu/specifications/common-specification">https://dilcis.eu/specifications/common-specification</a>
32	E-ARK Submission Information Package (SIP) 2.1.0 (II) <a href="https://dilcis.eu/specifications/sip">https://dilcis.eu/specifications/sip</a>
33	E-ARK Archival Information Package (AIP) 2.1.0 (II) <a href="https://dilcis.eu/specifications/aip">https://dilcis.eu/specifications/aip</a>
34	E-ARK Dissemination Information Package (DIP) 3.2.1 (II) <a href="https://dilcis.eu/specifications/dip">https://dilcis.eu/specifications/dip</a>
35	Metadata Encoding & Transmission Standard (METS) 1.12.1 (II) <a href="https://www.loc.gov/standards/mets/">https://www.loc.gov/standards/mets/</a>
36	Preservation Metadata: Implementation Strategies (PREMIS) Data Dictionary 3.0 (II) <a href="https://www.loc.gov/standards/premis/">https://www.loc.gov/standards/premis/</a>
37	Encoded Archival Description (EAD3) 1.1.1 (II) <a href="https://www.loc.gov/ead/">https://www.loc.gov/ead/</a>
38	FHIR R5 resurser (II) <a href="https://www.hl7.org/fhir/resourceList.html">https://www.hl7.org/fhir/resourceList.html</a>
39	FHIR R5 Clinical Module (II) <a href="https://www.hl7.org/fhir/clinicalsummary-module.html">https://www.hl7.org/fhir/clinicalsummary-module.html</a>
40	European Health Data Space (II) <a href="https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en">https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space_en</a>
41	Projektdirektiv Etablera Mellanarkiv <a href="https://vgregion.sharepoint.com/:w:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/Styrdokument/Projektdirektiv%20Etablera%20mellanarkiv%20utkast%20version%205%20240116.docx?d=wb18363c03dbb46e98acd0f6b72dfec3a&amp;csf=1&amp;web=1&amp;e=0mqnco">https://vgregion.sharepoint.com/:w:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/Styrdokument/Projektdirektiv%20Etablera%20mellanarkiv%20utkast%20version%205%20240116.docx?d=wb18363c03dbb46e98acd0f6b72dfec3a&amp;csf=1&amp;web=1&amp;e=0mqnco</a>

42	ADD Mellanarkiv Journal I <a href="https://iserver.vgregion.se/#!/document/viewer/020ee60d-83a2-4ccf-9933-8c2831b62d41">https://iserver.vgregion.se/#!/document/viewer/020ee60d-83a2-4ccf-9933-8c2831b62d41</a>
43	CITS eHealth1 Patient Medical Records 2.0.0 DRAFT (II) <a href="https://iserver.vgregion.se/#!/object/details?objectId=8420146b-2d29-48c0-a12a-b5bc31c3a103">https://iserver.vgregion.se/#!/object/details?objectId=8420146b-2d29-48c0-a12a-b5bc31c3a103</a>
44	Informationsklassificering CESAR <a href="https://vgregion.sharepoint.com/:w:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/Informationsklassning%20av%20CESAR.docx?d=w9222110936aa4a0f80cca6d1707c9bd9&amp;csf=1&amp;web=1&amp;e=Wtd94w">https://vgregion.sharepoint.com/:w:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/Informationsklassning%20av%20CESAR.docx?d=w9222110936aa4a0f80cca6d1707c9bd9&amp;csf=1&amp;web=1&amp;e=Wtd94w</a>
45	Utredning: behörighetskoncept för CESAR <a href="https://vgregion.sharepoint.com/:w:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/IAM%20-%20F%C3%B6lja%20PDL.docx?d=w4562f288850240b78cae5974c8728727&amp;csf=1&amp;web=1&amp;e=63k3Rq">https://vgregion.sharepoint.com/:w:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/IAM%20-%20F%C3%B6lja%20PDL.docx?d=w4562f288850240b78cae5974c8728727&amp;csf=1&amp;web=1&amp;e=63k3Rq</a>
46	Behörigheter i RODA (beslut) <a href="https://vgregion.sharepoint.com/:b:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/Beh%C3%B6righeter%20RODA.pdf?csf=1&amp;web=1&amp;e=vwvudI">https://vgregion.sharepoint.com/:b:/r/sites/sy-rs-projekt-etablera-mellanarkiv/Delade%20dokument/Beh%C3%B6righeter%20RODA.pdf?csf=1&amp;web=1&amp;e=vwvudI</a>
47	ADD SITHS NLS (1.1) <a href="https://iserver.vgregion.se/#!/document/viewer/0c9543aa-df8c-44ff-b7d1-6048d1d1e0b0">https://iserver.vgregion.se/#!/document/viewer/0c9543aa-df8c-44ff-b7d1-6048d1d1e0b0</a>

# Innehållsförteckning

<b>1</b>	<b>DOKUMENTETS SYFTE .....</b>	<b>11</b>
<b>2</b>	<b>INLEDNING .....</b>	<b>11</b>
<b>2.1</b>	<b>Bakgrund, syfte och mål .....</b>	<b>11</b>
2.1.1	Syfte och mål .....	12
<b>2.2</b>	<b>Intressenter.....</b>	<b>13</b>
<b>2.3</b>	<b>Omfattning.....</b>	<b>14</b>
<b>2.4</b>	<b>Antaganden, beroenden och begränsande faktorer .....</b>	<b>14</b>
2.4.1	Antaganden .....	14
2.4.2	Beroenden.....	15
2.4.3	Kvalitetskrav .....	15
2.4.4	Övriga begränsande faktorer .....	15
2.4.5	Risker .....	15
<b>2.5</b>	<b>Tillämpade arkitekturprinciper .....</b>	<b>15</b>
2.5.1	EA-principer.....	15
2.5.2	Andra principer .....	17
<b>3</b>	<b>NULÄGESARKITEKTUR.....</b>	<b>18</b>
<b>3.1</b>	<b>Verksamhet.....</b>	<b>18</b>
3.1.1	Organisation och funktion.....	18
3.1.2	Processbeskrivning .....	19
<b>3.2</b>	<b>Informations-/dataarkitektur .....</b>	<b>20</b>
3.2.1	Informationsmodell .....	20
3.2.2	Förteckning över informationsobjekt .....	20
3.2.3	Tillämpade kodverk .....	21
<b>3.3</b>	<b>Applikationsarkitektur.....</b>	<b>21</b>
3.3.1	Applikationsförteckning .....	21
3.3.2	Funktionell vy.....	21
3.3.3	Informationsflöden.....	22
3.3.4	Interaktioner .....	23
3.3.5	Integrationsarkitektur .....	23
<b>3.4</b>	<b>Infrastruktur.....</b>	<b>24</b>
3.4.1	Teknisk plattform.....	24
3.4.2	Driftplatser och miljöer .....	25
<b>3.5</b>	<b>Säkerhetsarkitektur.....</b>	<b>25</b>
<b>4</b>	<b>BESKRIVNING OCH UTVÄRDERING AV LÖSNINGALTERNATIV .....</b>	<b>26</b>
4.1.1	Referensarkitekturer .....	26
4.1.1.1	<i>OAIS referensmodell.....</i>	<i>26</i>
4.1.1.2	<i>Referensarkitektur för relationsdatabaser .....</i>	<i>29</i>
4.1.1.3	<i>Bevarande av databaser.....</i>	<i>30</i>
4.1.1.4	<i>SIARD.....</i>	<i>32</i>
4.1.1.5	<i>Database Preservation Toolkit .....</i>	<i>33</i>
4.1.1.6	<i>E-ARK.....</i>	<i>33</i>
4.1.1.7	<i>E-Ark CSIP .....</i>	<i>34</i>
4.1.1.8	<i>E-Ark eHealth1.....</i>	<i>35</i>
4.1.1.9	<i>E-Ark eHealth1 rev 2.0.....</i>	<i>37</i>

4.1.1.10	HL7 FHIR resurser.....	40
4.1.1.11	RIM för informationshantering (inom hälso- och sjukvård).....	40
4.1.1.12	European Health Data Space (EHDS).....	41
4.1.2	Målarkitekturer.....	42
4.1.2.1	Plan för digital informationshantering.....	42
4.1.2.2	ADD - ARK Strategisk målarkitekturför mellanarkivering.....	43
4.1.2.3	Strategi för datamigrering i Millennium.....	46
4.1.3	Standarder.....	47
4.1.4	Bedömning av möjlighet till återanvändning.....	48
4.1.5	Lösningalternativ - sammanfattning.....	48
4.1.6	Lösningalternativ 1 - RODA Community.....	49
4.1.6.1	Beskrivning av lösningalternativ 1.....	49
4.1.6.2	Utvärdering av lösningalternativ 1.....	50
4.1.6.3	Riskhantering av lösningalternativ 1.....	51
4.1.7	Lösningalternativ 2 - RODA Enterprise.....	51
4.1.7.1	Beskrivning av lösningalternativ 2.....	51
4.1.7.2	Utvärdering av lösningalternativ 2.....	57
4.1.7.3	Riskhantering av lösningalternativ 2.....	58
<b>4.2</b>	<b>Val av lösningalternativ.....</b>	<b>58</b>
<b>5</b>	<b>MÅLARKITEKTUR FÖR VALT LÖSNINGSALTERNATIV.....</b>	<b>59</b>
<b>5.1</b>	<b>Verksamhet.....</b>	<b>59</b>
5.1.1	Organisation och funktion.....	59
5.1.2	Processbeskrivning.....	61
<b>5.2</b>	<b>Informations-/dataarkitektur.....</b>	<b>62</b>
5.2.1	Informationsmodell för arkiverade patientjournaler.....	63
5.2.1.1	SIP modell.....	63
5.2.1.2	AIP modell.....	63
5.2.1.3	DIP modell.....	64
5.2.2	Förteckning över informationsobjekt.....	64
5.2.3	Tillämpade kodverk.....	65
<b>5.3</b>	<b>Applikationsarkitektur.....</b>	<b>66</b>
5.3.1	Informationssystem och masterdata.....	66
5.3.2	Applikationsförteckning.....	66
5.3.3	Funktionell vy.....	68
5.3.4	Informationsflöden.....	68
5.3.5	Förteckning över informationsflöden.....	69
5.3.6	Interaktioner.....	70
5.3.7	Integrationsarkitektur.....	70
5.3.8	Förteckning över integrationskomponenter.....	71
<b>5.4</b>	<b>Infrastruktur.....</b>	<b>72</b>
5.4.1	Teknisk plattform.....	72
5.4.1.1	Lokal DEV miljö.....	73
5.4.1.2	DEV miljö.....	73
5.4.1.3	QA miljö.....	74
5.4.1.4	PROD miljö.....	74
5.4.2	Driftplatser och miljöer.....	74
5.4.3	Förteckning över infrastrukturkomponenter.....	74
<b>5.5</b>	<b>Säkerhetsarkitektur.....</b>	<b>76</b>

5.5.1	Följsamhet till patientdatalagen (PDL) .....	76
<b>6</b>	<b>GAP-ANALYS .....</b>	<b>77</b>
<b>6.1</b>	<b>Verksamhet.....</b>	<b>77</b>
6.1.1	Organisation .....	77
6.1.2	Processer .....	77
<b>6.2</b>	<b>Informations-/dataarkitektur .....</b>	<b>77</b>
<b>6.3</b>	<b>Applikationsarkitektur .....</b>	<b>78</b>
<b>6.4</b>	<b>Infrastruktur .....</b>	<b>78</b>
<b>7</b>	<b>ROADMAP FÖR GENOMFÖRANDE .....</b>	<b>80</b>
<b>7.1</b>	<b>Strategi för implementation och migration .....</b>	<b>80</b>
<b>7.2</b>	<b>Beskrivning av övergångsarkitekturer .....</b>	<b>80</b>
7.2.1	Definition av övergångsstadier.....	80
7.2.2	Övergångsarkitekturer för respektive stadie .....	80
7.2.3	Hantering av risker kopplat till implementation och migration .....	80
<b>7.3</b>	<b>Roadmap diagram.....</b>	<b>81</b>
<b>8</b>	<b>AVVIKELSER.....</b>	<b>82</b>

# Ordlista

Begrepp	Beskrivning
Ansvar	Varje arkivbildare ansvarar för sina allmänna handlingar. Regionstyrelsen har det tekniska arkivansvaret för regiongemensamma system. Säker hantering av känsliga uppgifter krävs.
AIP	Archive Information Package, ett paketformat för arkivering av digitala handlingar
API	Application Programming Interface, ett gränssnitt som tillåter program att kommunicera med varandra.
Arkiv	En samling av handlingar som bevaras för att tjäna som bevis på en verksamhet eller händelse.
Arkiverad information	Ska vara öppen för konsumtion om den fyller en funktion.
Arkiverande informationshandlingar	Dokument och data som behöver arkiveras.
Arkivlag	En svensk lag som reglerar hantering av allmänna handlingar.
Arkivlagen	Ställer krav på informationshantering direkt när en handling skapas.
Arkivvård	Den verksamhet som syftar till att bevara och tillgängliggöra arkiv.
Avveckling	Processen att ta ur drift ett system eller en tjänst.
Data	Information som representeras i en form som kan bearbetas av en dator.
Digitalisering	Processen att omvandla information till digital form.
DIP	Dissemination Information Package, ett paketformat för arkivering av digitala handlingar.
E-arkiv	Ett arkiv som lagrar digitala handlingar.
ETL	ETL ( <i>Extract, Transform, Load</i> ) är en metod för att transformera data; Hämta data från källa (Extract), transformera data (Transform), och ladda data hos mottagare (Load).
Gallring	Processen att rensa bort handlingar som inte längre behöver bevaras.
Handling	En skriftlig eller digital information som dokumenterar en händelse eller transaktion.
Informationsdomäner för verksamhetssystem	Gruppering av liknande informationsmängder.
Informationshantering	Den samlade verksamheten för att skapa, använda, bevara och gallra information.
Informationshanteringsplan och diarieplan	Ska baseras på klassificeringsstrukturen.
Informationsmodell	En beskrivning av hur information är strukturerad och relaterad.
Informationssäkerhet	En viktig del av bevarandet av digital information.
Integration	Processen att sammanföra information från olika källor.
Interoperabilitet	Förmågan för olika system att arbeta tillsammans.

Klassificeringsstruktur	Kopplar handlingslag till processer.
Klinisk information	Information om en patients hälsa och vård.
Koncernstab digitalisering (KSD)	Ansvarar för digitala lösningar för mellanarkivering.
Konsumtion	Användning av information.
Kärnsystem	Det nya systemet, Millennium, som regionen övergår till.
Livscykelhantering av information	Informationens livscykel startar när den skapas och slutar när den rensas eller gallras. Information som går till slutarkivering bevaras permanent.
Långsiktig hållbarhet	Förmågan att fungera väl under en lång tid.
Mellanarkivering	Tillfällig lagring av handlingar innan de slutarkiveras.
Metadata	Information om information.
Metadata	Metadata
Millennium	Ett nytt IT-system som införs i Västra Götalandsregionen.
NPÖ	Nationell Patientöversikt
OAIS	Open Archival Information System, en referensmodell för arkivering av digitala handlingar.
Offentlighetsprincipen	Principen att allmänna handlingar ska vara tillgängliga för allmänheten.
PDL	Patientdatalag (SFS 2008:355)
Plandokument	Dokumentet som beskriver behovet av att hantera allmänna handlingar enligt Arkivlagen.
Processororienterad informationsredovisning	Riktlinjer styr hur information ska ordnas för att kunna förtecknas.
Processägare	Företräder arkivbildare som informationsägare.
R7-samarbetet	Ett tidigare samarbete som VGR nu har beslutat att gå ur.
Regional masterdata	Reglerar innehållet i informationsmängder och attribut.
Regionala arkivlösningar	System för att hantera allmänna handlingar.
Regionala tjänster	Ska ge åtkomst till klassificeringsstrukturer och informationsplaner.
Regionarkivet	Arkivet för Västra Götalandsregionen.
RIM	Reference Information Model, en modell för att beskriva informationens struktur och semantik.
SIP	Submission Information Package, en paketformat för inleverans av digitala handlingar till arkiv.
Slutarkivering	Permanent bevarande av handlingar.
SQL	Structured Query Language

Standard	En teknisk specifikation som används för att säkerställa interoperabilitet.
Strategisk målarkitektur	En vägledning för hur mellanarkivering praktiskt ska genomföras.
Tillgodose olika intressenters behov	Allmänhetens behov, kärnverksamhetens behov, forskningens behov.
Tillgänglighet	Förmågan att få tillgång till information.
XML	Extensible Markup Language, ett språk för att strukturera och märka upp digital information.
XSD	XML Schema Definition
Ärende	En samling av handlingar som rör en specifik fråga eller händelse.

Tabell 1.1 - Ordlista

# 1 Dokumentets syfte

Detta dokument (ADD) är en central leverans i arbetet med lösningsarkitektur och beskriver en föreslagen lösning samt den förändring av det befintliga IT-landskapet som krävs för att lösningen ska kunna realiseras.

Dokumentet ger bakgrund och kontext till förändringen, sammanställer de arkitekturella vyer som tagits fram inom arkitekturuppdraget och kompletterar dessa med förtydligande beskrivningar.

ADD spänner över alla arkitekturdomäner (Verksamhet, Data, Applikation och Infrastruktur) enligt VGR:s arkitekturramverk VARVET samt beskriver förändringen från nuläge till börilage och mellanliggande tillstånd.

Dokumentet lagras tillsammans med underliggande modeller, artefakter och övriga leverabler i VGR:s samlade arkitekturbibliotek.

## 2 Inledning

Under hösten 2023 har ett projekt initierats för etablering av ett långsiktigt, strategiskt mellanarkiv för information inom den regionala vårddomänen. Se projektdirektivet i bilaga [41] för mer info. Målarkitekturen i detta dokument beskriver projektets huvudleverabel.

Projektet genomförs agilt enligt SAFe med kvartalsvis PI-planering, och lösningen baseras på egenutvecklad mjukvara i kombination med opensource komponenter. Arkitekturuppdraget anpassas till detta genom löpande revisioner i samma kadens. Nästa revision blir således ca 15 maj 2024. Målet är att ADD i varje revision ska adressera relevanta frågor inför nästkommande PI-period.

Denna ADD ska därför inte ses som en färdig lösningsarkitektur. Det finns fortfarande frågor som behöver utredas, bl.a:

- Krav på åtkomst och behörighet till både mellanarkiv i sig och arkiverade journalhandlingar i olika användningsfall behöver färdigställas i projektet för att lösningens säkerhetsarkitektur kan färdigställas.
- Fysisk lagringsmodell för arkiverade patientjournaler och i synnerhet strukturerad data behöver beslutas i nära samråd med avvecklingsprojekt för Melior och Asynja Visph.

Flera delar av lösningsarkitekturen är därför formulerade på hög nivå och ska endast ses som indikativa; de kommer att utredas och beskrivas mer detaljerat i kommande revisioner.

### 2.1 Bakgrund, syfte och mål

Patient- och journalinformation i vårdsystemen *Medidoc* och *Journal3* har i tidigare projekt mellanarkiverats i den nationella arkivtjänsten *R7e-arkiv* med hjälp av verktygsapplikationen *Marshal OAS*.

I januari 2024 infördes en ny kortsiktig/temporär lösning som beskrivs i ”ADD Mellanarkiv Journal I” [42]. I bilagan beskrivs lösningsarkitekturen som en övergångsarkitektur i kapitel 7. I detta dokument har denna arkitekturbeskrivning flyttats till kapitel 3 för att beskriva nuläget.

Arkitekturuppdraget för det fortsatta långsiktiga uppdraget att etablera ett mellanarkiv [24] innehåller följande beskrivning: VGR har ett behov att etablera en komplett förmåga för att mellanarkivera alla typer av informationsmängder som finns i regionen där informationen behöver sparas tillgänglig även om systemet har avvecklats. VGR ska utveckla en integrerad, effektiv, hållbar och säker mellanarkivlösning för både klinisk och icke-klinisk information som optimerar hanteringen av informationsflödet inom VGR-regionen. Denna lösning syftar till att underlätta och möjliggöra snabb tillgång till viktig information när den behövs och samtidigt fullt ut uppfylla arkivlagens bestämmelser och krav.

Arkitekturuppdraget dokumenteras i detta dokument och kallas ”ADD Mellanarkiv Journal II”.

Plan för digital informationshantering [20] beskriver fem informationsdomäner för verksamhetssystem:

- Fristående dokument [...]
- Publiceringssystem [...]
- Processbaserad informationshantering
- Ekonomi, HR och administrativ ärendehantering
- Vårddomänen

Den sista, vårddomänen, adresseras nu i detta uppdrag. Vårddomänen är relevant för mellanarkivering av klinisk data och delas upp i fem underdomäner:

- Patientjournal
- Patientadministration
- Diagnostik, utredning och laboratoriesystem
- Registerdata vård
- Forskningsdata vård

Det finns ett starkt behov av en mellanarkivslösning för patientjournaler, patientadministrativ information samt diagnostik, utredning och laboratoriesystem. Behovet förstärks vid införande av Millennium genom att ett stort antal journalsystem och patientadministrativa system ska avvecklas.

Det finns enligt [20] i nuläget inget behov att mellanarkivera kvalitetsregister.

Ett lösningskoncept behöver även utvecklas för hantering av forskningsdata.

Mellanarkivering av övriga informationsdomäner kommer att beskrivas i framtida revisioner av detta dokument eller i separat ADD.

### 2.1.1 Syfte och mål

Syftet med arkitekturarbetet är att dels etablera en generell plattform för mellanarkivering av informationshandlingar, dels att möjliggöra mellanarkivering av klinisk information inom vårddomänen.

Mellanarkivering av klinisk information kännetecknas bl.a. av:

- Arkivfunktionen förvaltar informationen
- Verksamheten har begränsad direktåtkomst utifrån väl avgränsade syften
- PDL-stöd krävs om verksamheten ges direktåtkomst
- Verksamheten får endast redigera information inom ramen för väl avgränsade syften
- Informationen struktureras enligt särskild specifikation för paketstruktur (VGR)
- Mellanarkivet byts nästan aldrig ut men kringliggande applikationer kan bytas ut

Mellanarkivering av ytterligare informationsdomäner kommer att adresseras i framtiden.

Mål	Avseende
Huvudmål för fas 2	Att etablera ett mellanarkiv som följer OAIS standarden inkl. API för hantering av OAIS informationspaket samt exponering av en harmoniserad informationsmodell per domän/underdomän för läsning av data från operationella applikationer; t.ex. via uthopp
Delmål för fas 2	Möjliggör mellanarkivering av kliniska system som avvecklas efter införande av Millennium

Tabell 2.1 - Taktiska mål

Samtidigt är de övergripande målen för strategisk mellanarkivering enligt [16] följande:

- Arkiverad information ska kunna användas för praktiska tillämpningar inom områden som dataanalys, maskininlärning för AI-tillämpningar, forskning eller andra konsumtionsbehov som kärnverksamheten ser ett värde i.

- Allmänna handlingar ska hanteras i enlighet med Plandokumentets uttryckta färdriktning, bl.a. att Arkivlagen ska följas.
- Digitalt systemstöd ska säkerställa krav på livscykelhantering av information från informationens skapande till gallring eller slutarkivering.

En samlad målbild beskrivs nedan:

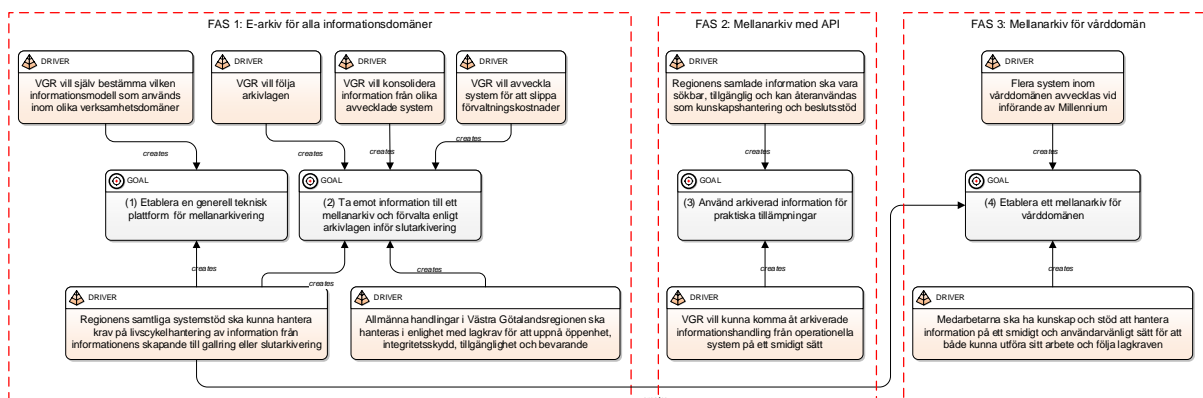


Diagram 2.1 - Övergripande för Mellanarkiv

Delmål kopplade till varje övergripande mål visas nedan. Dessa prioriteras som led i PI-planeringen.

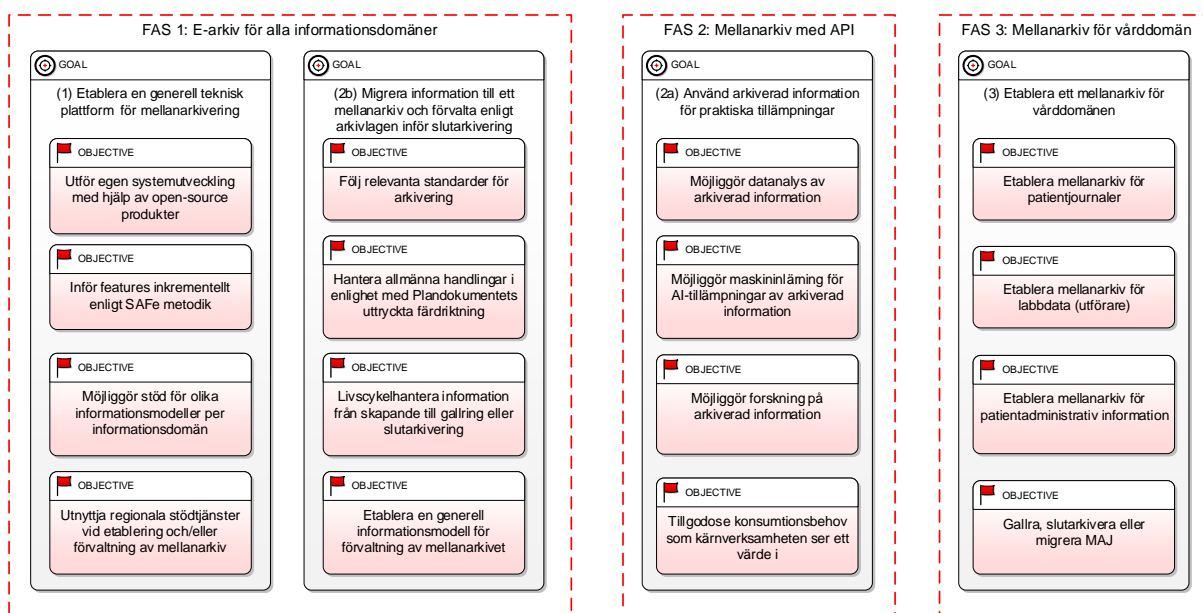


Diagram 2.2 - (Prioriterade) delmål för Mellanarkiv

## 2.2 Intressenter

En intressent är en person, en grupp av människor eller en organisation som på ett eller annat sätt påverkas av projektet ”Etablera mellanarkiv”, eller som kan påverka förutsättningarna för projektet, som till att börja med i detta fall är etablering av ett mellanarkiv för patientjournaler.

I tabellen nedan visas en första intressentanalys. Den kan behöva kompletteras med en mer detaljerad analys av vad ”ta del av [...] arkiverade journaler” innebär.

Intressent(grupp)	Kommentar
Vårdpersonal - VGR vårdgivare	Läkare, administrativ personal (medicinska sekreterare, kanslist, receptionister), sjuksköterskor, undersköterskor, paramedicinsk

	personal <sup>1</sup> , biomedicinska analytiker vill kunna ta del av arkiverade patientjournaler på ett enkelt och smidigt sätt via uthopp från Millennium.
Forskning och utbildning	Forskare, kliniska forskare vill kunna använda arkiverade patientjournaler som forskningsunderlag.
Patienter	Invånare vill kunna ta del av sina egna mellanarkiverade patientjournaler på ett enkelt och smidigt sätt via 1177.
Privata vårdgivare och kommuner	Samma som Vårdpersonal - VGR vårdgivare.
Andra vårdgivare	(i andra regioner) vill kunna ta del av mellanarkiverade patientjournaler på ett enkelt och smidigt sätt via NPÖ.
Koncernkontoret	Kansli och säkerhet, Koncernstab Digitalisering vill ha ett mellanarkiv för patientjournaler som är enkelt, smidigt och säkert att förvalta.
Ekonomi	Dataanalytiker, ekonomiassistenter vill kunna använda arkiverade patientjournaler som underlag för statistisk analys.
Övriga intressenter	Polis, socialtjänst, förvaltningsrätt, Inspektionen för vård och omsorg, försäkringsbolag, rättsmedicin, transportstyrelsen vill kunna ta del av mellanarkiverade patientjournaler på ett enkelt och smidigt sätt

Tabell 2.2 - Mellanarkivets intressenter (patientjournaler).

## 2.3 Omfattning

Arkitekturuppdraget är omfattande och kommer att genomföras i inkrement motsvarande den övergripande leveransplanen för mellanarkiv. Beslut har tagits om egenutveckling i stället för upphandling, så arkitekturarbetet kan fokusera på vad som är av mest nytta för det egna utvecklingsteamet.

Område (arkitekturdomän)	Omfattning
Verksamhet	Detaljerad kartläggning av baseline, target, gap-analys och plan för genomförande.
Information	Detaljerad kartläggning av baseline, target, gap-analys och plan för genomförande.
Applikation	Detaljerad kartläggning av baseline, target, gap-analys och plan för genomförande.
Infrastruktur	Detaljerad kartläggning av baseline, target, gap-analys och plan för genomförande.

Tabell 2.3 - Omfattning per arkitekturområde

## 2.4 Antaganden, beroenden och begränsande faktorer

### 2.4.1 Antaganden

- Integration med Millennium (och andra relevanta journalsystem) realiseras som befintlig lösning i form av uthopp.

<sup>1</sup> Arbetsterapeuter, sjukgymnaster, fysioterapeuter, fysiologer, kuratorer, logopedier, dietister m fl.

### 2.4.2 Beroenden

- Inleveransspecifikation för arkivering av patientjournaler tas fram i nära samarbete med avvecklingsprojektet för Melior. Specifikationen behöver ses över och vid behov revideras vid avveckling av övriga Meliorinstanser och andra journalsystem som ersätts av Millennium.
- Integration med Millennium behöver samordnas med Millenniumprogrammet. Se även antagande ovan.
- Relevant metadata publiceras till GIP datakatalog. Egen instans av Denodo kommer att användas för datavirtualisering. Mulesoft Anypoint kommer kanske att användas som API gateway. Alla ingår i GIP.
- Avställningen av LIS-system syftar till att säkerställa och vidmakthålla information från laboratorieinformationssystem som ställa av i samband med övergången till det Multi LIS system som VGR har intention att upphandla. Upphandlingsprojektet heter LISA och kommer att ställa krav på mellanlagring av laboratorieinformation från ett stort antal LIS-system som idag används inom VGR. Laboratorieinformationen ska i ett senare läge arkiveras till mellanarkivet.

