

Henna Mölsä

ASTUVANSALMEN KALLION
ONTOLOGISOINTI
Prosessikuvaus

Opinnäytetyö
Sähköisen asiointin ja arkistoinnin koulutusohjelma


Maaliskuu 2011




MIKKELIN AMMATTIKORKEAKOULU

Mikkeli University of Applied Sciences

KUVAILULEHTI

 MIKKELIN AMMATTIKORKEAKOULU Mikkelin University of Applied Sciences		Opinnäytetyön päivämäärä 11.3.2011
Tekijä(t) Henna Mölsä	Koulutusohjelma ja suuntautuminen Sähköisen asioinnin ja arkistoinnin koulutusohjelma	
Nimeke Astuvansalmen kallion ontologisointi : prosessikuvaus		
Tiivistelmä <p>Opinnäytetyön tavoitteena oli testata ontologisointia ja kuvata tämä prosessi. Tapauskohteeksi valittiin Astuvansalmen kallio ja sen kalliomaalaukset. Ontologisointi koski koko kalliota ja sen ympäristöä - ei ainoastaan kallion kalliomaalauksia. Ontologisointi toteutettiin aineistolähtöisesti ja vahvan semantiikan ontologisointiin pohjautuen CIDOC CRM -mallin avulla. Pääpaino tekemisessä oli oppimisessa tutkimusmenetelmien pohjautuessa reflektiiviseen oppimiseen ja dialogiin. Aiheen toimeksiantajana toimi Mikkelin ammattikorkeakoulun Viva3-hanke.</p> <p>Prosessi koostui aloitus- ja määrittelyvaiheesta, toteutusvaiheesta ja arviointivaiheesta. Toteutusvaihe sisälsi aineiston kartoituksen, kuvien ja kuva-aiheiden kartoituksen, ontologisten suhteiden miettimisen tulkintoihin, kuvailuun, käsitelmallinnukseen ja CIDOC CRM:ään liittyen sekä metatietoskeeman mallinnuksen. Ontologisointiprosessi dokumentointiin ja työvaiheet kuvattiin sanallisesti sekä hyödyntäen toimintamalleja ja kaaviota.</p> <p>Ontologisoinnissa mukana oleminen kehitti omaa asiantuntemusta eli oppimista tapahtui sekä ontologisointiin että kalliokuvaan liittyen. Yhteistyö eri asiantuntijatahojen kanssa oli antoisaa. Omalta osaltani pääsin tavoitteisiini, vaikka hankkeen osalta toiminnallisen prototyypin tekeminen ja testaus jäi tekemättä. Metatietoskeeman osalta päästiin välivaiheeseen, CIDOC CRM -laajennusehdotus tehtiin ja moniperiytyvyyden osalta luotiin ontologiadokumentti.</p> <p>Ontologisointiprosessin kuvaus antaa tietoa esimerkinomaisesti kuinka aineistolähtöinen vahvan semantiikan ontologisointi voi tapahtua ja mitä haasteita tällaiseen ontologisointiin liittyy. Tätä tietoa on mahdollista hyödyntää Mikkelin ammattikorkeakoulussa niin sähköisten palvelujen kehittämisessä kuin koulutuksen järjestämisessä. Tieto on sovellettavissa paikallisesti, kansallisesti kuin kansainvälisesti kulttuuriaineiston ontologisointiprosesseissa.</p>		
Asiasanat (avainsanat) ontologiat, metadata, semanttinen web, CIDOC CRM, prosessit, kulttuuriperintö, kallioitaide, Etelä-Savo, Ristiina, Astuvansalmi		
Sivumäärä 73 s. + 8 liitettä	Kieli fin	URN
Huomautus (huomautukset liitteistä)		
Ohjaavan opettajan nimi Mirja Lopenen	Opinnäytetyön toimeksiantaja Mikkelin ammattikorkeakoulu, Viva3-hanke	

DESCRIPTION

 <p>MIKKELIN AMMATTIKORKEAKOULU Mikkeli University of Applied Sciences</p>		Date of the master's thesis 11.3.2011
Author(s) Henna Mölsä	Degree programme and option Eservices and Digital Archiving	
Name of the master's thesis The Ontologization of Astuvansalmi Rock Art Site : Process Description		
Abstract <p>The purpose of this master's thesis was to test ontologization and describe this process. The main object was to learn how ontologization proceeded and what challenges there were. In this case, ontologization based on strong semantics and the CIDOC CRM model and the point of view was material-oriented. The target was Astuvansalmi rock art site, the best known Finnish rock paintings site situated in Eastern Finland, Ristiina. The subject was assigned by the Viva3 project of Mikkeli University of Applied Sciences.</p> <p>In general, ontologization process consisted of definition, execution and evaluation stages. More specifically, the execution stage included the search and investigation of materials, rock paintings and the subject matters of rock paintings. The execution also included the discovery of ontological relations, a concept model and a metadata schema. The ontologization process was documented and the stages of the work were described both in writing and with the use of function models and a figure.</p> <p>The learning process developed my own expertise both in the field of ontologization and also in rock paintings. Especially cooperation with different experts was interesting. On my behalf, I reached my goals, although all goals of the project were not reached. A functional prototype was not made and tested. The metadata schema reached only an intermediate stage but the CIDOC CRM extension proposal and ontology document were made.</p> <p>The description of the ontologization process gave example type information on carrying out material-based ontologization with strong semantics. This information can be used in the development of electronic services of Mikkeli University of Applied Sciences and also for education purposes. In principle, this information can be applied locally, nationally and also internationally in ontologization processes in the field of cultural heritage.</p>		
Subject headings, (keywords) ontologies, metadata, semantic web, CIDOC CRM, processes, rock paintings, cultural heritage, Eastern Finland, East Savo, Ristiina, Astuvansalmi		
Pages 73 p.	Language Finnish	URN
Remarks, notes on appendices		
Tutor Mirja Lopenen	Master's thesis assigned by Mikkeli University of Applied Sciences, Viva3 project	

Lyhenteet

BPMN	Business Process Modeling Notation http://www.bpmn.org/
CIDOC CRM	CIDOC Conceptual Reference Model http://cidoc.ics.forth.gr/
DC	Dublin Core http://dublincore.org/
EDM	Europeana Data Model http://version1.europeana.eu/web/europeana-project/technicaldocuments/
ESE	Europeana Semantic Elements http://www.europeana.eu/schemas/ese/ESE-V3.2.xsd
FRBR	Functional Requirements for Bibliographic Records http://www.ifla.org/en/publications/functional-requirements-for-bibliographic-records
FRBRoo	FRBR-object oriented http://archive.ifla.org/VII/s13/wgfrbr/FRBRoo_V9.1_PR.pdf
ICOM	The International Organisation Of Museums http://icom.museum/
IFLA	The International Federation of Library Associations and Institutions http://www.ifla.org/
JHS	Julkisen hallinnon suositus http://www.jhs-suositukset.fi/
KDK	Kansallinen Digitaalinen Kirjasto http://www.kdk2011.fi/
KM	Kansallismuseo
LIDO	Lightweight Information Describing Objects http://www.athenaeurope.org/index.php?en/149/athena-deliverables-and-documents
MARC	Machine-Readable Cataloguing http://www.loc.gov/marc/umb/
METS	Metadata Encoding & Transmission Standard

	http://www.loc.gov/standards/mets/
MIDAS Heritage	a data standard for information about the historic environment http://www.english-heritage.org.uk/professional/archives-and-collections/nmr/heritage-data/midas-heritage/
MODS	Metadata Object Description Schema http://www.loc.gov/standards/mods/
OWL	Web Ontology Language http://www.w3.org/2007/OWL/
ONKI	Kansallinen Ontologia Palvelu http://www.yso.fi/onki/yso/
RDA	Resource Description and Access http://www.rda-jsc.org/rda.html
RDF	Resource Description Framework http://www.w3.org/RDF/
RDF(S)	RDF Schema http://www.w3.org/TR/rdf-schema/
SHFA	Svenskt Hällristnings Forsknings Arkiv http://www.shfa.se/
SPARQL	SPARQL Protocol and RDF Query Language http://www.w3.org/TR/rdf-sparql-query/
SPECTRUM	Standard Procedures for Collections Recording Used in Museums http://www.rcip.gc.ca/English/Standards/metadata_documentation.html
TEI	Text Encoding Initiative http://www.tei-c.org/
TIFF	Tagged Image File Format http://partners.adobe.com/public/developer/tiff/index.htm
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
URN	Uniform Resource Name
VMF	Vocabulary Mapping Framework http://cdlr.strath.ac.uk/VMF/index.htm
W3C	The World Wide Web Consortium http://www.w3.org/

XML

eXtensible Markup Language

<http://www.w3.org/XML/>

YSO

Yleinen Suomalainen Ontologia <http://www.yso.fi/>

SISÄLTÖ

1	JOHDANTO	1
2	TUTKIMUSAIHE JA SISÄLTÖ: KALLIOSTA ONTOLOGIAAN.....	2
2.1	Aiheen rajaus ja kehittämistoimenpiteet.....	3
2.2	Tutkimusmenetelmät	6
3	ASTUVANSALMEN KALLIOMAALAUSTEN DOKUMENTOINTI.....	8
3.1	Kalliomaalausten dokumentointi	9
3.2	Astuvansalmen kalliomaalaukset.....	11
4	SEMANTTINEN WEB, MALLINTAMINEN JA ONTOLOGIAT	13
4.1	Semanttisen webin tekniikat ja mallintaminen.....	15
4.2	Ontologiat	17
4.2.1	Ontologioiden käsitteet	18
4.2.2	Ontologioiden luokittelu	20
4.2.3	CIDOC CRM	24
4.3	Ontologisoinnin prosessit	27
5	PROSESSIT JA PROSESSIEN KUVAAMINEN.....	30
6	ONTOLOGISOINTI JA TOIMINTAMALLIT: TYÖVAIHEIDEN KUVAUS..	32
6.1	Aloitus- ja määrittelyvaihe	34
6.2	Toteutusvaihe.....	38
6.2.1	Aineiston kartoitus	38
6.2.2	Kuvien ja kuva-aiheiden kartoitus	40
6.2.3	Kuvailu ja käsittemallinnus.....	44
6.2.4	Ontologisten suhteiden miettiminen: CIDOC CRM.....	50
6.2.5	Tunnisteet.....	53
6.2.6	Metatietoskeeman mallinnus	54
6.3	Arviointivaihe	59
6.4	Yhteenvedo työvaiheiden kuvauksesta ja ontologisointiprosessikaavio	60
7	TUTKIMUSPROSESSIN ARVIOINTI JA KEHITTÄMISIDEAT	62
8	LOPUKSI.....	65
9	LÄHTEET	67

LIITTEET

- 1 Työryhmän esittely
- 2 Vaatimusmäärittely
- 3 Toimintamallit
- 4 Kuvailumallin mukainen esimerkkikuvailu
- 5 Käsitelliesimerkki
- 6 Metatietoskeemaesimerkki – excel
- 7 Metatietoskeemaesimerkki – word
- 8 Muistilista huomioitavista asioista

1 JOHDANTO

Nykyisessä informaatioyhteiskunnassa on tarve tehdä Internetistä älykkäämpi, jotta informaation saatavuus tehostuisi ja tiedonhankinta helpottuisi. Ihmiset kamppailevat informaatiopaljouden kanssa ja tarve löytää oleellista tietoa informaatiomäärästä kasvaa koko ajan. Tarve automaattiselle kieli- ja kulttuurirajat ylittävälle informaation vaihdolle ja yhdistelylle on lisääntynyt (Kaario 2008, 93). Yhtenä ratkaisuna tähän ongelmaan ovat ontologiat, jotka pyrkivät saamaan informaation merkitykset muotoon, jota tietokoneet ja ohjelmistot pystyvät ymmärtämään. Ontologioita käyttävät tavalla tai toisella sekä tiedonhakijat, tiedon luetteloijat että ontologioiden kehittäjät.

Ontologioiden historia on kuitenkin pitkä ja ulottuu antiikin aikaan, jolloin Aristoteles määritteli filosofian alalla ontologian olemassaolon opiksi. Tietojenkäsittelyn alalla tämä tarkoittaa informaation tekemistä koneymmärrettäväksi. Koneymmärrettävyys vaatii informaation merkitysten ja suhteiden säilyttämistä koneympäristössä. Sisällöllisten merkitysten säilyttäminen vaatii eri alojen sisällöiltä ja palveluilta yhteen toimivuutta, jossa korostuvat standardien mukainen tapa toimia ja jaetut tunnistet.

Suomessa ontologioita on kehitetty systemaattisesti vuodesta 2003 lähtien FinnONTO-hankkeessa. Hanke on tuottanut useita ontologioita, kansallisen ontologiapalvelu ONKIn, metatietostandardeja sekä erilaisia työvälineitä. Suomessa ontologioita koskevat tutkimukset ovat hyvin pitkälti sidoksissa FinnONTO-hankkeessa tehtyyn työhön, joka perustuu asiasanastojen muuttamiseen heikon semantiikan ontologioiksi.

Opinnäytetyöni toimeksiantajana toimii Mikkelin ammattikorkeakoulun Viva3-hanke. Työn tavoitteena on kokeilla vahvan semantiikan ontologisointia ja kuvata tätä prosessia. Ontologisoinnissa päädyttiin käyttämään valmista ontologiaa CIDOC CRM:ää. Ontologisointikohteeksi valikoitui Astuvansalmen kallio Ristiinassa, Etelä-Savossa. Aiheen koettiin herättävän mielenkiintoa kansainvälisestikin. Astuvansalmen kallio on tunnetuin suomalainen kalliomaalauspaikka ja sillä on suurin yhtenäinen kalliomaalauspinta. Ehkä yllättäen kalliomaalauksilla ja ontologioilla on myös jotain yhteistä. Molemmat ovat tapa jäsentää todellisuutta ja maailmaa, jossa elämme.

2 TUTKIMUSAIHE JA SISÄLTÖ: KALLIOSTA ONTOLOGIAAN

Mielenkiintoni ontologioita kohtaan lähtee työelämästä. Kirjastoalalla ontologioista on puhuttu jo jonkin aikaa ja toteutuksia on alkanut syntyä FinnONTO-hankkeen myötä esim. yleisen suomalaisen asiasanaston muuntamisessa (YSA) yleiseksi suomalaiseksi ontologiaksi (YSO). Kirjasto-, arkisto- ja museoalojen yhteistyönä ollaan myös tuottamassa yhteinen ylätason ontologia museopuolelta tutun CIDOC CRM:n avulla. Tavoitteena on päästä hyödyntämään hankkeessa opittuja tietoja työelämässä.