### 2.4.3 Kvalitetskrav

<beskrivande text>

### 2.4.4 Övriga begränsande faktorer

<beskrivande text>

### 2.4.5 Risker

Informationsklassificering samt risk- och sårbarhetsanalys har utförts tidigare för Mellanarkiv I.

Skyddsvärt		Hot		Sårbarheter	Konsekvensbeskrivning	Riskbedömning innan åtgärd			Fortsatt analys?	Åtgärder
Skyddsvärda tillgångar relevanta för analysen		Möjlig, oönskad händelse med negativa konsekvenser		Problem/brister/forsaker som ligger till grund för hoten	Beskrivning av de troliga konsekvenserna om hotet inträffar.	Konsekvens	Sannolikhet	Risikvärde	Vilka hot som ska vidare till steg 3?	Vad kan göras för att eliminera, begränsa eller beakta riskerna och dess sårbarheter?
ID	Tillgång	ID	Hot	Sårbarhet	Konsekvens					
1	Mellanarkiverad patientjournal	H01	Informationen stjäls / röjs	Bristande behörighetskontroll	Sekretessbelagd känslig information röjs. Brott mot PDL	3	2	6	Nej - Acceptera risk	Följ etablerade rutiner. Begränsa. LoA3 inloggning och aktivitets / åtkomstlogg. Bevaka. Kontrollera åtkomst regelbundet.
2	Mellanarkiverad patientjournal	H02	Informationen förfalskas / ändras av obehörig	Otillräcklig informationssäkerhet	Sekretessbelagd känslig information förfalskas. Brott mot PDL. Patientsäkerheten hotas om journalen förfalskas	3	1	3	Nej - Acceptera risk	
3	Mellanarkiverad patientjournal	H03	Informationen är otillgänglig. Hotet inträffar även om lokal loggning är otillgänglig	Driftstörning / sabotage	Patientsäkerheten hotas när journalen är otillgänglig	3	2	6	Nej - Acceptera risk	
4	Logg åtkomst / aktivitet	H04	Informationen stjäls / röjs	Bristande behörighetskontroll	Obehörig får åtkomst till logg	2	1	2	Nej - Acceptera risk	
5	Logg åtkomst / aktivitet	H05	Informationen förfalskas / ändras av obehörig	Otillräcklig informationssäkerhet	Spårbarheten tappas. Brott mot PDL	3	1	3	Nej - Acceptera risk	
6	Logg åtkomst / aktivitet	H06	Informationen är otillgänglig	Driftstörning / sabotage	Granskning av loggar enligt PDL försenas	2	2	4	Nej - Acceptera risk	
7	Administrativ data i systemkonfiguration	H07	Informationen stjäls / röjs	Bristande behörighetskontroll	Delar av konfigurationen för MAJ röjs med försumbar konsekvens	1	2	2	Nej - Acceptera risk	
8	Administrativ data i systemkonfiguration	H08	Informationen förfalskas / ändras av obehörig	Otillräcklig informationssäkerhet	Systemet skyddar inte informationen enligt tänkt behörighetskontroll	3	1	3	Nej - Acceptera risk	
9	Administrativ data i systemkonfiguration	H09	Informationen är otillgänglig	Driftstörning / sabotage	Patientjournalen blir otillgänglig	3	2	6	Nej - Acceptera risk	

Tabell 2.4 - Risktabell

## 2.5 Tillämpade arkitekturprinciper

### 2.5.1 EA-principer

Principnamn	Principformulering	Påverkan på arkitekturen	Referens
-------------	--------------------	--------------------------	----------

Flera aktörer - en region	Samverkan sker inom regionen och med externa parter för att skapa kvalitet och effektivitet inom berörda verksamheter.  Gränserna mellan olika aktörer ska utifrån invånarnas och organisationers perspektiv suddas ut. Genom samverkan ökar förutsättningarna att de tjänster som erbjuds uppfattas som meningsfulla.	Information från hela regionen som ska bevaras i ett mellanarkiv, sparas till ett gemensamt mellanarkiv.	EA-principer [1]
Prioritera gemensamma arbetssätt	Regionala och gemensamma arbetssätt ska prioriteras för en enhetlig och effektiv verksamhet.	En gemensam hantering av mellanarkivet tas fram.	EA-principer [1]
Prioritera digitalisering	Digitalisera regionens utbud för att möta samhällets och invånarnas förväntan på digitala tjänster.  Digitaliseringen är också en nödvändig del i utvecklingen och effektiviseringen av regionens verksamhet för att kunna möta ett ökande vårdbehov. (ökande behov av tjänster från regionens verksamheter).	Mellanarkivet lagrar all information digitalt.	EA-principer [1]
Information är en strategisk tillgång	Den information som Västra Götalandsregionen hanterar ska behandlas och vårdas som en strategisk, taktisk och operativ tillgång.	Informationen som mellanarkiveras fortsätter att vara sökbar och tillgänglig.	EA-principer [1]
Information ska vara tillgänglig	Verksamheten, invånare och externa aktörer har tillgång till den information de behöver för att utföra sina uppgifter.  Information delas mellan olika verksamhetsfunktioner, organisatoriska enheter och ska tillgängliggöras för externa parter i så stor omfattning som möjligt, med beaktande av sekretess- och integritetsaspekter	Information i mellanarkivet ska göras tillgänglig till verksamheten, invånare och externa aktörer via API.	EA-principer [1]
Information utgår från gemensamma begrepp och standarder	Regionens information skall så långt som möjligt bygga på en gemensam utveckling av informationsstrukturer och terminologier, vilka ska utgå från regionala beslut.	Referensinformationsmodeller framtagna inom VGR ska användas som referens.	EA-principer [1]
Informationssäkerhet baseras på informationens värde	Värdet på informationen ligger till grund för vilket skydd, nivå av tillgänglighet och behov av riktighet som krävs för hantering av informationen.	Informationsklassificering utförs enligt gällande rutiner.	EA-principer [1]
Minimera antalet verksamhetsapplikationer	Eftersträva att hålla nere antalet verksamhetsapplikationer samt konsolidera där det är möjligt.  Ambitionen är att det för varje verksamhetsfunktion endast ska finnas en verksamhetsapplikation dvs. minimera förekomster av flera verksamhetsapplikationer som utför samma sak samt flera driftsatta versioner av samma verksamhetsapplikation.	Bli mer aktuellt i framtida versioner i takt med att flera informationsdomäner mellanarkiveras.	EA-principer [1]

	Ett komplext applikationslandskap med många applikationer är både kostnadsdrivande och påverkar ledtider vid förändringar.		
Återanvänd före köp före utveckling	Applikationer, infrastruktur- och systemkomponenter ska i första hand återanvändas, i andra hand inköpas om det behövs och i sista hand byggas och utvecklas endast om det finns unika behov eller krav som inte kan uppfyllas på annat sätt.	Projektet har utrett köp vs utveckla själv och beslutat att egenutveckling är att föredra i det här fallet.	EA-principer [1]
Teknisk interoperabilitet genom standardisering	Mjukvara och hårdvara ska följa regionalt beslutade standarder för att på ett säkert sätt kunna utbyta information med den kvalitet som parterna i ett informationsutbyte kommit överens om för att uppnå teknisk interoperabilitet.	Standardisering används i hög grad i lösningsarkitekturen. Beskrivs närmare i kap. 4.	EA-principer [1]

## 2.5.2 Andra principer

Komplettera med andra principer från ny bilaga [16] i nästa revision.

Principnamn	Principformulering	Påverkan på arkitekturen	Referens

## 3 Nulägesarkitektur

VGR har under hösten 2023 etablerat en temporär/enkel/snabb lösning med övergripande mål att ersätta R7 e-arkiv från 2024-01-01. Lösningen är inte ett komplett mellanarkiv; det saknas t.ex. stöd för att arkivera ny information.

### 3.1 Verksamhet

#### 3.1.1 Organisation och funktion

I nuläget avgränsas relevanta systemroller till regionala aktörer. Detta förenklar organisationskartan i förhållande till målarkitekturen.

Kommunala- och privata vårdgivare visas i diagrammet, men är inte aktörer i mellanarkivet i nuläget.

Roller i organisationen:

- Förvaltare - administrerar och förvaltar mellanarkivet
- Vårdpersonal - ger hälso- och sjukvård till patienter
- Loggranskare - kontrollerar accessloggar till patientjournaler

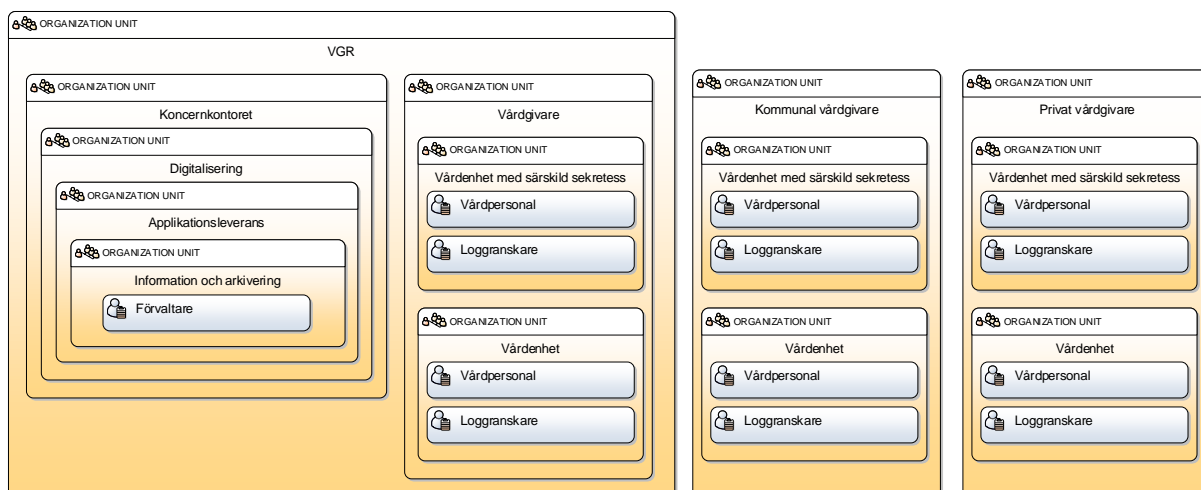


Diagram 3.1 - Organisationskarta (baseline)

I nuläget utförs inga nya avvecklingar och ingen ny data kan överföras till mellanarkivet. Det finns därför ingen process för att migrera/arkivera information till mellanarkivet.

Funktioner:

- Informations- och arkiveringshantering - ansvarar för att mellanarkivet hanteras på ett korrekt sätt enligt gällande lagar
- Journalföring inom hälso- och sjukvård - ansvarar för att dokumentera relevanta händelser och aktiviteter inom hälso- och sjukvård på ett korrekt sätt enligt gällande lagar

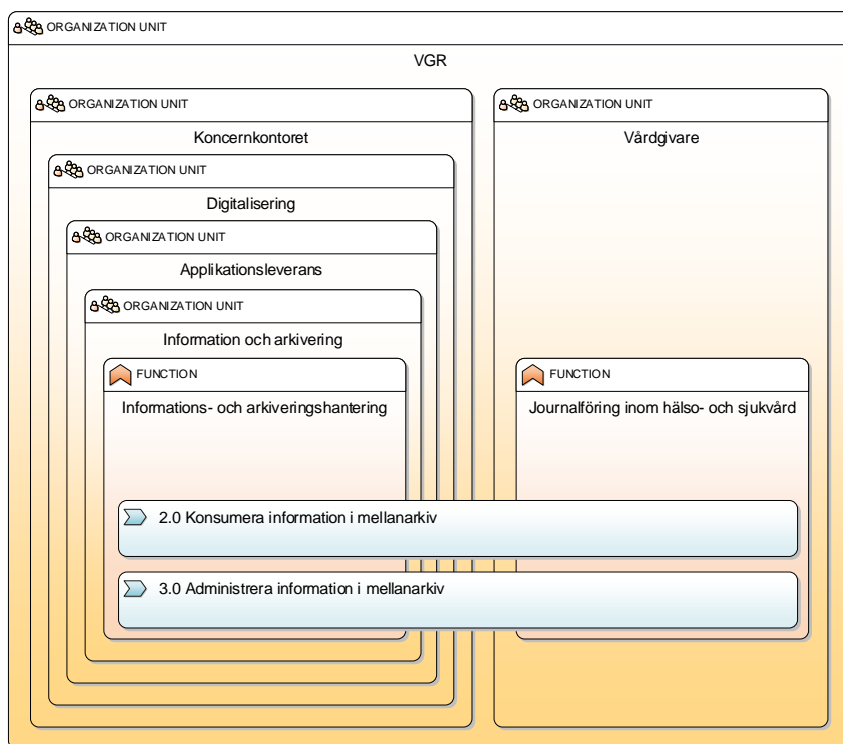


Diagram 3.2 - Functional Decomposition Diagram (baseline)

### 3.1.2 Processbeskrivning

I nuläget utförs inga nya avvecklingar och ingen ny data kan överföras till mellanarkivet. Det finns därför ingen process för att migrera information till mellanarkivet.

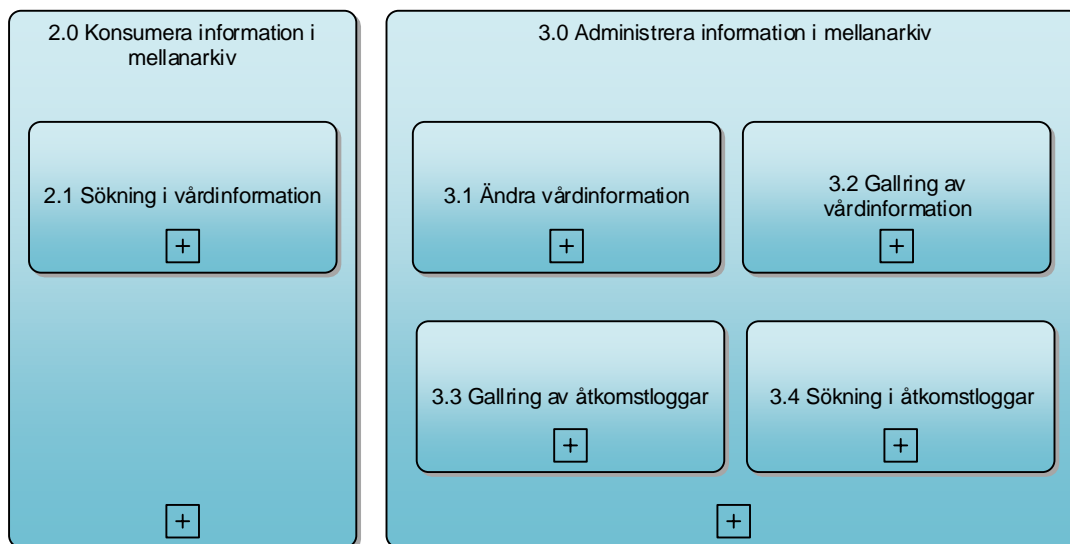


Diagram 3.3 - Processkarta (baseline)

Delprocess *2.1 Sökning i vårdinformation* använder en webbapplikationen för att söka efter arkiverade journalhandlingar. Behörig användare (vårdpersonal) initierar en webbapplikationen genom uthopp från befintliga journalsystem<sup>2</sup> för att se arkiverade journalhandlingar för specifik patient.

Delprocess *3.4 Gallring av åtkomstloggar* och *3.5 Sökning i åtkomstloggar* använder Ineras Loggrapporttjänst.

<sup>2</sup> Melior, Asynja Visph och T4

## 3.2 Informations-/dataarkitektur

### 3.2.1 Informationsmodell

Informationsarkitekturen representeras av den informationsmodell som tidigare användes i Marshall OASIS och R7 e-arkiv [19]. Informationsmodellen realiseras som XML-filer.

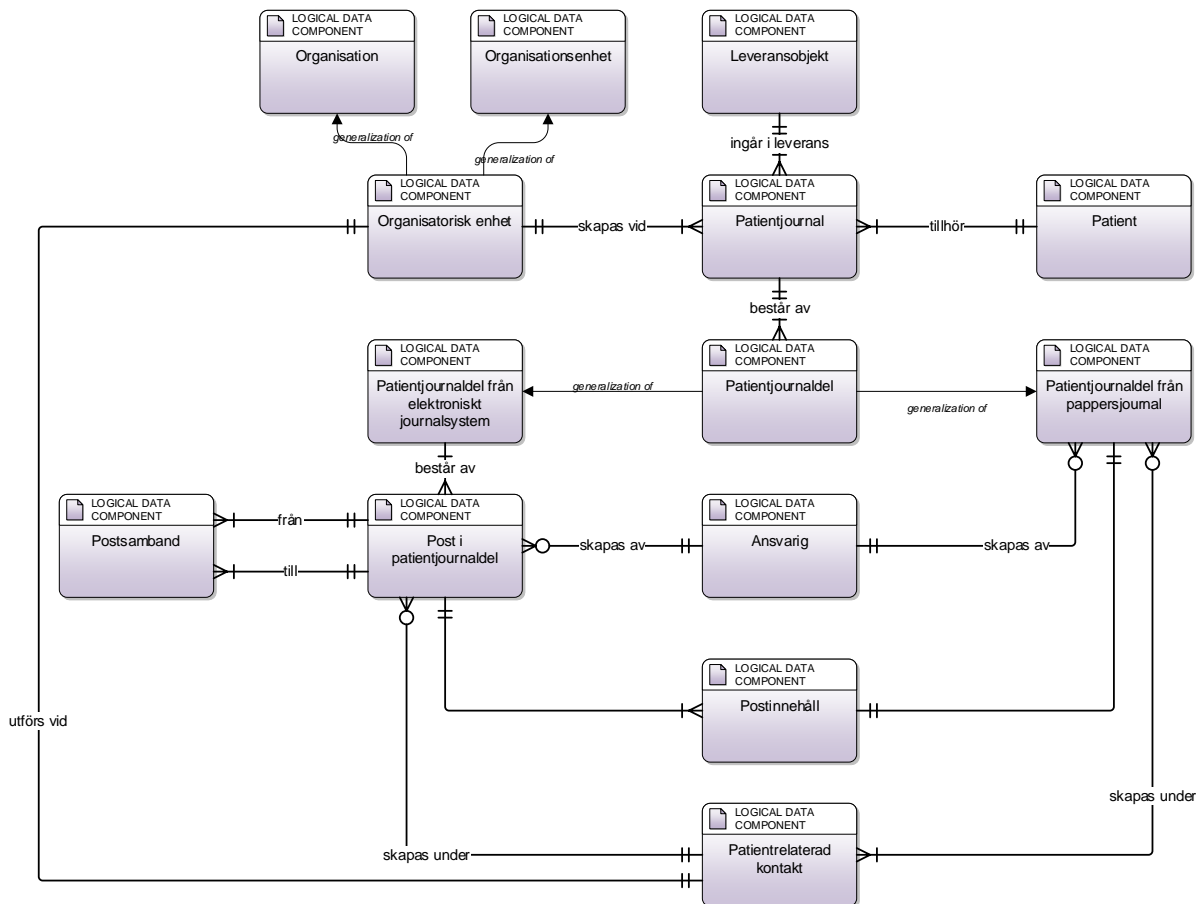


Diagram 3.4 - Informationsmodell för MAJ (baseline)

### 3.2.2 Förteckning över informationsobjekt

Logisk Datakomponent	Beskrivning
Patient	
Patientjournal	
Patientjournaldel	
Patientjournaldel från pappersjournal	
Patientjournaldel från elektroniskt journalsystem	
Post i patientjournaldel	
Postinnehåll	
Postsamband	
Patientrelaterad kontakt	
Organisation	
Organisationsenhet	

Organisatorisk enhet	
----------------------	--

Tabell 3.1 - Data Component Catalog (baseline)

### 3.2.3 Tillämpade kodverk

Har ej utretts. Utifrån informationsmodellen gör vi antagandet att kodverk som används i arkiverade journalhandlingar inte har några externa kopplingar/beroenden till kodverk som hanteras i t.ex. GIP Datakatalog.

Kodverk (namn)	Tillämpning

Tabell 3.2 – Tillämpade kodverk (baseline)

## 3.3 Applikationsarkitektur

I applikationsarkitekturen är befintliga Marshal applikationer för arkivförvaltning oförändrade sedan R7. Ny applikation för åtkomst/sökning i mellanarkivet har utvecklats för att ersätta R7.

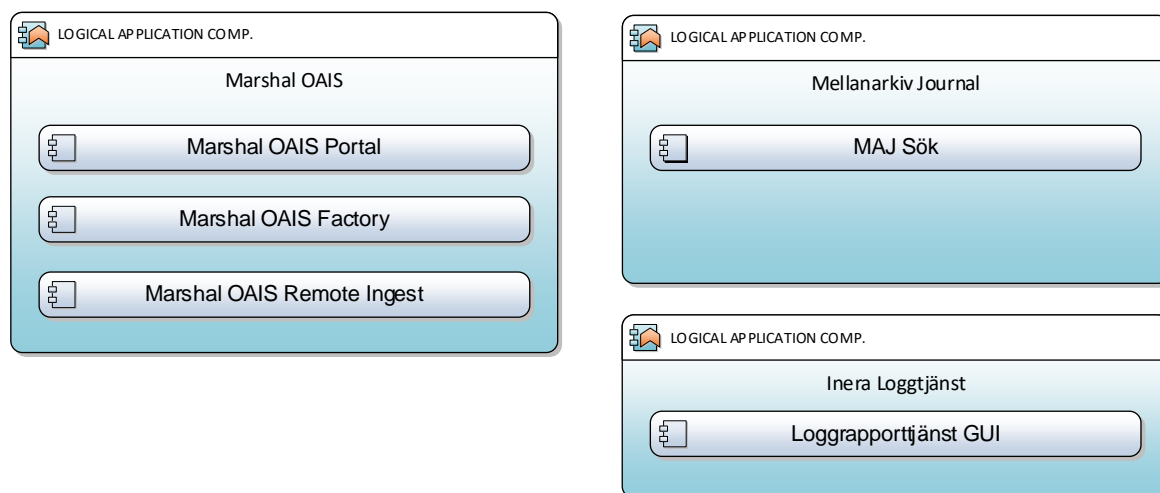


Diagram 3.5 - Applikationer (baseline)

### 3.3.1 Applikationsförteckning

Fysisk applikation	Logisk applikation	Logisk applikation (kategori)	Beskrivning
Marshal OAIS Portal, Factory & Remote Ingest	Marshal OAIS	Mellanarkiv	Används för att förvalta arkivet enligt OAIS standard.
MAJ Sök	Mellanarkiv Journal	Mellanarkiv Journal applikation	Används för att söka fram arkiverade journalhandlingar
Loggrapporttjänst GUI	Inera Loggtjänst	Mellanarkiv Journal applikation	Används för att skapa och förvalta åtkomstloggar enligt PDL

Tabell 3.3 - Application Component Catalog (baseline)

### 3.3.2 Funktionell vy

I den funktionella vyn ses kopplingen mellan aktörer/roller, processer och applikationer.

Systemförvaltare deltar i *3.0 Administrera information i mellanarkiv* med IT-stöd i Marshal OAIS. Granskning av åtkomstloggar till mellanarkivet sker via Ineras Loggrapporttjänst GUI.

Vårdpersonal deltar i 2.0 Konsumera information i mellanarkiv genom uthopp från befintliga patientjournalssystem.

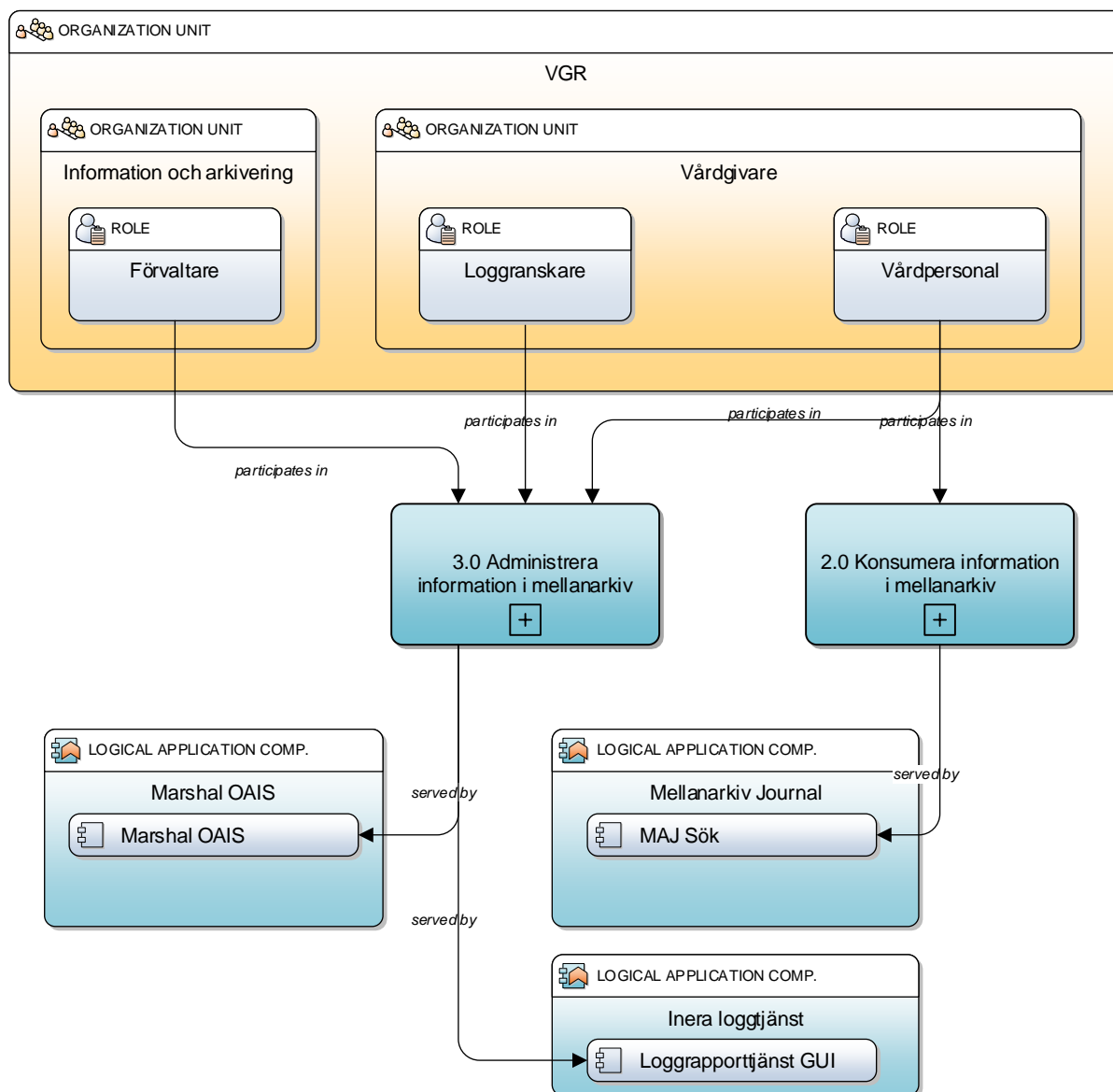


Diagram 3.6 - Application Usage Diagram (baseline)

### 3.3.3 Informationsflöden

I informationsflödesdiagrammet nedan redovisas de huvudsakliga informationsöverföringarna i lösningsarkitekturen.

Sammanfattningsvis:

- Befintliga journalssystem gör uthopp till MAJ Sök för sökning i mellanarkiv
- Identifiering och autentisering av användare sker med Lokal IdP
- Behörighetskontroll sker med hjälp av attribut i HSA-katalogen
- MAJ Sök läser spärrinfo från Ineras nationella spärrtjänst

Om en spärrad journal behöver nödöppnas enligt PDL:s riktlinjer, skall detta loggas. Däremot behöver inte Ineras spärrtjänst uppdateras med information om tillfällig öppning av spärr etc. eftersom de arkiverade journalerna lagras lokalt och kan läsas oavsett.

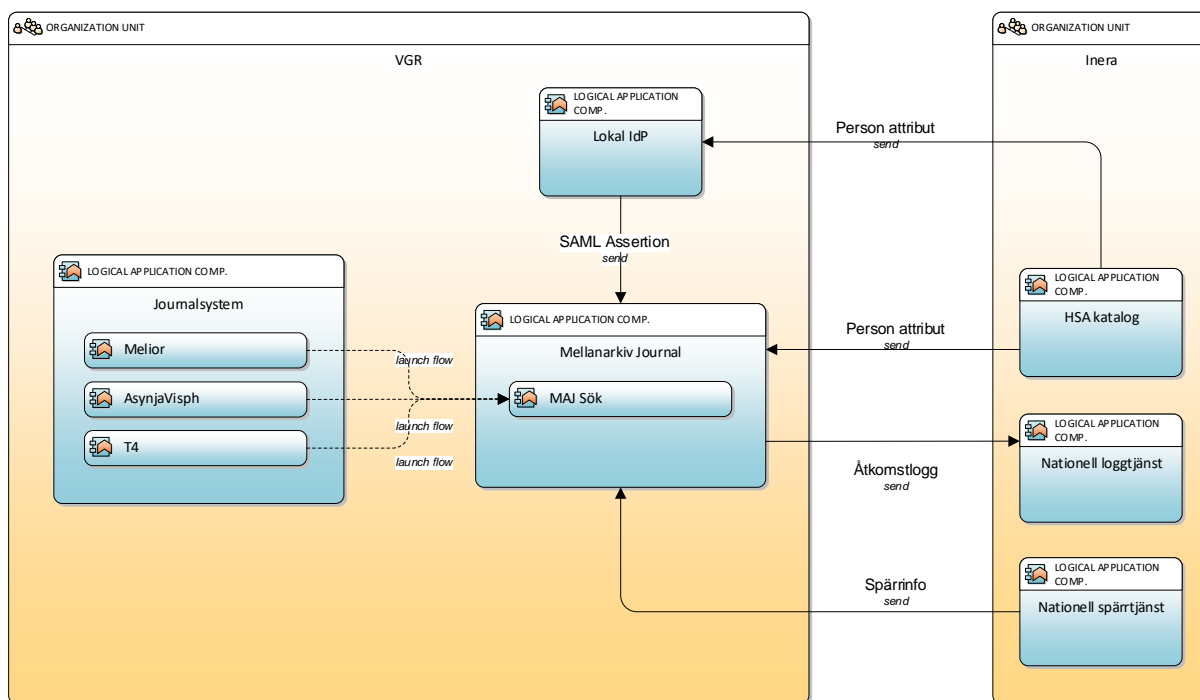


Diagram 3.7 - Informationsflöden (baseline)

Informationsöverföringarna redovisas även i tabellform nedan.

Från Applikation	Till Applikation	Informationsmängd	Beskrivning
Journalssystem	Mellanarkiv Journal	Patient ID	Kontext för sökning
Lokal IdP	Mellanarkiv (via webbläsare)	Åtkomstbiljett	SAML Assertion som överförs som HTTP cookie
HSA-katalog	Lokal IdP	Användarens identitet	För säker inloggning
HSA-katalog	Mellanarkiv Journal	Användar- och organisationsattribut	Används vid behörighetskontroll av åtkomst till arkiverad journalhandling
Mellanarkiv Journal	Nationell Loggtjänst	Åtkomstlogg	För att logga åtkomst och aktiviteter
Nationell spärrtjänst	Mellanarkiv Journal	Spärrinfo	Används vid behörighetskontroll av åtkomst till arkiverad journalhandling

Tabell 3.4 - Förteckning över informationsflöden (baseline)

### 3.3.4 Interaktioner

Utreds ej. För sekvensdiagram över inloggningsförfarande med SITHS eID hänvisas till bilaga [42].

### 3.3.5 Integrationsarkitektur

Integrationsarkitektur använder samma mönster för alla integrationer till nationella tjänster: anrop använder nationella tjänstekontrakt och går via RTjP och NTjP.

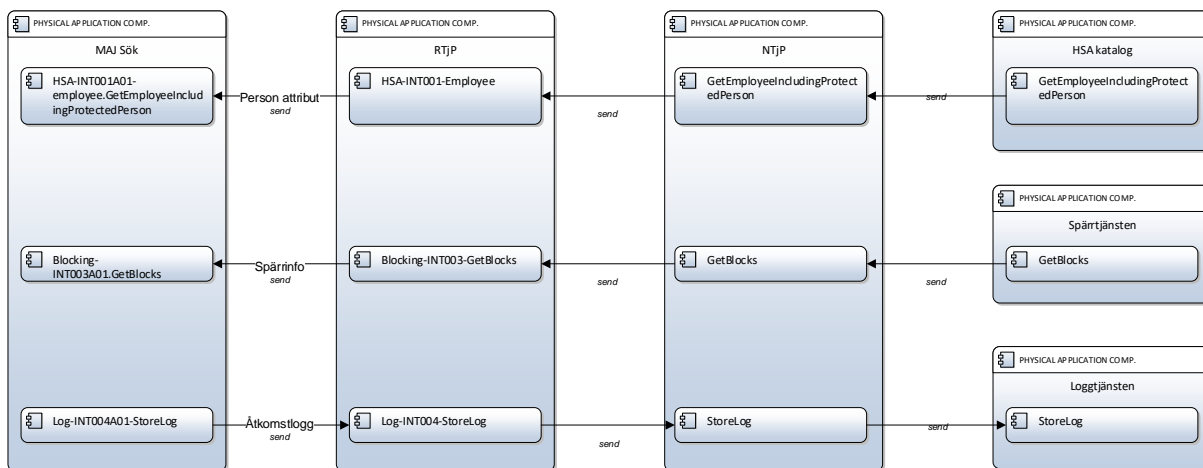


Diagram 3.8 - Integrationer (baseline)

### 3.4 Infrastruktur

#### 3.4.1 Teknisk plattform

Befintlig infrastruktur för Marshal OAIS och egenutvecklade MAJ applikationen omfattar Webbserver, SQL Server databas samt NAS lagring.

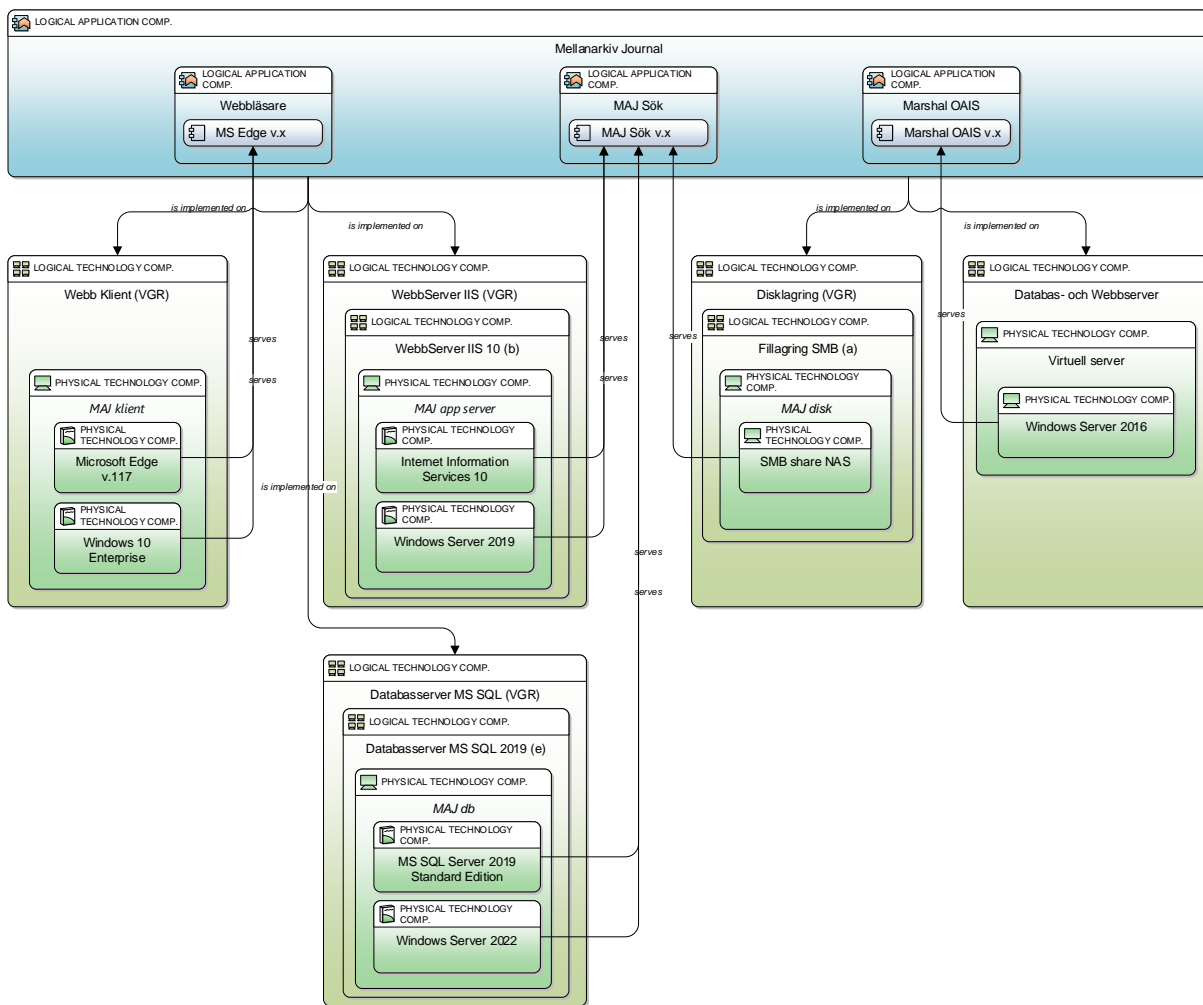


Diagram 3.9 - Processing diagram (baseline)



## 4 Beskrivning och utvärdering av lösningsalternativ

Detta avsnitt avser att klargöra vilka (inte beskriva i detalj) mål- och referensarkitekturer, lösningsmönster, standarder, produkter, teknik etc. som är relevanta för lösningsförslaget. Utgångspunkten är att använda de mål- och referensarkitekturer samt standarder som är antagna av VGR inom olika områden där så är tillämpligt.

### 4.1.1 Referensarkitekturer

Lösningalternativen avgränsas för närvarande till mellanarkivering av patientjournaler. Det finns ett stort antal referensarkitekturer och standarder som måste beaktas i lösningsarkitekturen. Dessa sammanfattas i det följande och slutsatser/observationer inramas i rött.

#### 4.1.1.1 OAIS referensmodell

Referensmodellen för ett öppet arkivinformationssystem (OAIS) utgör bilaga [23]. Modellen har utvecklats för att underlätta ett brett, disciplinoberoende, samförstånd om kraven på ett arkiv eller förvar för att tillhandahålla långsiktigt bevarande av digital information. Det var också avsett att stödja utvecklingen av ytterligare digitala arkivstandarder.

Ett OAIS är ett arkiv som består av en organisation av människor och system som har accepterat ansvaret för att bevara information och göra den tillgänglig för en specifik målgrupp. Standarden definierar en uppsättning ansvarsområden som ett OAIS-arkiv måste uppfylla och detta gör att ett OAIS-arkiv kan särskiljas från andra användningar av termen arkiv.

Sedan den antogs som en ISO-standard har OAIS-referensmodellen anammats i stor utsträckning av praktiskt taget alla typer av digitala arkiv. De flesta moderna digitala arkiv refererar till standarden. Den har också använts i stor utsträckning av organisationer för att informera om deras implementeringar av nya eller upgraderade arkiv.

OAIS beskriver tre aktörer:

- **Producer/producent**, som sammanställer och lämnar in information till ett OAIS för arkivering
- **Consumer/Konsument**, som begär att få ut arkiverad information från ett OAIS
- **Management/Ledning**, som förvaltar ett OAIS

En viktig roll som återfinns inom Management/Ledning är arkivarien, som arbetar med att utreda, vårda, ordna och förteckna arkivbestånd samt att ge råd och service till arkivets användare och besökare.

OAIS beskriver 6 funktionsmoduler (functional entities) och en generell informationsmodell. Funktionsmodulerna beskrivs kortfattat nedan; se [23] för en detaljerad beskrivning.:

- ***Ingest***: hanterar mottagning av Submission Information Packages (SIP) från producenter och säkerställer att informationen lagras och förvaltas i arkivet. I funktionen ingår kvalitetskontroll, sammanställning av ett Archival Information Package (AIP), extraktion av beskrivande information (metadata) till arkivdatabasen samt uppdateringar av *Archival Storage* och *Data Management* funktioner.
- ***Archival Storage***: hanterar lagring, förvaltning och hämtning av AIP:er. I funktionen ingår mottagning och persistent lagring av AIP:er, uppdatering av lagringshierarkin, hämtning av AIP:er från arkivdatabasen, felkontroll samt katastrofåterställning (recovery) av arkivdatabasen.
- ***Data Management***: hanterar beskrivande information för arkivinhåll och administrativ information som används för att förvalta arkivet.
- ***Administration***: hanterar den övergripande driften av arkivet. I funktionen ingår hantering av avtal med producenter, teknisk monitorering och optimering av arkivet, samt att inventera, rapportera och migrera innehållet i arkivet.
- ***Preservation Planning***: hanterar den långsiktiga arkivmiljön för att säkerställa att innehållet fortsätter att vara tillgänglig och kan förstås av organisationen som konsumerar arkiverad information.

- **Access:** hanterar information om existens, lokalisering, beskrivning och tillgänglighet av information som lagras i OAIS. Funktionen tillåter konsumenter att begära och ta emot arkiverade informationspaket med accesskontroll. Information levereras som en Distribution Information Package (DIP).

OAIS använder en generell informationsmodell för att hantera alla typer av (arkiverad) information. Huvudkonceptet är att ett informationsobjekt består av både ett dataobjekt (som antingen är digitalt eller fysiskt) och ett representationsinformationsobjekt, som gör att dataobjektet kan tolkas och förstås korrekt. Digitala objekt består av en bit-sekvens och representationsinformationsobjektet beskriver hur bitarna ska tolkas.

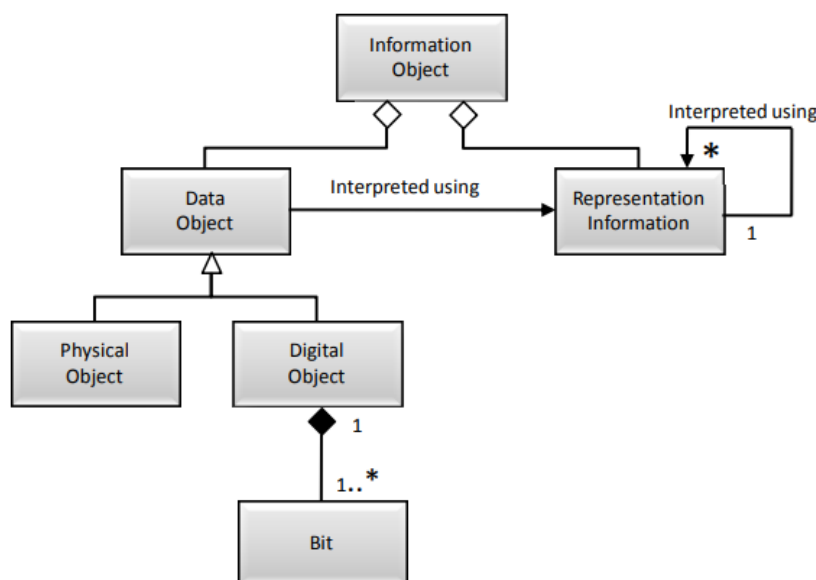


Diagram 4.1 - OAIS Informationsobjekt

Representationsinformationsobjektet beskriver med hjälp av strukturell information hur dataobjektets bit-sekvens ska tolkas, t.ex. som en sekvens av bokstäver som bildar ord i en text. Men oftast räcker inte detta - man måste t.ex. även veta vilket språk texten är skriven i, vilket beskrivs i semantisk information. Även annan information kan behövas för att beskriva dataobjektet; t.ex. mjukvara som krävs för att läsa en databasfil.

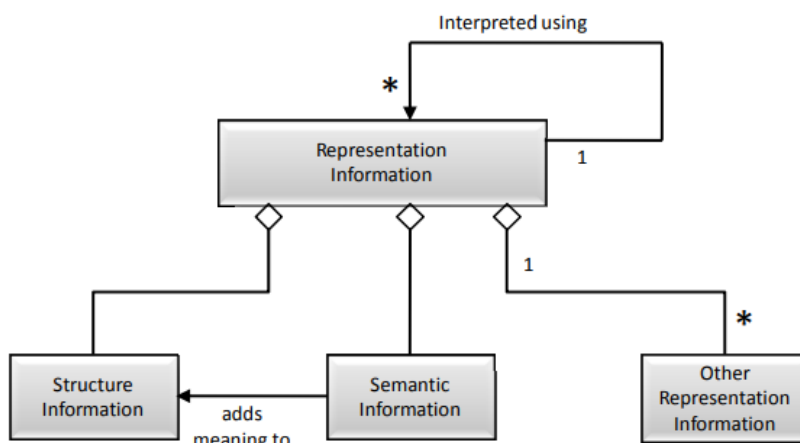


Diagram 4.2 - OAIS Representationsinformationsobjekt

Representationsinformationsobjektet är även en typ av informationsobjekt som kan behöva ett eget representationsinformationsobjekt för att tolkas - därav självreferensen i modellen.

Ett OAIS hanterar olika klasser av informationsobjekt, som alla representeras med ett dataobjekt och ett eller flera representationsinformationsobjekt. Taxonomin av informationsobjekt visas nedan.

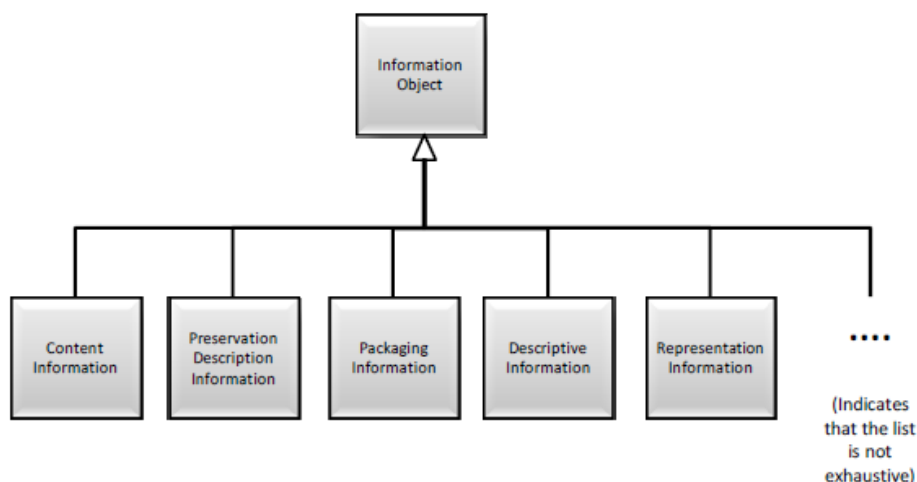


Diagram 4.3 - OAIS informationsobjektstaxonomi

Den första klassen, *Content Information*, motsvarar den huvudsakliga informationen som arkiveras. T.ex. i ett mellanarkiv för patientjournaler lagras patientjournaler som *Content Information*.

Utöver *Content Information* behöver ett AIP innehålla information som är underlag för bl.a. förtroende (trust), access och kontext under en oändlig tid. Den kollektiva benämningen på dessa informationsobjekt är *Preservation Description Information* (PDI). PDI måste innehålla all information som krävs för att bevara tillhörande *Content Information* i arkivet. Exempel på information som ingår i PDI visas nedan. Se [23] för en mer detaljerad beskrivning av PDI.

- *Reference Information* identifierar en eller flera mekanismer för att hantera unik identifiering av *Content Information*.
- *Context Information* beskriver hur *Content Information* relaterar till den omkringliggande informationsmiljön.
- *Provenance Information* är en typ av *Context Information* som beskriver informations ursprung och historik.
- *Fixity Information* hanterar dataintegritet t.ex. genom kryptoteknik för att säkerställa informationens äkthet.
- *Access Rights Information* beskriver hur åtkomst till informationen ska begränsas, t.ex. juridiskt, med licensavtal och via accesskontroll.

*Packaging Information* används för att koppla informationsobjektet till en specifik fysisk instans (en ZIP fil, ett XML dokument, etc.), och *Descriptive Information* används för indexering och effektiv åtkomst i arkivet. Se [23] för en mer detaljerad beskrivning.

Den generella/logiska informationsmodellen som beskrivs ovan realiserar i ett OAIS med hjälp av informationspaket. Ett informationspaket innehåller *Content Information* och PDI samt relationer till *Packaging-* och *Descriptive Information* enligt nedan.

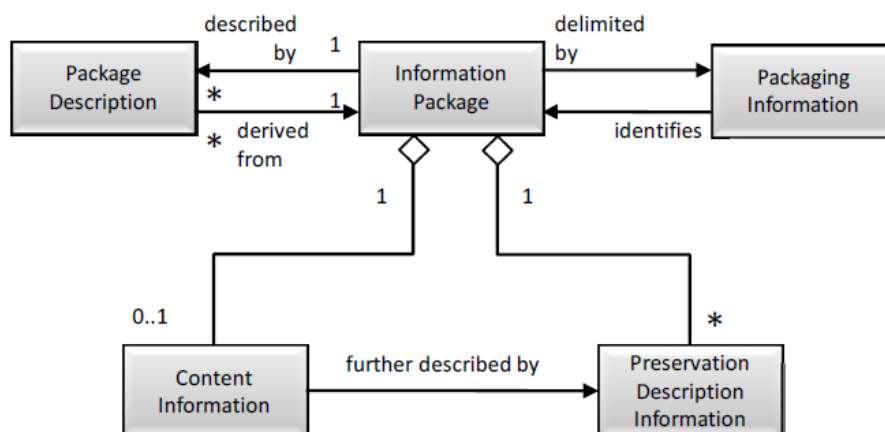


Diagram 4.4 - UML modell för OAIS informationspaket

Det finns tre typer av informationspaket: SIP, AIP och DIP. SIP och DIP är transienta medan AIP är det som bevaras i arkivet.

Slutsatser om OAIS som referensarkitektur: VGR:s mellanarkiv för vårddomänen ska vara ett OAIS. Det är en etablerad "best practice" och det finns många internationella och europeiska standarder som med fördel kan användas vid realisering av informationspaket m.m.

#### 4.1.1.2 Referensarkitektur för relationsdatabaser

Digitala patientjournaler använder som regel (i princip alltid) en relationsdatabas för att hantera data, och det är därför relevant att utreda hur data som lagras i relationsdatabaser kan mellanarkiveras. Temat "Bevarande av databaser" (database preservation) är inte ovanlig och har utretts av många parter internationellt. Innan vi kartlägger relevanta tekniker och lösningsmönster ges en kort beskrivning av en referensarkitektur för relationsdatabaser.

I en relationsdatabas lagras data i ett proprietärt format i datafiler och exponeras i ett semantiskt lager (SQL) i form av bl.a. tabeller med kolumner. Åtkomst till det semantiska lagret sker via både proprietära och standardiserade API (ODBC, JDBC), som realiserar en client-server arkitektur via ett eller flera enkapsulerade nätverksprotokoll som i praktiken begränsas till TCP/IP.

Det interna dataformatet i relationsdatabasen är proprietärt och skiljer sig från de flesta programmeringsspråk. T.ex. lagras ofta både hel- och flyttal på samma sätt med bibehållen binär sortering och variabel precision/längd. Ofta krävs en konvertering i det semantiska lagret för att leverera data i ett format som kan hanteras av klienten.

En relationsdatabas måste ha s.k. ACID egenskaper (atomicity, consistency, isolation, durability), vilket kan realiserar på olika sätt, men oftast med bl.a. transaktionskoncept, låsningsmekanismer och flera samtidiga versioner av data. Implementationen kräver stöd i både det semantiska- och datalagret.

Relationsdatabaser har i grunden en klient-server arkitektur, men den kan även realiserar i en N-tier arkitektur där databasklienten agerar applikationsserver för flera klienter. Serverarkitekturen varierar från olika leverantörer, men ofta delas den upp i flera serverprocesser/trådar för att t.ex. underlätta asynkrona skrivningar av data medan skrivningar till transaktionsloggen alltid sker synkront.

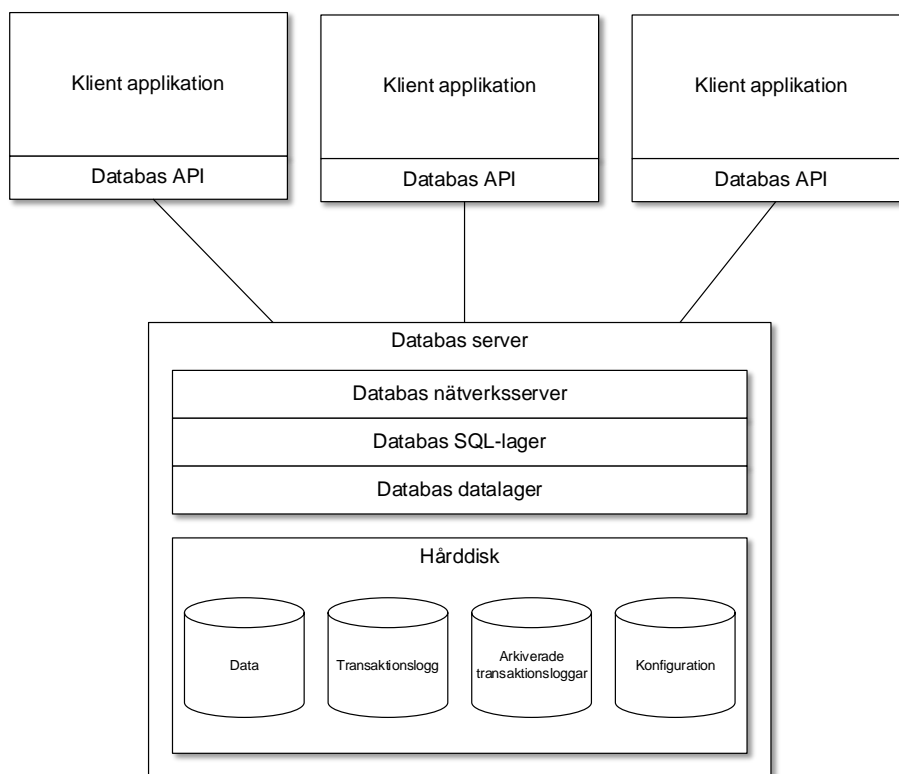


Diagram 4.5 - Referensarkitektur för en relationsdatabas

Slutsatser om referensarkitektur för relationsdatabaser: en relationsdatabas är inte i första hand en fysisk lagring (i databasfiler) utan en samling applikationer som realiserar funktionalitet som karakteriserar en relationsdatabas. Det finns standarder inom databasvärlden (SQL språket och vissa API), men stora delar är fortfarande proprietära.

Det bedöms orimligt att anta att ett relationsdatabassystem kan hållas i drift under en längre tid (decennier). Detta måste hanteras på strategisk nivå; t.ex. genom att etablera ett förvaltningsteam inkl. R&D kompetens för att säkerställa långsiktighet i en "VGR relationsdatabas" baserad på öppen källkod.

#### 4.1.1.3 Bevarande av databaser

Nederländernas regering och riksarkiv m.fl. bildade för drygt 20 år sedan en forskningsgrupp kallad "The Digital Preservation Testbed" (*Testbed* fortsättningsvis) med uppdrag att utreda hur man rent praktiskt bevarar digital information och samtidigt tillhandahåller digital åtkomst i framtiden. Testbed har utrett hur man kan bevara olika typer av information, och i bilaga [21] "Preserving Databases" har de utrett hur man kan bevara information som lagras i databaser.

Först definieras begreppet digital informationshandling (digital record) som en kombination av hårdvara, mjukvara och datafil:

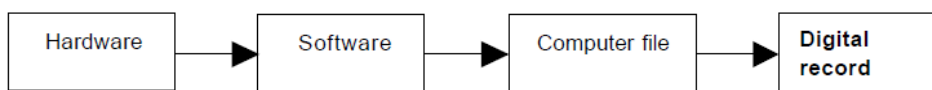


Diagram 4.6 - Definition av digital handling

Den digitala informationshandlingen skapas genom att kombinera specifik hårdvara och mjukvara och lagras sedan kodat på sett specifikt sätt i en datafil. Specifika applikationer kan läsa och förstå datafilen för att kunna rendera/visualisera den digitala handlingen på en dataskärm (eller bearbeta den på annat sätt). Oftast kan en digital informationshandling endast tolkas korrekt genom att använda denna specifika kombination av hårdvara, mjukvara och datafil. T.ex. om den digitala handlingen reproduceras i en annan datormiljö kan den se ut och beté sig på ett helt annat sätt än i den ursprungliga miljön.

Testbed identifierar *autenticitet* som ett nyckelkoncept vid bevarande av digitala informationshandlingar:

- *Integritet* behövs för att säkerställa att informationshandlingen är intakt och inte har ändrats eller korrumpierats på ett sätt som kan leda till en felaktig tolkning av innehållet.
- *Verifiering* (eller *autentisering*) behövs för att säkerställa att den digitala informationshandlingen verkligen är det som den utgör sig för att vara.

Vid bevarande av databaser krävs tekniker som tillåter kontroll av integritet och verifiering av innehållet av de arkiverade digitala informationshandlingarna.

Vid bevarande av pappershandlingar identifierar Testbed fyra karakteristiker som tillsammans bidrar till handlingens autenticitet: kontext, innehåll, struktur och presentation (appearance). Även digitala informationshandlingar har dessa karakteristiker, men de är lösare kopplade och är starkt beroende av hur mjukvaran tolkar den digitala informationshandlingen. Dessutom finns en femte karakteristik för digitala informationshandlingar: *betéende*.

- *Kontext* är den ursprungliga miljön där informationshandlingen användes. För att kunna tolka en arkiverad digital informationshandling krävs viss information om det ursprungliga kontextet.
- *Innehåll* är information som faktiskt lagras i en databas.
- *Struktur* beskriver den logiska hierarkin och inbördes relationer mellan den digitala informationshandlingen olika delar. Strukturen realiserar med hjälp av standardtekniker i relationsdatabaser, och om detta går delvist förlorat vid arkivering/migration kan det resultera i en databas som inte längre kan presentera information på rätt sätt.
- *Presentation* beskriver hur en digital informationshandling presenteras på en dataskärm. Här ingår t.ex. typsnitt, punktstorlek samt bruk av understrykning, fetstil, osv. I en databas är presentationen huvudsakligen knutet till applikationer med GUI som hanterar informationen.
- *Betéende* beskriver den interaktiva karakteristiken av en digital informationshandling. I en databas är beteendet huvudsakligen knutet till applikationen som hanterar informationen.

Den individuella vikten av de olika karakteristikerna av en digital informationshandling har ett starkt beroende till vilka verksamhetsprocesser som använder informationen. För att bevara information kopplat till de olika karakteristikerna av en digital informationshandling vid arkivering, samt för att säkerställa att autenticiteten bevaras, används *metadata*.

Testbed beskriver tre tänkbara strategier, var och en med olika variationer, för att bevara databaser:

1. Migrering av digitala informationshandlingar från en hårdvaru- och mjukvarumiljö till en annan
2. Konvertering av digitala informationshandlingar till XML
3. Emulering av hårdvaru- och mjukvarumiljö

I korthet innebär migrering att man flyttar hanteringen av informationshandlingar från en specifik kombination av applikation, databas, operativsystem och hårdvara till en ny kombination. Ibland kan flytten vara relativt smärtfri (när nya versioner är 100% bakåtkompatibla), men ibland kan det krävas mer eller mindre handpåläggning inklusive kodändringar i applikationer för att uppnå en lyckad migrering. Beroende på mängden applikationskod inkl. beroenden till öppen källkod måste man räkna med frekventa migreringar - oftast flera gånger om året.

XML innebär att man konverterar informationshandlingar från en databas till (ett) XML dokument. XML är en öppen standard som förvaltas av Worldwide Web Consortium (W3C) och innebär att i princip all information lagras i ett för människor läsbart textformat. XML dokumentet tar över databasens ansvar för innehåll och struktur, men saknar i viss mån databasens förmåga att hantera t.ex. referentiell integritet i strukturen.

Emulering innebär att man bevarar hela systemstacken bestående av hårdvara, operativsystem, databas och applikation(er) och använder emuleringsteknologi för att säkerställa att systemet fortsätter att fungera även långt efter att ingående systemkomponenter fortsätter att fungera.

Fördelar och nackdelar med dessa tre strategier sammanfattas nedan.

Strategi	Fördelar	Nackdelar
Migrering	<ul style="list-style-type: none"> <li>• Effektiv arkivering av data som lagras i databas</li> <li>• Hög prestanda</li> <li>• Fortsätt bra stöd för operationell data</li> <li>• Tillämpad harmoniserad informationsmodell per informationsdomän</li> <li>• Relativt stort arbete att genomföra en migrering</li> <li>• SQL</li> <li>• Skalbart</li> </ul>	<ul style="list-style-type: none"> <li>• Kräver hög kompetens inom databasteknologi</li> <li>• Inte anpassat för arkivering</li> <li>• Måste migrera/skriva om applikationer</li> <li>• Affärslogik i applikationslagret kan behöva migreras</li> </ul>
XML	<ul style="list-style-type: none"> <li>• Öppen standard</li> <li>• Kan tillämpa harmoniserad informationsmodell per informationsdomän</li> <li>• Kan läsas av människor</li> <li>• Bra för dokumenthantering</li> <li>• Bra som export/exchange format</li> </ul>	<ul style="list-style-type: none"> <li>• Svårt att hantera stora datasamlingar</li> <li>• Inte effektivt för hantering av stora samlingar operationell data</li> <li>• Begränsat stöd för relationer</li> </ul>
Emulering	<ul style="list-style-type: none"> <li>• Kan arkivera hela systemet inkl. applikationer</li> <li>• Kan vara enkelt och snabbt att realisera</li> </ul>	<ul style="list-style-type: none"> <li>• Inga kostnadsbesparingar för SW licenser</li> <li>• Kan innebära stora säkerhetsrisker för system som har nått end-of-life</li> <li>• Inte anpassat för arkivering</li> </ul>

Tabell 4.1 - Strategier för bevarande av databaser

Slutsatser om bevarande av databaser är följande:

- Migrering är en lämplig strategi för mellanlagring av data, som behöver läsas i en operationell kontext efter att ett system avvecklas. Migrering är inte en lämplig strategi för att realisera ett OAIS.
- XML är en lämplig strategi för mellanarkivering av data enligt OAIS standarden. OAIS informationspaket kan lämpligtvis implementeras med XML. Mellanarkiverade data kan fortfarande läsas operationellt (t.ex. via REST API), men det kan finnas praktiska begränsningar och signifikant sämre prestanda vid aggregerade bearbetning (t.ex. statistisk analys).
- Emulering är varken lämpligt för mellanlagring eller mellanarkivering av data, men kan vara en kort- till mellansiktig lösning när applikationer måste bevaras.