Opinnäytetyön toimeksiantajana toimii Mikkelin ammattikorkeakoulun ESR-rahoitteinen Viva3-hanke. Hankkeen tehtävänä on digitoinnin prosessien kehittäminen, alan osaamisen vahvistaminen ja palvelutuotannon toimintamallien luominen. Hanke keskittyy 3D-mallinnuksiin ja liikkuvaan kuvaan ja tekee yhteistyötä Kansalliskirjaston Ephemera-hankkeen kanssa. Ephemera- ja Viva3-hankkeilla on mm. yhteinen metatietotyöryhmä, jonka yhtenä toimeksiantona on ottaa kantaa ontologioihin ja kokeilla niiden käyttöä. (Viva3 projektisuunnitelma 2008; Viva3 ja Ephemera –projektien yhteiset työryhmät 2009.)

Tutkimuksen aiheena on Astuvansalmen kallion ontologisointi ja sen ontologisointiprosessin kuvaaminen. Tutkimusmenetelmät perustuvat pääasiassa reflektoivaan oppimiseen ja dialogiin sekä grounded theoryyn perustuvaan aineistolähtöiseen tarkastelutapaan.

Ensin esittelen hankkeessa olleen työryhmän ja kerron tutkimusaiheesta tarkemmin. Teoriaosuus alkaa kalliokuvien – ja erityisesti Astuvansalmen kalliomaalausten – esittelyllä luvussa Astuvansalmen kalliomaalausten dokumentointi, jonka jälkeen kerron tarkemmin ontologioista ja niiden laajemmasta viitekehyksestä. Luku pitää sisällään semanttisen webin tekniikoita ja mallintamista sekä ontologioiden käsitteistöä ja luokittelua sekä erilaisia ontologisoinnin prosesseja. Teoriaosuus päättyy prosessien kuvaamisen teoriaan. Varsinaisen työn kuvaus alkaa tehtävänkuvauksella, joka koostuu aloitus- ja määrittelyvaiheesta, toteutusvaiheesta ja arviointivaiheesta. Ontologisoinnin vaiheet vedän yhteen yksinkertaiseksi ontologisointiprosessikaavioksi ja kuvaan työnkulkua toimintamallien avulla. Lopuksi arvioin tehtyä tutkimusprosessia ja sen lopputuloksia sekä esittelen kehittämisideoita.

2.1 Aiheen rajaus ja kehittämistoimenpiteet

Viva3-hankkeen yhteyshenkilöinä ovat toimineet projektipäällikkö Juhani Grönhagen Mikkelin ammattikorkeakoulusta, projektityöntekijä Miia Herrala Mikkelin ammattikorkeakoulusta ja asiantuntija Mika Nyman Synapse Computing Oy:stä (liite 1). Muita työskentelyssä mukana olleita henkilöitä ovat Mikkelin ammattikorkeakoulun paikkatietoasiantuntija Esa Hannus, kirjailija Pekka Kivikäs ja Jyväskylän yliopiston projektitutkija Karoliina Järvinen ja akatemiaturkija tohtori Antti Lahelma. Asiantuntijuuden merkitys on korostunut ontologisointiprosessissa.

Mikkelin ammattikorkeakoulun Viva3-hanke ehdotti opinnäytetyön tekoa hankkeelle. Alun pitäen vaihtoehtoina olivat ontologisointi tai RDA-mallin hyödyntäminen hankkeessa. RDA (Resource Description and Access) ei kuitenkaan ollut vielä valmis opinnäytetyön aiheen valintavaiheessa, joten päädyin ontologisointiin. Ontologisoinnilla tuolloin tarkoitettiin prosessia, jossa RDF-pohjaisesta aineistosta tehdään ontologisoitu versio, ja selvitystä, mitä ontologisointi käytännössä voi olla ja miten YSO:n kaltaisia sanasto-ontologioita voidaan konkreettisesti hyödyntää. Jo tässä vaiheessa todettiin, että työn tarkemmat tavoitteet tarkentuvat ja mahdollisesti muuttuvat työn edetessä, joten niitä on hankala määritellä tarkasti etukäteen. (Herrala 2009.)

Aihetta tarkennettiin seuraavissa palavereissa. Lähtökohdaksi otettiin Kansallisen digitaalisen kirjaston standardisalkku, koska yhteiset standardit helpottavat aineistojen yhdistämistä laajempien ylätasen ontologioiden kanssa. Ontologisoinnilla tarkoitettiin tässä vaiheessa heikon semantiikan ontologisoinnin sijaan vahvan semantiikan ontologisointia eli informaation merkitysten ja suhteiden tulisi säilyä hyvin koneympäristössä. KDK:n standardisalkusta päädyttiin käyttämään valmista ontologiaa CIDOC CRM:ää, jota on mahdollista käyttää paikallisissa toteutuksissa sen tyyppi-käsitteen vuoksi ja siinä ilmenee myös aineistojen tapahtumapohjaisuus. CIDOC CRM:ää käytetään kansainvälisesti, joten se antaa mahdollisuuden yhteistyölle ja käsitteiden jakamiselle kansainvälisesti. Muitakin siirtoformaatteja ja metatietoformaatteja projektin kuluessa käsiteltiin. Tämä opinnäytetyö keskittyy kuitenkin ainoastaan Astuvansalmen kallon ontologisoinnin malliksi valittuun CIDOC CRM:n esittelyyn ja käyttöön projektissa, koska tällöin on mahdollista saavuttaa jossain määrin myös syvyyttä.

Ontologisointikohteeksi valittiin Astuvansalmen kallio Ristiinassa, Etelä-Savossa. Astuvansalmi oli yksi Viva3-hankkeen tutkimuskohteista ja siitä oltiin jo tekemässä 3D-mallia, jota varten Mika Nyman teki tunnistepalvelua. Koettiin hyväksi toimia näiden samojen aineistojen parissa. Astuvansalmen kallio ja sen kalliomaalaukset edustavat hyvin Suomen kalliomaalauksia ja Astuvansalmi on tunnetuin suomalainen kalliomaalauspaikka. Astuvansalmen kalliomaalausten osalta tärkeä raja-alue oli ontologisoinnin koskeminen koko kalliota ja sen ympäristöä ei ainoastaan kallion kalliomaalauksia. Kohteen valintaa puolsi sopivan kokoinen dokumentaatioaineisto ja aiheen – kalliomaalausten ontologisoinnin - koettiin herättävän mielenkiintoa kansainvälisesti.

Tärkeintä tekemisessä on oppiminen ja oppimisprosessi sekä prototyypin testaus. Ontologisointiprosessin dokumentointi ja kuvaus ovat merkittävä osa prosessia. Opinnäytetyö on käytännön selvitys siitä, kuinka ontologisointi tapahtuu ja mitä ongelmia tekemiseen liittyy. Prosessikuvaus on työn lopputulos ja muutoin painopiste on tekemisessä. Prosessikuvauksen tarkoituksena on pääasiassa auttaa hahmottamaan ja ymmärtämään ontologisointiin liittyvää kokonaisuutta. Aineistoista, käsitelmällisistä, metatietoskeemoista, ontologioista, metatietostandardeista ja prosesseista muodostuu prosessuaalinen kokonaisuus. Ontologisointiprosessia kuvataan sekä kirjallisesti työvaiheiden kuvauksen yhteydessä että prosessikaavioiden avulla visualisoiden. Koska ontologisointi on ollut yhteistyön tulosta, esittelen seuraavana karkeasti prosessin vaiheet ja niiden vastuuhenkilöt.

Alun määrittelyvaiheessa prosessiksi sovittiin:

1. Aineistojen kartoitus (vastuu Henna Mölsä)
2. Pienimuotoisen tietokannan mallinnus
3. Tietokannan luominen Protégélla (vastuu Mika Nyman)
4. Ontologisten suhteiden miettiminen (vastuu Henna Mölsä, Miia Herrala)
5. Prosessikuvauksen tekeminen dokumentoinnista (vastuu Henna Mölsä)

Lopulta prosessi noudatti pikemminkin seuraavaa kaavaa:

1. Aineiston kartoitus (vastuu Henna Mölsä)
2. Kuvien ja kuva-aiheiden kartoitus (vastuu Henna Mölsä)
3. Käsitelmällisyys ja kuvailu (vastuu Henna Mölsä)

4. Metatietoskeeman mallinnus, välivaihe (vastuu Henna Mölsä)
5. Ontologisten suhteiden miettiminen: CIDOC CRM (vastuu Miia Herrala, Mika Nyman)
6. CIDOC CRM:n laajentaminen projektin tarpeisiin (vastuu Mika Nyman)
7. Yksinkertaisen prosessikuvauksen tekeminen työvaiheista (vastuu Henna Mölsä)

Aihetta päädyttiin alun määrittelyn jälkeen muuttamaan aineistolähtöisemmäksi. Tämä toteutui mm. kartoitukseen ja mallinnukseen keskittymällä ja unohtamalla työvälineen Protégén. Prosessin työvaiheet selvisivät prosessin kuluessa ja muuttuivat yksityiskohtaisemmiksi osiksi. Tässä vaiheessa tuli myös yllätyksiä alkuperäisen suunnitelman muuttuessa ja selkeytyessä. Käsitelmän osalta todettiin, että siinä tuli näkyä tulkinnat ja uskomukset. Kolme ontologista kenttää on riittävä tavoite ja sisällöt näihin riittävät perustiedoilla.

Ajanpuutteen vuoksi toiminnallisen prototyypin teosta luovuttiin. Tästä syystä tarkempi määrittely, sen suhteen mitä tietokannalla täytyy pystyä tekemään, jäi vaillinaiseksi. Alustavasti prototyypin toiminnallisuuksiksi oli ajateltu seuraavia näkökohtia. Toiminnallisen prototyypin tarkoitus oli selventää aineistolähtöisyyttä. Yhtenä tavoitteena oli mahdollistaa useat tulkinnat. Tulkintojen tuli sisältää myös mahdollisten käyttäjien oma tulkinta eli sosiaalisen median näkökohdan huomioiminen. Prototyypissä oli tarkoitus tuottaa fasettipohjainen haku. Fasetit olisivat määräytyneet CIDOC CRM:n perusteella. Tietojen syöttöä varten olisi ollut mahdollista rakentaa www-lomake. Mahdollisesti olisi myös annotoitu eli kuvailtu automaattisesti kuvat. Kuvista ja kuvasarjoista olisi ollut mahdollista merkata yksittäisiä osia TEIn avulla.

Arviointivaihe jäi prosessissa keskeneräiseksi. Prototyypin jälkeen arvioitavana kohteena olisi ollut CIDOC CRM -pohjainen metatietoskeema, mutta ennen kuin skeeman luonnissa päästiin välivaihetta pidemmälle, päätettiin laajentaa CIDOC CRM:ää. Joka tapauksessa ontologisoinnissa on tärkeää arvioida mihin tapahtumiin skeeman luokat kytkeytyvät, kuinka hyvin ominaisuudet ovat periytyneet ja että suhteet vastaavat aineistoa. Ontologia auttaa siis tekemään kysymyksiä aiheesta ja prosessi on tässä mielessä kehämäinen. Koska hankkeessa on kyse testauksesta,

prosessissa ei päästy ylläpitotasolle ja siten toteuttamaan kehämäisyyttä kokonaisuudessaan.

Kuitenkin lopputulokset ovat toissijaisia verrattuna oppimiseen. Fyysisiä lopputuloksia omalta osaltaan ovat tehdyt listaukset: kartoitetut aineistot, kuva-aiheiden ryhmittelyt, käsittemalli ja metatietoskeeman välivaihe, jotka tehtiin prosessin mukaisesti. Kartoitus tuli kattamaan sekä aineistot että näiden aineistojen sisältämän kuva-aiheiden kartoituksen. Aineistojen kartoituksessa ei saatu käyttöön kaikkia aineistoja osittain johtuen kustannuksista. Prototyyppejä ei tehty, joten sen perusteella tehdyt mallinnuksien arvioinnit jäivät tekemättä.

Merkittävä välietappi työskentelyssä oli osallistuminen konferenssiraportin kirjoittamiseen CIDOC:in joka toinen vuosi järjestettävään konferenssiin.

2.2 Tutkimusmenetelmät

Työskentelymenetelmänä on refleктоiva oppiminen sekä dialogi eri asiantuntijatahojen kanssa. Tavoitteena on oppiminen, oivallukset ja uuden tiedon luominen sekä ajan myöten aito ymmärrys aiheesta. Uuden tiedon oppiminen antaa mahdollisuuden tiedon edelleen opettamiselle. Dialogilla tässä tapauksessa tarkoitetaan ontologisointiryhmän prosessia, jossa ryhmän jäsenet keskenään ja itsensä kanssa keskustelevat samasta aiheesta. Tarkoitus on luoda yhteistä tietopääomaa, jota ei yksin ole mahdollista saavuttaa. Ihmisten eri taustoista johtuen dialogissa on mahdollista löytää eri näkökulmia ja haastaa kunkin omia uskomuksia ja lopulta synnyttää yhteinen kokonaisnäkemys asioista. Refleктоivalla oppimisella puolestaan tarkoitetaan omakohtaista kokemusta oppimisprosessissa. Lähtökohtana oppimisessa ovat kunkin yksilön ennakkokäsitykset ja havaitsemiseen liittyvät tietoiset ja tiedostamattomat toiminnot. Tätä melko intuitiivista toimintaa pyritään sitten tietoisesti ymmärtämään, kyseenalaistamaan ja käsitteellistämään. Lopputuloksena kokemusta hyödynnetään soveltavassa toiminnassa. Opinnäytetyön tuotos on kokonaisuudessaan dialogin tulos ja kooste yhteisesti opituista asioista. (Heikkilä & Heikkilä 2000; Ruohotie 2000, 137 – 155.)

Aihetta eli Astuvansalmen kalliota on lähdetty tarkastelemaan aineistolähtöisesti. Yhtenä teoreettisena viitekehysenä on ollut grounded theory. Käytännössä tämä on

tarkoittanut aineiston kartoitusta, jonka jälkeen aineisto on analysoitu. Analyysi on toteutettu viitekehyksen puitteissa, joka on toiminut kuvailun apuna. Grounded theoryn periaatteisiin kuuluu ennakkoluulottomuus aineistoa kohtaan. Aineiston annetaan itsessään luoda teoria ja kuvastaa siten paremmin todellisuutta. (Strauss & Corbin 1998, 12 – 13.)

Grounded theoryn ennakkoluulottomuuden näkökulma vaatii luovuutta. Erilaisille näkemyksille ja suunnanmuutoksille on oltava tilaa prosessissa. Luovuus on vakiinnuttanut paikkansa nyky-yhteiskunnan keskustelussa kuten Opetusministeriön (2006) luovuusstrategian loppuraportista ja Kauppa- ja teollisuusministeriön Luovien alojen yrittäjyyden kehittämistrategiasta 2015 (2007) voi päätellä. Pekka Uotila (2009) on kirjoittanut puolestaan ammattikorkeakoulujen projektitoiminnan luovuudesta. Haasteena on yhdistää hyvä projektinhallinta ja kontrolli luovuuden kanssa. Uotilan mukaan nykyinen tietotekniikkaa hyödyntävä projektien hallinnoiminen koetaan byrokraattisena. Luovuus vaatii omalta osaltaan kokeilemistä, epäonnistumisen sietoa, harhailua, kuvittelua ja odottamattomuutta. Projektin perustan luo valmisteluvaihe, joka mahdollistaa luovuuden – tilan, jossa kuvitellaan tulevaa.