#### 4.1.1.4 SIARD

SIARD är en förkortning av "Software Independent Archiving of Relational Databases". Det är en specifikation som togs fram av schweiziska riksarkivet 2007. SIARD anger format för ett SIARD arkiv, som är en ZIP-fil innehållande all data (en XML-fil per tabell) och metadata (en XML-fil som beskriver alla tabeller) som har exporterats från en databas.

Ett SIARD 1.0 arkiv utgör inte en operationell databas, utan är ett exportformat som kan användas för att återskapa den arkiverade databasen genom import till en ny databas som stödjer SQL:1999 standarden.

SIARD 1.0 har uppgraderats till SIARD 2.0 med bl.a. följande förbättringar:

- Stöd för SQL:2008 inkl. user-defined types (UDT)
- Mer explicita valideringsregler för datatyper
- Stöd för lagring av externa objekt utanför SIARD arkivet (t.ex. BLOBS)

*SIARD saknar stöd för transformering till en ny informationsmodell. I standard ETL-mönstret för arkivering är det alltså endast E(xtract) och L(oad) som erbjuds.*

Slutsatser om SIARD är att det är ett verktyg som kan skapa en komplett databasexportfil, som sedan (åtminstone i teorin) kan användas för att återskapa en kopia av den ursprungliga databasen. Ett problem med SIARD är att SQL standarden har reviderats 4 gånger sedan SQL:2008 och att stöd för databastriggers, lagrade procedurer och funktioner saknas (i SQL standarden).

Som ett alternativ har samtliga databaser egna proprietära import/export funktioner som kan användas för att skapa exakta och fullständiga kopior av databasen i fråga, men det fungerar endast när den exporterande och den importerande databasen kan använda samma verktyg; i praktiken att de är från samma leverantör.

#### 4.1.1.5 Database Preservation Toolkit

År 2006 startade Portugals riksarkiv RODA projektet (Repository of Authentic Digital Objects) för att arkivera informationshandlingar tillhörande portugisiska myndigheter. I projektet utvecklades ett verktyg kallad Database Preservation Toolkit (DPT), som automatiskt exporterar en databas till SIARD format.

Genom att använda en connector-arkitektur kan DPT användas för att automatiskt flytta data mellan olika leverantörers databaser.

*Liksom SIARD saknar DPT stöd för transformering till en ny informationsmodell. I standard ETL-mönstret för arkivering är det alltså endast E(xtract) och L(oad) som erbjuds.*

Slutsatser om DPT är att det är ett verktyg som använder SIARD för att skapa en komplett databasexportfil, som sedan (åtminstone i teorin) kan användas för att återskapa en kopia av den ursprungliga databasen. DPT har samma problem och alternativ som beskrivs för SIARD ovan.

#### 4.1.1.6 E-ARK

E-ARK är en förkortning av ”European Archival Records and Knowledge Preservation” och var ett EU-finansierat multinationellt projekt under tiden 2014-2017, som hade som mål att utveckla tekniska specifikationer och verktyg för en integrerad arkiveringsinfrastruktur.

E-ARK projektet publicerade följande (översätts ej):

- Guidelines on pan-European e-archiving as part of EC e-infrastructure
- Open archival products (tools, services, framework, metadata specifications)
- Open technical products (tools, services, metadata specifications)
- Open operational products (ingest and access tools, services, metadata specifications)
- Open access tools, services, metadata specifications, including data mining tools for business intelligence
- Open interfaces from tools, services, metadata specifications to existing systems products
- Outcomes of a legal study
- Outcomes of pilots, especially where similar archival material to that under consideration was processed
- Project papers on the integration work undertaken

Publicerade dokument återfinns på E-ARKs hemsida [22].

Av särskilt intresse för VGR finns E-ARK SIP standarden som är en europeisk standard för inlämningsformat till OAIS arkiv.

Slutsatser om E-ARK är att det är relevant för ett OAIS där man vill använda pan-europeiska standarder för t.ex. OAIS informationspaketformat. Området är stort med många specifikationer som inte har utretts i detalj. Här behövs ett principbeslut om VGR ska följa E-ARK.

#### 4.1.1.7 E-Ark CSIP

E-Ark CSIP är EU standard som beskriver den fysiska strukturen för alla typer av informationspaket som hanteras i ett (OAIS) e-arkiv [31]. CSIP beskriver gemensamma aspekt, medan specialiserade informationspaket beskrivs i egna specifikation för SIP [32], AIP [33] och DIP [34]. Det är små, detaljerade variationer på låg nivå (hur metadata beskrivs i XML) och för omfattande att utreda här.

E-Ark CSIP beskriver även i detaljkrav för metadatahantering i informationspaket, i synnerhet nyttjande av METS och PREMIS standarderna.

Syftet med CSIP är att uppnå en grad av interoperabilitet mellan olika sorters informationspaket från olika arkiv, som gör det möjligt för standardverktyg som utvecklas för hantering av CSIP paket att användas utan ytterligare anpassningar. Detta uppnås genom:

- Etablering av gemensamma krav för interoperabel hantering av informationspaket
- Etablering av en gemensam bas för utveckling av specialiserade informationspaket inom e-arkiv disciplinen
- En detaljerad XML-baserad implementation baserad på etablerade standarder för bevarande av information i e-arkiv

Målet är att använda E-Ark CSIP oavsett informationstyp och sedan använda en domänspecifik E-Ark Content Information Type Specification (CITS) för att kunna hantera krav på validering, bevarande och förvaltning för specifika informationstyper. En relevant CITS för det här projektet är *eHealth1* som beskrivs i nästa kapitel.

E-Ark CSIP specifikationen [31] innehåller följande kapitel:

1. Introduktion, bakgrund, syfte, mål
2. Behovet av CSIP
3. Principer för interoperabla informationspaket
  - a. Generella principer
  - b. Identifiering av informationspaket
  - c. Struktur av informationspaket
  - d. Metadata för informationspaket
4. CSIP struktur enligt principerna
5. Nyttjande av metadata
6. Implementationsövervägande

Principerna är logiska och okontroversiella. T.ex. att varje informationspaket måste tydligt identifieras som SIP, AIP eller DIP, och att CITS för data och metadata är tydligt angivet. En annan princip för hantering av metadata är att man måste följa en standard. I nuläget är de relevanta standarderna för metadatahantering METS [35] och PREMIS [36] (se även kap. 4.1.3 Standarder).

Standarden lämnar utrymme för implementatören att välja olika CSIP strukturer: både minimalistiska och mer sofistikerade. Överväganden och val görs i målarkitekturen kap. 5.2 Informationsarkitektur.

Slutsatser om E-ARK CSIP inkl. SIP, AIP och DIP är att det är ytterst relevanta standarder som bör följas vid etablering av ett mellanarkiv. Standarden beskriver både den gemensamma och den paketspecifika fysiska strukturen för lagring av data och metadata för alla sorters informationspaket i arkivet.

#### 4.1.1.8 E-Ark eHealth1

E-Ark eHealth1 är en E-Ark CSIP Content Information Type Specification (CITS) för arkivering av patientjournaler [30]. Syftet med E-Ark CSIP är att ge teknisk vägledning i hur data och metadata ska formateras och placeras i ett CSIP informationspaket för att åstadkomma interoperabilitet vid informationsutbyte. Syftet med eHealth1 CITS är att ge teknisk vägledning i hur data och metadata ska paketeras för digitala patientjournaler.

Innehållet i ett eHealth1 arkivinformationspaket är skiktat:

- CSIP utgör det yttersta generiska lagret
- CSIP innehållet är paketerat som SIP, AIP eller DIP
- SIP/AIP/DIP innehållet är av en viss CITS, i det här fallet eHealth1
- eHealth1 strukturen kan sedan anpassas lokalt (av VGR)

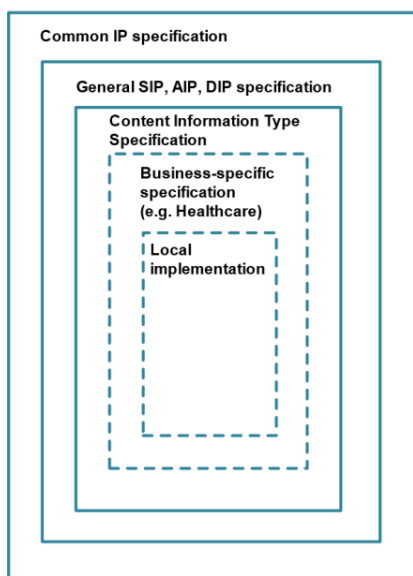


Diagram 4.7 - eHealth1 datamodell

Arkivpaket som följer eHealth1 innehåller digitaliserade patientjournaler i form av innehåll (data) och beskrivning av innehåll (metadata). Vid inleverans till arkivet används ett submission information package (SIP) innehållande patientjournalavsnitt för en eller flera patienter med tillhörande metadata.

CITS eHealth1 förespråkar att HL7 FHIR (resurser) används för att beskriva patientens person- och kliniska data på ett standardiserat sätt för interoperabilitet, men det är tillåtet att använda egna lokala implementationer så länge de är dokumenterade med hänvisning i SIP paketets metadata (Submission Agreement).

Som exempel på FHIR resurs visas UML-modell för Patient nedan:

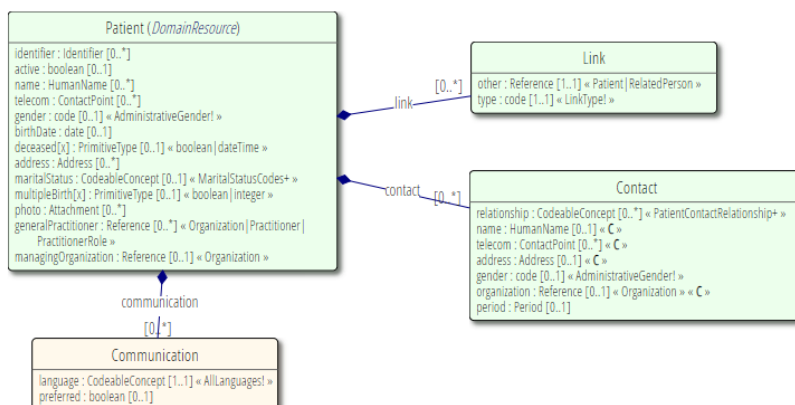


Diagram 4.8 - FHIR resurs: Patient

En fördel med att använda HL7 FHIR är att man använder standardiserade kodverk som t.ex. ICD för att beskriva sjukdomar och SNOMED CT för kliniska termer. Detta förutsätter förstås att källsystemet använder kodverken.

I CITS eHealth1 kan ett journalavsnitt innehålla ett eller flera dokument med gemensam metadata och inbördes länkar. T.ex. ett PDF dokument med attachment, eller ett dokument med separat digital signering.

Ett journalavsnitt kan även bestå av bild/video data och representeras ofta i DICOM formatet. Även andra format kan användas sålänge de kan identifieras med en MIME-typ.

Vid inleverans till arkivet antas journalavsnitt från flera patienter kunna aggregeras i ett SIP paket, även om det inte nödvändigtvis är så. Antalet patienter som ingår i varje AIP är ett implementationsbeslut, men det norska Arkiverket har beslutat att varje AIP ska innehålla information om en patient från en arkiverande myndighet/aktör. Dvs. om det t.ex. finns journalhandlingar för en patient som har arkiverats av 5 olika sjukhus, finns detta arkiverat i 5 olika AIP.

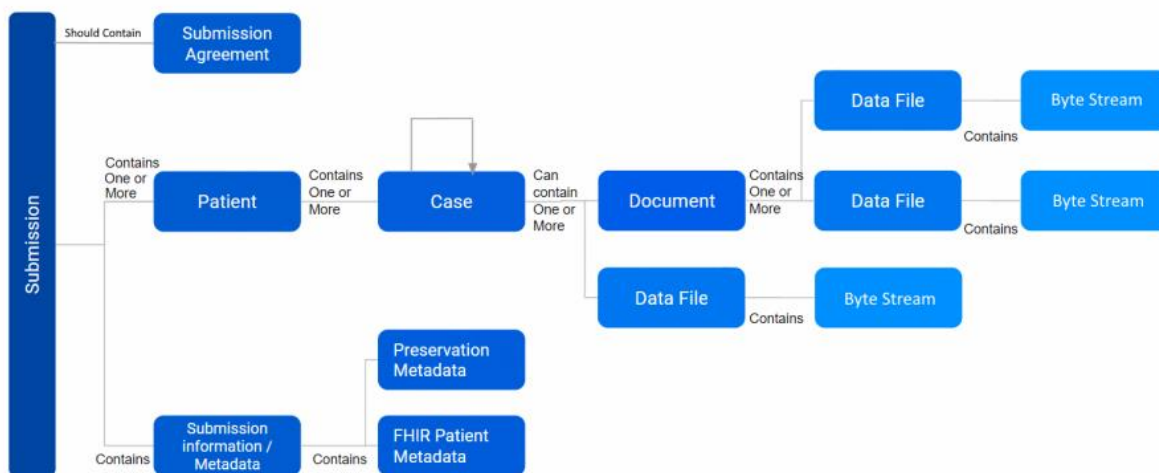


Diagram 4.9 - eHealth1 SIP datamodell

Det finns ytterligare val för hur information struktureras i arkivet. I CITS specifikationen [30] ges tre exempel på hur patientjournaler kan lagras i ett arkiv:

I det första exemplet innehåller ett informationspaket ett komplett utdrag för en patient från ett journalsystem i form av t.ex. ett PDF dokument. I det här fallet innehåller ett informationspaket en patient, ett case, ett dokument och en datafil.

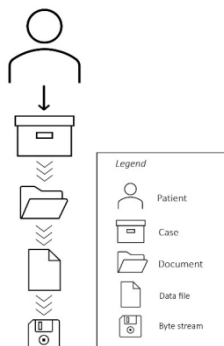


Diagram 4.10 - Arkiverad patientjournal som en fil

I det andra exemplet innehåller ett informationspaket flera utdrag för en patient motsvarande hur informationen organiserades tematiskt i journalsystemet; dvs. som journalhandlingar. Varje informationspaket innehåller en patient, ett eller flera case, ett dokument per case och en datafil per dokument.

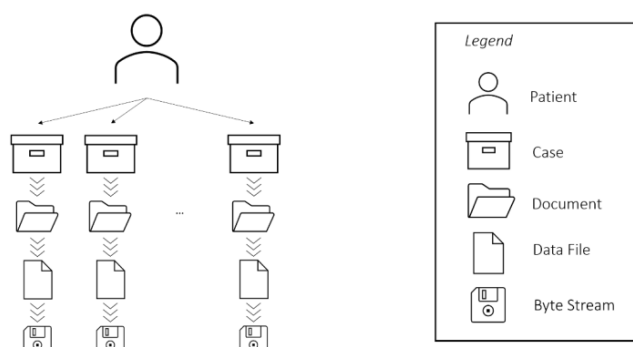


Diagram 4.11 - Arkiverad patientjournal som tematiska filer

I det tredje exemplet innehåller ett informationspaket liksom exempel två flera case/handlingar, men varje handling kan bestå av flera dokument (ev. med inbördes länkar), som i sin tur kan innehålla flera datafiler i olika mediaformat (ljud, bild, etc.).

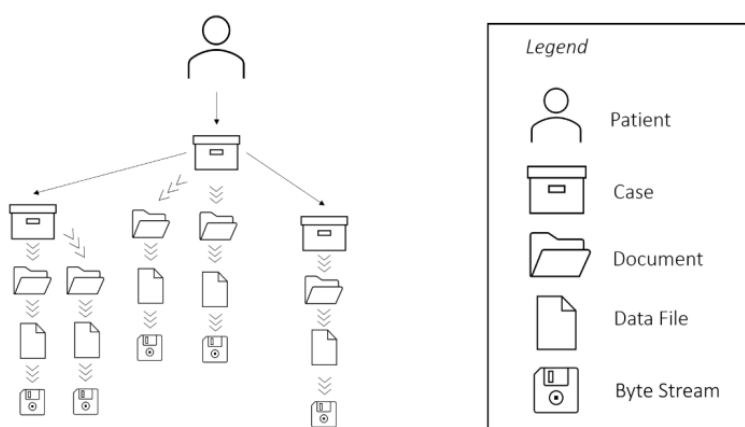


Diagram 4.12 - Arkiverad patientjournal som en samling dokument per case eller sub-case

Utöver den arkiverade informationen som lagras i en eller flera *representationsmappar*, innehåller ett informationspaket metadata i form av METS-, PREMIS och EAD filer, samt strukturerad data om den arkiverade information som beskrivande metadata. Se CITS specifikationen [30] för en detaljerad beskrivning.

Slutsatser om CITS eHealth1 är att det är en lämplig standard för VGR att följa vid arkivering av patientjournaler. Journalhandlingar kan arkiveras på olika sätt, men huvudsakligen som ostrukturerad data. Det strukturerade datat som man vill bevara för att möjliggöra t.ex. sökning/filtrering och API exponering lagras i en separat XML fil i arkivpaketet som beskrivande metadata; t.ex. i form av FHIR resurser.

#### 4.1.1.9 E-Ark eHealth1 rev 2.0

Under våren 2024 har Dilcis (förvaltare av E-Ark standarder) reviderat eHealth1 standarden och har samlat in feedback fram till mitten av april. Draft specifikationen bifogas som bilaga [43]. Ny standard har i skrivande stund inte publicerats. Det som är ändrat i standarden sammanfattas nedan.

#### Större ändringar

- Ändring i hur patient records aggregeras i representationer
- Uppdaterad representations METS profil

### Mindre ändringar

- Borttagning av dubblerade krav för CSIP och paketspecifikationer
- Mindre rättelser i METS krav kardinalitet och nivåer
- Rättade exempel
- Uppdaterade kravformat för att matcha CSIP
- Uppdaterad METS exempel för att matcha CSIP
- Uppdaterad standardtext för att vara i linje med CSIP
- Nya diagram
- Uppdaterad ordlista
- Nya generella krav
- Nya principer
- Ny förklaring för användningsfallet där patientjournaler arkiveras från en databas med hjälp av SIARD
- Ytterligare förklaringar kring hur patientjournaler kan aggregeras i inleveranspaket av arkivbildare
- Kontextuell och förklarande innehåll har flyttats till separat Guideline dokument
- Nytt exempel i Guideline dokumentet om piql E-Ark3 eHealth1 SIP verktyg
- Motivering av principer flyttas till Guideline dokumentet
- Uppdaterad rot METS profil

I EAD3 [37] används en generell aggregeringsstruktur för arkiv bestående av *class*, *collection*, *file*, *fonds*, *item*, *otherlevel*, *recordgrp*, *series*, *subfonds*, *subgrp* och *subseries*. Denna struktur anser eHealth1 inte är lämplig för patientjournaler. I den nya eHealth1 standarden [43] ändras aggregeringsstrukturen i arkivpaketen till följande nivåer:

- Patient är en individ som har mottagit hälsovård
- Case är en journaldel
- Sub-case är en tillåten men inte obligatorisk indelning av journaldelar i underdelar
- Document är en eller flera inbördes relaterade datafiler med gemensam metadata
- Data-file är en komponent med en MIME-typ
- Byte-stream är en komponent med en MIME-typ. Flera byte-streams kan lagras tillsammans i en data-file som i så fall är ett containerformat som t.ex. MP4, DICOM eller Matroska.

Strukturen med inbördes kardinalitet i *representationsmappen* visas nedan. Detta är den ena stora ändringen i eHealth1.

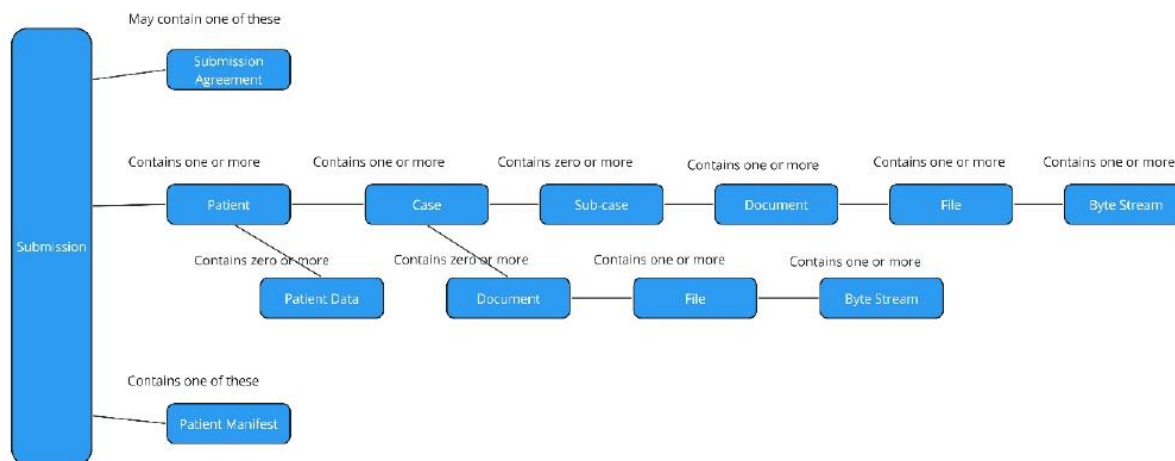


Diagram 4.13 - eHealth1 SIP struktur med kardinalitet

Journaldelar tillhörande en patient arkiveras i en eller flera datafiler enligt denna struktur. På översta nivå finns en rot-METS fil som beskriver innehållet i SIP paketet samt fyra mappar med innehåll:

- metadata innehåller administrativ metadata och metadata för bevarande. Det administrativa metadatat innehåller åtminstone ett manifest över alla patienter vars journaldelar finns i representationsmappen.
- representation innehåller en enda representation som följer aggregeringstrukturen ovan med patient/case/sub-case/data-file/byte-stream. På översta nivå finns enligt CSIP en METS metadatafil som beskriver representationen. För varje patient kan det finnas strukturerad data i form av FHIR resurser (eller liknande).
- schemas innehåller XML schemadefinition för validering av XML filer i arkivpaketet. För bevarande metadata används PREMIS och därför inkluderas premis.xsd. I exemplet nedan bevaras två FHIR resurser per patient, *FHIR patient* och *FHIR condition*. Därför inkluderas även XSD filer för dessa.
- documentation innehåller Submissionagreement.pdf, som är avtalet mellan arkivbildare (som skickar in SIP paket) och arkivförvaltare (som tar emot SIP paket) och bör bl.a. beskriva exakt vad som måste ingå i ett SIP paket.

Exemplet nedan är saxat från [43] och är tyvärr av dålig kvalitet. En bättre version förväntas när eHealth1 rev 2.0 publiceras.

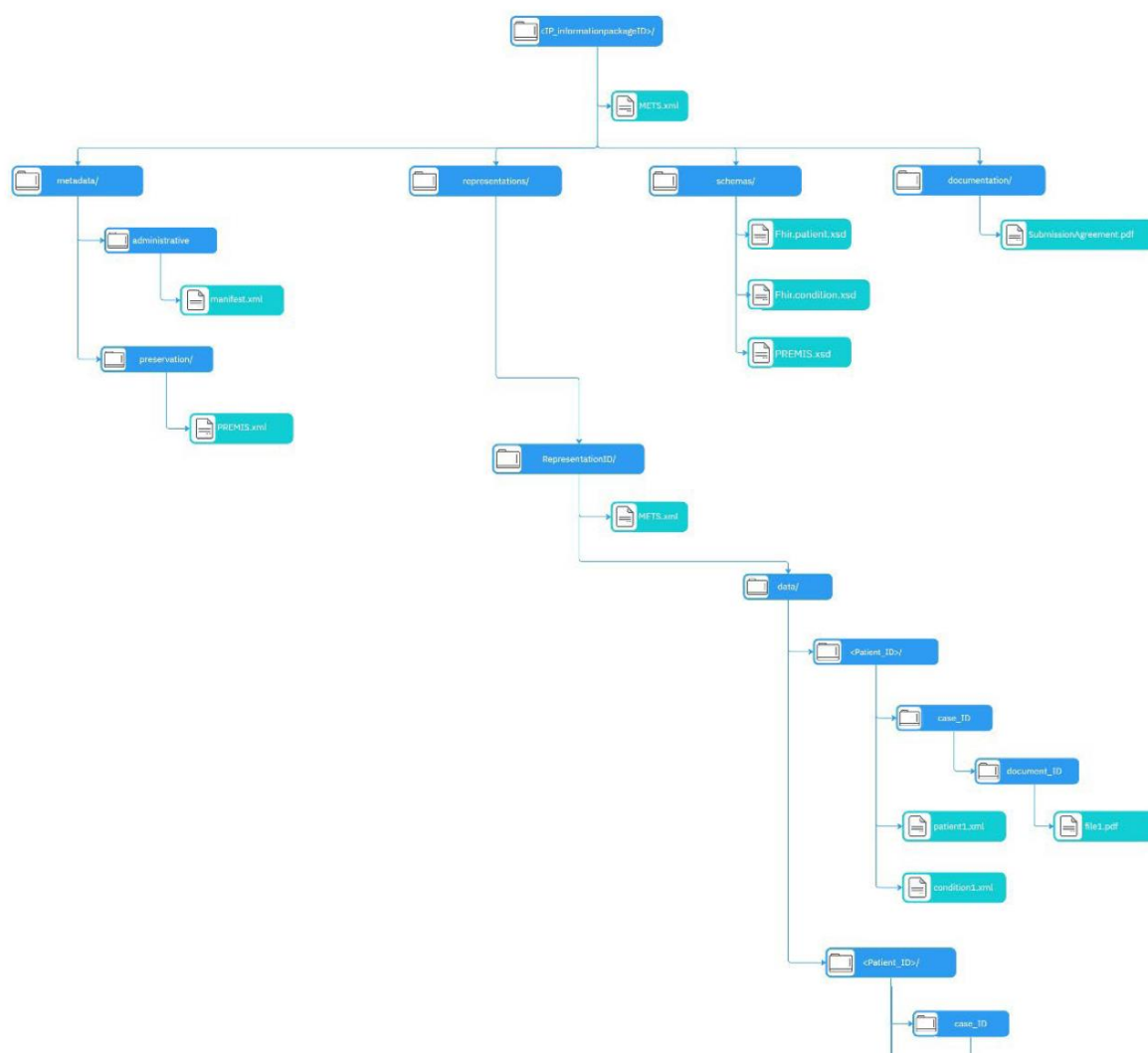


Diagram 4.14 - Exempel på eHealth1 SIP struktur

Den andra stora ändringen i eHealth1 är innehållet i METS filen i representationsmappen. Placering och innehåll följer nu E-Ark CSIP specifikationen [31]. För en detaljerad beskrivning se bilaga [43].

Slutsatser om CITS eHealth1 rev 2.0 är att det fortfarande är en lämplig standard för VGR att följa vid arkivering av patientjournaler. Journalhandlingar kan arkiveras på olika sätt, men huvudsakligen som ostrukturerad data. Det strukturerade datat som man vill bevara för att möjliggöra t.ex. sökning/filtrering och API exponering lagras i en eller flera separata XML filer i arkivpaketet som beskrivande metadata; t.ex. i form av FHIR resurser.

#### 4.1.1.10 HL7 FHIR resurser

Standarden Fast Healthcare Interoperability Resources (FHIR, uttalas "fire") är en uppsättning regler och specifikationer för utbyte av elektronisk sjukvårdsdata. Den är utformad för att vara flexibel och anpassningsbar, så att den kan användas i ett brett spektrum av miljöer och med olika IT-system inom sjukvården. Målet med FHIR är att möjliggöra ett sömlöst och säkert utbyte av hälso- och sjukvårdsinformation, så att patienter kan få bästa möjliga vård. Standarden beskriver dataformat och element (så kallade "resurser") och ett applikationsprogrammeringsgränssnitt (API) för utbyte av elektroniska hälsojournaler (EHR). Standarden skapades av Health Level Seven International (HL7) hälso- och sjukvårdsstandardorganisation.

Det finns 157 resurser definierade i FHIR [38], som grupperas i moduler; bl.a. *Clinical Module*, *Diagnostics Module*, och *Medications Module*. E-Ark eHealth1 standarden identifierar *Clinical Module* [39] som lämplig modell för att spara strukturerad metadata i ett arkiv. För varje resurs finns olika representationer; bl.a. en UML modell. Nedan visas *AllergyIntolerance* resursen som exempel:

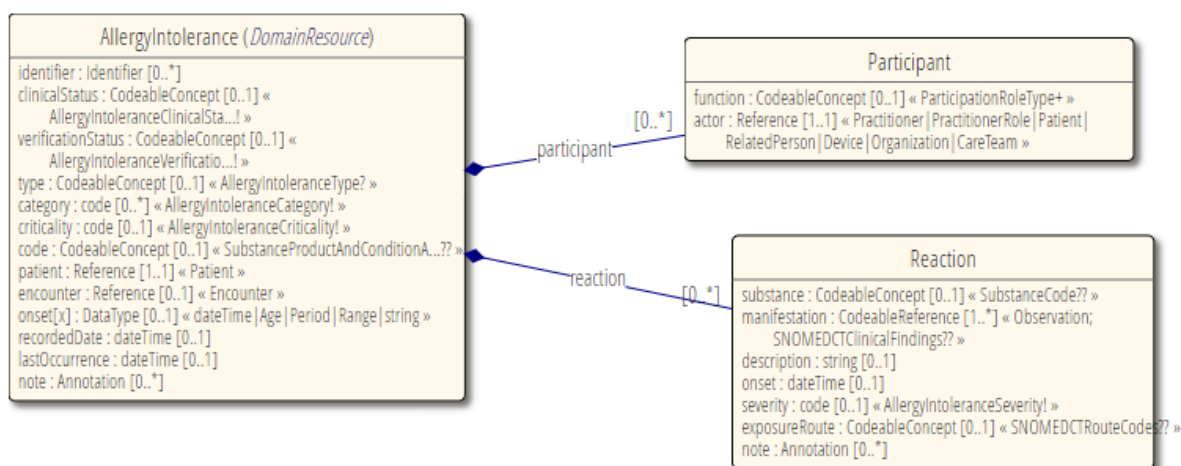


Diagram 4.15 - FHIR resurs *AllergyIntolerance*

Slutsatser om FHIR resurser som val för beskrivande metadata i mellanarkivet är att det är en ytterst lämplig standard att följa. Det finns två saker som måste beaktas vid val av en eller flera FHIR resurser som metadata i mellanarkivet: för det första om underlaget finns tillgängligt vid arkivering, och för det andra om det är önskvärt. Det man främst uppnår med detta är att FHIR resurserna kan exponeras som API enligt standard.

#### 4.1.1.11 RIM för informationshantering (inom hälso- och sjukvård)

Inom VGR har det tagits fram en informationsmodell för informationshantering inom hälso- och sjukvård - se bilaga [25]. Modellen har tagits fram i form av ett klassdiagram, där relationer mellan klassobjekt och relevanta attribut redovisas.