Teoreettisesti on hyödynnetty myös Mika Nymanin asiantuntemusta ontologioiden saralla. Hän on hyödyntänyt omassa työssään mm. neurobiologiasta peräisin olevaa perception – action cyclea, jonka teoriaan palataan työvaiheiden kuvauksen yhteydessä. Perception - action cycleä hyödynnettiin teoreettisena pohjana hankkeessa luodulle kuvailumallille, jota käytetään aineistojen kuvailun apuna.

Astuvansalmen ontologisoinnin osalta alussa ideoitiin mahdollisuuksia ja mahdollisia tarpeita, joista toteutuskelpoisimmat valikoituivat käytännössä ajan ja resurssien puitteissa. Tilaa luovuudelle jätettiin koko prosessin ajaksi, koska mitään ei lyöty lopullisesti lukkoon, vaan mahdollisuus suunnanmuutoksille oli koko ajan olemassa. Käytännön työ sisälsi paljon epätietoisuutta suurelta osin johtuen tekijöiden kokemattomuudesta ja lähtökohdasta, että asiantuntijana toiminut Mika Nyman ei tehnyt työtä projektin puolesta, vaan tarjosi näkökulmia, jotka jäivät tekijöitten pohdittaviksi. Uotilan (2009) mukaan luovuutta on myös myöntää jälkikäteen prosessin epäonnistumiset, mikäli ne poikivat uusia mahdollisuuksia. Uotila esittää, että projektit sisältäisivät luovuutta, mikä ilmenee luottamuksena, tarkoituksen korostamisena, yksilöiden ja ryhmien pyrkimyksenä vaikeasti määriteltävään asiaan,

aikana jatkuvaan ja kannustavaa arviointiin sekä mahdollisuutena muuttaa tavoitteita, lähtökohtia ja keinoja. Arvioinnissa tulisi korostaa koettujen tekojen kuvaksia ja uusia tulkintoja. Ontologisoinnin vaikeasti määriteltävyyttä lisäsi ero heikon ja vahvan semantiikan ontologioiden välillä ja tämän seikan ymmärtäminen tehdessä. Suomessa yleisesti ontologioista puhuttaessa tarkoitetaan heikon semantiikan ontologioita, mikä aiheuttaa, että on vaikea alkuvaiheessa ymmärtää mitä vahvan semantiikan ontologiolla oikeastaan tarkoitetaan ja miten se ilmenee. Näihin ontologioiden luokituksiin palataan kuitenkin tarkemmin luvussa Semanttinen web, mallintaminen ja ontologiat.

Projektien prosesseissa mukana olevista tulee kehittäjiä ja osallistujia. Arviointitutkimuksessa tätä kutsutaan prosessievaluaatioksi, jota kutsutaan myös auttavaksi malliksi tai konsultoinniksi. Tuloksia voidaan tulkita suhteessa toimintaympäristöön. Toimintatapoja on mahdollista jakaa pienempiin osiin. Prosessievaluaatiossa on hankala löytää yhtä menetelmää eli on hyödynnettävä erilaisia metodeja ja sovittaa niitä yhteen. (Jalava & Virtanen 1998, 129 – 131.)

Tiedonhakuprosessi on ollut kaksitahoinen. Toisaalta on kartoitettu aineistoa koskevat olemassa olevat kirjat ja kuvat yms. sekä toisaalta on hankittu teoretietoa tekemisen tueksi ja opinnäytetyön lähteiksi. Tiedonhaussa käytettäviä Yleisen suomalaisen asiasanaston (YSA) mukaisia asiasanoja ovat mm. ontologiat, metadata, semanttinen web, tietokannat, Protégé, prosessit, kalliotaide, Astuvansalmi, Ristiina. Yhtenä osana tiedonhakua on ollut asiantuntijalähteet kuten kirjailija Pekka Kivikäs ja valokuvaaja Ismo Luukkonen, joiden aineistoja on ollut käytettävissä. Erityisesti näissä tapauksissa tekijänoikeuksiin on tullut kiinnittää huomiota, että heidän oikeutensa tulevat huomioiduksi.

3 ASTUVANSALMEN KALLIOMAALAUSTEN DOKUMENTOINTI

Akateemikko, professori Abram Stolyarin määritelmä kalliotaiteesta:

”Kalliotaide on ennen kaikkea kommunikaatiota. Se kantaa informaatiota ja siirtää sitä uusille sukupolville. Kalliotaide toimii samanaikaisesti opetuksen ja muistin apuvälineenä. Se on myös palvelun kohde. Lisäksi se on voimakas ärsykkeiden ja

tunteiden synnyttäjä [stimulus] sekä tärkeä maailmanselittämisen malli” (Kivikäs 2008, 28.)

Yleensä kalliokuvista puhutaan kalliotaiteena kuten Stolyarin määritelmässään. Kalliotaiteen sijasta olisi osuvampaa puhua kalliokuvista, koska kaikkia kalliomaalaukset eivät välttämättä yllä taiteen tasolle, vaan ovat olleet nimenomaan viestinnän välineitä. Kalliokuvat sisältävät sekä maalauksia että piirroksia. Piirrookset puolestaan voi olla tehty kaivertamalla, uurtamalla tai raaputtamalla. (Kivikäs 1995, 11 – 15.)

3.1 Kalliomaalauksen dokumentointi

Astuvansalmen kallion tutkimuksia ovat tehneet Pekka Sarvas vuonna 1968, Juhani Grönhagen 1988, 1990 - 1992, Heino Vänskä 2001 ja Ismo Luukkonen 2001. Sarvas on tehnyt myös tarkistuskaivauksia myöhemmin alueella. Grönhagen on tehnyt alueella vedenalaisia sukellustutkimuksia, Vänskä jäkäläkartoituksen ja Luukkonen valokuvausdokumentoinnin. Kalliomaalauksen lisäksi paikalta on löytänyt meripihkariipuksia Juhani Grönhagen. Näitä riipuksia kutsutaan nimillä Astuvan ukko (1990), Astuvan akka (1991), Astuvan poika (1991) sekä meripihkaisen karhunpään (1993). (Muinaisjäännösrekisteri.)

Hankkeessa kerättäviin ja käsiteltäviin aineistoihin ja niiden tulkintoihin vaikuttavat oleellisesti käytetyt dokumentointimenetelmät. Kalliomaalauksia on dokumentoitu Suomessa mm. kopioimalla teoksia eri tavoin. Ensimmäiset dokumentoinnit tehtiin valkoisella liidulla kalliomaalauksen viivojen väliin, josta viivat piirrettiin läpinäkyvälle paperille. Näihin dokumentointeihin liittyi lisäksi sanalliset kuvaukset ja mustavalkokuvat. Dokumentointia on tehty myös siten, että kuviot on jäljennetty myös ennen liituaamista. Menetelmää kehitettiin ottamalla käyttöön eri värisävyjä, joita tuotiin käyttämällä punaista ja keltaista huopakynää värittömälle muovikankaalle. Kuvat jäljennettiin kahdesti – ensin ilman liitua ja sitten liitujen kanssa. Myöhemmin dokumentointia tehtiin liituaamalla ääri viivoja ja kopioimalla kuvat huopakynällä muoville. Astuvansalmen kuvien kohdalla ei ole alkuvaiheessa mm. dokumentoitu epäselviä kuvia. 1960-luvun jälkeen dokumentointitapaa ei ole välttämättä enää mainittu eikä niistä ole enää sanallisia kuvauksia. Dokumentoinnissa alettiin käyttää ns. pistetekniikka, jossa kuvat jäljennetään tussilla muovikalvolle käyttäen pieniä

pisteitä tai vetoja. Menetelmän hyötynä on ollut erottaa sävyeroja kuvista sekä kuvan läheisyydessä olevien ja heikompien viivojen esittämisen mahdollisuus. Mustavalkokuvien rinnalle tulivat dia- ja värikuvat. Väri negatiivien pohjalta tehtiin kopioita asemointikalvolle, monistettiin paperille ja viimeisteltiin vesivärein värierot huomioiden. Hyötynä tässä menetelmässä on ollut kallion muotojen saaminen mukaan dokumentointiin sekä päällekkäisten värikerrosten ja –sävyjen erottelumahdollisuus. Dokumentointien yhteydessä kalliopintaa on voitu puhdistaa sammalista ja jäkälisiä eri tavoin, että kalliomaalaukset on saatu paremmin näkyville. (Taskinen 2007, 123 – 129, 134.)

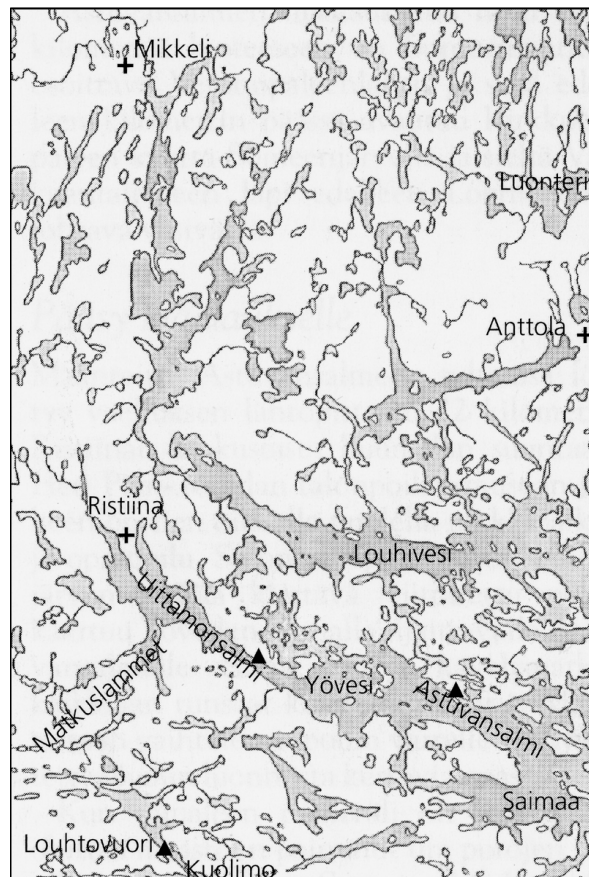
2000-luvulta alkaen on hyödynnetty kuvankäsittelytekniikkaa dokumentoinnissa. Kuvissa on vahvistettu punaväriä tai muutettu kokonaan värejä. Kuvankäsittelyä hyödyntämällä on saatu esiin uusia kuvia, vahvistettu ja varmennettu epätäydellisiä kuvia sekä saatu selvitettyä puuroutuneita punavärialueita. Valokuvaus menetelmänä ei vahingoita itse kalliota kuten liituaminen. Museoviraston tekemät kuvaukset 2001 – 2005 on toteutettu ympäri Suomea samoilla menetelmillä eli valokuvaamalla kalliopinta 40x30 cm paloissa samalta etäisyydeltä ja suoraan edestä huomioiden koko kallio. Kuvauksissa on käytetty telineitä ja usein kuvauksia on ollut mahdollista tehdä ainoastaan talvisin jäältä, koska kalliomaalauspaikat sijaitsevat veden ääressä. Talvikuvaus on kannatettavaa myös sen vuoksi, että kuvat näkyvät tällöin kaikkein selkeimmin. Kuvatessa on käytetty mittatikkua mittakaavan tarkastamista varten sekä harmaata vertailuväriskaalaa värisävyn säätämistä varten. Astuvansalmen kalliomaalauksen osalta Museoviraston valokuvausdokumentoinnin on toteuttanut Ismo Luukkonen vuonna 2001. Samassa yhteydessä Astuvansalmella tehtiin jäkäläinventointi. Hankkeen käytössä oli Luukkosen vuonna 1996 ottamia kuvia. (Taskinen 2007, 129 – 134, 136.)

Viva3-hankkeen 3D-mallissa Astuvansalmen kalliomaalaukset dokumentoitiin valokuvaamalla maalikuussa 2010. Valokuvista rakennetun 3D-mallin avulla on helppo tutkia ja mitata kalliomaalauksia ja malli huomioi pelkkiä valokuvia paremmin tietenkin syvyysulottuvuuden. Kerran tehdyn valokuvausdokumentoinnin pohjalta 3D-malli on mahdollista saattaa kaikkien halukkaitten käyttöön. Eniten 3D-mallia voi hyödyntää, jos dokumentointi tehdään jonkin koordinaattijärjestelmän mukaisesti korkeus huomioiden, mitä kutsutaan myös georeferoinniksi. Georeferointi mahdollistaa puolestaan yksilöityjen tunnisteen luonnin ja lisätietojen liittäminen

helpottuu. Mallin tietoja voidaan käyttää myös muissa palveluissa kuten Google Earthissa. Tunnisteiden luonnista ja merkityksestä kerrotaan vielä tarkemmin sekä semanttisen webin tekniikoita käsiteltäessä että työvaiheiden kuvauksen yhteydessä. (Hannus ym. 2010.)

3.2 Astuvansalmen kalliomaalaukset

Tarina alkaa Saimaalta Yöveden rannalta Ristiinan Astuvansalmelta kivikaudella tyypillisen kampakeramiikan kaudella 4200-2900 eaa. Kampakeraaminen pyyntiväestö asettautui asumaan Suursaimaan rannoille, jonka vedenkorkeus oli 10,5 metriä nykyistä korkeammalla. Ensimmäiset kalliomaalaukset tehtiin Astuvansalmen kallion pinnalle Vuoksen puhkeamisen aikoihin n. 4000-3900 eaa. Saimaan vesi alkoi laskea nykyiselle tasolle. Nuorimmat kalliomaalaukset ajoittunevat kivikauden lopulle n. 2000-1500 eaa. Ylimmät kalliomaalaukset ovat n. 11,8 metriä Saimaan nykyisen vedenpinnan yläpuolella ja suurin osa n. 8,3-9,8 metrin kohdalla. (Jussila 1999; Jussila 2004; Kivikäs 1995, 33, 49, 52.)



KUVA 1. Astuvansalmen kalliomaalauksen sijaintipaikka (Kivikäs 2009, 175)

Mikä sai pyyntiväestön valitsemaan nimenomaisesti Astuvansalmen kallion pyhäksi paikakseen? Kalliomaalauspaikkojen sijainnille (kuva 1) on tyypillistä vesireittien risteymäpaikat sekä kolmen elementin eli veden, maan ja taivaan läheisyys. Paikat ovat usein korkeita ja vaikeasti saavutettavia. Kallio toimii maamerkkinä korkeutensa puolesta. Kalliot ovat lisäksi jyrkkiä ja tasaisia sekä niissä on suojaava lippa, jolloin kuvien tekeminen on helpompaa ja säilyminen varmempaa. Tärkeitä valintakriteerejä ovat ympäristön hyvät metsästys- ja kalastuspaikat. Erityistä Astuvansalmelle on kallion ihmismäinen muoto. (Lahelma 2008, 22.)