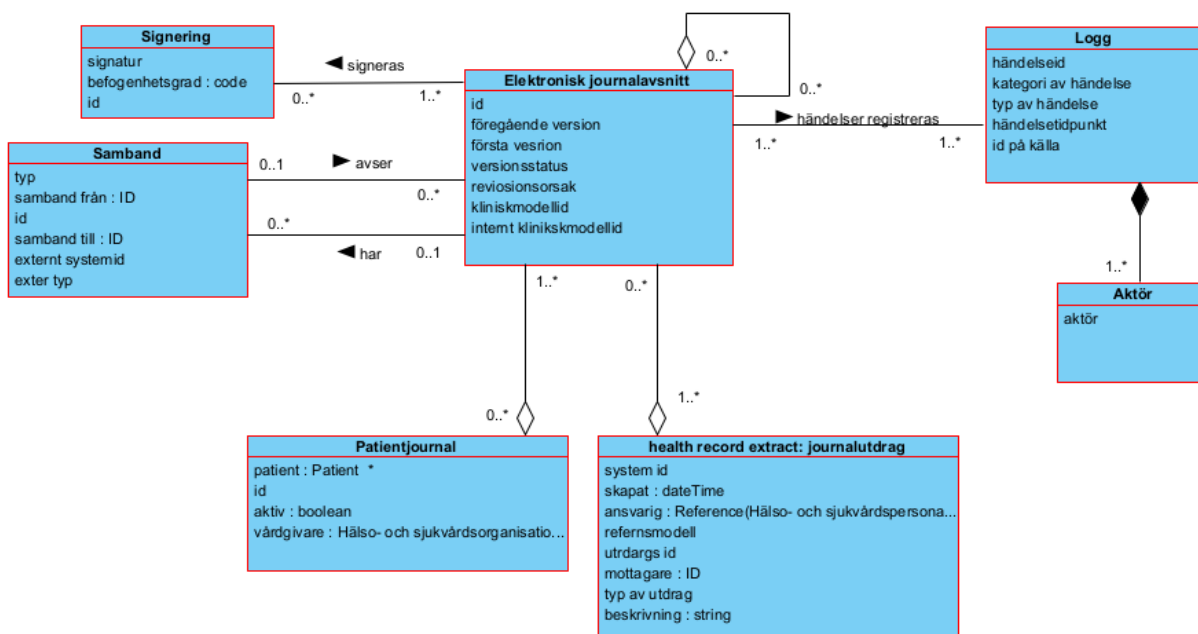


Diagram 4.16 - Klassdiagram för informationshantering inom hälso- och sjukvård

Informationsmodellen används som referens vid framtagande av en tillämpad informationsmodell för hantering av arkiverade patientjournaler. I målarkitekturen görs detta i kap. 5.2 i det efterföljande.

Name	Description
Samband - Class	Beskrivning av samband som går mellan elektroniska journalavsnitt
Signering - Class	All information kring signeringen av det elektroniska journalavsnittet.
Aktör - Class	
Logg - Class	Information av händelser i systemen.
health record extract: journalutdrag - Class	
Elektronisk journalavsnitt - Class	
Patientjournal - Class	Klassen beskriver genrellt attribut som gäller för hela patientjournalen

Tabell 4.2 - Tabell över klasser

Slutsatser om VGR:s RIM är att den ska användas som referens vid framtagning av en tillämpad informationsmodell för arkiverade informationshandlingar inom informationsdomänen patientjournaler. Detta görs med stöd av informatiker.

#### 4.1.1.12 European Health Data Space (EHDS)

EHDS [40] är en ny EU lag som presenterades för EU kommissionen 2022. En godkänd EU lag träder omedelbart i kraft i hela unionen och behöver inte ratifieras i varje medlemsland. EHDS utgör ett ramverk för hur en EU invånares hälsodata ska kontrolleras av patienten själv och kunna delas över gränser till hälso- och sjukvårdspersonal.

EU lagen kan få betydelse för hur patientjournaler, inkl. arkiverade patientjournaler, lämnas ut till både invånare, vårdpersonal och forskare utanför Västra Götaland och Sverige.

Slutsatser om EHDS är att Sverige behöver ta ett samlat grepp om både operationella och arkiverade patientjournaler; ett initiativ som faller utanför avgränsningarna för detta projekt.

## 4.1.2 Målarkitekturer

### 4.1.2.1 Plan för digital informationshantering

Plan för digital informationshantering [20] beskriver fem informationsdomäner för verksamhetssystem:

- Fristående dokument [...]
- Publiceringssystem [...]
- Processbaserad informationshantering
- Ekonomi, HR och administrativ ärendehantering
- Vårddomänen

Den sista, vårddomänen, är relevant för mellanarkivering av klinisk data och delas upp i fem underdomäner:

- Patientjournal
- Patientadministration
- Diagnostik, utredning och laboratoriesystem
- Registerdata vård
- Forskningsdata vård

Nedan följer korta beskrivningar av varje subdomän samt beskrivning av behov av mellanarkivering. Texten är huvudsakligen saxad från [20].

#### Patientjournal

En journalhandling definieras som information som upprättas eller inkommer i samband med vården av en patient och som innehåller uppgifter om patientens hälsotillstånd eller andra personliga förhållanden eller om vidtagna eller planerade vårdåtgärder. Flera sådana journalhandlingar om samma patient utgör en patientjournal.

*Behov av mellanarkivering:* I samband med utvecklingen av en ny vårdinformationsmiljö införs vårdsystemet Millennium, som ska ersätta merparten av de gamla specialiserade vårdsystemen. Detta leder till att ett stort antal system behöver avvecklas och vårdinformationen behöver tas om hand och tillgängliggöras utan att förlora publiceringsmöjligheter till Nationell patientöversikt (NPÖ) och 1177 Journalen på nätet.

#### Patientadministration

Viktiga delar inom patientadministration är tidbokning och hantering av kassa samt akut-, öppen- och slutenvårdskontakter. Ett patientadministrativt system behöver kopplingar till ekonomisystem och konsumerar metadata från olika kodverk, tex diagnoskoder och organisationskoder, för att informationen ska klassas rätt.

*Behov av mellanarkivering:* Stor del av informationen i ett patientadministrativt system kan sannolikt gallras efter 7–10 år om informationen finns i patientens journal, gallringsfrist för ekonomimaterial har gått ut och det inte finns annan information som är av vikt.

#### Diagnostik, utredning och laboratoriesystem

Det finns system som hanterar laborietest på kroppsvätskor eller vävnad från patienter. Det klassiska sättet att arbeta är att vårdverksamheten skickar en remiss till en utförare, vanligtvis internt, som gör en mätning/undersökning av patienten. Beställning och resultat finns som utförares exemplar i systemet. Det skickas ett utlåtande till remitterande vårdenhet, som sedan läggs i patientens journal. Detta är beställarens exemplar och är en journalhandlingen.

*Behov av mellanarkivering:* Man har kommit fram till att båda exemplaren ska bevaras. En utmaning är att den information som skapas ofta förekommer i format som saknar godkända arkivformat.

Bild- och funktionsregistret (BFR) är en regiongemensam lagring och distribution av medicinsk bild och information, mellan producenter och konsumenter, framför allt inom Bild- och funktionsmedicin och Folk tandvården. Det är ett s.k. VNA, Vendor Neutral Archive, ett arkiv som lagrar digitala röntgenbilder i ett standardformat med användargränssnitt som är öppet för alla.

*Behov av mellanarkivering:* Som VNA fungerar BFR som mellanarkiv och tanken att all medicinsk bild ska hanteras via BFR bör stödjas. Det är oklart hur och var rörlig bild, till exempel längre videoinspelningar av patienter, ska bevaras.

#### Registerdata vård

Nationella kvalitetsregister samlar på ett systematiskt sätt data om olika patientgrupper. En del register gäller alla patienter med en viss diagnos. Andra fokuserar på vissa behandlingar. Registren innehåller uppgifter på individnivå om bakgrundsfaktorer, diagnoser, behandlingar och vilket utfall som vården har gett för varje patient. Syftet med kvalitetsregistren är att de ska användas för löpande förbättringsarbete i vården.

*Behov av mellanarkivering:* I VGR är huvudregeln att nationella kvalitetsregister ska bevaras. När man slutar att rapportera in ny information och kommer fram till att uppgifterna inte längre behövs för kvalitetsgranskning, bör uppgifterna slutarkiveras för framtida forskning. Det har inte uppstått behov av mellanarkiv.

#### Forskningsdata vård

Den dokumentation som forskning inom regionen baserar sig på är ofta samlade journalkopior, testprotokoll och enkätmaterial på patienter, ibland avidentifierat men märkt med koder som kan identifieras via register i en prövarpärm. Materialet kan vara både pappersbaserat och digitalt, och omfattas delvis av sekretess. Det skapas också en hel del databassystem för unika forskningsändamål.

*Behov av mellanarkivering:* Det är rimligt att kunna erbjuda samarbetsytor med koppling till mellanarkiv till forskningsprojekt där en myndighet inom VGR är huvudman, så att det går att arkivera handlingarna. För att det ska vara möjligt krävs att externa parter kan bjudas in till slutna samarbetsytor, eftersom informationen tillhör informationsklass 3.

Slutsatser om *Plan för digital informationshantering* är att det inom vårddomänen finns ett starkt behov att mellanarkivera data för patientjournaler, patientadministrativ information samt diagnostik, utredning och laboratoriesystem. Behovet förstärks vid införande av Millennium genom att ett stort antal journalsystem och patientadministrativa system ska avvecklas.

Det finns i nuläget inget behov att mellanarkivera kvalitetsregister.

Ett lösningskoncept behöver utvecklas för hantering av forskningsdata. Huvudmål är att kunna erbjuda en delmängd av mellanarkiverad information motsvarande varje individuell forskares behov och behörighet. För enklare fall kan det räcka med en delad dokumenttyta, men för t.ex. statistisk bearbetning av större datamängder krävs databaser. Beroende på forskningens syfte och mandat kan det även vara nödvändigt att anonymisera eller pseudonymisera person- och/eller patientdata, och det kan därför vara mest lämpligt att generera transienta forskningsdatabaser eller dataset "on-demand" från mellanarkivet.

#### 4.1.2.2 ADD - ARK Strategisk målarkitektur för mellanarkivering

Arkitekturarbetet [16] beskriver de digitala lösningar med avseende på mellanarkivering som behöver etableras för att tillmötesgå verksamhetens/organisationens mål beskriven i Plan för digital informationshantering [20].

Kravbilden är en digital miljö där:

- Alla allmänna handlingar hanteras i enlighet med gällande lagstiftning.
- Allmänna handlingar är tillgängliga för behöriga användare.
- Informationen är strukturerad och samlad på ett sätt som gör det enkelt att hitta och använda den.

Informationshantering sträcker sig enligt lagstiftning och praxis bortom dokumenthantering och omfattar all typ av information, inklusive:

- Inlägg i kommunikationskanaler
- Databasinnehåll
- HTML-baserad information
- Diagnostik
- Inspelat ljud och video
- Medicinsk bild från Bild- och Funktionsmedicin
- Mätdata från medicintekniska produkter
- Genomik-data

Målbilden är en uppsättning arkivlösningar där arkiverad information är lättillgänglig och kan användas för olika ändamål. Informationen ska vara strukturerad och samlad på ett sätt som gör det enkelt att hitta och använda den.

I nuläget har varje styrelse, nämnd, förvaltning och majoritetsägt bolag inom ramen för Västra Götalandsregionen ansvar för den information som produceras och ska säkerställa adekvat hantering av handlingar som upprättas. Det innebär att informationsägarskap är ett decentraliserat ansvar. Det kan också förekomma lokala arkivfunktioner som ansvarar för t ex utlämnande av handlingar som man ansvarar för.

Regionarkivet är arkivmyndighet för Västra Götalandsregionen och ansvarar för slutarkivering av handlingar. Regionarkivet tillhör rent organisatoriskt Göteborgs Stad och har således ansvar för såväl Göteborgs Stad som Västra Götalandsregionen. Förutom slutarkivering är Regionarkivet att betrakta som tillsynsmyndighet och uttolkare av de riktlinjer som tas fram av Riksarkivet.

Följande illustration visar översiktligt skillnaden mellan källsystem, mellanlager, mellanarkiv och slutarkiv och hur hantering av informationshandlingar ska fungera vid olika tillstånd för informationshandlingar:

Verksamhetssystem		Arkiv	
Källsystem	Mellanlager	Mellanarkiv	Slutarkiv
<ul style="list-style-type: none"> <li>• Verksamheten förvaltar informationen</li> <li>• Verksamheten har direktåtkomst</li> <li>• PDL-stöd krävs</li> <li>• Verksamheten får redigera information</li> <li>• Valfri informationsstruktur (domän)</li> <li>• Systemet byts ut ofta utifrån verksamhetens behov och LOU</li> </ul>	<ul style="list-style-type: none"> <li>• Verksamheten förvaltar informationen</li> <li>• Verksamheten har direktåtkomst</li> <li>• PDL-stöd krävs</li> <li>• Verksamheten får redigera information</li> <li>• Valfri informationsstruktur (domän)</li> <li>• Mellanlagret byts ut sällan utifrån infrastruktur</li> </ul>	<ul style="list-style-type: none"> <li>• Arkivfunktionen förvaltar informationen</li> <li>• Verksamheten har begränsad direktåtkomst utifrån väl avgränsade syften</li> <li>• PDL-stöds krävs om verksamheten ges direktåtkomst</li> <li>• Verksamheten får endast redigera information inom ramen för väl avgränsade syften</li> <li>• Informationen struktureras enligt särskild specifikation för paketstruktur (VGR)</li> <li>• Mellanarkivet byts nästan aldrig ut men kringliggande applikationer kan bytas ut</li> </ul>	<ul style="list-style-type: none"> <li>• Arkivmyndigheten förvaltar informationen</li> <li>• Verksamheten saknar direktåtkomst</li> <li>• Slutarkivet byts nästan aldrig ut</li> <li>• PDL-stöd krävs inte</li> <li>• Informationen struktureras enligt arkivmyndighetens specifikationer (<a href="#">core</a>)</li> </ul>

Tabell 4.3 - Jämförelse mellan informationshandlingar förvarade i olika system

För konsumtion av arkiverade informationshandlingar utgår arkitekturarbetet från att arkiverade informationshandlingar exponeras som dataprodukter och kan konsumeras via exponerade API:er. Man kan med fog säga att den strategiska målarkitektur för mellanarkivering är en specialisering av det som framkommit i målarkitekturen för GIP (Gemensam informationsplattform) där utgångspunkten av konsumtion av arkiverade allmänna handlingar görs enligt de principer som framgår av den<sup>3</sup>.

Arkitekturarbetet beskriver två huvudsakliga lösningsalternativ och rekommenderar alternativ 2, som grundar sig på att bygga vidare på de arkivlösningar som redan är etablerade och komplettera med de arkivlösningar som fattas. De arkivlösningar som redan är etablerade behöver ses över och eventuellt vidareutvecklas för att nå upp till säkerhetskrav och krav på adekvat arkivvård.

I lösningsalternativet identifieras 6 olika lösningsmönster för att etablera arkivlösningar baserade på OASIS-modellen. Alla 6 valda för etablering av specialiserade arkiveringslösningar för olika syften.

Rekommendationen är att bygga arkiveringslösningar på antingen RODA eller Archivemata, som båda är vedertagna Open Source-produkter med god förutsättningar att skräddarsys till specifika behov.

Målarkitekturen beskriver 6 lösningsmönster. Se bilaga [16] för en detaljerad beskrivning av dessa.

1. OASIS-arkiv med utlämningsförfarande (den klassiska OASIS-modellen). Lösningsalternativ 1 bygger på ett OASIS-arkiv som hanterar inkommande DIP-paket, transformerar och lagrar SIP-paket och lämnar ut DIP-paket med generella arkivapplikationer för mottagning, arkivlagring, åtkomst, datahantering, bevarande och administration.
2. OASIS-arkiv utökad med intern cache för realtidsåtkomst. Som lösningsmönster 1 med skillnaden att man skapar ett internt sekundärformat som kan exponeras som dataprodukter och konsumeras via API:er.
3. OASIS-arkiv utökad med intern cache för realtidsåtkomst och lagring i XML-format i NOSQL databas. Som lösningsmönster 1 med två skillnader: (1) man skapar ett internt sekundärformat som kan exponeras som dataprodukter och konsumeras via API:er och (2) man lagrar arkiverade allmänna handlingar i XML-representation och lagrar dessa i en XML-databas (t ex en NOSQL-databas) som potentiellt kan konsumeras som dataprodukter via API:er mer direkt.
4. OASIS-arkiv utökad med intern cache för realtidsåtkomst och lagring i arkivdatabasformat. Som lösningsmönster 1 med två skillnader: (1) man skapar ett internt sekundärformat som kan exponeras som dataprodukter och konsumeras via API:er och (2) man lagrar arkiverade allmänna handlingar i relationsdatabasformat. I princip skulle det kunna vara en kopia av källsystemets databasstruktur och innehåll, men kompletterad med metadata för mellanarkivering, slutarkivering och eftersökning.
5. OASIS-arkiv utökad med intern cache för realtidsåtkomst och transformerad till gemensam RIM. En variant på lösningsmönster 2, 3 och 4 där data transformerar i enlighet till en referensinformationsmodell (RIM), t ex OpenEHR.
6. Konsolidering av handlingar från olika källsystem. Detta är ett specialfall där allmänna handlingar från olika instanser av ett system sammanförs till en gemensam homogen arkivlösning.

Samtliga lösningsmönster kan tillämpas med stöd i regionala stödtjänster:

- Autentiserings- och behörighetstjänster:
  - Autentiseringstjänst - för att autentisera slutanvändare (skulle kunna vara via eSITHS eller Mobilt bankid t ex)
  - Behörighetstjänst - för att säkerställa att slutanvändare har rätt behörighet (skulle kunna vara behörighetsgrupper i HSA-katalog t ex)
  - Spärrtjänst - Om en klinisk handling är spärrad av en patient.
  - Samtycketjänst - Om en patient gett samtycke till åtkomst till andra vårdgivare.
  - Underskriftstjänst - För att validera underskrifter.

---

<sup>3</sup> I det här kontexten är GIP en vision som den beskrivits i målarkitekturen för GIP, inte den inriktning som Data & Analys väljer.

- Loggtjänst - för loggning av åtkomst till handlingar, men också till arkivvårdande aktiviteter.
- Söktjänster
  - Masterdatatjänst - för att vid transformation av handlingar inför arkivering hämta underlag för metadatasättning.
  - Hitta-i-diariet - för att ge åtkomst till allmänna handlingar via hitta-i-diariet (om diarieförda handlingar arkiverats i mellanarkiv)
  - Hitta-i-arkivet - för att ge åtkomst till allmänna handlingar via hitta-i-arkivet
- Stödtjänster för mellanarkivering
  - Utlämnings tjänst - För begäran om utlämning av handlingar till forskningsprojekt eller på begäran från allmänheten.
  - Överföringstjänst - För att samla ihop, paketera och överföra handlingar till Regionarkivet.
  - Pseudonymiseringstjänst - för att i samband med utlämning av handlingar till forskningsprojekt (eller andra konsumtionsbehov som kräver det) pseudonymiseras identiteter på personer eller patienter. Borde också fungera som en ren anonymiseringstjänst.

Slutsatser om ADD - ARK Strategisk målarkitektur för mellanarkivering [16]:

- RODA eller Archivemata rekommenderas som teknisk plattform
- OAIS används som standard för mellanarkivet
- De 6 lösningsmönster som beskrivs har beaktats i målarkitekturen för mellanarkivering av patientjournaler. Lösningsmönster 2, 5 och 6 kombineras i målarkitekturen.
- Relevanta regionala stödtjänster bör användas i målarkitekturen

#### 4.1.2.3 Strategi för datamigrering i Millennium

Dokumentet *Strategi datamigrering Millennium* [29] är styrande för program Millennium och programägaren för Millennium är ägare av strategin. Strategin beskriver följande principer för datamigrering från befintliga system till Millennium (se bilaga [29] för en detaljerad beskrivning av varje princip):

1. Patientsäkerhet styr hur data hanteras
2. Begränsa de data som migreras till ett minimum
3. Information som kräver historisk organisation migreras inte
4. Manuell migrering väljs när det är kostnadseffektivt
5. Avlidna patienters information migreras inte
6. Information med äldre reservnummerformat hanteras separat
7. Varje huvudman har eget ansvar för säkerställande av data för migrering
8. Varje huvudman har eget ansvar för historik och avveckling

Princip 2, 3 och 5 styr vad som inte migreras till Millennium och som följaktligen måste gallras, sparas till mellanlagring eller sparas till mellanarkiv (eller slutarkiv).

Princip 8 identifierar huvudintressenten vid mellanarkivering av enskilda system.

Strategin beskriver i kap. 6 även principer för mellanlagring, där mellanlagring definieras enligt följande:

*”Mellanlagring innebär att informationen i de gamla systemen flyttas till en regiongemensam lagringslösning som bland annat ska fungera som mellanarkiv med läsåtkomst till historisk information och med bevarande enligt arkivlagen. Lagringslösningen ska också fungera som ett datalager med stöd för analys, uppföljning och forskning”.*

I dokumentet likställs mellanlagring med mellanarkivering.

Strategin beskriver följande principer för mellanlagring (se bilaga [29] för en detaljerad beskrivning av varje princip):

1. Förenkling bör styra turordning
2. Balansen mellan nytta, risk och kostnad bör beaktas vid val kring hur informationen mellanlagras och kopplas till funktionalitet
3. Dokumenthanteringsplanerna ska följas
4. Information med bevarandekrav men utan verksamhetsbehov bör slutarkiveras
5. Mellanlagring ersätter inte slutarkivering

Princip 2 slår fast att allt som arkiveras inte behöver exponeras via läsårskomst (från andra applikationer).

Princip 4 och 5 slår fast att information som verksamheten inte har behov av ska slutarkiveras eftersom det är mindre kostsamt. Beslut om detta bör tas som del av ett avvecklingsprojekt.

Slutsatser för *Strategi datamigrering Millennium* är att princip 2 för mellanlagring har stort inflytande på hur strukturerad data som ska kunna exponeras via API lagras i mellanarkivet. Princip 4 och 5 har stort inflytande på lagringsvolymerna i mellanarkivet. Alla principer måste beaktas vid införande av mellanarkivlösning.

### 4.1.3 Standarder

Standarder som är relevanta för hantering av mellanarkiv<sup>4</sup>:

- OAIS - Open Archival Information System
  - En OAIS är ett arkiv som består av en organisation av människor och system som har accepterat ansvaret för att bevara information och göra den tillgänglig för en specifik målgrupp. Standarden definierar en uppsättning ansvarsområden som ett OAIS-arkiv måste uppfylla och detta gör att ett OAIS-arkiv kan särskiljas från andra användningar av termen arkiv.
- METS - Metadata Encoding and Transmission Standard
  - Metadata Encoding and Transmission Standard (METS) är en datakodnings- och överföringsspecifikation, uttryckt i XML, som tillhandahåller medel för att förmedla de metadata som krävs för både hanteringen av digitala objekt inom ett arkiv och utbytet av sådana objekt mellan arkiv (eller mellan arkiv och deras användare). Detta gemensamma objektformat har utformats för att uppmuntra samarbete vid utveckling av verktyg/tjänster för informationshantering och för att underlätta interoperabelt utbyte av digitalt material mellan institutioner (inklusive leverantörer) [35].
- EAD - Encoded Archival Description
  - Encoded Archival Description (EAD) är en XML-standard för kodning av hjälpmedel för arkivsökningar, som underhålls av den tekniska underkommittén för kodade arkivstandarder i Society of American Archivists, i samarbete med Library of Congress [37].
- DC - Dublin Core
  - Dublin Core, även känd som Dublin Core Metadata Element Set (DCMES), är en uppsättning av femton huvudmetadataobjekt för att beskriva digitala eller fysiska resurser. Dublin Core har formellt standardiserats internationellt som ISO 15836 av International Organization for Standardization (ISO) och som IETF RFC 5013 av Internet Engineering Task Force (IETF), samt i USA som ANSI/NISO Z39.85 av National Information Standards Organisation (NISO).
- E-ARK Information Package specifications

---

<sup>4</sup> Översättning från engelska har gjorts med Google translate. Det är inte vackert men det uppfyller behovet.

- OAIS-ramverket definierar tre typer av informationspaket (IPs) som finns i det digitala bevarandecosystemet: Submission Information Packages (SIPs), Archival Information Packages (AIPs) och Dissemination Information Packages (DIPs). Dessa tre IP-typer används för att skicka in data och metadata till digitala arkiv, lagra dem långsiktigt och leverera dem till konsumenter. Huvudmålet med E-ARK har varit att identifiera och standardisera de gemensamma aspekterna av IP:er som är lika relevanta och implementerade av någon av de funktionella enheterna i den övergripande digitala bevarandeprocessen som presenteras i OAIS (d.v.s. förintag, intag, arkivlagring, datahantering och åtkomst). Den gemensamma specifikationen för informationspaket [31] åtföljs av motsvarande utökade implementeringsspecifikationer för E-ARK SIP [32], E-ARK AIP [33] och E-ARK DIP [34].
- PREMIS - Preservation Metadata
  - PREMIS Data Dictionary for Preservation Metadata är den internationella standarden för metadata för att stödja bevarandet av digitala objekt och säkerställa deras långsiktiga användbarhet. PREMIS har utvecklats av ett internationellt team av experter och implementeras i digitala bevarandeprojekt runt om i världen, och stödet för PREMIS är inorporerat i ett antal kommersiella och öppen källkod för digitala bevarandeverktyg och system. PREMIS redaktionskommitté koordinerar revisioner och implementering av standarden, som består av Data Dictionary, ett XML-schema och stödjande dokumentation [36].

Standarder som kan vara relevanta för hantering av operationell data:

- SQL är ett funktionellt språk som är en ISO och ANSI standard för att söka i och bearbeta data som lagras i relationsdatabaser. SQL är effektivare och snabbare än andra programmeringsspråk därför att bearbetningar alltid utförs av databasservern, som därmed slipper överföra data via nätverk under bearbetning. Även all metadata lagras i databasen och kan bearbetas med SQL.
- XML är ett markupspråk som är en W3C standard för att serialisera data och datastrukturer som text tillsammans med beskrivande metadata. Största fördelen med XML är att data lagras i läsbara textdokument och att det är ett lämpligt format för serialisering och överföring av data mellan system; t.ex. vid SOAP eller REST API anrop. Största nackdelen med XML är att det är ineffektivt; speciellt vid bearbetning av stora mängder data.
- XSD är en standard för att definiera ett "namespace" i XML med tillhörande regler för syntax och semantik för att kunna validera att ett XML dokument är korrekt utformat. XSD möjliggör programmatisk validering av XML dokument.

#### 4.1.4 Bedömning av möjlighet till återanvändning

Delar av implementation som beskrivs i [42] kan återanvändas i målarkitekturen.

#### 4.1.5 Lösningalternativ - sammanfattning

Ett grundläggande "build vs. buy" vägval rekommenderas med motivering att flera andra svenska myndigheter och regioner redan har utfört tekniska utredningar och kommit fram till samma slutsatser:

1. Det är bättre att själv utveckla en lösning för ett långsiktigt digitalt arkiv än att genomföra återkommande offentliga upphandlingar.
2. Open-source produkten RODA används som teknisk plattform vid egenutveckling av digitalt arkiv.

Ytterligare motiveringar sammanfattas nedan vid utvärdering av lösningalternativ.

Två lösningalternativ beskrivs i det efterföljande. Det ena är baserat på RODA Community Edition och det andra är baserat på RODA Enterprise Edition. Bägge alternativ innehåller samma open-source RODA server, men i Enterprise versionen ingår licenser för kommersiella plugins och tilläggsprogram som inte är open-source. Dessa kan dock vid behov licensieras individuellt.

## 4.1.6 Lösningalternativ 1 - RODA Community

### 4.1.6.1 Beskrivning av lösningalternativ 1

Lösningalternativet består av en gemensam mellanarkivplattform, som utnyttjar relevanta regionala och nationella stödtjänster, samt en eller flera front-end applikationer per informationsdomän. Initialt läggs fokus på ett subdomän till vårddomänen; nämligen patientjournaler. Även VGR:s befintliga infrastrukturskomponenter för datavirtualisering och API hantering förväntas ingå i plattformen. Detta utreds närmare i kap. 5.

Logisk applikation för mellanarkivering av patientjournaler kommer att analyseras i mer detalj i kap. 5.3 Applikationsarkitektur, men kommer att bestå av flera logiska och/eller fysiska applikationer för att realisera domänspecifika aspekter av de två huvudanvändningsfallen vid mellanarkivering av patientjournaler:

- Använd arkiverad information för praktiska tillämpningar (t.ex. MAJ sök)
- Migrera information till ett mellanarkiv och förvalta enligt arkivlagen inför slutarkivering

Generella aspekter (dvs ej knutna till någon specifik informationsdomän) av de två användningsfallen förväntas kunna realiseras direkt i den underliggande mellanarkivplattformen. Plattformen utnyttjar även relevanta regionala och nationella stödtjänster, vilket utreds i större detalj i kap. 5.3.5, 5.3.6 och 5.3.7.

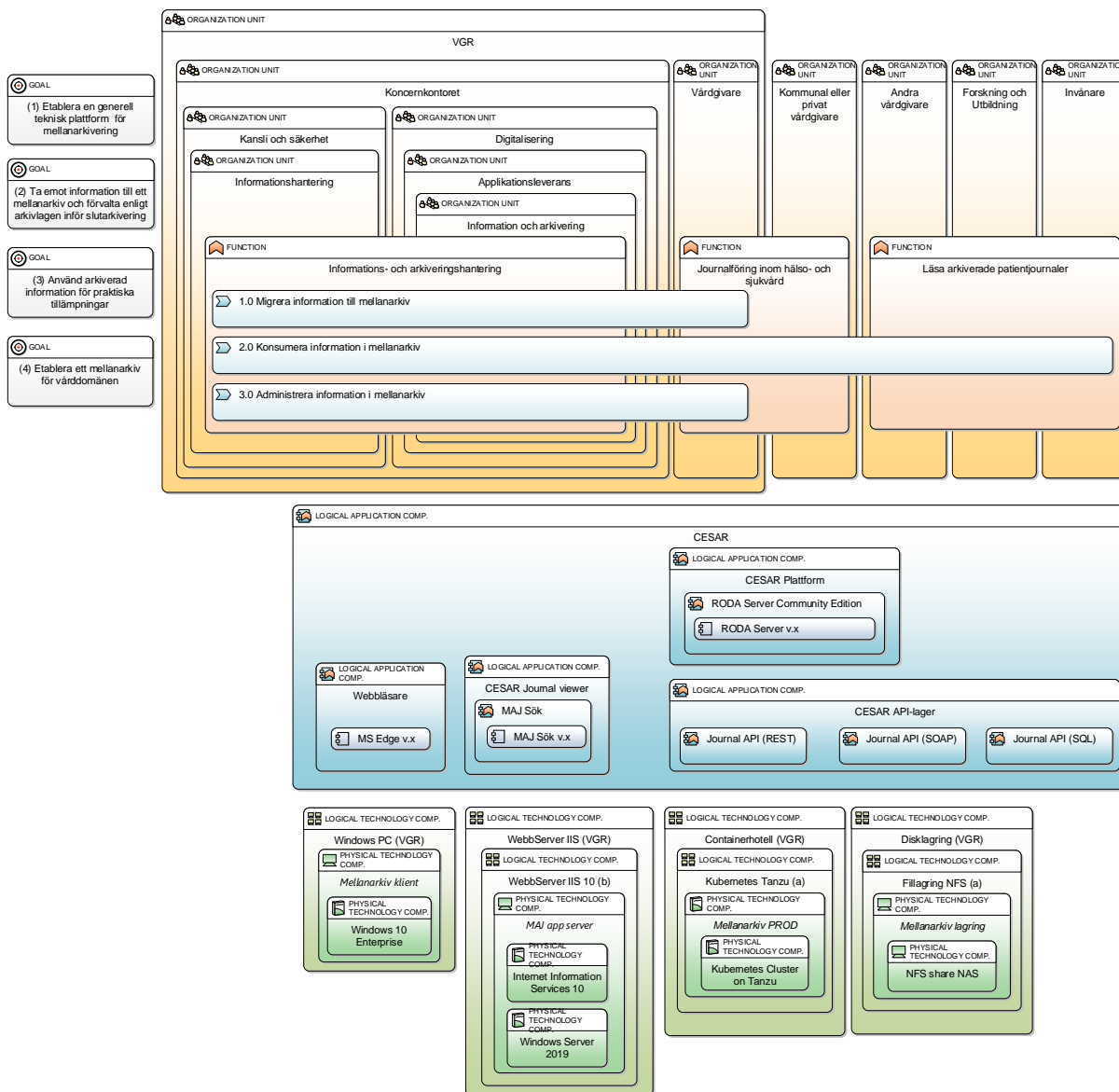


Diagram 4.17 - Lösningalternativ 1

4.1.6.2 Utvärdering av lösningalternativ 1

Försäkringskassan, Tullverket och Region Skåne har alla utfört omfattande utredningar av lösningalternativ för digital arkivering under perioden 2021-2023. Alla tre valde att utveckla egna arkivlösningar baserat på öppen källkod framför upphandling:

- Detta ger större flexibilitet med att forma systemet och samarbete med andra system
- Möjlighet till en långsiktig strategisk upplägg samt möjligheter att sprida kostnaderna
- Inga kontinuerliga upphandlingar
- Samverkan med andra myndigheter/regioner

Alla tre valde RODA som plattform:

- Infriar flest verksamhetskrav
- Kan få teknisk support inom EU vid behov (leverantör är portugisisk)
- Förvaltningskostnader synliggörs tydligt genom egenutveckling
- Samverkan med andra myndigheter/regioner

Detaljerade slutrapporter från Försäkringskassan [26], Region Skåne [27] och Tullverket [28] bifogas som bilagor. Rapporterna har i sig även ett antal bilagor som återfinns på Sofia projektytan för ”Projekt Etablera Mellanarkiv”.

#### 4.1.6.3 Riskhantering av lösningsalternativ 1

De risker som har identifierats i kap. 2.4.5 kräver att lösningen har säker inloggning, säker behörighets- och åtkomstkontroll, loggning samt hög driftstillförlitlighet (high availability). Detta förväntas kunna uppnås med det valda lösningsalternativet.

#### 4.1.7 Lösningsalternativ 2 - RODA Enterprise

##### 4.1.7.1 Beskrivning av lösningsalternativ 2

Lösningsalternativet är identiskt med lösningsalternativ 1, men använder en kommersiell licens ”RODA Enterprise” för den underliggande arkivplattformen.

RODA Enterprise innehåller, utöver samma RODA server som ingår i RODA Community, två typer av kommersiella mjukvarukomponenter: fristående tilläggsprogram (kallad Components) och pluggbara moduler (kallad Plugins) som laddas av RODA servern för att uppnå tilläggsfunktionalitet.

Varken Components eller Plugins är open-source utan måste licensieras kommersiellt; antingen bundlat i RODA Enterprise licensen eller individuellt licensierade. De beskrivs nedan enligt respektive produktblad från Keep Solutions (översätts ej från engelska).

Följande Components finns:

Component	Beskrivning (en)	Behövs för Mellanarkiv	Kommentar
Drop Folders	The drop folders component facilitates the automated, unsupervised ingestion of submission information packages via shared folders, which is crucial for a smooth integration with other data production systems.	Nej	Alternativ finns; t.ex. FTP överföring av SIP filer.
Folder Scattering	The Folder Scattering component ensures that files are evenly distributed across the file system to prevent over-utilization of any specific storage area and allows for the extension of storage capacity without interruption to ongoing operations.	Nej	Hanteras av VGR lagringstjänst
Remote Replication	The Remote Replication component allows a high availability setup that replicates information packages to remote data centres. This ensures that in the event of an unexpected event, such as natural disasters or cyber-attacks, critical data is still available, minimising the disruption to the business operations.	Nej	Hanteras av VGR lagringstjänst
Inventory Report Comparator	The Inventory Report Comparator is a tool that helps users compare inventory reports from different time periods or produced by different systems. This enables the easy identification of any changes or discrepancies in the repository's content over time or to cross check if the entire content of an ingest batch has been correctly archived. This tool simplifies the process of validating the completeness and correctness of the repository content by comparing it to previous inventory reports generated during pre-ingest.	Nej	Mest användbart vid slutarkivering.
Identity and Access Management	Supports the integration of various authentication protocols, including LDAP, SAML, OAuth, Azure AD, EU Login, Citizen Card, and more. This allows organisations to authenticate both users and services in a secure and flexible manner, ensuring that only authorised individuals and systems can access sensitive information.	Ja	Kan utvecklas av VGR
Reporting	Customizable dashboards for organizations to support informed decision-making and strategic planning. This feature allows for tracking of key	Önskvärt	

	performance indicators (KPI), risk analysis, and other important data assessment tasks, providing management with the information necessary to make informed decisions and drive the success of the preservation process.		
Monitoring	Load and storage monitoring capabilities with advanced dashboard functionality. Customizable alert notifications allow for proactive monitoring and ensure optimal performance and efficient resource utilization. This provides organizations with valuable insights into the health and performance of their system, allowing them to take proactive measures to maintain high levels of availability and efficiency.	Önskvärt	

Tabell 4.4 - RODA Components

Följande Plugins finns:

Område	Plugin	Beskrivning (en)	Behövs för Mellanarkiv	Kommentar
Characterization	File Feature Extractor and Office Documents Text Extractor	The File Feature Extractor is a powerful plugin that allows users to extract technical metadata from a wide range of file formats, making it an essential tool for digital curators. The tool supports numerous file formats, including images, raw image formats from digital cameras from several manufacturers, DICOM metadata, Photoshop tags, metadata from audio and video formats, and metadata from office documents such as Microsoft Word, Excel, PowerPoint, or OpenDocument formats.	Nej	Hanteras i pre-ingest.
	Optical Character Recognition Extractor	The Optical Character Recognition Extractor is a powerful plugin designed to extract typed or printed text from digitalized images, making it an essential tool for professionals in various fields, including data analysis, document management, and research.	Nej	Hanteras i pre-ingest.
Conversion	Audio Converter	The Audio Converter is a highly effective plugin that leverages the capabilities of 'SoX' (Sound eXchange tool) a versatile cross-platform tool for audio file format conversion. With this plugin, users can convert audio files from one format to another and apply a variety of advanced effects such as volume adjustments, equalization, reverb, delay, chorus, flanging, tempo and pitch changes.	Nej	SoX är open-source GPL-2.0. Hanteras i pre-ingest.
	Image Converter	The Image Converter plugin harnesses the power of ImageMagick, a leading image manipulation tool, to effortlessly convert between over 200 different image formats including PNG, JPEG, JPEG-2000, GIF, TIFF, DPX, EXR, WebP, Postscript, PDF, and SVG.	Nej	Imagemagick är open-source. Hanteras i pre-ingest.
	Office Documents Converter	The Office Documents Converter is a versatile plugin that utilizes the 'unoconv' (Universal Office Converter) technology to convert a wide range of office file formats. The supported formats include Open Document Format (odt), Microsoft Word (doc), Microsoft Office Open/Microsoft OOXML (ooxml), Portable Document Format (pdf), HTML (html), XHTML (xhtml), Rich Text Format (rtf), Docbook (docbook), and many others. This tool provides a convenient and efficient solution for	Nej	Unoconv is deprecated open-source GPL-2.0. Hanteras i pre-ingest.

Område	Plugin	Beskrivning (en)	Behövs för Mellanarkiv	Kommentar
		converting office documents to a desired format.		
	PDF to PDF/A Converter	The PDF to PDF/A Converter plugin transforms standard Portable Document Format (PDF) files to PDF/A by leveraging the power of the "ghostscript" tool. PDF/A is an ISO-standardized version of the Portable Document Format (PDF) specialized for use in the archiving and long-term preservation of electronic documents. PDF/A differs from PDF by prohibiting features ill-suited to long-term archiving, such as font linking (as opposed to font embedding) and encryption. The ISO requirements for PDF/A file viewers include colour management guidelines, support for embedded fonts, and a user interface for reading embedded annotations.	Nej	Ghostscript är open-source GPL. Hanteras i pre-ingest.
	Video Converter	The Video Converter is a powerful plugin that leverages the capabilities of 'avconv,' a high-speed video and audio conversion tool. This converter can perform arbitrary sample rate conversions and resize video in real-time with a high-quality polyphase filter. The plugin allows for the conversion of files containing a variety of different stream types, including video, audio, subtitles, attachments, and data. The number and type of streams are determined by the container format, and the plugin can enforce these limitations as necessary.	Nej	Libav is an abandoned free software project, forked from FFmpeg in 2011. GPL-2.0. Hanteras i pre-ingest.
Database preservation	Dabase Visualization	The Database Visualization plugin creates a special purpose DIP and loads SIARD 2 files into an instance of DBPTK for visualisation purposes. DBPTK (Database Preservation Toolkit) is an open-source tool for archiving and browsing the contents of relational databases. It is designed to preserve and provide access to databases that are part of a digital archive. The viewing tool also provides a range of features such as database export, saved searches, records concentration diagrams, etc.		
Digital signature	DIP Digital Signature Creator	The DIP Digital Signature Creator plugin is a powerful tool that enables users to generate a new Dissemination Information Package (DIP) for a specified Archival Information Package (AIP). The DIP contains all the files from the AIP, digitally signed with the repository's digital certificate. By digitally signing the files in the DIP with the repository's digital certificate, the plugin helps ensure that the files are authentic and trustworthy. Depending on regional legislation, the files in the DIP may be considered certified copies of the original digital assets, further enhancing their authenticity and reliability. Using the DIP Digital Signature Creator plugin, users can create a secure and reliable package of digital assets that can be easily disseminated to others. This is especially important for organizations and individuals who need to ensure that their digital assets are trustworthy and can be relied upon for legal or regulatory purposes.	Önskvärt	

Område	Plugin	Beskrivning (en)	Behövs för Mellanarkiv	Kommentar
	Digital Signature Expiry Date Extender	The Digital Signature Expiry Date Extender uses a technique called Long-Term Validation (LTV) to ensure the integrity and authenticity of digital objects over an extended period of time. LTV works by verifying the digital signature of a digital object at regular intervals to ensure that the object has not changed since its original creation. The plugin uses the mLTV service provided by Multicert to extend the validity of document's digital signatures by applying timestamps and integrated status information to the content of a PDF document. The mLTV is an external webservice provided by Multicert. As such it requires a service contract from Multicert to be used.	Önskvärt	
	Digital Signature Expiry Date Extractor	The Digital Signature Expiry Date Extractor plugin obtains expiration dates from qualified digital signatures embedded in PDF files and saves them in metadata. The extracted dates provide a convenient means of sorting and selecting files based on signature expiration date, enabling proactive management of signature validity, and ensuring the ongoing authenticity of the digital content in combination with the Digital Signature Extension Plugin.	Önskvärt	
	Digital Signature Validator	The Digital Signature Validator performs a comprehensive evaluation of embedded digital signatures within files to ascertain their validity. This task supports verification of digital signatures in the following file formats: PDF, Microsoft Office Formats (such as .docx, .xlsx, and .pptx), and OpenDocument formats (such as .odt, .ods, and .odp). Upon completion of this process, three significant outcomes are achieved: Firstly, the outcome of the verification is recorded and documented in a PREMIS event. Secondly, the extracted digital signatures are securely stored within the Archival Information Package (AIP) in the designated 'metadata/other' folder. Finally, the original files with the digital signatures removed are securely stored in a new representation within the AIP.	Önskvärt	
eArchiving	E-ARK AIP Manifest Updater	For performance reasons, RODA does not keep updated versions of the METS manifest prescribed by the E-ARK AIP specification. The E-ARK AIP Manifest Updater plugin creates, or updates METS manifest files based on AIP information found in the storage layer.	Önskvärt	
	E-ARK AIP Validator	The E-ARK AIP Validator plugin provides a comprehensive evaluation to ensure that AIPs meet the requirements outlined in the E-ARK specification, version 2.0.4.	Önskvärt	
	E-ARK DIP Manifest Updater	For performance reasons, RODA does not keep updated versions of the METS manifest prescribed by the E-ARK DIP specification. The E-ARK DIP Manifest Updater plugin creates, or updates METS manifest files based on DIP information found in the storage layer.	Önskvärt	

Område	Plugin	Beskrivning (en)	Behövs för Mellanarkiv	Kommentar
Maintenance	Find and Replace	The Find and Replace plugin is a powerful tool for quickly and efficiently updating the descriptive metadata of multiple records at once. Whether you need to make a global change to your metadata or simply correct a typo, this plugin can help you save time and improve the consistency and accuracy of your metadata. With the Find and Replace plugin, you can select all records or a specific set of records in search and replace the content of descriptive metadata attributes. You can choose to replace the entire attribute or search for specific words or phrases to replace. Advanced users can also use regular expressions to match on the content of an attribute, extract groups, and use them in the replacement text.		
	Permissions Manager	The Permissions Manager plugin enables users to update permissions of selected AIPs by adding or removing groups, users or setting specific permissions such CREATE, READ, UPDATE, DELETE, GRANT.		
Risk assessment	Incomplete File Format Detector	The Incomplete File Format Detector plugin verifies if a file has a format name attribute adequately filled with a MIME type, PRONOM ID, or a Format designation. If this information is missing, it creates a new risk entry titled 'File(s) not comprehensively characterized' in the Risk Register and assigns the file in question to that risk. Furthermore, by creating a new risk entry for files with missing format information, the plugin helps organizations track and manage the risks associated with incomplete file format information. This can help organizations prioritize their efforts and resources for addressing these risks and improving the completeness and accuracy of their digital assets representation information.		
	Incomplete Representation Information Detector:	The Incomplete Representation Information plugin is a powerful tool that can help ensure the completeness and accuracy of the representation information for digital files. The plugin verifies that each file has its format described in the Representation Information Network, and that it is marked as a known or supported format. Additionally, the plugin checks that all configured attributes for Archival Information Packages (AIPs), Representations, and Files are accurately described in the Representation Information Network.		
	Representation Information Broken Links Verifier	The Representation Information Broken Links Verifier plugin is a valuable tool for verifying the accuracy and accessibility of external links referenced in Representation Information Records. Representation Information can point to external links for additional information, but these web pages can cease to exist, leading to broken links and potentially lost or inaccessible information. The plugin checks if the links being pointed by the Representation Information Records are still accessible and		

Område	Plugin	Beskrivning (en)	Behövs för Mellanarkiv	Kommentar
		optionally removes them in case they are broken. By identifying broken links and removing them, the plugin helps ensure that the Representation Information is accurate and accessible for future use and preservation. Ensuring the accuracy and accessibility of Representation Information is critical for the long-term preservation and usability of digital assets. Broken links can cause confusion, errors, and lost information, making it difficult or impossible to interpret or use digital assets correctly.		
Validation	Digitization profile validator for TIFF images	This plugin checks if the images produced through digitization processes meet the expectations defined in a digitization profile. The digitization profile typically outlines rules and guidelines for minimum DPI resolution, compression type, photometric interpretation, and other technical aspects of the image file format. The goal of a digitization profile validator is to ensure that digital images are of sufficient quality and have the necessary technical attributes for preservation and accessibility over time.		
	Format Validator for PDF/A	The Format Validator for PDF/A is a specialized tool designed to ensure compliance with the ISO-standardized Portable Document Format (PDF) specification for archival and long-term preservation of electronic documents. This plugin validates PDF files against the PDF/A specification, which imposes restrictions and requirements on the 'base' PDF standards, including PDF 1.4 for PDF/A-1 and ISO 32000 for PDF/A-2 and PDF/A-3, as well as a set of additional third-party standards. The PDF/A specification prohibits features that are not conducive to long-term archiving, such as font linking, and encryption.		

Lösningalternativ 2 visas nedan. De flesta, men inte alla, RODA komponenter och plugins är med i diagrammet.

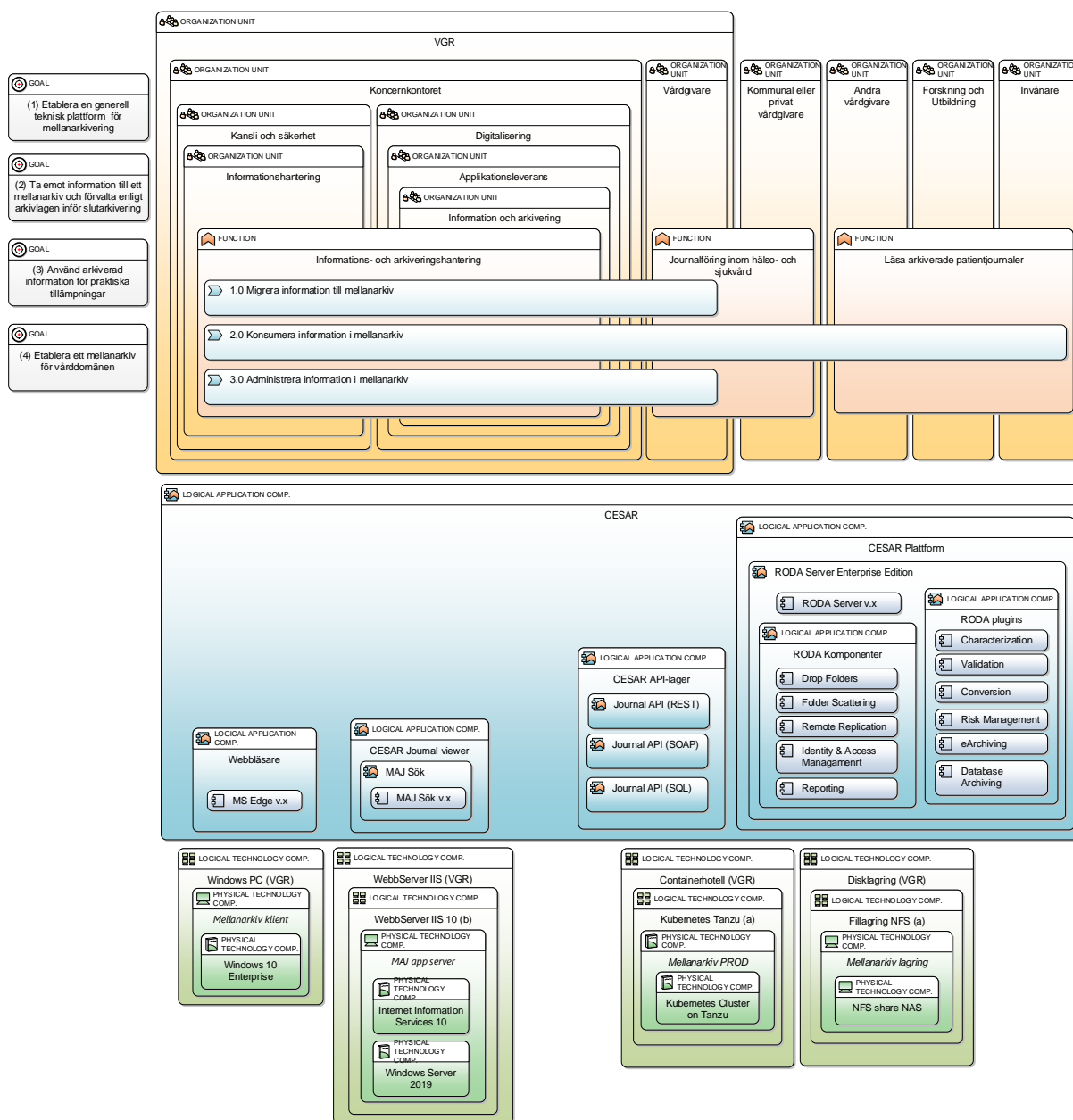


Diagram 4.18 - Lösningalternativ 2 - RODA Enterprise

#### 4.1.7.2 Utvärdering av lösningalternativ 2

Lösningalternativ 2 tillför betydligt mer funktionalitet men också betydligt högre kostnad i form av mjukvarulicenser samt löpande underhållsavgifter för dessa.

Trots att en del tillägsprogram och plugins kan anses vara önskvärda vid etablering av ett mellanarkiv, finns det ingen anledning att börja med en Enterprise licens - denna kan alltid införskaffas vid ett senare tillfälle eller så kan individuella licenser inhandlas.

Det går bra att påbörja etableringen av ett mellanarkiv med RODA Community för att senare byta till RODA Enterprise om det behövs.

#### 4.1.7.3 Riskhantering av lösningsalternativ 2

De risker som har identifierats i kap. 2.4.5 kräver att lösningen har säker inloggning, säker behörighets- och åtkomstkontroll, loggning samt hög driftstillförlitlighet (high availability). Detta förväntas kunna uppnås med det valda lösningsalternativet.

## 4.2 Val av lösningsalternativ

I stället för att utföra ytterligare en utredning på eget bevåg, har projektet beslutat att gå samma väg som Försäkringskassan, Tullverket och Region Skåne med beslut att bedriva egenutveckling av mellanarkiv samt att basera mellanarkivet på teknikplattformen RODA. Beslutet har förankrats i styrgrupp.

Antagandet är att om VGR i linje med andra myndigheter och regioner väljer att själv utveckla en lösning för mellanarkivering, kommer detta att ske agilt och inkrementellt enligt gällande SAFe metodik. Det antas därför inte tillföra något större värde att försöka detaljutreda alla tekniska eller arkitekturella alternativ i förväg, vilket annars är kutym i vattenfallsprojekt.

I stället förlitar vi oss på att olika alternativ beträffande både funktionalitet (features) och IT-arkitektur utreds och hanteras löpande i projektet som en naturlig del av både PI-planering och sprintar. Lösningens arkitekturen blir därmed inte statisk utan kommer att revideras löpande för att reflektera detta.

För valet mellan lösningsalternativ 1 och 2 är det ett beslut som inte behöver tas just nu; utan det kan skjutas på framtiden. Därför rekommenderas lösningsalternativ 1 (initialt).

## 5 Målarkitektur för valt lösningsalternativ

### 5.1 Verksamhet

#### 5.1.1 Organisation och funktion

Följande organisatoriska enheter deltar i funktionerna informations- och arkiveringshantering, journalföring inom hälso- och sjukvård samt att läsa arkiverade patientjournaler:

1. VGR/Koncernkontoret/Kansli och säkerhet/Informationshantering
2. VGR/Koncernkontoret/Digitalisering/Applikationsleverans/Information och arkivering
3. Alla vårdgivande myndigheter/förvaltningar inom VGR
4. Kommunala eller privata vårdgivare inom Västra Götalandsregionen
5. Andra vårdgivare utanför Västra Götalandsregionen
6. Forskare<sup>5</sup>
7. Invånare

I funktionerna utförs tre huvudprocesser:

- *1.0 Migrera information till mellanarkiv.* Här deltar organisatoriska enheterna 1-3 ovan. Processen utförs ofta som en del av *avvecklingsprocessen*, men kan även utföras på andra sätt (t.ex. löpande). Hos kommunala vårdgivare finns det även en arkivfunktion, men de har eget arkiveringsansvar och kan ha egna mellanarkiv. Hos privata vårdgivare som använder Millennium finns det ett arkivansvar som kan vara kortare än det offentliga.
- *2.0 Konsumera information i mellanarkiv.* Här deltar samtliga enheter ovan, men processtegen för att läsa arkiverade journalhandlingar kan variera. T.ex. behöver forskare kunna läsa många journaler medan invånare endast får läsa den egna journalen.
- *3.0 Administrera information i mellanarkiv.* Här deltar samma enheter som i *1.0 Migrera...*

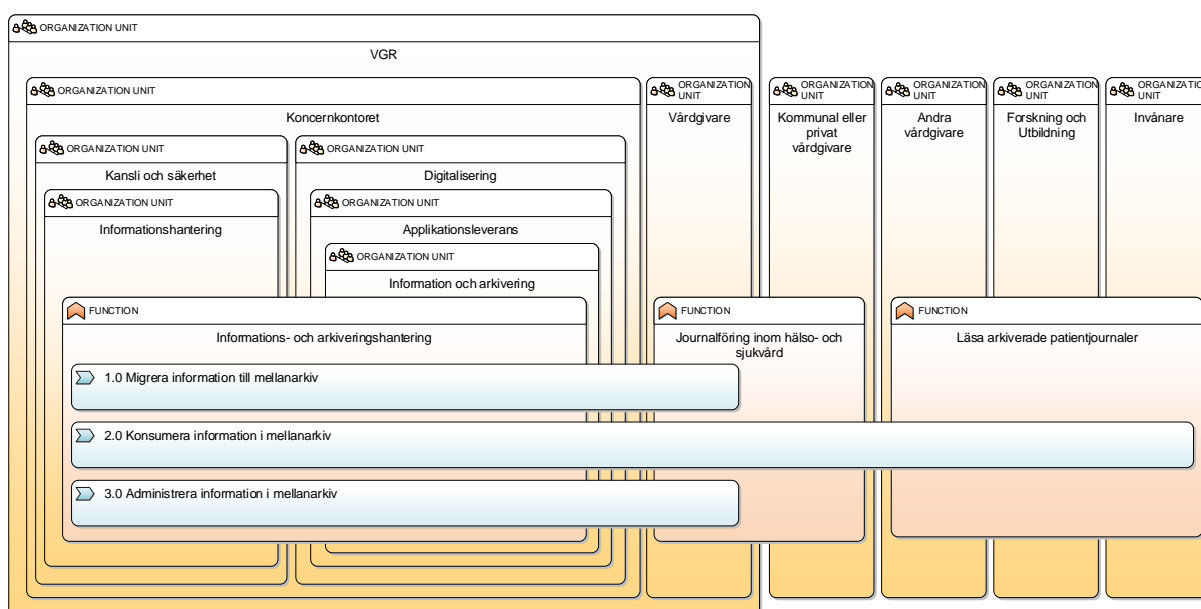


Diagram 5.1 - Verksamhetsfunktioner (target)

Följande roller deltar i processerna ovan:

<sup>5</sup> Forskning och utveckling och Invånare är egentligen inte organisatoriska enheter, men visas som detta i diagrammet för att indikera att de ändå deltar i vissa processer.

- *Arkivarie*: förvaltar arkivet och information som finns i arkivet. Ska kunna utföra samtliga funktioner i mellanarkivet. Förvaltare ska kunna göra journalförändring/-förstöring.
  - Arkivarie ska kunna se all information som inte är sekretessbelagd inom sin myndighets arkiv eller de myndigheter som arkivarien har ansvar för.
  - Arkivarien ska inte kunna se alla journalhandlingar om arkivarien inte fått medarbetaruppdrag för att kunna ta del av enskilda journaler. Det bör styras av samma åtkomsträttigheter som andra användare har. Det behöver finnas en behörighetsroll som kan göra journalförstöring, som kan läggas till den person/funktion som bäst kan utföra den.
  - Arkivarien ska kunna märka om AIP med metadata som beskriver handlingarna.
  - Arkivarien ska kunna märka AIP med metadata slutarkiv och därmed lämna över ansvaret för handlingarna till Arkivmyndigheten.
- *Administratör*
  - *Systemadministratör* administrerar och livscykelhanterar system
  - *Vårdadministratör* lämnar ut journal på begäran (t.ex. från medborgare), loggranskning, användningsstatistik och övriga rapporter.
  - *Arkivadministratör* ska kunna söka och skriva ut handlingar för expediering till medborgare utanför organisationen och internt.
- *Vårdpersonal*: omfattar både klinisk personal som t.ex. läkare och sjuksköterskor, men även vårdadministrativ personal. Ska kunna söka, läsa och skriva ut arkiverade patientjournaler.
- *Produktägare*: företräder verksamheten för att prioritera systemunderhåll och roadmap
- *Loggranskare*: utför två saker: dels granskning av aktivitetsloggar enligt PDL, dels utlämning av aktivitetsloggar, t.ex. på begäran av IVO
- *Forskare*: utför statistisk analys på arkiverad information som del i sitt forskningsuppdrag
- *Invånare*: människor huvudsakligen bosatta i Västra Götaland, men det finns undantag, t.ex. när invånare från andra regioner har vårdats här eller har flyttat till annan region
- *Handläggare*: administration personal på IVO som handlägger inkomna ärenden angående vård
- *Informationsägare*: formellt ansvarig för den information som arkiveras. Ansvarig för att informationsklassificering finns
- *Arkivansvarig*: ska kunna tilldela personal rollen arkivarie.

Inom VGR finns *vårdenheter med särskild sekretess*. Arkiverade journaler från dessa enheter är sekretessbelagda och deras existens får inte framgå av sökningar som utförs av obehörig personal.

NPÖ har en egen accesslogg som granskas via Ineras LoggranskningsGUI.

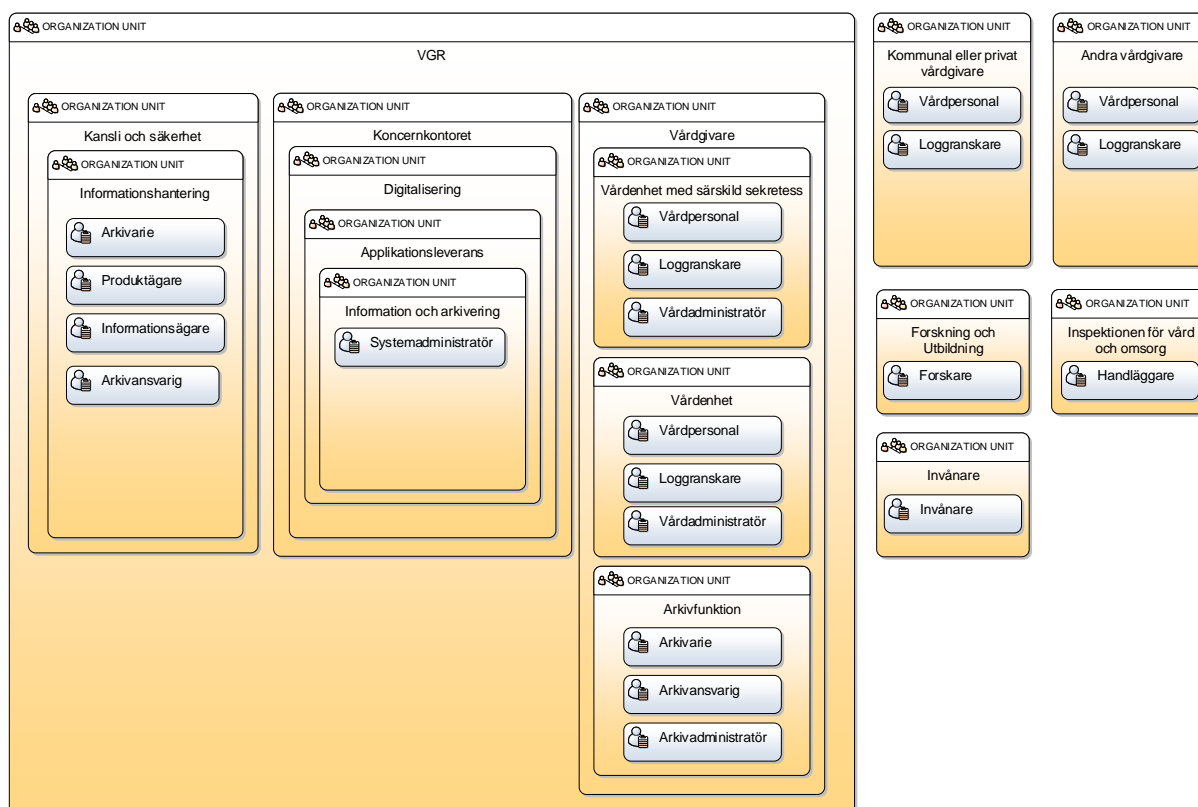


Diagram 5.2 - Organisationskarta (target)

## 5.1.2 Processbeskrivning

Processkartan delas in i tre områden:

- Huvudprocess 1.0 handlar om att förbereda och paketera information för arkivering, och det utförs som regel i ett avvecklingsprojekt eller löpande av en arkivbildande myndighet<sup>6</sup>
- Huvudprocess 2.0 handlar om att förvalta ett *generellt* mellanarkiv enligt OAIS standarden. Här ingår administrativa uppgifter som att skapa avtal för inleverans inkl. specifikation för fysisk lagring av data och metadata i arkivet.
- Huvudprocess 3.0, 4.0 och 5.0 handlar om använda, administrera och förvalta mellanarkiverade patientjournaler enligt relevanta regulatoriska krav; t.ex. PDL.