Oma mystiikkansa liittyy itse kalliomaalauksiin. Mikä niiden tarkoitus ja mitä niillä on haluttu sanoa jälkipolville? Tulkintoja on niin paljon kuin tulkitsijoitakin, mutta jotain voinee pitää todennäköisenä. Suomen kalliomaalauksille tyypilliset hirvieläimet on liitetty selkeästi metsästyksen – metsästysonnen varmistamiseen ennen metsästystä ja kiitoksenosoitukseen saaliista metsästyksen jälkeen. Saaliin saamisen lisäksi kalliomaalauksilla on jäsenetty omaa ympäristöä, merkitty pyyntialue, edistetty ryhmän menestymistä ja synnytysten onnistumista, torjuttu sairauksia ja uhkia, selviydytty elämänmuutoksista ja varauduttu tulevaan. Tärkeää on ollut toisten ryhmien tapaaminen ja sopimusten teko sekä oman alkuperän tunteminen ja identiteetin vahvistaminen. Kalliomaalaukset ovat toimineet rajapintana sisäiseen maailmaan ja ilmentäneet ulkoista maailmaa. (Kivikäs 1997, 20 – 21; Lahelma 2008, 45.)



KUVA 2. Astuvansalmen kallio (Kivikäs)

Astuvansalmen kalliomaalaukset ovat poikkeuksellisen rikkaita (kuva 2). Sydänpisteellisten ja pisteettömien hirvieläinten lisäksi kalliopinnalta löytyy muita keskeisiä kuvamerkkejä kuten erilaisia ihmismäisiä hahmoja, sarvi- ja veneaiheita, painamia ja erilaisia geometrisia kuvioita. Onpa eri tulkintojen mukaan löydetty kaloja, karhuja ja majavia. Näitä kuvamerkkien jaotteluja on useampia. Mm. Antti Lahelma jaottelee skeemassaan kuvat hirvieläimiin, ihmismäisiin hahmoihin, veneisiin, kämmenpainamiin, ei hirvieläimiin ja geometrisiin kuvioihin (Lahelma 2008, 194). Ihmismäiset hahmot koostuvat Astuvansalmella ihmisistä, sarvipäisistä ihmishahmoista ja jousikäisestä naisesta. Mm. jousikäisestä naisesta ollaan montaa mieltä, onko nainen itse asiassa ollenkaan nainen eli onko naisen merkeiksi tulkitut rinnat oikeastaan rinnat vai aivan erillinen merkki, piteleekö ihminen jousta vai jotain äänentuottamiseen tarkoitettua välinettä jne. Myöhemmin työvaiheiden kuvauksessa perehdymme tulkinnan ongelmiin tarkemmin.

4 SEMANTTINEN WEB, MALLINTAMINEN JA ONTOLOGIAT

Kalliotaide määriteltiin erityisesti kommunikaatioksi – informaation kantamiseksi ja siirtämiseksi - sekä muistin apuvälineeksi. Yhteys kalliotaiteen ja ontologioiden välillä näyttää tästä näkökulmasta ilmeiseltä. Jäsennämme omaa maailmaamme itsellemme helpommin ymmärrettäväksi. Ontologioiden tarkoitus on nimenomaan helpottaa informaation löytämistä. Nykyistä maailmaamme on jo vaikea kuvitella ilman tietokoneita ja nopeita nettiyhteyksiä. Tämä on kuitenkin aiheuttanut ongelman informaation paljouden takia. Kommunikaatiossa ja viestinnässä oleelliseksi muodoksi tulee tiedonhankinta ja sen laatu (Vakkari 2000, 11). Kuinka saamme tietokoneilla käytettävän informaation jäsennettyä itsellemme ymmärrettäväksi ja kuinka on mahdollista löytää oleellista tietoa verkkoviidakosta?

Ontologioissa on kyse tiedosta ja sen jäsentämisestä. Tietoa käytetään yleensä yleismääritelmänä, mutta perinteisen jaottelun mukaan tieto jaetaan dataan, informaation, tietoon ja tietämykseen. Data on merkkejä, mitä koneet yleensä ymmärtävät. Ontologioiden kohdalla data viedään uudelle asteelle. Tässä työssä puhutaan pääasiassa informaatiosta eli aineistoilla viestittävistä sisällöistä, mutta myös tietämyksestä eli kalliomaalauksen tulkinnoista. Myöhemmin käy selväksi tiedon

epätarkkuus ja epämääräisyys. Tiedon kuvaaminen liittyy lähes aina käsitteisiin, mikä tekee siitä suhteellista. Käsitteet vaikuttavat merkityksiin, joita asioille annamme. Tulkitsemme ja pyrimme ymmärtämään maailmaa erilaisten käsitteiden kautta. Todellisuutta yritetään kuvata näiden kieleen yhteydessä olevien käsitteiden kautta, jotka eivät kuitenkaan todellisuutta tosiasiasa ole. Todellisuuden parempaa välittämistä sellaisena kuin se on pyrkivät ontologiat omalta osaltaan tuomaan tietojenkäsittelyn puolelle. Ontologia tulee sanoista ”onta” ja ”logos”. Onta tarkoittaa todellisuutta tai todellisuudessa olevia asioita ja logos tarkoittaa oppia jostakin. Filosofian alalla Aristoteles määritteli ontologian olemassaolon opiksi. Ontologioiden yksi osa on metafysiikka, jossa pyritään vastaamaan kysymyksiin mitä on olemassa ja millaista olemassa oleva on. Ontologioissa ei käsitellä pelkkää dataa tai informaatiota vaan myös niiden suhteita ja ominaisuuksia. Poli (2003, 1) erottelee tiedon semioottiseen, semanttiseen ja ontologiseen. Semioottinen ja semanttinen eivät ole hänen mukaansa ontologista tietoa. Tässä työssä semanttisuutta ei ajatella merkitysteorianaan vaan ylipäätään merkityksinä ja niiden säilymisinä koneympäristössä. Tietojenkäsittelyn määritelmään ontologioista palataan tarkemmin myöhemmin. (Guarino & Giaretta 1995, 2; Kettula 2009, 35; Kotkavirta & Nyysönen 2005, 29, 64; Kuronen 1995, 194.)

Tietojenkäsittelyn puolella tiedosta puolestaan käytetään yleensä termiä metatieto eli tietoa tiedosta. Kansallinen digitaalinen kirjasto –hankkeessa metatieto on määritelty ”tiedon kontekstia, sisältöä ja rakennetta sekä niiden hallintaa ja käsittelyä koko elinkaaren ajan kuvaavaksi tiedoksi. Tätä tietoa voidaan käyttää mm. aineiston hakuun, paikallistamiseen ja tunnistamiseen”. Ontologisointiprosessissa on pyritty mahdollistamaan myös sosiaalisen metatiedon hyödyntäminen eli käyttäjien tuottama tieto. Sosiaalinen metatieto voi olla sekä lisätietoa että aivan uutta tietoa. Koska tässä ontologisointiprosessissa ylipäätään tukeudutaan KDK-hankkeen standardisalkkuun ja työhön on loogista käyttää heidän metatietomäärittelynsä metatiedosta ja sosiaalisesta metatiedosta. KDK:aineistot tulevat siirtymään Europeanaan, jonka tehtävänä on Euroopan kulttuuri- ja tiedeaineistojen saattaminen kaikkien kansalaisten saataville. Europeanan kautta tulevat huomioon otettaviksi sen metatietoformaattit eli ESE (Europeana Semantic Elements) ja EDM (Europeana Data Model). Metatieto ei sisällä yleensä tarpeeksi tietoa kohteesta ja olemassa oleva tieto on puolestaan riippuvainen tulkinnoista eikä sen avulla myöskään pystytä takaamaan tiedon merkitysten menettämistä järjestelmästä toiseen siirryttäessä vaan tieto köyhtyy. Tiedon syöttö

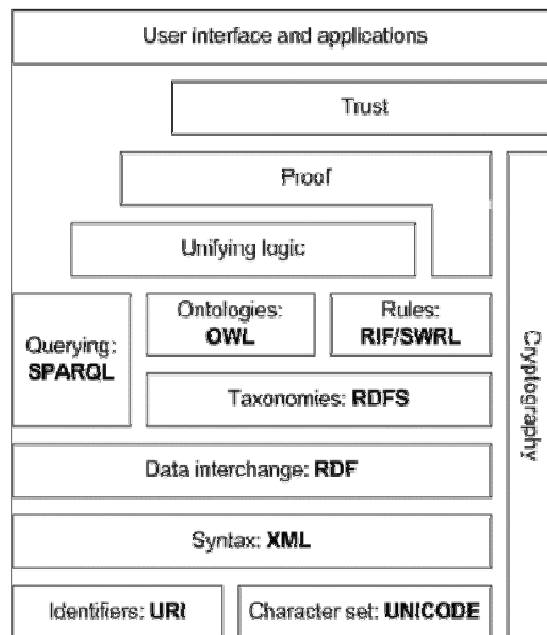
järjestelmiin on tulkinnanvaraista. Yksilöivät, katkeamattomat ja jaetut tunnisteet mahdollistavat datan semanttisen rikastamisen. Yleinen käsitelmä/metatietoskeema puolestaan mahdollistaa tiedon systemaattisen lisäämisen. Tämän takia tarvitaan formaaleja ontologioita pelkän metatiedon sijaan. (Hannus ym. 2010; KDK 2010b)

Laajempaan viitekehykseenä ontologioille toimivat semanttinen web sekä Web 2.0 -ilmiö. Semanttisessa webissä tehdään www-sisällöistä merkityksellisiä siten että sekä ihminen että koneet niitä voivat ymmärtää. Semanttinen web juontaa juurensa termistä semantiikka, joka tarkoittaa kielitieteessä merkitysoppia. Filosofian puolella se puolestaan tarkoittaa merkkien ja niiden kulloistenkin tarkoitteiden suhteita koskevaa tutkimusta. Semanttisen webin koneluettavuutta helpottamaan on kehitetty standardeja, teknologiaratkaisuja ja ontologiatyökaluja, kuten RDF ja OWL. Web 2.0 puolestaan antaa ihmisille mahdollisuuden osallistua sisällöntuotantoon ja -jakeluun, minkä yhtenä osa-alueena on sosiaalinen metatieto. Semanttisen webin toinen merkittävä osa-alue koneymmärrettävyyden lisäksi on saada eri alojen sisällöt ja palvelut keskenään yhteen toimiviksi semanttisesti eli sisällölliset merkitykset säilyttäen. Yhteinen kieli mahdollistaa siis eri ontologioiden yhdistelyn ja hyödyntämisen semanttisella tasolla. Tekniikoilla pyritään ennakoimaan käyttäjien tarpeita paremmin mm. suosittelulla, visualisoinnilla, työprosessien automatisoinnilla tai muilla käsitteiden välisiä yhteyksiä hyödyntävillä toiminnoilla (Karhulampi 2008, 1). (Hyvönen 2008a, 3; KDK 2010b; Koukkunen 1990, 509.)

4.1 Semanttisen webin tekniikat ja mallintaminen

Astuvansalmen kohdalla keskitytään mallintamiseen. Toisaalta voidaan puhua käsitteellisestä mallintamisesta ja toisaalta loogisesta mallintamisesta. Käsitteellinen mallintaminen on järjestelmäriippumaton. Loogisessa mallintamisessa on kuitenkin otettava kantaa tietokantatyypistä. Ontologiatietokannat perustuvat yleensä ns. kolmikerrosmallille tai kolmitasoarkkitehtuurille. Ylimpänä on käyttöliittymä, joka on yleensä selainpohjainen. Sovelluspalvelin puolestaan välittää informaatiota käyttöliittymän ja tietokantapalvelimen välillä. Tietokantapalvelin välittää tietoa tietokantaan ja tietokannasta. (Hovi ym. 2005, 17 – 18, 76.)

Semanttinen web puolestaan koostuu tietyistä tekniikoista, joita yleisimmin käytetään. Tim Berners-Lee on luonut semanttisen webin arkkitehtuurikaavion (kuva 3), johon on koottu nämä erilaiset tekniikat. Muitakin mahdollisuuksia on kuitenkin olemassa.



KUVA 3. Tim Berners-Leen semanttisen webin arkkitehtuuri (Semantic web stack 2010)

Ontologiaa tehtäessä herää kysymys miten pystytään tunnistamaan ontologian eri oliot eli esim. sanat, käsitteet, kuvat jne. toisistaan. Esim. kuinka erottaa paikkakunta Nokia yritys Nokiasta. Olioiden yksilöiminen vaatii toimiakseen tunnisteita. Semanttisen webin puolella yksilöiminen on ratkaistu URI-tunnisteiden avulla. URI (Uniform Resource Identifier) on merkkijono, joka kertoo tietyn tiedon paikan (URL) tai yksikäsitteisen nimen (URN). Tim Berners-Leen semanttisen webin arkkitehtuurissa URI-tunnisteet sijoittuvat semanttisen webin peruskiveksi, jonka päälle muuta semantiikka lähdetään rakentamaan. (Hyvönen 2008a, 2.)

Semanttisen webin muut tekniikat perustuvat pitkälti XML-kuvauskieleen (eXtensible Markup Language). XML kuvauskieli, jolla tiedon merkitys saadaan kuvattua tiedon sekaan. XML:n pohjalta on kehitetty erityisesti metatietoja ajatellen RDF (Resource Description Framework) ja siitä puolestaan RDF(S) eli RDF-skeema. RDF määrittää yleiset säännöt suhteiden kuvaamiseen, kun RDF(S) määrittää suhteiden tarkemmat tyypit. RDF ja RDF(S) kuitenkin pystyvät kuvaamaan vain yksinkertaista olioperustaista semantiikkaa, minkä takia ontologioita varten niistä on lähdetty

kehittämään OWL-kieltä (Web Ontology Language). OWLissa on mahdollista kuvata enemmän ominaisuuksia ja luokkia eli se tarjoaa paremman mahdollisuuden tiedon merkityksen säilymiselle ja siirtämiselle. OWLin käyttäminen ei kuitenkaan ole välttämätöntä ontologioita tehdessä. Yhteenvetona kielet voidaan siis jakaa kahteen osaan eli kuvailukieliin, joita ovat XML ja RDF, sekä skeemakieliin, joita ovat XML(S), RDF(S) ja OWL. (W3C 2004; W3C 2009.)