Det finns även en avslutande process för när information avslutar sin livscykel i mellanarkivet och ska flyttas till ett slutarkiv. Denna process är i princip identisk med huvudprocess 1.0, men där SIP filer ska tillverkas enligt slutarkivets specifikationer.

Huvud- och delprocesser utreds/modelleras inte i närmare detalj utan kravställs i det agila arbetet med user-stories<sup>7</sup>. Om/när behov uppstår kommer delprocesser att modelleras i framtida revisioner av ADD.

<sup>6</sup> I det här projektet kommer journalhandlingar från Melior och Asynja Visph att arkiveras i separat(a) avvecklingsprojekt

<sup>7</sup> Detta görs i Azure Devops

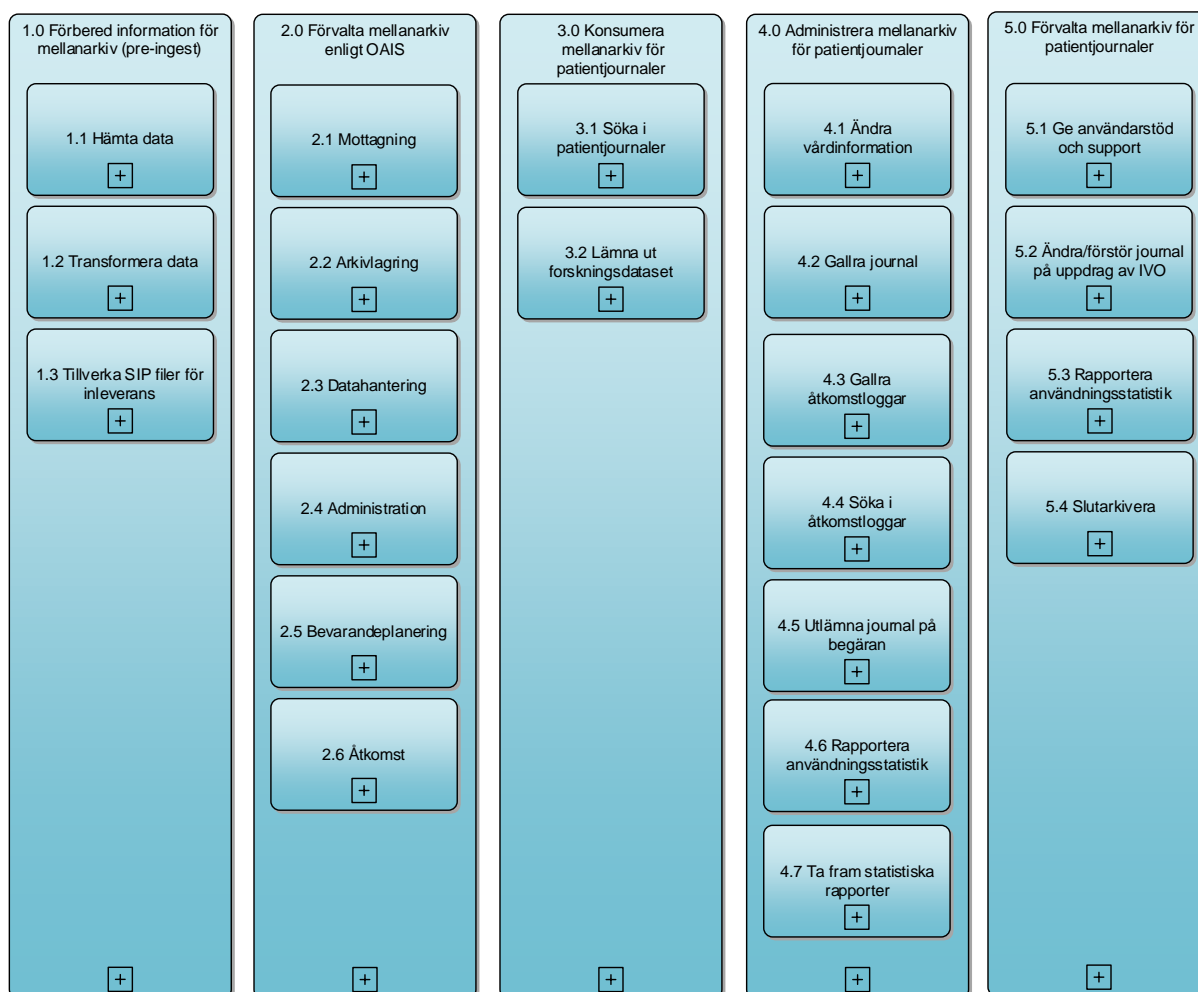


Diagram 5.3 - Processkarta (target)

## 5.2 Informations-/dataarkitektur

Mellanarkivet ska använda VGR:s referensinformationsmodeller för att skapa en enhetlig struktur för informationen i Vårddomänen. De tillämpade modellerna ligger sedan till grund för hur informationen hanteras i olika paketeringar, lagring samt vid konsumtion.

Mellanarkivet ska använda följande riktlinjer för fysisk datalagring:

- E-Ark CSIP standarden [31] som gemensamt format för alla informationspaket i arkivet
- E-Ark SIP [32], AIP [33] och DIP [34] standarder som format för inleverans, förvaltning och utleverans av informationspaket
- E-Ark eHealth1 standarden [30] som format för arkiverade patientjournaler

Detta innebär att informationspaket som hanteras av mellanarkivet (SIP/AIP/DIP) består av en folderstruktur innehållande (oftast ostrukturerad) data i diverse format (t.ex. PDF, MS Office, Dicom<sup>8</sup>, Jpeg, osv.), metadata i XML format som följer META och PREMIS standarder, samt strukturerad data som lagras i XML format som följer EAD standarden.

För att kunna exponera (via API) strukturerad data utöver arkiverade journalhandlingar i dokumentformat, kan FHIR resurser lagras i beskrivande metadata dokument (en FHIR resurs per XML dokument) enligt EAD standarden. Detta behöver utredas under parallellt pågående avvecklingsprojekt för Melior och i nära samarbete med Informatik.

<sup>8</sup> I nuläget mellanarkiveras Dicom filer i BFR och det finns inget behov att ändra på detta. Men möjligheten finns dock att arkivera Dicom filer i mellanarkivet om behovet uppstår i framtiden.

## 5.2.1 Informationsmodell för arkiverade patientjournaler

### 5.2.1.1 SIP modell

Ett SIP paket kan innehålla journaler/journalhandlingar och metadata för många patienter.

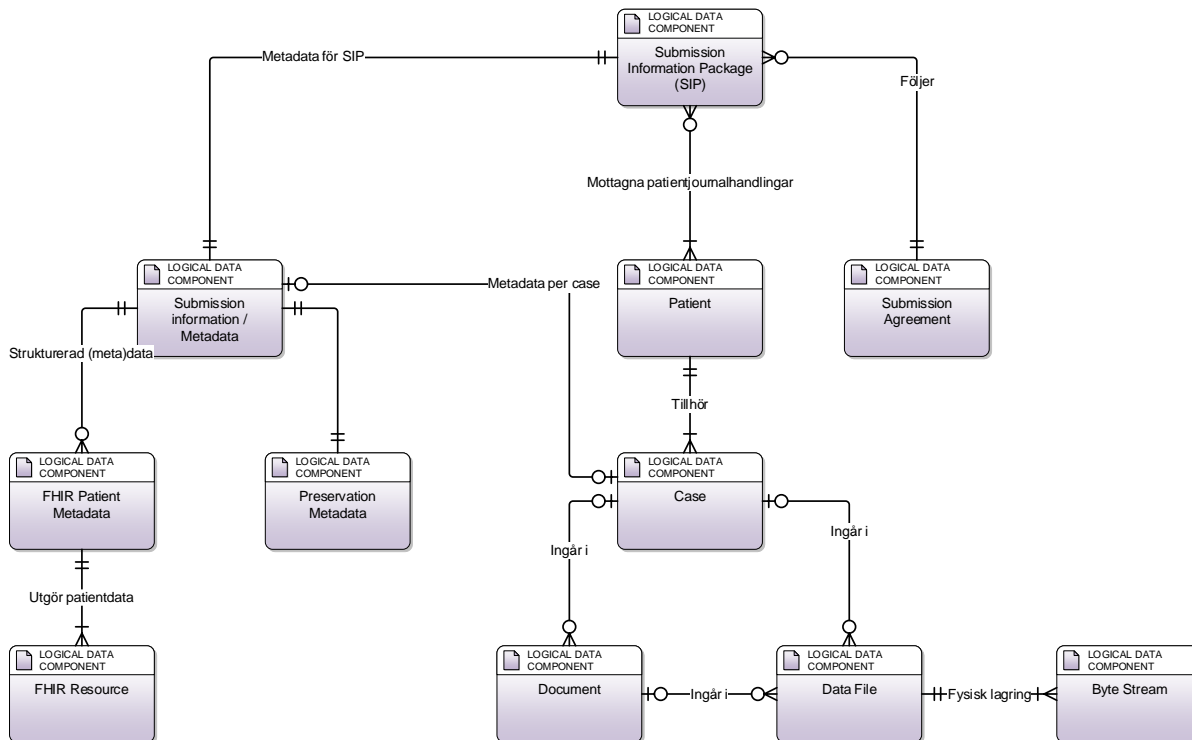


Diagram 5.4 - Informationsmodell för inleverans

### 5.2.1.2 AIP modell

I mellanarkivet rekommenderas att endast lagra journalhandlingar från en arkivbildande aktör/myndighet och för en patient per arkivinformationspaket (AIP). Detta innebär att SIP paket kan behöva brytas isär vid inleverans. Modellen för AIP är nästan identisk med viss skillnad i relationerna.

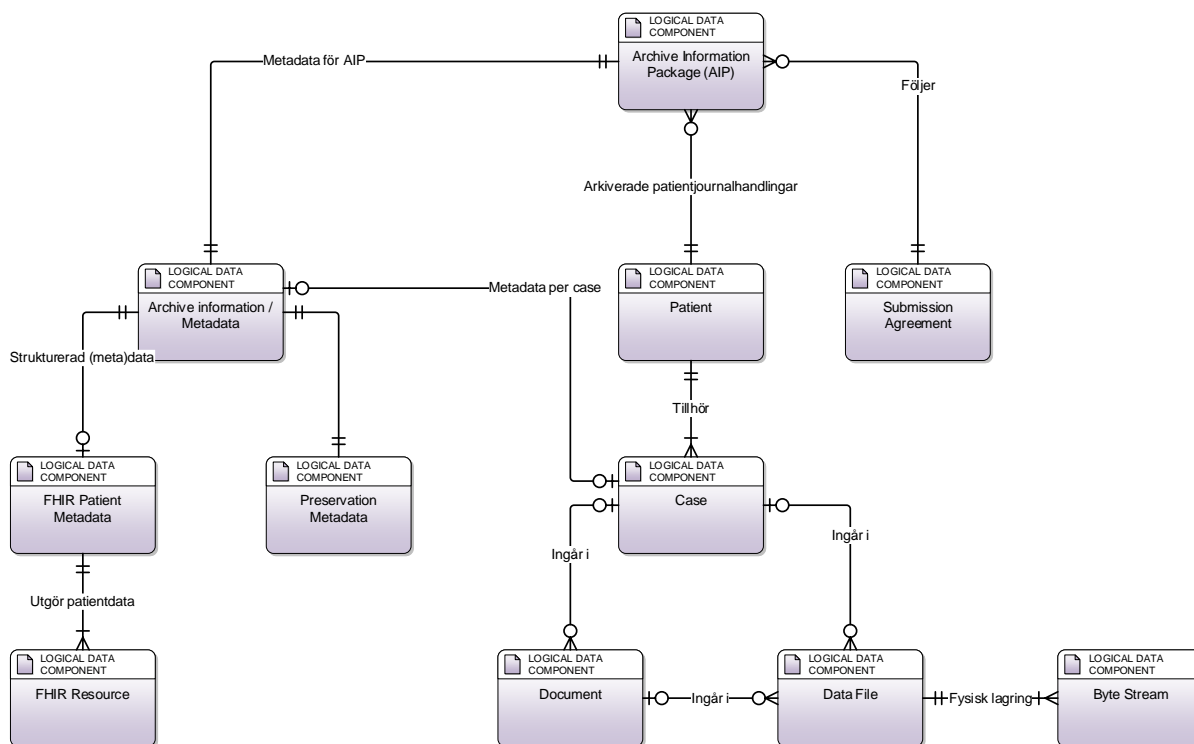


Diagram 5.5 - Informationsmodell för mellanarkivet

### 5.2.1.3 DIP modell

Vid utlämning av information från mellanarkivet sker detta med ett paket som i princip är identisk med ett AIP, men där onödigt metadata har rensats bort. Vad som är onödigt är en subjektiv bedömning och utreds inte närmare här.

### 5.2.2 Förteckning över informationsobjekt

Informationsmodellen är en fysisk lagringsmodell, och många logiska komponenter är i praktiken foldrar (mappar) i ett filsystem<sup>9</sup>. Detta markeras i listan nedan med [+] för att indikera att komponenten realiserar som en folder och inte en fysisk fil.

Logisk Datakomponent	Beskrivning
AIP	Archive Information Package. En folderstruktur som innehåller data och metadata och förvaltas i ett arkiv. Målarkitekturen använder E-Ark CSIP [31] och AIP [33] standarder.
DIP	Dissemination Information Package. En folderstruktur som innehåller data och metadata som lämnas ut från ett arkiv. Målarkitekturen använder E-Ark CSIP [31] och DIP [34] standarder.
SIP	Submission Information Package. En folderstruktur som innehåller data och metadata som lämnas in till ett arkiv. Målarkitekturen använder E-Ark CSIP [31] och SIP [32] standarder.
Patient [+]	En folder i ett AIP/SIP/DIP som innehåller data och metadata för en patient
Case [+]	En folder i ett AIP/SIP/DIP som innehåller data och metadata för en journal eller journalhandling för en patient
Document [+]	En folder i ett AIP/SIP/DIP som innehåller en eller flera datafiler som ingår i en journal eller journalhandling för en patient
Data File	En datafil som innehåller en eller flera byteströmmar (t.ex. video- och audioströmmar i en digital videofilm)
Byte Stream	En binär sekvens som tolkas av mjukvara

<sup>9</sup> Alla foldrar har inte tagits med i modellen. T.ex. ska *Submission Agreement* egentligen ligga i en mapp som heter *Documentation*.

Submission Agreement	Dokumenterar villkor för arkiveringsprocessen; t.ex. vilka parterna är och vilka format som ska användas för metadata.
Archive Information / Submission Information metadata	Metadata som tillämpas för hela informationspaketet
Preservation metadata	Metadata relevant för hantering i arkivet
FHIR Patient Metadata	Innehållsförteckning över FHIR resurser
FHIR Resource	FHIR resurs

Tabell 5.1 - Data Component Catalog

### 5.2.3 Tillämpade kodverk

E-Ark eHealth1 standarden använder FHIR resurser, som i sin tur använder SNOMED CT och ICD. Huruvida dessa kodverk kan användas är dock även beroende av om källsystemet använder dem.

Kodverk (namn)	Tillämpning
SNOMED CT	SNOMED (Systematized Nomenclature of Medicine), är ett system för systematisk representation av medicinsk terminologi som täcker alla områden av klinisk information som sjukdomar, undersökningsfynd, diagnostiska metoder, behandlingsmetoder, mikroorganismer, läkemedel etc.
ICD	International Statistical Classification of Diseases and Related Health Problems (ICD) är Världshälsoorganisationens (WHO:s) klassificeringssystem för olika diagnoser. ICD är den accepterade metoden att klassificera sjukdomar för epidemiologiska, statistiska och diagnostiska syften.

Tabell 5.2 – Tillämpade kodverk, Målarkitektur

## 5.3 Applikationsarkitektur

### 5.3.1 Informationssystem och masterdata

På övergripande nivå hanterar mellanarkivet tre sorters logiska informationsobjekt: SIP, AIP och DIP. SIP och DIP är endast relevant vid inlämning respektive utlämning. Förvaltning av data i mellanarkivet utgår från AIP.

Den valda applikationsplattformen för mellanarkivet har ett rikt API som exponerar all data och metadata som finns i AIP; både för läsning och skrivning (med rätt behörighet). Ett API-lager etableras för att kunna implementera olika gränssnitt och informationsmodeller, som sedan kan konsumeras av olika applikationer både internt och externt (via RTjP).

Information					
	SIP	AIP	DIP	Journal-API	
Applikation					
Hantera SIP	●				Pre-Ingest
Hantera AIP	○	●			CESAR plattform
Hantera DIP		○	●		CESAR API-lager
Hantera Journal-API		○		●	CESAR API-lager
				○	MAJ Sök

Tabell 5.3 - Informationshanteringsmatris

### 5.3.2 Applikationsförteckning

CESAR utgörs av webbapplikationen RODA Community Edition, ett API-lager och front-end applikationen MAJ Sök för att titta på arkiverade journaler.

CESAR API-lager används för att konvertera innehållet i arkivpaket till lämpliga datastrukturer för konsumtion av andra applikationer. Följande användningsfall finns:

1. Initierat från *MAJ Sök*: utför sökning i arkivet via REST gränssnitt och returnera resultat för en patient bestående av journaldelar enligt t.ex. VGR RIM.
2. Initierat från *NPÖ* eller *1177 Journalen* via RTjP: utför sökning i arkivet via SOAP gränssnitt och returnera resultat för en patient enligt tillämpade nationella tjänstekontrakt; t.ex. GetCareDocumentation.
3. Initierat från *IRIS*: utför sökning i arkivet via SQL gränssnitt (JDBC/ODBC) och returnera bulk-resultat för många patienter bestående av journaldelar som SQL resultat-set.

De två första fallen realiserar genom egenutveckling av REST- och SOAP tjänster.

För att exponera ett SQL gränssnitt till arkivdata som kan exportera till forskningsplattformen *IRIS* antas en egen instans av datavirtualiseringsplattformen *Denodo* kunna användas. Denna ansluter till ett eller flera av RODA plattformens API:er<sup>10</sup>.

<sup>10</sup> Enligt systemdokumentationen för RODA exponeras alla arkivfunktioner och all arkivdata och metadata via API

För läsning av arkiverade patientjournaler via uthopp från *Millennium* krävs en uppgraderad *MAJ Sök* applikation, som ansluter till *CESAR API-lager* för att läsa (arkiverade) journalhandlingar. Alla specialiserade åtkomst- och behörighetskontroller som krävs för att följa PDL hanteras av *CESAR API-lager*. Här ingår t.ex. kontroll av medarbetaruppdrag, spärrar, vårdenheter med särskild sekretess osv.

För läsning av arkiverade patientjournaler via *NPÖ* och *1177 Journalen* implementeras ett eller flera nationella tjänstekontrakt som SOAP gränssnitt. Specialiserade åtkomst- och behörighetskontroller som krävs för att följa PDL hanteras av *NPÖ* och *1177*, dvs inte av *CESAR*.

För loggning av aktiviteter mot patientjournaler används *Ineras Loggtjänst*, och deras applikation för granskning och framtagning av aktivitetsrapporter ingår därför även i *CESAR*.

För att kunna generera forskningsdatabaser används *IRIS* plattformen som tas fram i ett annat projekt. Ett SQL gränssnitt för bulköverföring av patientjournaler är under utredning (lila i diagrammet) och kan ev. implementeras med *Denodo* datavirtualiseringsplattformen.

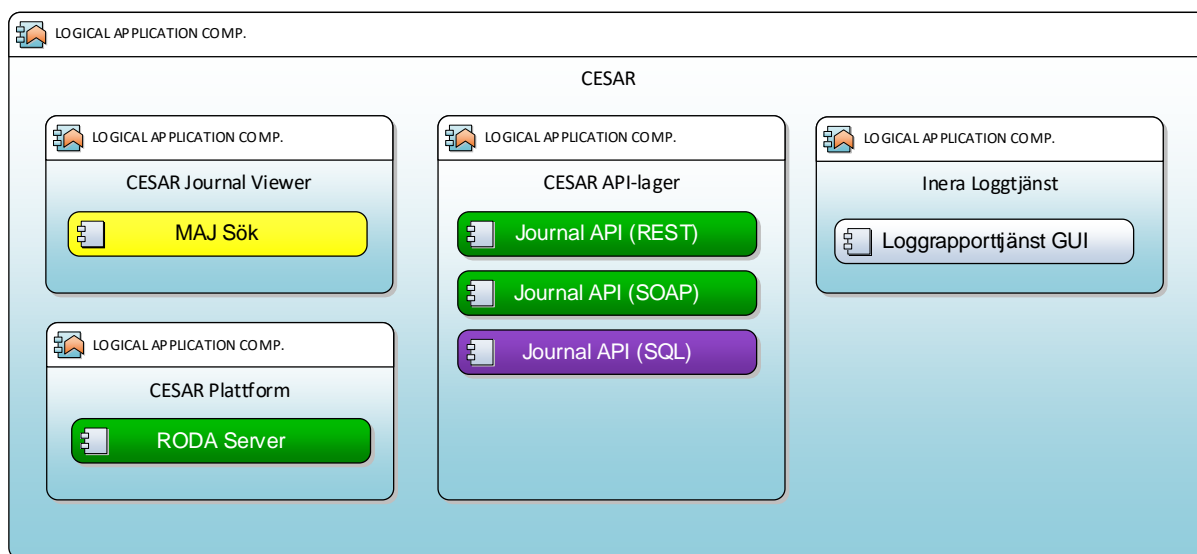


Diagram 5.6 - Applikationer (target II)

Det som är nytt eller ändrat i målarkitekturen är:

- Ny RODA server för att realisera CESAR plattformen
- Nytt API-lager med ett antal implementationer av Journal API
- Ny datavirtualiseringsplattform Denodo<sup>11</sup> för att tillhandahålla Journal SQL gränssnitt
- Uppdaterad MAJ Sök applikation som läser journaler från RODA via API-lager

Applikationerna sammanställs i tabellen nedan.

Fysisk applikation	Logisk applikation	Logisk applikationskategori	Beskrivning
Roda Server v.x	CESAR Plattform	CESAR	Ett OAIS kompatibelt e-arkiv
Journal API (REST) v.x	CESAR API-lager	CESAR	API för att läsa journaler
Journal API (SOAP) v.x	CESAR API-lager	CESAR	API som realiserar ett eller flera nationella tjänstekontrakt för att läsa journaler
Journal API (SQL) v.x	CESAR API-lager	CESAR	Exponerar data via SQL. Huruvida Denodo kan användas är under utredning.

<sup>11</sup> En egen instans rekommenderas, separat från befintlig installation som används främst för dataanalys. Detta för att SLA behov är olika, att BI-bearbetningar inte bör påverka prestanda i mellanarkivet, samt att mellanarkivets datavirtualisering måste kunna livscykelhanteras oberoende av andra system.

MAJ Sök v.x	CESAR Journal Viewer	CESAR	Visar arkiverade patientjournal(er); t.ex. efter uthopp från operationellt journalsystem
Loggrapporttjänst GUI	Inera Loggtjänst	CESAR	Används av loggranskare för att granska aktivitetsloggar

Tabell 5.4 - Applikationskomponenter

### 5.3.3 Funktionell vy

Den funktionella vyn visar kopplingar och därmed beroenden mellan aktörer/roller, processer och applikationer. Flera av rollerna återfinns i olika organisationer; t.ex. arkivarier finns både på koncernkontoret och på sjukhusen. Vårdpersonal finns både hos regionala, kommunala och privata vårdgivare. För att förenkla diagrammet har därför organisatoriska enheter inte tagits med.

Diagrammet visar att alla identifierade roller deltar i minst en process, samt att varje process använder minst ett IT-stöd.

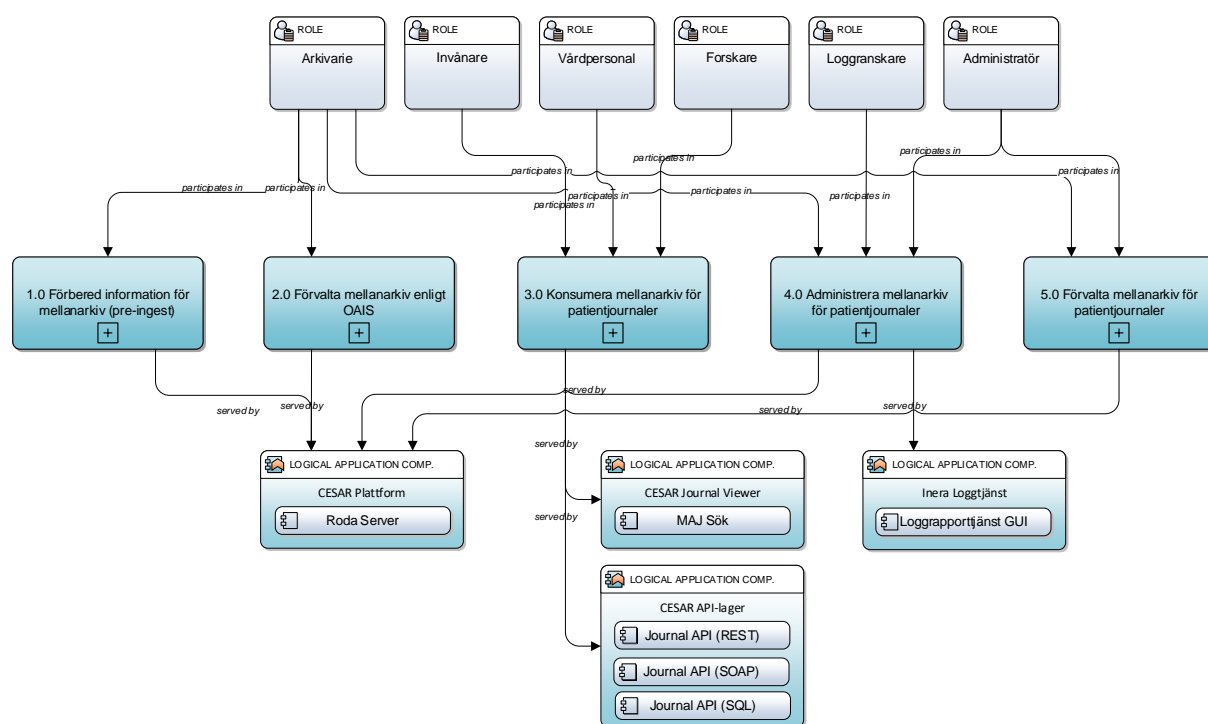


Diagram 5.7 - Applikationsanvändning

Det som är nytt eller ändrat i målarkitekturen är:

- Nya roller: Invånare, forskare
- Ny huvudprocess för pre-ingest
- Ny huvudprocess för förvaltning av OAIS arkiv
- Administration av patientjournaler uppdelat i två huvudprocesser för administration och förvaltning

### 5.3.4 Informationsflöden

Informationsflöden i systemet är betydligt mer komplicerat i lösningsarkitekturen jämfört med baseline. Detta beror i första hand pga införande av ett OAIS arkiv samt stöd för nya användningsfall för andra vårdgivare, forskare och invånare.

Informationsflöden mellan MAJ Sök, Lokal IdP, HSA katalog, Nationell spårertjänst samt Nationell loggtjänst är i princip oförändrade i målarkitekturen.

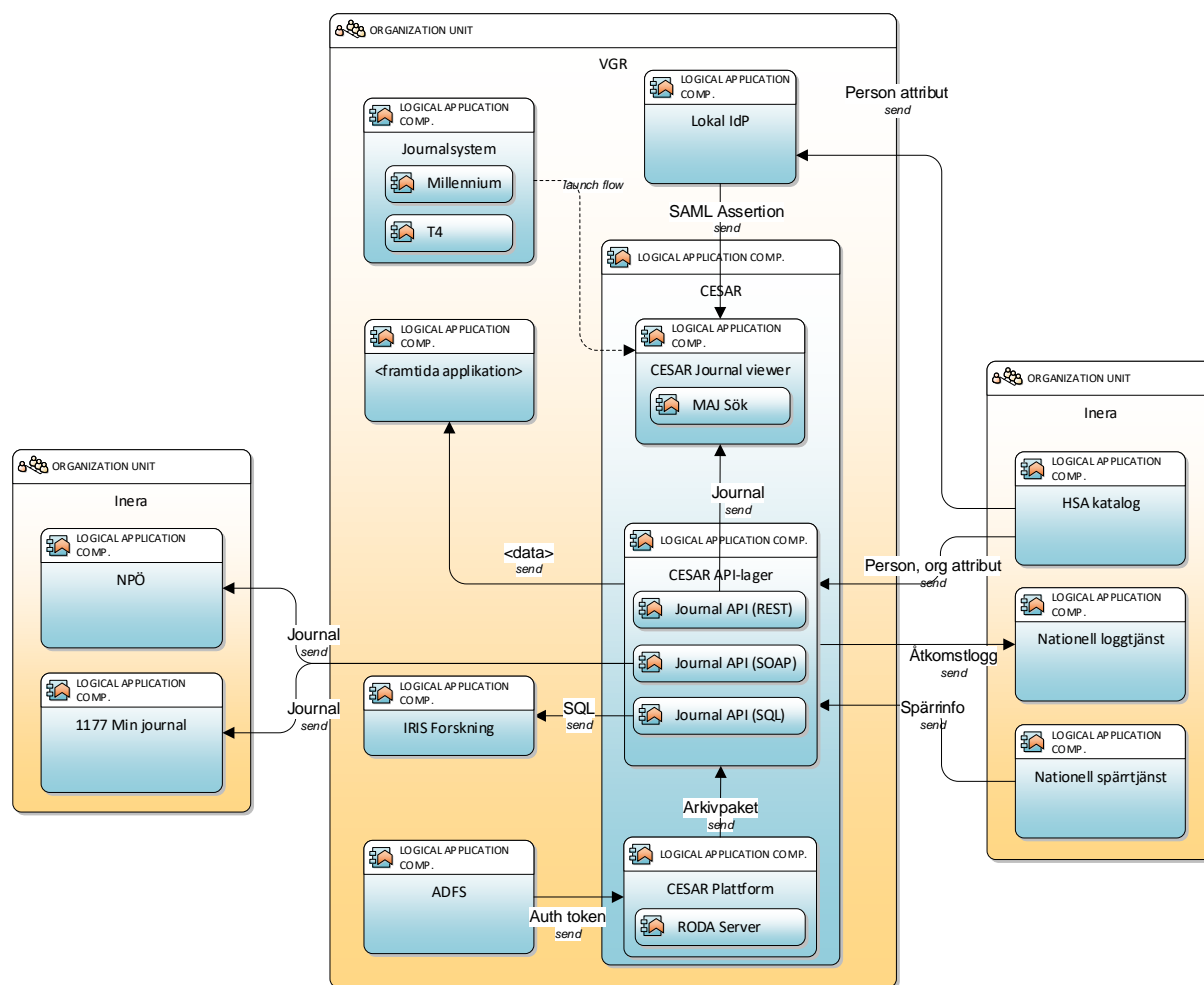


Diagram 5.8 - Informationsflöden

Det som är nytt eller ändrat i lösningsarkitekturen är:

- Nytt journalsystem Millennium ersätter Melior och Asynja Visph för att initiera uthopp till MAJ Sök
- MAJ Sök läser journaler från CESAR API-lager, som hämtar arkivpaket från CESAR plattform
- Hantering av Inera stödtjänster för att säkerställa att PDL följs hanteras av CESAR API-lager
- Nya applikationer som läser journaler från CESAR API-lager: NPÖ, 1177 Journalen, IRIS
- Framtida applikationer som läser journaler från CESAR API-lager: dataanalys, maskininlärning för AI
- Inloggning till RODA server sker med tillitsnivå 3 via ADFS

### 5.3.5 Förteckning över informationsflöden

Informationsflöden redovisas i tabellen nedan. Applikationer som inte redan finns och/eller kan finnas i flera varianter är oidentifierade och namnges i kursiv.