Suomessa ontologioiden kehitystyötä on tehty järjestelmällisesti vuodesta 2003 lähtien FinnONTO-hankkeissa. Tarve kansallisille ja suomenkielisille ontologioille oli olemassa mm. suomalaisten paikannimien, henkilöiden ja organisaatioiden muodossa. FinnONTO-hanke lähti kehittämään kansallisesti semanttisen webin ontologista sisältöinfrastruktuuria, jota julkiset ja yksityiset palvelut voivat hyödyntää. Sisältöinfrastruktuuri pitää nykyään sisällään ontologiat ja niiden peilaukset (mapping), jotka muodostavat KOKO-kokonaisontologian. Tämän lisäksi sisältöihin kuuluvat ONKI eli kansallinen ontologiapalvelu, jonka avulla ontologioita ja metatietostandardeja kehitetään, julkaistaan ja hyödynnetään. Kolmantena sisältönä ovat metatietostandardit ja neljäntenä erilaiset työvälineet ja parhaat käytännöt. Suomessa tehdyn ontologioiden kehitystyön lähtökohtana ovat olleet sanastot ja niiden muuttaminen ontologioiksi. FinnONTO-hankkeen tavoitteena on täsmentää sanasto-ontologioiden semantiikkaa ja rakennetta, että ne palvelisivat paremmin mm. päättelyä ja visualisointia (Hyvönen 2008b, 21). Sanastopohjaisten ontologioiden ja käsitteistöjen tekemistä varten löytyy helposti teoretietoa esim. Sanastokeskuksen Sanastotyön oppaan ja erilaisten sanastojen muodossa (Suonuuti 2006; Sanastokeskus). (Hyvönen 2008a, 1, 2 – 5.)

4.2 Ontologiat

Ontologioita on tähän mennessä käsitelty laajemmassa viitekehyksessä ja filosofiselta kannalta. Tässä luvussa ja alaluvuissa keskitytään erityisesti tietojenkäsittelyn puolella annettuihin määrittelyihin ontologioista, niiden rakenteesta ja luokitteluista. Tietojenkäsittelyn näkökulmasta ontologioiden merkitys korostuu erityisesti tiedonhaussa ja siksi pyritäänkin tehostamaan löydettävyyttä ontologioiden avulla.

Suomessa KDK-hanke ja FinnONTO-hanke ovat ontologiamäärittelyssään käyttäneet Gruberin määritelmää ”an ontology is a formal, explicit specification of a shared

conceptualization” eli ontologia on formaali ja eksplisiittinen määrittely yhteisestä käsitteistöstä. Gruberin määritelmä perustuu hänen aiempaan klassiseen määrittelyyn “Ontology is an explicit specification of a conceptualization”. Suomenkielisessä määrittelyssä käsite conceptualization on käännetty käsitteistöksi, kun Gruber sen itse määrittää pikemminkin tarkoittavan abstraktia, yksinkertaistettua näkemystä maailmasta. Hän ei siis rajoita sitä käsitteisiin. Formaalisuus määritelmässä puolestaan tarkoittaa koneymmärrettävyyttä ja eksplisiittisyys mallin tarkkaa ja selkeää määrittelyä. FinnONTO-hankkeessa Yleinen suomalainen ontologia (YSO) on määritellyt ontologian myös koneymmärrettävässä muodossa olevaksi tiedoksi, joka ilmenee käsitteinä ja käsitteiden välisinä suhteina. Tietojenkäsittelyssä ontologialla tarkoitetaan siis tiedon kuvailua mielenkiinnon alueista, jonka aihe on koneymmärrettävä tarkka määrittely, johon kuuluu formaalisti määritelty merkitys (Hitzler ym 2009, 2). Astuvansalmen kallion osalta ontologisointi viedään pidemmälle kuin pelkälle käsitetasolle, joten edellä esitellyt Suomessa yleisesti käytetyt määritelmät – tai ainakin niiden käännökset - eivät ole riittäviä. Palaan tähän Astuvansalmen kallion ontologian määrittelyyn selvitettyäni ensin ontologisointiin liittyviä käsitteitä ja erilaisia ontologioita tarkemmin. (Gruber 1993; Hyvönen 2001, 6; Hyvönen 2007, 1; KDK 2010b; Kettula 2009.)

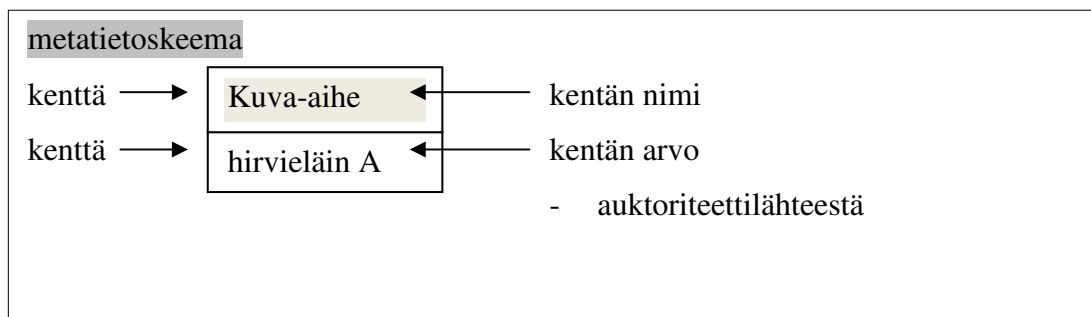
4.2.1 Ontologioiden käsitteet

Ontologioiden määrittelyssä törmää väistämättä käytettyihin käsitteisiin, jotka vaativat omaa määrittelyään. Erityisesti tärkeää on määrittellä käsitteet, joita kussakin ontologiassa käytetään. Asiaan perehtymättömälle on vaikea erottaa millaisesta ontologiasta kulloinkin puhutaan, koska samoja termejä käytetään molemmista. Tässä työssä käytettyjen käsitteiden nimiksi on valittu ensisijaisesti suomenkieliset vastineet CIDOC CRM -käsitteille. CIDOC CRM on ollut mukana työskentelyssä koko prosessin ajan, joten on ollut loogista käyttää sen käsitteitä sekavuuden välttämiseksi. Toisaalta suomenkielinen vastine CIDOC CRM –käsitteelle on tehnyt asiasta ymmärrettävämmän. Esim. CIDOC CRM entity on suomeksi entiteetti, mutten ole tyytynyt tällaiseen helppoon suomennokseen vaan suomennokseksi on valittu suomenperäinen sana kuten luokka tässä tapauksessa. Koska erilaisten käsitteiden määrä on suuri ontologioista puhuttaessa, tässä tapauksessa taulukkoon 1 on koottuna käytetyt käsitteet Astuvansalmen kallion ontologisoinnin osalta.

TAULUKKO 1. Ontologioihin liittyvät käsitteet ja niiden käyttö Astuvansalmen kallion ontologisoinnissa (Mukaillen Denny 2002, 1)

käytetyt termit	CIDOC CRM -termi	muut vaihtoehtoiset termit
luokka - alaluokka - yläluokka	entiteetti (entity) - subclass - superclass	käsite (concept), kategoria (category), tyyppi (type), termi (term), joukko (set), asia (thing)
ominaisuus	ominaisuus (property)	piirre (feature), predikaatti (predicate), lokero (slot), attribuutti (attribute), kriteeri (criterion)
ilmentymä	instanssi (instance)	yksilö (individual), resurssi (resource), laajennus (extension), kuvaus (description), objekti (object), olio (object)

Astuvansalmen ontologisoinnissa käytettävät käsitteet (kuva 4) ovat metatietoskeemasta puhuttaessa kentän nimi ja kentän arvo. Metatietoskeeman teko on mallintamisvaihetta, jolloin skeemaa ei ole vielä muutettu konekieliseksi. Metatietoskeeman muutto konekieliseksi tapahtuu muuttamalla kentän nimet CIDOC CRM –luokiksi, jotka ilmaistaan URI-tunnisteilla.



KUVA 4. Astuvansalmen kallion ontologisoinnissa käytettyjä käsitteitä

Näiden lisäksi ongelmallisia ovat olleet ns. yleiskäsitteet kuten skeemat, formaatit ja käsitemallit. Näitä käsitteitä käytetään hyvin vaihtelevasti ja tarkoittamaan välillä samaa asiaa. Ontologisointiprosessissa nämä on eroteltu toisistaan. Tähän erotteluun palaan Ontologioiden luokittelu –kappaleessa kulttuuriperinnön dokumentoinnin osista puhuttaessa.

4.2.2 Ontologioiden luokittelu

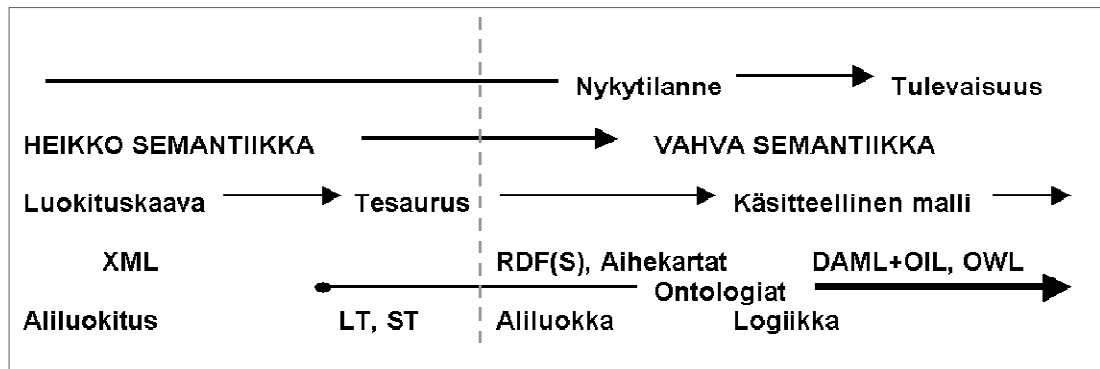
Ontologiat tarjoavat informaation jäsentämiseen yhden vastauksen semanttisen webin alalla. Ontologioita on kuitenkin olemassa hyvin erilaisia ja eritasoisia, kuten myös niistä tehtyjä luokituksia. Yksi näistä luokitteluista on Lassilan ja McGuinnessin (2001) tekemä ontologia-asteikko, jossa heikon semantiikan ontologiat ovat janan vasemmalla puolella ja vahvan semantiikan ontologiat janan oikealla puolella. Janalla oikealle päin siirryttäessä vaatimukset koneymmärrettävyydelle kasvavat ja toisaalta informaatioon liittyvät merkitykset säilyttävät paremmin todellisuutta vastaavat yhteytensä. Lassila ja McGuinnessin mukaan varsinaisia ontologioita ovat formaalit ontologiat, jotka noudattavat tiukkoja osa-kokonaisuussuhteita, sisältävät ominaisuudet ja niiden arvorajoitukset sekä yleiset loogiset rajoitukset. Tesauukset eivät täytä edellä mainittuja vaatimuksia (kuva 5). Astuvansalmen kallion ontologisoinnissa pyrittiin vahvaan semantiikan säilyttämiseen, joten se sijoittuu asteikossa oikealle.

TESAURUS VS.	ONTOLOGIA
Kontrolloitu sanasto (sovitut representaatiot)	Reaalimaailmaan viittaavat representaatiot
Sanojen organisointi	Käsitteiden organisointi
Termit (Käsitys / Käsite)	Käsitteiden merkitys ja määrittely
Syntaksi: Symbolit	Pragmatiikka: Käyttö
Yksinkertaiset suhteet (ST – LT)	Monipuoliset suhteet (myös omavaltaiset)
Heikko semantiikka	Rikas semantiikka
Luokitus käsitteistä, erottelu	Ihmisen tiedollisen mallin simulaatio
Tiedonhaun apuväline ihmisille	"Ymmärtämisen" väline tietokoneille
Termit, hierarkia	Oliot, identiteetti

KUVA 5. Tesauksen ja ontologian keskeisten piirteiden vertailua (Vakkari 2004, 27)

Lassilan ja McGuinnessin (2001) janan vasemmalla puolella ovat sanastot ja yksinkertaiset käsitteistöt, jossa tieto on järjestetty hierarkkisesti. Keskellä ovat tesauukset, joiden sanastot sisältävät termien ja käsitteiden välisiä suhteita. Janan oikealle puolelle sijoittuvat loogiset teorit, jotka sisältävät sääntöjä, jotka varmistavat formaaliuden ja loogisten perusteiden paikkansapitävyyden. Tällaisesta ontologiasta esimerkkinä on CIDOC CRM. Sanastoja, käsitteistöjä ja tesauksia kutsutaan usein erityisesti Suomessa FinnONTO-hankeen puitteissa ontologioiksi, vaikka tällaiset heikon semantiikan ontologiat eivät kykenekään säilyttämään asioiden välisiä suhteita tarpeeksi hyvin konemaailmassa. Kuvasta 6 näkyy kootusti semantiikan

vahvistumiseen liittyvät kuvauskielet ja suhteet. (Kaario 2008, 83, 92; Lassila & McGuinness 2001, 4 – 5.)



KUVA 6. Heikko ja vahva semantiikka (Vakkari 2004, 13)

Lassila ja McGuinness (2001) ovat määritelleet ontologioiden vaatimukset. Ensinnäkin ontologian tulisi sisältää rajallinen kontrolloitu sanasto. Toisekseen luokkien ja termien suhteet tulisi olla tulkittavissa yksiselitteisesti. Kolmanneksi ontologiassa tulisi olla tarkat hierarkkiset alaluokkasuhteet luokkien välillä. Ontologioille tyypillisiä, mutteivät välttämättömiä ovat ominaisuuksien tarkka määrittely luokkaperusteisesti, yksilöllinen sisällyttäminen ontologiassa ja arvorajoitteiden tarkka määrittely luokka-perusteisesti. Toivottavaa olisi myös, että ontologian irralliset luokat olisi määritelty ja että sattumanvaraiset loogiset yhteydet termien välillä olisi määritelty. (Lassila & McGuinness 2001, 4 – 5.)

Poli (2003, 1) erottaa ontologioista puhuttaessa semioottisen ja semanttisen tiedon ontologisen tiedon ulkopuolelle. Hän tekee eron myös kunnollisen ontologisen tiedon ja näennäisontologisen tiedon välillä. Polin mukaan ontologian pitäisi pystyä kertomaan, että tietty objekti on sijoitettu jonnekin tai että tapahtuma on tapahtunut tiettyä aikana. Ontologia ei pakota valitsemaan tiettyä kieltä tai koordinaattijärjestelmää. Ontologiat hän jaottelee kolmeen ryhmään eli kuvailevaan, luokittelevaan ja formalisoituun ontologiaan. Nämä eivät ole kuitenkaan toisistaan erillisiä, vaan vaikuttavat eri tavoin toisiinsa. Kuvailevat ontologiat kattavat elävät ja elottomat asiat, toiminnot, prosessit ja niiden tuotokset, ajatukset, tunteet ja päätökset, kielet ja tavat jne. Luokittelevat ontologiat järjestävät eri tavoin kuvailevien ontologioiden tuloksia. Luokkia voivat olla esim. asia, muoto, kokonaisuus, osa. Nämä luokat kuvaavat todellisuuden tyypejä tai ulottuvuuksia ilman, että ne vielä liittyvät tiettyyn formaaliuteen. Formalisoidut ontologiat vaativat sopivan formaalin

kodifikaation. Tämä liittyy ontologian formaaliuden ilmaisukyvyyn, koneymmärrettävyyden ja tietoisuuden arviointiin. (Poli 2003a).

Guarino & Giaretta (1995, 1) määrittelevät ontologiat seitsemälle tasolle, jotka liittyvät sekä filosofiaan että tietojenkäsittelyyn. Ontologia voidaan heidän mukaansa tulkita seuraavin tavoin: filosofian tieteenala, epäformaali käsitesysteemi, formaali semanttinen kuvaus, käsitteellistämisen määrittely, loogiseen teoriaan liittyvä käsitesysteemi, loogisen teorian käyttämä sanasto ja loogisen teorian metatason määrittely.

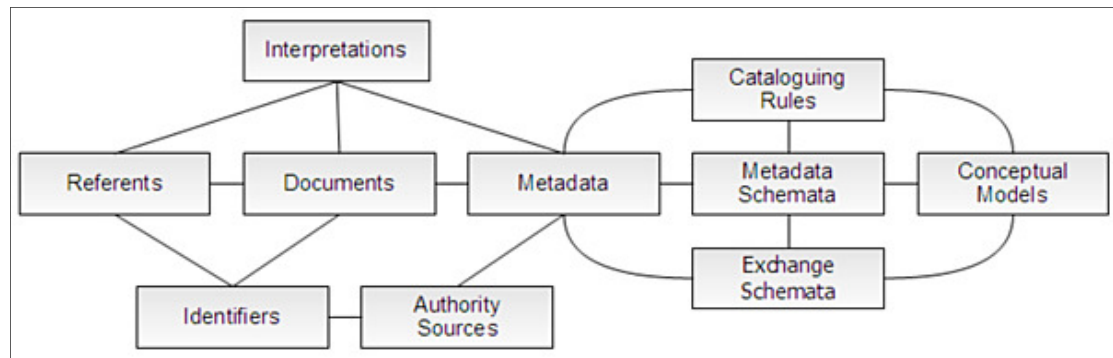
Projektin aluksi hahmoteltiin ontologioiden maailmaa. Mika Nymanin mukaan ontologioita voi käsitellä neljällä kerrostumalla. Ensinnäkin puhutaan käsitelmalleista, joita voidaan pitää ontologioina. Ne ovat käsitteellinen kuvaus suhteista. Toisaalta käsitelmalleihin liittyvät toiminnalliset määrittelyt (functional requirements) eli käytännössä kuvataan käsitelmää, joka kertoo mitä missäkin pitäisi olla. Tästä yksi esimerkki on FRBR (Functional Requirements for Bibliographic Records). Sekä käsitelmalleja että toiminnallisia määrittelyjä on kuvattu ER:llä (entity relations). Kirjastojen tuleva kuvauskieli RDA (Resource Description and Access) on tästä esimerkkinä. ER:n lisäksi on oliopohjaisia malleja. OWL on ontologioiden kuvauskieli, joka pohjautuu RDF-kuvauskieleen. Ontologiat eivät kuitenkaan edellytä OWLin käyttöä. (Viva3-Ephemera metatietoryhmän kokous 2009.)

Yhteenvedon erillisiä malleja ovat tekstipohjaiset käsitelmällit, ER, oliopohjaisuus ja formalisoitu OWL. Toisena kerrostumana on XML-pohjaiset (eXtensible Markup Language) metatietostandardit esim. METS (Metadata Encoding & Transmission Standard), MODS (Metadata Object Description Schema), TEI (Text Encoding Initiative). Kolmas kerrostuma ovat keräily- ja siirtoformaatit ja neljäs kerrostuma semanttinen web. (Viva3-Ephemera metatietoryhmän kokous 2009.)

Ontologioiden ideana on kertoa miten asiat liittyvät toisiinsa eli ne tarjoavat mahdollisuuden yhdistää toisiin sidoksissa olevia kohteita ja sisältöjä toisiinsa. Tähän tarkoitukseen XML ja Dublin Core ovat liian rajoittuneita tapoja kuvata asioita. Semantisoidun järjestelmän elementtinä on ontologia, joka on käsitteellinen kuvaus. OWL voi toimia kuvauskielenä ja Protegé välineenä. Yksilöivien ja mielellään globaalien tunnisteiden avulla pystytään linkittämään eri järjestelmiä ja saadaan

linkitettyä eri tahojen samat aiheet keskenään. Kyselykielenä on SPARQL. (Viva3-Ephemera metatietoryhmän kokous 2009.)

Koska ontologiat ovat hyvin erilaisia, on hyvä huomata, että ne pohjautuvat periaatteessa hyvin samanlaiseen perustaan. Projektissa on pyritty selittämään näiden kulttuuriperinnön dokumentoinnin osien ja ontologisointiin liittyvien erilaisten asioiden yhteyksiä (kuva 7). Tämä auttaa samalla selkeyttämään käytettyjä käsitteitä.



KUVA 7. Kulttuuriperinnön dokumentoinnin osat (Hannus ym. 2010)

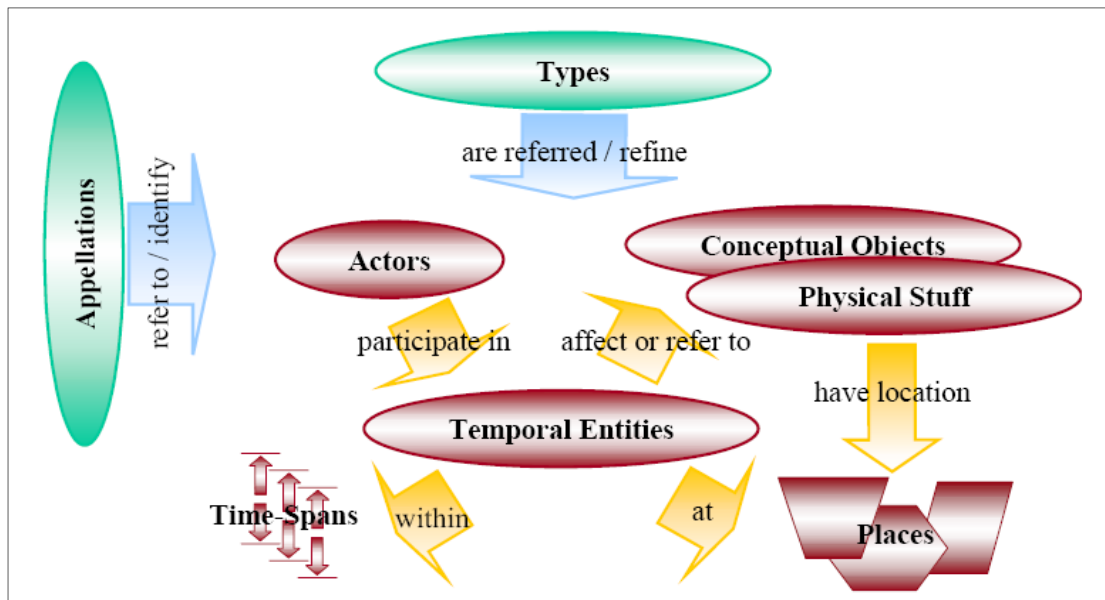
Astuvansalmen kallion osalta kuvan 7 mukaisia käsitteitä on sovellettu seuraavasti. Kartoitettu aineisto tullaan käsittemallintamaan (conceptual models). Käsittemallin perusteella tehdään metatietoskeema (metadata schema), joka muutetaan konekieliseksi. Metatietoskeeman perusteella tietokantaan syötettävät tiedot ovat metatietoa (metadata). Tietojen syöttämisessä on hyvä käyttää yhteisiä sääntöjä eli luettelointisääntöjä (cataloguing rules). Syötettävät metatiedot tulevat osittain auktorisoiduista lähteistä (authority source) eli esim. sovitusta käsitteistöistä, kontrolloiduista asiasanastoista tai auktoriteettitietueesta. Auktorisoidulla lähteellä tulee olla yksilöivä tunniste (identifier) eli useimmiten käytetään URI-tunnistetta.

Ontologioiden olisi mielestäni hyvä saavuttaa se mitä kirjastoissa voi tehdä saapumalla tietyn hyllyluokan kohdalle ja selaamalla aineistoja eli tiedon etsiminen ilman tietoista päämäärää. Eli verkossa olevien tietojärjestelmien tulisi saavuttaa samankaltainen suhteitten näennäisesti sattumanvarainen löytäminen. Linkitykset ja viittaukset eivät enää riitä vaan tarvitaan aihepiireihin liittyviä kokonaisuuksia ja aiheiden välisiä erilaisia yhteyksiä. (Koivunen 1995.)

Ontologioille on esitetty myös kritiikkiä. Kritiikki on kohdistunut menetelmän työläyteen, teoreettisia lähtökohtia on kyseenalaistettu, on myös mietitty voiko koneille tarkoitetut ontologiat olla myös ihmiselle ymmärrettäviä, olemassa olevat ontologiat ovat omalta osaltaan keskeneräisiä ja virheellisiä eli kokonaisuus ei ole kunnossa.

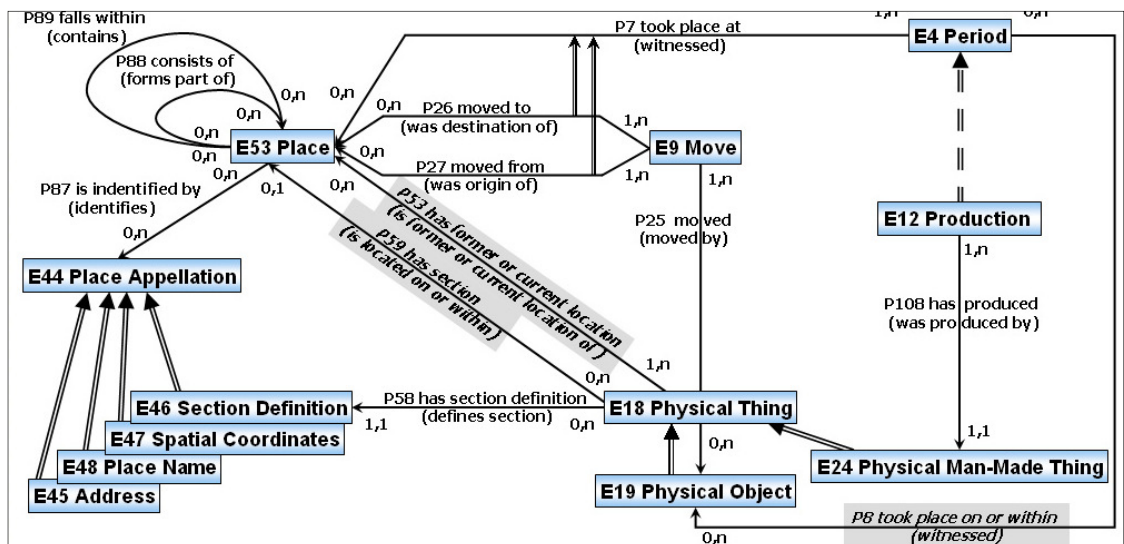
4.2.3 CIDOC CRM

CIDOC CRM on kehitetty, koska on ollut tarve mallille, joka kykenee ilmaisemaan monimutkaisia merkitysyhteyksiä ilman tiedon menettämistä. CIDOC CRM on formaali ontologia kulttuuriperinnön tallentamista varten. Siitä puhutaan myös kuvailun metadataformaattina ja siitä on olemassa kansainvälinen semanttisen webin dokumentointistandardi CIDOC CRM ISO 21127:2006. CIDOC CRM-malli on tapahtuma- ja oliopohjainen. CRM koostuu entiteeteistä (Entity) eli luokista ja ominaisuuksista (Property). Luokkia on 90 kpl ja ominaisuuksia 148 kpl. Ominaisuudet määrittävät kahden luokan väliset tietäntyyppiset suhteet. Ominaisuudet periytyvät yläluokilta alaluokille. Idea on saada CRM:n avulla kuvailusta tiedosta esille tapahtumat, joita ei alkuperäiseen kohteeseen ole tallennettu. Nämä merkitykset pysyvät myös samana tietoa järjestelmästä toiseen siirrettäessä mikäli formaatti pysyy samana. CIDOC CRM voidaan käyttää siis skeemojen ja formaattien luomisen apuvälineenä, olemassa olevan tiedon yhdistämisen kielenä ja tiedonsiirtoformaattina. CIDOC CRM:n tavoitteena on tarjota valmiit luokitukset ja semanttiset määritelmät kulttuuriperintötiedon saamiseksi erilaisista paikallisista järjestelmistä yhdenmukaiseksi kansainväliseksi tietomuodoksi. Toimijat, kohteet, paikat ja ajankohdat yhdistetään toisiinsa tapahtumilla. CIDOC CRM:n tärkeimpiä luokkia on esitelty kuvassa 8. (Definition of the CIDOC Conceptual Reference Model 2010; Hannus ym. 2010; Simovaara 2004, 52.)



KUVA 8. CIDOC CRM:n laadullinen metaskaema (Doerr ym. 2009)

Kuvasta 9 voi nähdä esimerkinomaisesti kuinka CIDOC CRM:n avulla voi kuvata paikkatietoja. E18 Physical Thing eli esim. kallio, joka on E24 Physical Man-Made Thingin eli esim. kalliomaalausten yläluokka, liittyy tiettyyn paikkaan ominaisuuden P53 tai P59 kautta. Kalliomaalausten tuottaminen (P12) puolestaan liittyy tiettyyn aikaan (E4) ja siten tiettyyn paikkaan. Kuhunkin luokkaan liittyy scope note.



KUVA 9. Paikkatiedot CIDOC CRM:ssä kuvattuna (CIDOC CRM 2010)

CIDOC CRM:n ohjeistuksessa kustakin luokasta käy ilmi sen ylä- ja alaluokat, käyttölaajuus, käyttöesimerkit sekä luokkaan soveltuvat ominaisuudet. Kuvassa 10 on

esimerkinomaisesti fyysisen ihmisen tekemän asian (D24) ohjeistus. Erityisesti scope note eli käyttölaajuus kertoo tarkemmin kyseisen luokan soveltamisesta.