Från Applikation	Till Applikation	Informationsmängd	Beskrivning
CESAR API-lager	1177 Journalen	Journal	Överförs via RTjP
CESAR API-lager	NPÖ	Journal	Överförs via RTjP
CESAR API-lager	<i>Dataanalys</i>	Journal	
CESAR API-lager	<i>Maskininlärning för AI</i>	Journal	
CESAR API-lager	MAJ Sök	Journal	

CESAR API-lager	IRIS	Journaler	Bulköverföring med SQL
Millennium, T4	MAJ Sök	Kontext (användare, patient)	Manuellt uthopp
Lokal IdP	MAJ Sök	Åtkomstbiljett (SAML assertion)	Flödet är förenklat och går delvist via webbläsare. Se diagram 5.20, 5.21, 5.22 i [47] för utförligare beskrivning
RODA Server	VGR ADFS	Inloggning eller åtkomstbiljett	
VGR ADFS	Roda Server	Åtkomstbiljett, attribut (roller) från AD	Flödet är förenklat och går delvist via webbläsare. Se diagram 5.23, 5.24 i [47] för utförligare beskrivning
HSA Katalog	CESAR API-lager	Attribut (medarbetaruppdrag, HSA-ID) för användare och organisation	Används för behörighetskontroll
Nationell spärrtjänst	CESAR API-lager	Information om spärrade journalhandlingar	Överförs via RTjP
CESAR API-lager	Nationell loggtjänst	Aktivitetslogg	Överförs via RTjP

Tabell 5.5 - Informationsflödestabell

Det som är nytt eller ändrat i målarkitekturen beskrivs i föregående avsnitt.

### 5.3.6 Interaktioner

Utreds ej i denna version.

### 5.3.7 Integrationsarkitektur

Målarkitekturen innehåller ett stort antal informationsöverföringar. De som realiseras med hjälp av RTjP visas nedan.

De tre första integrationerna hanteras nu i API-lagret i stället för på applikationsnivå (i MAJ Sök). Nya flöden visas för att överföra journal till annan vårdgivare via NPO och till invånare via 1177 Journalen. I bägge fallen används flera nationella tjänstekontrakt.

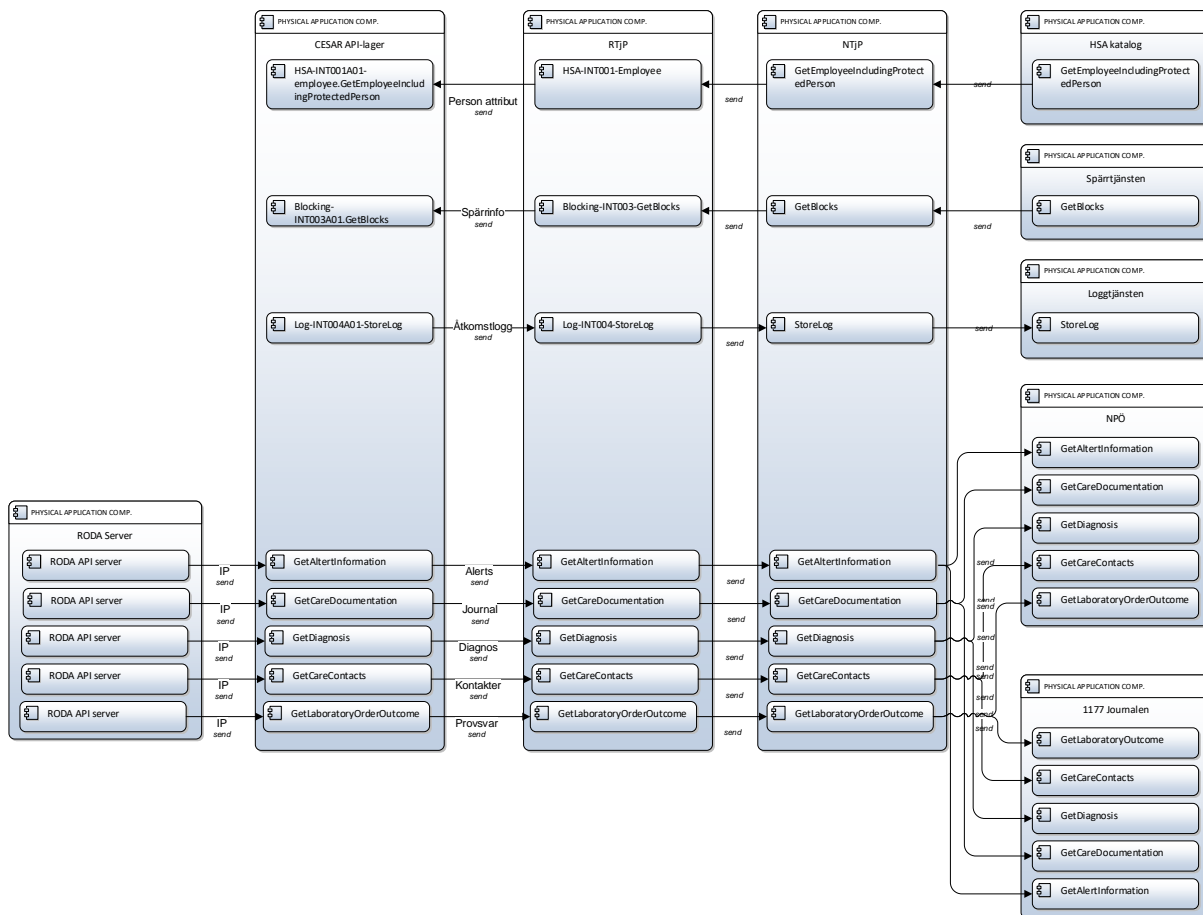


Diagram 5.9 - Integrationsdiagram

**Behöver lägga till engagemangsindex. Hur modellera filter?**

Det som är nytt eller ändrat i målarkitekturen är:

- Nya integrationer för överföring av journal till NPÖ
- Nya integrationer för överföring av journal till 1177 Journalen

5.3.8 Förteckning över integrationskomponenter

Fysisk Applikation	Avtal/Integration	Typ	Informationsmängd
RODA Server	RODA API server	REST	AIP
CESAR API-lager	GetAlertInformation	SOAP	Retunerar uppmärksamhetsinformation för en patient
CESAR API-lager	GetCareDocumentation	SOAP	Retunerar hälso- och sjukvårdsdokument för en patient
CESAR API-lager	GetDiagnosis	SOAP	Retunerar registrerade diagnoser för en patient
CESAR API-lager	GetCareContacts	SOAP	Retunerar vårdkontakter som finns dokumenterade för en patient
CESAR API-lager	GetLaboratorOrderOutcome	SOAP	Retunerar multidisciplinära laboratoriesvar för en patient
RTjP	GetAlertInformation	SOAP	
RTjP	GetCareDocumentation	SOAP	
RTjP	GetDiagnosis	SOAP	
RTjP	GetCareContacts	SOAP	

RTjP	GetLaboratorOrderOutcome	SOAP	
------	--------------------------	------	--

Tabell 5.6 – Integrationskomponenter

## 5.4 Infrastruktur

### 5.4.1 Teknisk plattform

CESAR plattform driftas i Docker containrar med tillhörande lagring, och Kubernetes används för orkestrering och elasticitet. Standard NFS lagring används (i stället för SMB) eftersom RODA är en Linux applikation.

CESAR API-lager innehåller 3 olika gränssnitt som är specifika för patientjournaler:

- *Journal API (REST)* konsumeras av *MAJ Sök* och ev. framtida (interna) applikationer på VGR.
- *Journal API (SOAP)* konsumeras av *NPÖ* och *1177 Journalen* via RTjP, och som implementerar relevanta nationella tjänstekontrakt.
- *Journal API (SQL)* är ett (bulk) SQL-gränssnitt som konsumeras av *IRIS* plattformen vid skapande av forskningsdatabaser. Gränssnittet implementeras med *Denodo*.

CESAR Journal viewer är en webbapplikation som utvecklas av VGR och driftas på en standard Webbserver IIS komponent.

Som klient används en standard webbläsare. MS Edge visas i modellen, men även Google Chrome, Mozilla Firefox m.fl. kan användas.

Koppling mellan applikationer och infrastruktur visas nedan.

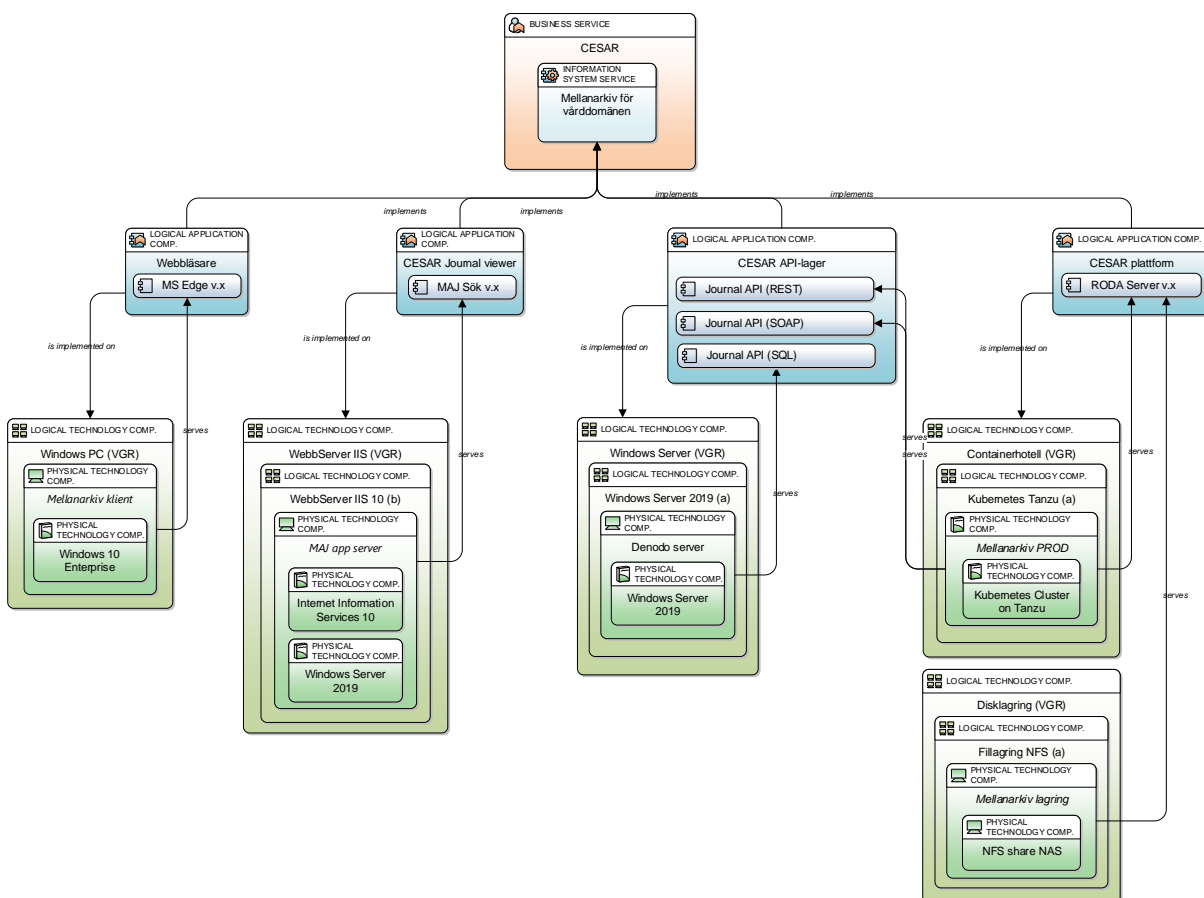


Diagram 5.10 - Processing Diagram (översikt)

Vi saknar kravställning på SLA som kan styra val och design av driftmiljö och infrastruktur:

- Patientjournaldelar. 100 kb (?), 20 filer, det behövs en modell för hur många filer (antal) och lagringsbehov (TB). Tillväxttakt.
- Trafikmodell: antal samtidiga användare, transaktionsstorlek, transaktioner per sekund (peak), IO krav (stora datamängder)
- Tillgänglighet: tid att återställa, återställning av lagring kan ta lång tid. Relativt komplext (många infrastrukturskomponenter) kan ta tid att driftsätta vid återställning.
- 

#### 5.4.1.1 Lokal DEV miljö

DEV miljön är personlig och används för att utveckla och anpassa funktionalitet. Enhetstester utförs av utvecklare för att säkerställa att de nyutvecklade komponenterna fungerar innan det checkas in i gemensamt code-repository.

<diagram>

#### 5.4.1.2 DEV miljö

TEST miljön är gemensam och används för att kompilera/bygga komponenter som hämtas från code-repository och sedan testas tillsammans. Externa gränssnitt realiserar som mock-interface. I TEST miljön utförs (intern) integrationstest, systemtest och regressionstest.

*Integrationstest* säkerställer att systemets olika komponenter fungerar ihop. *Systemtest* verifierar att systemet motsvarar ställda funktionella krav i aktuell leverans. *Regressionstest* utförs för att säkerställa befintlig funktionalitet så att fel inte har blivit introducerade i oförändrade delar av systemet.

I DEV miljön finns 10 docker containers

RODA docker container

- RODA webb applikation
- RODA API server

SOLR Operator container

- SOLR Operator

SOLR Cloud container (elastisk)

- SOLR

Zookeeper Operator

- Zookeeper operator

Zookeeper container (elastisk)

- Zookeeper 3 stk

Siegfried container (identifierar filtyp)

- Siegfried

Clam anti-virus container

- Clam

API-lager container för Inera

- Apache Camel som plattform
- 5-8 nationella tjänstekontrakt
- Egenutvecklat

API-lager internt

- Apache Camel





## 5.5 Säkerhetsarkitektur

Ny informationsklassificering för arkiverade patientjournaler har utförts enligt gällande rutiner [44]. Informationen bedöms ha säkerhetsklass 3.

Informationstillgång (benämning samt beskrivning av tillgången)	Informationsägare	Krav (markera med "X" om tillämplig eller ange)													Informationsklass (ange informationsklass och motivera)				
		Offentlighets- och sekretesslag (2009:400)	Tryckfrihetsförordningen (1949:105)	Arkivlagen (1990:782)	Lag om tillgänglighet... (2016:1937)	Lag om tystnadsplikt vid... (2020:914)	NIS (2018:1174)	Dataskyddsförordningen	Hälsa- och sjukvårdslag (2017:30)	Patient säkerhetslag (2010:659)	Patientdatalag (2008:355)	Lag om sammanhållen vård... (2022:913)	Socialtjänstens föreskrifter... (HSLF-FS 2016:40)	Andra rättsliga eller interna krav	Konfidentialitet (0-4)	Riktighet (0-4)	Tillgänglighet (0-4)	Motivering (motivera klass)	
Mellanarkiverad patientjournal	Johan Flarup	X	X	X	X	X	X	X	X	X	X	X	X	X	Andra rättsliga eller interna krav	3	3	3	Stor påverkan för liv och hälsa ifall det finns brister i tillgänglighet, riktighet och konfidentialitet
Loggar för åtkomst/aktivitet	Johan Flarup	X	X	X			X			X			X	VGR har krav på 10-års bevarandetid	3	2	2	Loggranskning sker genom Ineras loggtjänst och i <b>Roda</b> . Personuppgifter kan påverka en högre klassning på konfidentialitet.	
Administrativ (ej ekonomisk eller persondata) data i systemkonfiguration	Johan Flarup													VGR praxis rörande hantering och möjliggörande av dolda enheter	1	3	2	Om ändring (omedvetet eller medvetet) görs är det allvarligt, men lindrigt om vi kan spåra när och av vem det gjordes. Ex. Risk för fel för enheter med särskild sekretess.	
Arkivpaket	Johan Flarup	x	x											Endast som del av arkivförvaltning (ej vård)	3	3	2	Konfidentialitet kan variera.	

Tabell 5.8 - Informationsklassificering

Ny risk- och sårbarhetsanalys för arkiverade patientjournaler planeras och utförs i projektet.

### 5.5.1 Följsamhet till patientdatalagen (PDL)

Behörighetskoncept för CESAR har utretts separat [45] och sammanfattas nedan. Behov av att följa PDL är olika när arkiverade patientjournaler hanteras av vårdpersonal och arkivpersonal:

En applikation som ansluter till befintlig och framtida IT-tjänsten Mellanarkiv journal för att hämta och presentera en arkiverad journaldel för vårdpersonal måste säkerställa att följande realiserar för att följa PDL:

1. Säker identifiering av användaren med SITHS eID med hjälp av Lokal IdP.
2. Kontroll av ev. spärrar med hjälp av Ineras spärrtjänst
3. Kontroll av användarens medarbetaruppdrag med hjälp av Ineras HSA katalog
4. Loggning av all aktivitet med hjälp av Ineras loggtjänst

Inom VGR finns även ett antal vårdenheter med särskild sekretess. För dessa gäller ytterligare en kontroll, nämligen att:

5. Existensen av arkiverade journaldelar från enheter med särskild sekretess får inte röjas till användare utan medarbetaruppdrag på den specifika vårdenheten

Processägare har fattat följande beslut om hur PDL tillämpas för arkivpersonal [46]: när arkivpersonal förvaltar arkiverad information i mellanarkivet med hjälp av webbapplikationen RODA (eller liknande) gäller följande:

1. Då direktaccess till RODA gäller ett fåtal användare (20?) som skall utföra arkivvård, kan ett undantag från PDL göras enligt nedan beskrivning. För dessa användare gäller mycket strikt behörighetsstyrning. Rutin och process för behörighetshantering skall finnas.
2. Accessloggning skall ske i RODA (behöver inte integreras med Ineras loggtjänst). Rutiner för loggranskning och eventuellt utlämning av loggar skall finnas.
3. Spärr av information enligt PDL behöver inte finnas. Dolda enheter behöver ej vara dolda.

## 6 Gap-analys

### 6.1 Verksamhet

#### 6.1.1 Organisation

Den organisatoriska modellen har utökats ganska mycket i målarkitekturen. Det är i huvudsak befintliga roller i verksamheten som inte var inblandade i TMA (temporärt mellanarkiv).

Target Architecture	Vårdpersonal	Loggransare	Arkivarie	Produktägare	Informationsägare	Arkivansvarig	Administratör	Forskare	Invånare	Handläggare (IVO)	Eliminated
Baseline Architecture											
Förvaltare											Ersätts av andra roller
Vårdpersonal	Match										
Loggransare		Match									
New			Ny roll i målarkitekturen	Ny roll i målarkitekturen	Ny roll i målarkitekturen	Ny roll i målarkitekturen	Ny roll i målarkitekturen	Ny roll i målarkitekturen	Ny roll i målarkitekturen	Ny roll i målarkitekturen	

Tabell 6.1 - GAP analys, Roller (target II)

#### 6.1.2 Processer

Processkartan har utökats mycket i målarkitekturen; dels genom att lägga till huvudprocesser för hela mellanarkivets livscykel, dels genom att utreda vilka delprocesser som behövs av de nya aktörerna. Uppdelning i huvudprocess 4.0 *Administrera* och 5.0 *Förvalta* har gjorts enligt input från produktägare, men dessa skulle kunna slås ihop till en huvudprocess.

Target Architecture	1.0 Förbered information för mellanarkiv (pre-ingest)	2.0 Förvalta mellanarkiv enligt OAIS	3.0 Konsumera mellanarkiv för patientjournaler	4.0 Administrera mellanarkiv för patientjournaler	5.0 Förvalta mellanarkiv för patientjournaler	Eliminated
Baseline Architecture						
2.0 Konsumera information i mellanarkiv						Ersätts av 3.0
3.0 Administrera information i mellanarkiv						Ersätts av 4.0
New	Ny huvudprocess i målarkitekturen	Ny huvudprocess i målarkitekturen	Ny huvudprocess i målarkitekturen	Ny huvudprocess i målarkitekturen	Ny huvudprocess i målarkitekturen	

Tabell 6.2 - GAP analys, Processer (target II)

## 6.2 Informations-/dataarkitektur

Målarkitekturen förespråkar att E-Ark CSIP standard för SIP, AIP och DIP informationspaket används tillsammans med E-Ark eHealth1 som *content information type* för patientjournaler. Detta är således en helt ny datamodell (fysisk lagringsmodell) som ersätter den befintliga strukturen från R7 och det temporära mellanarkivet MAJ.

Det finns dock tydliga kopplingar mellan objekt som patient, journaldel osv, vilket indikeras i GAP matrisen genom "konceptuell match", som färgas gult för att indikera att det sker en ändring i målarkitekturen.

Anledningen till att en del objekt kopplas till flera objekt i målarkitekturen är att det finns utrymme för VGR att anpassa eHealth1 modellen. T.ex. kan "Case" både vara en komplett journal, en ostrukturerad journaldel eller en komposit journaldel med multipla datafiler och inbördes länkar/relationer.

Utformning av anpassningar görs med hjälp av Informatik under införande och kommer att dokumenteras i kommande revisioner av denna ADD.

Target Architecture	Informationspaket (SIPA/IPIDIP)	Informationspaket metadata	Submission agreement	Preservation metadata	FHIR Patient metadata	FHIR Resource	Patient	Case	Document	Data file	Byte stream	Eliminated
Baseline Architecture												
Leveransobjekt	Konceptuell match, men ny struktur											
Patient					Konceptuell match, men ny struktur	Konceptuell match, men ny struktur	Konceptuell match, men ny struktur					
Patientjournal								Konceptuell match, men ny struktur				
Patientjournaldel från pappersjournal								Konceptuell match, men ny struktur	Konceptuell match, men ny struktur			
Patientjournaldel från elektroniskt journalsystem								Konceptuell match, men ny struktur	Konceptuell match, men ny struktur			
Post i patientjournaldel								Konceptuell match, men ny struktur	Konceptuell match, men ny struktur			
Postinnehåll										Konceptuell match, men ny struktur	Konceptuell match, men ny struktur	
Postsamband		Konceptuell match, men ny struktur										
Ansvarig												
Patientrelaterad kontakt					Konceptuell match, men ny struktur	Konceptuell match, men ny struktur						
Organisatorisk enhet					Konceptuell match, men ny struktur	Konceptuell match, men ny struktur						
New	Nytt informationsobjekt		Nytt informationsobjekt	Nytt informationsobjekt								

Tabell 6.3 - GAP analys, Datamodell

### 6.3 Applikationsarkitektur

I målarkitekturen införs två nya applikationer: CESAR plattform och CESAR API-lager

För läsning av patientjournaler uppdateras MAJ Sök. Befintlig applikation för loggranskning förblir oändrad.

Target Architecture	CESAR Journal viewer	Inera Loggrapporttjänst GUI	CESAR plattform	CESAR API-lager	Eliminated
Baseline Architecture					
MAJ Sök	Behöver uppdateras				
Inera LoggRapporttjänst GUI		Match			
Marshal OAIS					Ingår ej i målarkitektur
New			Ny applikation i målarkitekturen	Ny applikation i målarkitekturen	

Tabell 6.4 - GAP analys, Applikationer

### 6.4 Infrastruktur

I målarkitekturen utgår Marshal OAIS med tillhörande infrastruktur. Huruvida Marshal ETL fortsatt används inom avvecklingsprojekt ligger utanför avgränsningen för detta uppdrag.

Nya infrastrukturskomponenter behövs för CESAR plattform och CESAR API-lager.

Target Architecture	Webb klient	Webbserver IIS	Databasserver MS SQL	Kubertenes cluster	Disklagring (NFS)	Windows Server	Eliminated
Baseline Architecture							
Webb klient	Match						
Webbserver IIS		Match					
Databasserver MS SQL			Match				
Disklagring (SMB)							Marshal O AIS
Databas och Webbserver							Marshal O AIS
New				Ny komponent i målarkitekturen	Ny komponent i målarkitekturen	Ny komponent i målarkitekturen	

Tabell 6.5 - GAP analys, Infrastruktur

## 7 Roadmap för genomförande

Roadmapen beskriver övergripande förflyttningen från nuläge (Baseline Architecture) till börläge (Target Architecture) för de olika aspektområdena (Architecture Domains) verksamhet, information, applikation och infrastruktur.

Implementations- och migrationsplanen beskriver på ett mer detaljerat sätt övergången mellan nuläge och börläge. Den är tänkt att kunna användas som ett planeringsunderlag för genomförandefasen för ett projekt.

### 7.1 Strategi för implementation och migration

Lösningens arkitektur kommer att realiseras under 2024-2027 genom sprintar enligt SCRUM, och feature roadmapen kommer att realiseras via PI-planering enligt SAFe. Roadmap för genomförande kan därmed inte sammanställas i detalj, men en indikativ plan kan ändå formuleras för mellanarkivering av journaler och redovisas i kap 7.3 nedan.

### 7.2 Beskrivning av övergångsarkitekturer

#### 7.2.1 Definition av övergångsstadier

Utreds vid behov.

#### 7.2.2 Övergångsarkitekturer för respektive stadie

Utreds vid behov.

#### 7.2.3 Hantering av risker kopplat till implementation och migration

Utreds vid behov

## 7.3 Roadmap diagram

En roadmap för 2024-2025 visas nedan. Inom informationsarkitekturen behöver följande delar utvecklas:

- CESAR IP är beslut om E-ARK CSIP som format för paketstruktur SIP/AIP/DIP i mellanarkivet.
- CESAR Journal CITS är beslut om E-ARK eHealth1 v2.0 som format för patientjournaler i mellanarkivet.
- CESAR Journal SIP (aka ”inleveransspecifikation”) följer av de två föregående punkterna och är inleveransavtalet (submission agreement) för arkivering av journaler. Avvecklingsprojekt för olika journalsystem ska använda denna specifikation vid arkivering.

Inom applikationsarkitekturen behöver applikationer motsvarande varje MVP beskrivning utvecklas:

- MVP I (2024-Q4) - *Första version av CESAR journaltjänst*
  - Hantera paket som följer standarden eHealth1 v2
  - Arkivering av journal enligt utkast till VGR:s journalformat
  - Journaltjänst kan söka fram vilka journaler som finns för en patient i arkivet
  - Konsumtion av journal med hjälp av testklient
- MVP II (2025-Q1) - *Konsumtion av arkiverad journal via webbgränssnitt*
  - Visa alla journalposter för en patient
  - Webbgränssnitt för patientjournal
  - Avgränsningar
    - Ej NPÖ, 1177
    - Ej rättighetskontroll
    - Ej spärllistor
    - Ej loggning till Inera
- MVP III (2025-Q2) - *Säker konsumtion av arkiverad journal via webbgränssnitt*
  - Inloggning
  - Rättighetskontroll
  - Spärrhantering
  - Loggning till Inera
  - Skyddade personuppgifter
  - Avgränsningar
    - Ej NPÖ, 1177
- MVP IV (2025-Q3) - *CESAR Journal har samtliga funktioner som dagens MAJ*
  - Utskrift av journal
  - Utlämnande av journal
  - Uthoppsfunktion
- MVP V (2025-Q4) - *Säker konsumtion av arkiverad journal via nationella tjänster*
  - Integration mot Inera: NPÖ, Journalen (1177)

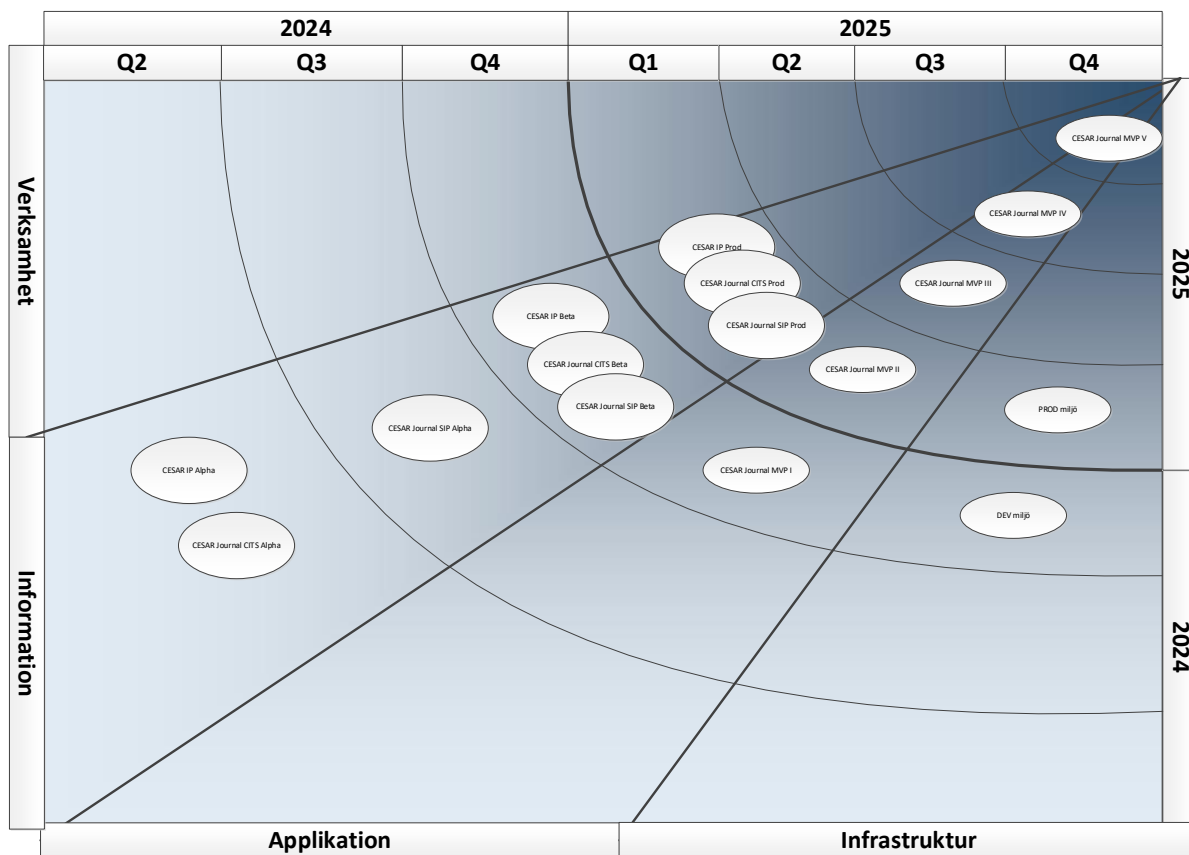


Diagram 7.1 - Roadmap Diagram

## 8 Avvikelser

*Här dokumenteras restpunkter över arkitektur som inte blivit implementerad efter genomförandefasen:*

*Efter genomförandet av implementationen av lösningsarkitekturen kan det uppkommit avvikelser från den planerade lösningsarkitekturen, t.ex på grund av beroenden till andra initiativ som gjort att förutsättningarna för viss arkitektur inte var på plats. Är avvikelserna acceptabel ny målarkitektur så uppdateras målarkitekturbeskrivningen. Är dock avvikelserna att betrakta som tillfälliga så finns en rest som behöver utföras av t.ex. förvaltning eller efterföljande projekt. Beskriv dem restpunkterna här.*