E24 Physical Man-Made Thing	
Subclass of:	E18 Physical Thing E71 Man-Made Thing
Superclass of:	E22 Man-Made Object E25 Man-Made Feature E78 Collection
Scope Note:	This class comprises all persistent physical items that are purposely created by human activity. This class comprises man-made objects, such as a swords, and man-made features, such as rock art. No assumptions are made as to the extent of modification required to justify regarding an object as man-made. For example, a “cup and ring” carving on bedrock is regarded as instance of E24 Physical Man-Made Thing.
Examples:	<ul style="list-style-type: none"> ▪ the Forth Railway Bridge (E22) ▪ the Channel Tunnel (E25) ▪ the Historical Collection of the Museum Benaki in Athens (E78)
Properties:	<p>P62 depicts (is depicted by): E1 CRM Entity (P62.1 mode of depiction: E55 Type)</p> <p>P65 shows visual item (is shown by): E36 Visual Item</p> <p>P128 carries (is carried by): E73 Information Object</p>

KUVA 10. E24 Physical Man-Made Thing (Definition of the CIDOC Conceptual Reference Model 2010)

CIDOC CRM on joustava ja laajennettavissa. Se ei ota kantaa kenttien arvoihin eikä dokumentoinnin kohteeseen. Se myös huomioi kohteen koko elämänkaaren. Astuvansalmen tapauksessa CIDOC CRM:ää hyödynnetään luokkatasolla. Oleellista on valita sopivat luokat haluttujen asioiden ilmaisemiseksi. Mika Nyman tekee Viva3-hankkeessa lisäksi erillisen raportin CIDOC CRM:n käyttöönnotosta sekä Astuvansalmen kallion ontologisoinnin osalta jatkaa työtä CIDOC CRM:ää laajentamalla. (Herrala 2009.)

Loogiset ontologiat kuten CIDOC CRM mahdollistavat siis fyysisen kohteen, sen kontekstin ja ominaisuuksien sekä esim. kohteesta luotujen kirjallisten ja kuvallisten aineistojen liittäminen toisiinsa. Lisäksi kohteeseen voidaan liittää esimerkiksi nimi tai tunniste, tyyppi ja jokin käsitteellinen sisältö. Kohteen ympärille punoutuu tällä tavalla semanttinen verkko. Verkkorakenteelle on ominaista, että on olemassa navigointimahdollisuus jokaisesta solmusta mihin tahansa toiseen solmuun. CIDOC CRM:n avulla voidaan myös luoda ns. oikoteitä. Navigoinnin ei tarvitse noudattaa mitään tiettyä järjestystä. (Nyman 2011f.)

4.3 Ontologisoinnin prosessit

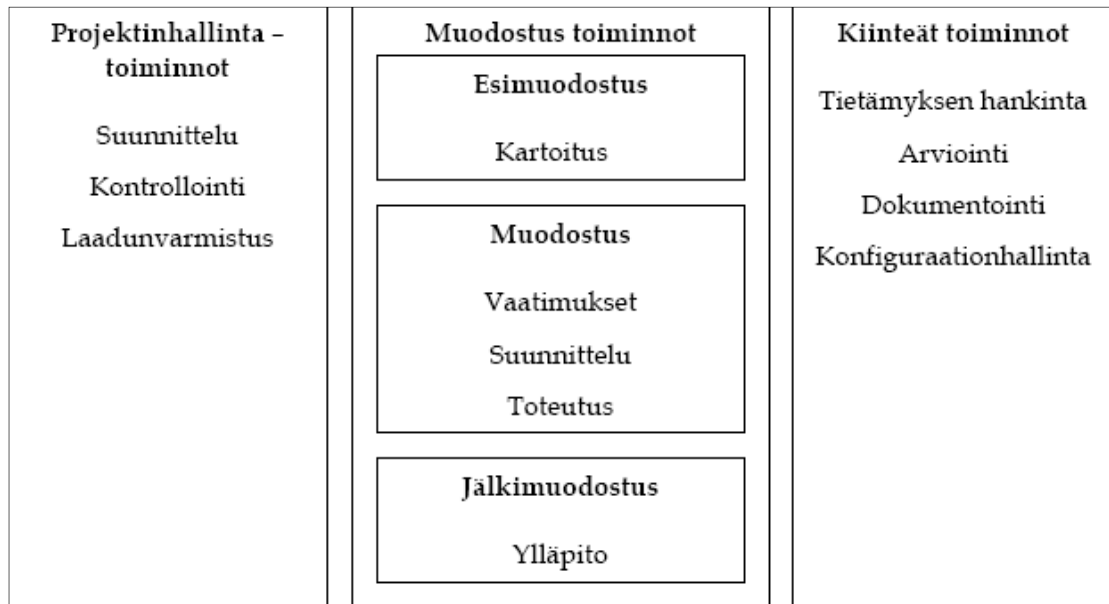
Opinnäytetyön tarkoituksena on kuvata ontologisoinnin prosessia. Ontologioiden erilaisuudesta johtuen myös muodostusprosessit ovat erilaisia. Esittelen erilaisten ontologisointien yhteydessä käytettäviä prosesseja ja arvioin niiden hyödynnettävyyttä Astuvansalmen ontologisoinnissa.

Kaarion (2008, 95, 112) ontologiasuunnittelussa lähtökohtana ovat taksonomiat. Hänen mukaansa ontologisointi voidaan jakaa seuraaviksi tehtäviksi: ontologian suuruuden rajaaminen ja sovelluskohteen määrittely, selvitys olemassa olevien ontologioiden hyödyntämismahdollisuuksista, kuvattavien käsitteiden luettelu, taksonomian määrittely, käsitteiden suhteiden määrittely, käsitteiden rajoitusten määrittely, ilmentymien määrittely ja ontologian yhtenäisyyden tarkastus. Kaario pitää tärkeänä erilaisten välineiden hyödyntämistä suunnittelussa, että ajatuksia ja ideoita saadaan testattua. Varsinaisilla ontologiamallinnustyökaluilla voidaan hioa lopputulosta.

FinnONTO-projekti perustuu sanastoille ja tesauruksille, joten Kaarion ontologiasuunnitteluprosessia voitaneen pitää myös sen lähtökohtana. Projektissa on kuitenkin kehitetty omia työkaluja prosessin mahdollistamiseksi ja ontologioiden käyttöönoton helpottamiseksi. FinnONTO:n sanastopohjaiset ontologiat perustuvat jaetun tunnisteiden ympärille. Kaikki käsitteet saavat siis jaetun URI-tunnisteen. URIt löytyvät projektin luoman ONKI-palvelun kautta. ONKI toimii myös julkaisukanavana muille sanastopohjaisille ontologioille. Käsitteiden väliset suhteet mallinnetaan ontologiaksi. Kuvailun (annotointi) avuksi on luotu niin puoliautomaattisia kuin automaattisiakin välineitä mm. selainpohjaisen annotointityövälineen SAHAN. Sanaston vieminen ONKI-palveluun tapahtuu muuntimen avulla. (Hyvönen 2008a, 2.)

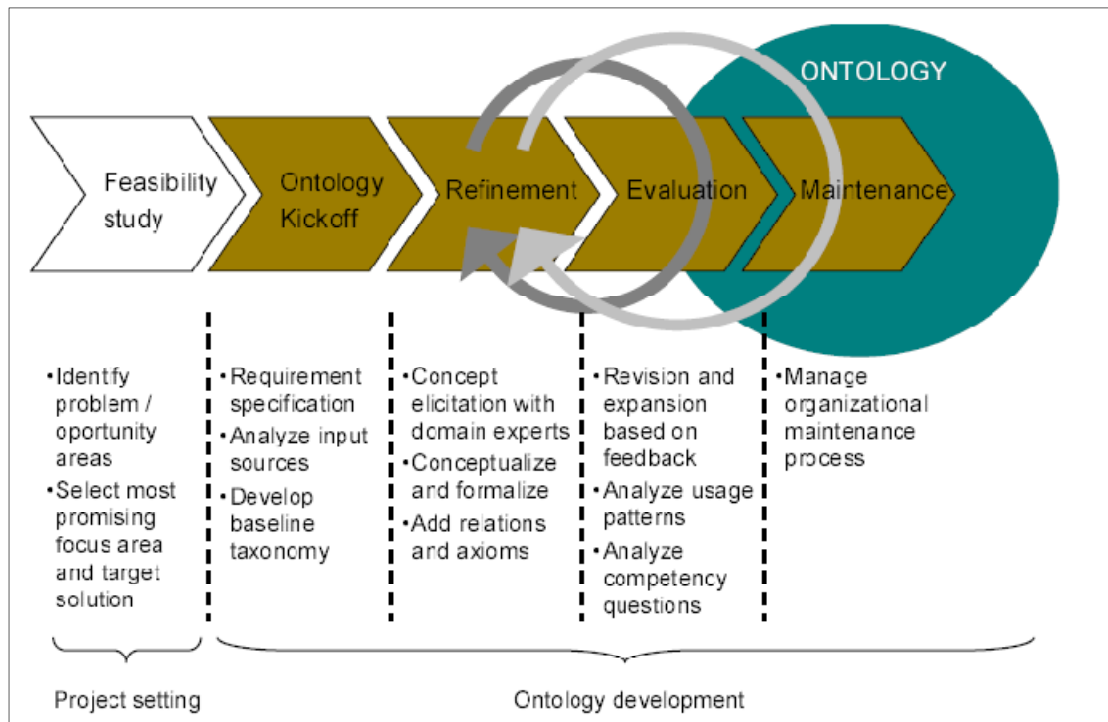
Edelliset esimerkit kuvaavat Suomen tilannetta. Näissä sanastopohjaisissa ontologioissa yhdistyvät sekä sanastotyön prosessit että tietojen mallintamisen prosessit. Tässä työssä ei ole tarpeen käsitellä näitä heikon semantiikan prosesseja tarkemmin.

Fernández-López (1999) on hahmotellut ontologisointiprosessia erilaisista toiminnoista koostuvaksi (kuva 11). Fernández-Lópezin prosessi perustuu ohjelmistojen kehitysprosessistandardiin IEEE 1074–1995. Astuvansalmen kohdalta keskitytään seuraamaan erityisesti muodostustoimintoja. (Ruotsalainen 2008, 19.)



KUVA 11. Ontologioiden muodostusprosessiin liittyvät toiminnot (Ruotsalainen (2008, 19) mukaillen Fernández-Lópezia (1999))

Staab ym. (2000) ontologisointiprosessi koostuu toteutettavuustutkimuksesta, aloitus- ja määrittelyvaiheesta, kehitysvaiheesta, arviointivaiheesta sekä ylläpitovaiheesta (kuva 12). Astuvansalmen osalta toteutettavuustutkimusta ei tehty, vaan mentiin suoraan aloitusvaiheeseen, jossa kylläkin toteutettavuuden suhteen keskusteltiin eri näkökulmia ja ontologisoinnin kohteen valinnasta. Aloitus- ja määrittelyvaihe sisältää Astuvansalmen osalta siis osia toteutettavuustutkimuksesta ja osia aloitusvaiheesta. Staab ehdottaa toteutettavuustutkimuksen suorittamista varten CommonKADS – tutkimusmenetelmää.



KUVA 12. Staabin ontologisointiprosessi (The Knowledge Meta Process) (Staab ym. 2000)

Staabin ym. (2000) aloitusvaiheeseen kuuluu ontologiavaatimusten määrittelydokumentti (ORSD, ontology requirements specification document). Se sisältää ontologian tavoitteet, aiheen, kohteen ja alan, ontologiassa tuetut sovellukset, tietolähteet, mahdolliset käyttäjät ja käyttäjäennusteet, tietojärjestelmän kykykyselyn sekä mahdolliset uudelleenkäytettävät ontologiat. Astuvansalmen osalta ei ollut olemassa ennestään järjestelmää vaan lähdettiin alusta liikenteeseen, minkä takia sekä tietojärjestelmäkysely että entiset ontologiat olivat poissuljettuja määrittelydokumentista.

Staabin ontologisointiprosessin kehitysvaihetta olen itse kutsunut toteutusvaiheeksi Astuvansalmen työvaiheistuksessa. Staabin kehitysvaiheessa tulisi aloitusvaiheen alustavan taksonomian perusteella luoda ydinontologia. Tässä työskentelyssä hyödynnetään asiantuntijoita, asioita käsitteellistetään ja tehdään konekieliseksi, jolloin ydinontologiasta tehdään tapauskohtainen ontologia. (Staab ym. 2000.)

Staabin ym. (2000) mukaan arviointi tapahtuu vertaamalla asioita vaatimusmäärittelydokumenttiin, testaamalla halutussa sovelluksessa sekä käyttökokemuksia hyödyntämällä. Arviointi ei ole irrallinen osa vaan yhteydessä kehitysvaiheeseen. Näitä kahta vaihetta toistetaan tarvittava määrä, että saavutetaan

haluttu lopputulos. Ylläpitovaiheessa voi tulla myös osaltaan asioita mm. käyttäjien toimesta esille, joiden pohjalta kehitystyötä ja arviointia tehdään ja ontologiaa muutetaan tarvittaessa. Ylläpitovaihe on voimassa niin kauan kuin ontologiaa käytetään.

Staabin prosessi on ollut keskeinen tekijä mm. Stakesin (nykyinen THL) rahoittamassa sähköisten palvelujen mahdollisuuksia koskevassa projektissa (Hotti 2008). Staabin prosessi on sovellettavissa sekä heikon että vahvan semantiikan ontologisoointeihin, vaikkakin siinä puhutaan pääasiassa taksonomialähtöisesti. Tässäkin opinnäytetyössä Staabin prosessia on käytetty alustavana lähtökohtana teolle ja asioiden jäsentelylle sen sovellettavuuden ja yksikertaisuuden takia. Prosessia on noudatettu pääasiassa yläotsakkeiden tasolla, ei kaikilta osin Staabin mukaan niiden sisältämällä asioilla, joita tosin on hyödynnetty tarpeen mukaan. Myöhemmin esiteltävä työvaiheiden kuvaus mukailee Staabin jaottelua.

Kuten edellä on esitetty, ontologioita on mahdollista sekä kuvata että tehdä hyvin eri tavoin. Seuraavassa kappaleessa keskitytään erityisesti prosessien kuvaamisen vaihtoehtoihin.

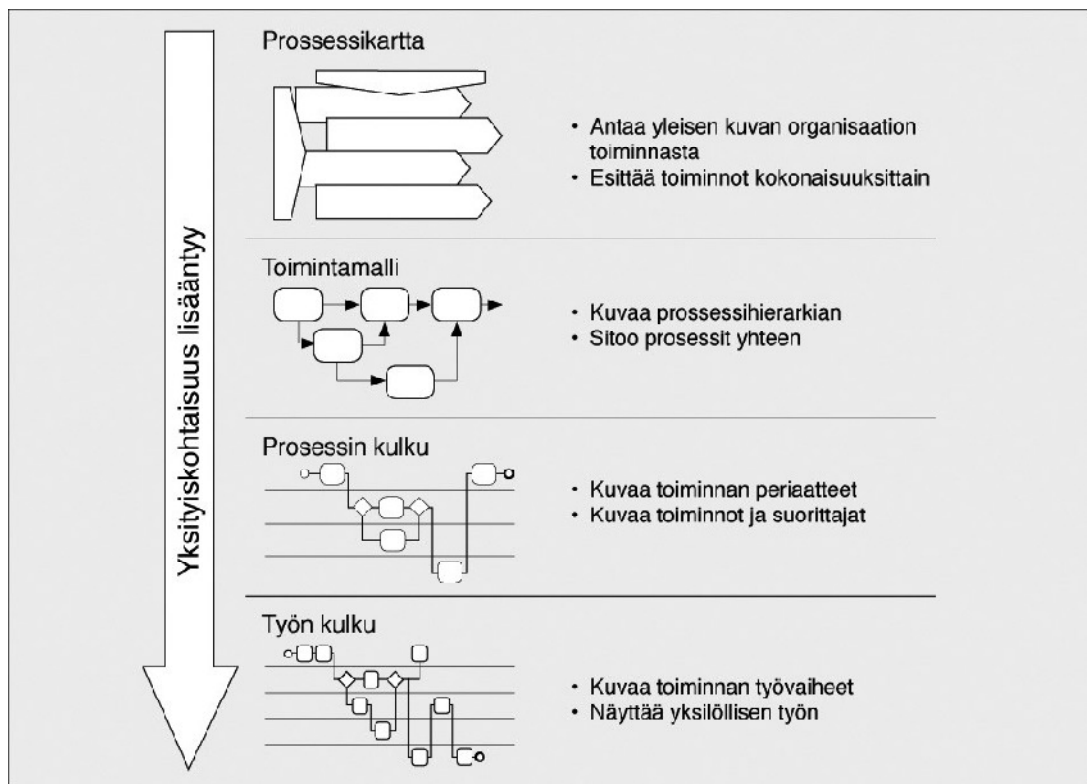
5 PROSESSIT JA PROSESSIEN KUVAAMINEN

Prosesseja kehitetään organisaation suunnittelua ja kehittämistä varten ja ne pohjautuvat aina organisaation visioihin ja strategioihin (JHS 152 2002). Viva3 – hankkeen tavoitteena on ollut testata ontologioiden käyttöä. Tämän prosessin dokumentointia ja mahdollista jatkokehittämistä varten kuvataan ontologisoinnin prosessi. Hankkeen tulisi palvella myös ympäröivää koulutusorganisaatiota ja sen opetuksen kehittämistä. Prosessikuvausta on toivottavasti mahdollista käyttää tulevaisuudessa myös opetuksen apuna ontologisointiin perehdyttäessä.

Prosesseja voidaan kuvata hyvin eri tavoin tilanteesta riippuen. Julkishallinnon puolelle on Suomessa olemassa omat suosituksensa prosessien kuvaamisesta. Yksityisellä puolella erityisesti liiketalouden käyttöön on olemassa suuri valikoima erilaisia prosessimallinnustapoja. Mikkelin ammattikorkeakoulussa virallisia prosesseja on kuvattu Microsoft Vision avulla yksisivuisina prosesseina, joista käy

ilmi taustamateriaalit, prosessit, vastuuhenkilöt ja tuotetut dokumentit. Edellisessä luvussa käsiteltiin myös ontologisoinnissa käytettyjä prosesseja.

Julkisen hallinnon tietohallinnon neuvottelukunnan prosessien mallinnuksessa (JHS 152 2002) sovelletaan BPMN-määrittystä (Business Process Modeling Notation), joka määrittelee kuvauksessa käytettävät symbolit. Käytettävien symbolien määrittely yhdenmukaistaa organisaation käytäntöjä. Määrittely ja sen dokumentointi omalta osaltaan kertovat ja avaavat prosesseja ja kuinka niitä tulisi tulkita. Prosesseja voidaan kuvata eri tasoilla. JHS 152 -suosituksen mukaan nämä tasot koostuvat prosessikartasta, toimintamallista, prosessin kulusta ja työnkulusta (kuva 13).



KUVA 13. Prosessien kuvaustasot (JHS 152 2002, 7)

Mikä sitten on sopiva kuvaustapa Astuvansalmen kalliota varten? Hankkeen edetessä Astuvansalmen kallion prosessikuvauksesta keskustellessa toiveena oli yksinkertaisuus ja asian selkeyttäminen asiasta tietämättömille. Prosessikuvauksen tulisi siis palvella prosessista oppimista. Apilo ym. (2007, 116 – 117) mukaan organisaation oppiminen koostuu erilaisista tasoista. Alkuvaiheessa tietoa kerätään ja omaksutaan, ollaan mukana ja jaetaan kokemuksia. Tämän pohjalta tulkitaan kyseessä olevaa asiaa keskustelemalla, ideoimalla ja arvioimalla, mikä luo ryhmän osaamisen ja yhteisen ymmärryksen asioista. Osaamista ja tietoa yhdistetään, jonka jälkeen

osaaminen vakiinnutetaan toimintatapojen, ohjeiden ja prosessikuvausten avulla. Astuvansalmen kallion ontologisointiprosessi olisi hyvä huomioida nämä asiat, että oppimisen eri tasot pääsisivät esille ja että oppiminen toteutuisi prosessin avulla. JHS 152 –suositusta tullaan käyttämään Astuvansalmen ontologisoinnin ohjenuorana ja soveltamaan sen teoreettista taustaa prosessien kuvaamisessa. Suosituksen hyödyntäminen on perusteltua, koska se määrittelee Suomen tasolla pitkälti suositeltavat tavat kuvata prosesseja. Kansainvälisesti hyväksytyt symbolit ovat helposti tulkittavissa myös Suomen ulkopuolella. Toinen merkittävä prosessin kuvausta määrittävä dokumentti on Staabin ontologisointiprosessi, koska se mielestäni tuo hyvin esille prosessin kehämäisyyden ja kehittämisen ja se on tehty erityisesti ontologisointia varten eli sovellusalue on sama (Staab ym. 2000).

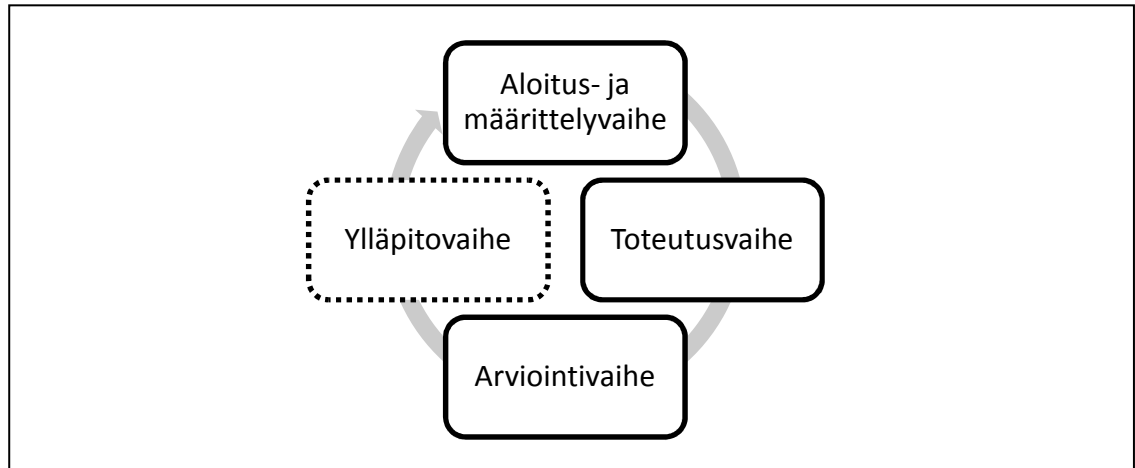
JHS 152 –suositus (2002, 8) määrittelee toimintamallitason sisältävän prosessihierarkian, jossa prosessit jakautuvat osaprosesseiksi. Toimintamalli kuvaa prosessien kulun, prosessien väliset riippuvuudet, omistajat, vastuut, tavoitteet ja tulokset. Toimintamallin avulla on mahdollista hahmottaa kokonaisuus ja sitoa eri prosessien osat yhteen, mikä onkin sen hyöty juuri Astuvansalmen kallion ontologisointiprosessin kuvauksessa.

Astuvansalmen kallion ontologisointiprosessi tullaan kuvaamaan toisaalta yksinkertaisen kaavion avulla mukailien Staabia ja toisaalta toimintamallitasolla hyödyntäen JHS 152 –suositusta (Staab ym. 2000). Toimintamallin kuvaamisen tavoitteena on dokumentoida juuri tämä nimenomainen Astuvansalmen kallion ontologisointi. Kaavio puolestaan yleistää prosessikuvauksen laajemmin käytettäväksi. Kunkin toiminnon toimintamalli löytyy seuraavassa kappaleessa esiteltävän tehtävänkuvauksen lopusta.

6 ONTOLOGISOINTI JA TOIMINTAMALLIT: TYÖVAIHEIDEN KUVAUS

Kuten Staab ym. (2000) aiemmin määritteli, ontologisointiprosessi alkaa aloitus- ja määrittelyvaiheella, josta siirrytään toteutus-, arviointi- ja ylläpitovaiheeseen (kuva 14). Käsittelen asiat pääsääntöisesti jäsentäen ne näiden vaiheiden kautta. Tosin ylläpitovaihetta ei tässä työssä käsitellä. Monet asiat ovat todellisuudessa tapahtuneet päällekkäin ja joihinkin asioihin on palattu myöhemmin. Eheän näkemyksen

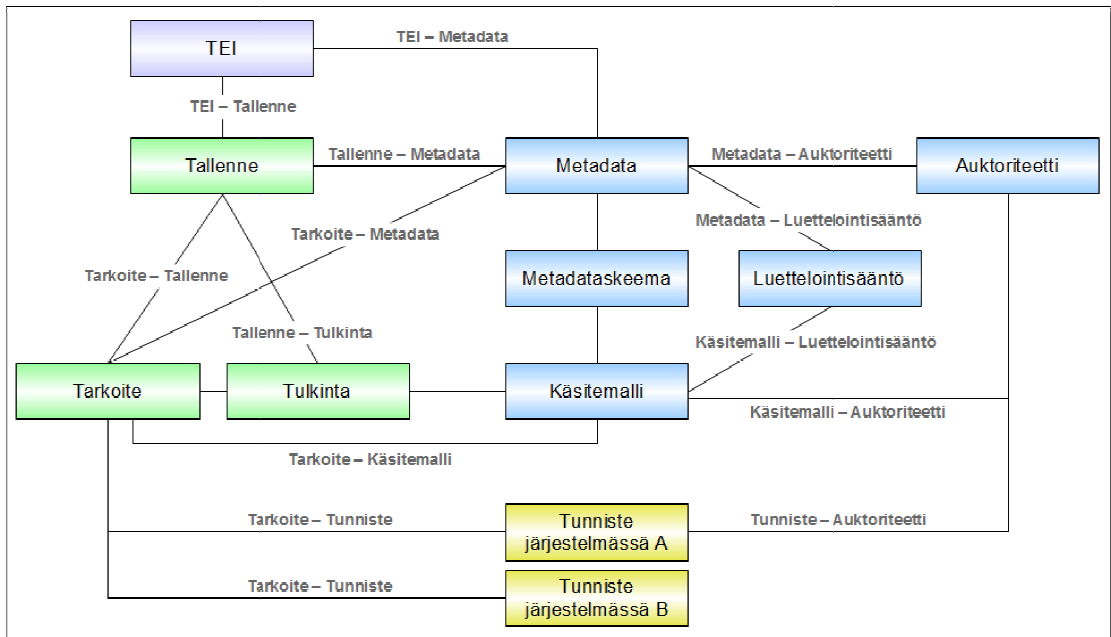
saamiseksi prosessin vaiheista joitain asioita on koottu saman otsikon alle, vaikka asiat eivät prosessuaalisesti olisi tapahtuneet näin selkeästi. Näistä poikkeamista on ainakin osittain maininnat tekstissä. Kaikilta osin eheyteen ei ole päästy, koska prosessin aikana on ilmennyt uusia asioita, joihin on reagoitu.



KUVA 14. Toimintamallin vaiheet

Prosessin kuluessa Mika Nyman (2010b) esitteli ontologioiden kokonaisuutta kuvaavan kaavion (kuva 15), joka selventää ontologisointiprosessin eri osa-alueita. Prosessin osalta esittelen myöhemmin erityisesti käsitelmän että metatietoskeeman luomiset. Tästä opinnäytetyöstä erillisenä, mutta hankkeeseen muutoin oleellisesti kuuluvana on tunnisteiden luominen. Tunnisteita on mahdollista saada joko auktoriteettikannasta tai muualta. Kuvioista käy selkeästi esille erilaisten tulkintojen dokumentointi tallenteisiin eli kukin kirjailija tekee omia tulkintojaan kalliokuvista ja tällöin tallenteena toimii kirja.

TEI-standardi (Text Encoding Initiative) on mielenkiintoinen lisä prosessissa, josta keskusteltiin, mutta jota ei lähdetty toteuttamaan. TEI mahdollistaa esim. tallenteen tekstin tai kuvien pilkkomisen isompiin tai pienempiin osasiin. Kalliomaalauksesta on mahdollista mm. erottaa jousi ihmisen kädestä erilleen tai tekstistä saadaan viittauksia tietyille sivuille, kappaleille, riveille tai jopa sanatasolle. TEI mahdollistaa tekstin rakenteisuuden, joka puolestaan helpottaa tiedonhakua. TEI olisi myös antanut mahdollisuuden sosiaalisuudelle tietokannassa mahdollistaen käyttäjien tulkintojen tekemisen esim. kalliokuvista. (Nyman 2010b.)



KUVA 15. Kokonaisuus (Nyman 2010a)

6.1 Aloitus- ja määrittelyvaihe

Prosessi alkoi ensimmäisellä kokoontumisella kesäkuussa 2009. Kokoontumisessa Mika Nyman kertoi ontologisoinnin mahdollisuuksista ja mistä asioista se koostuu. Samana päivänä järjestettiin Viva3- ja Ephemera -hankkeiden ensimmäinen yhteistyöpalaveri ja hankkeiden tavoitteita määriteltiin.

Viva3-hankkeen tavoitteena oli kokeilla ontologisointia. KDK:n (Kansallinen Digitaalinen Kirjasto) standardisalkkua pidettiin hyvänä ja riittävänä lähtökohtana työskentelylle, joka varmistaa osaltaan ontologisoinnin olevan linjassa muualla tehtävän työn kanssa. Rajapintojen luominen yhteisillä standardeilla osaltaan helpottaa aineistojen yhdistämistä laajempien ontologioiden kanssa.

Tekemisen pääpainoksi määriteltiin oppimisprosessi ja prototyypin testaus. Tarkoituksena oli kokeilla asioita, joiden toiminnasta ei ollut varmuutta. Kokeilu konkretisoitui oppimisena.

Käsitemallin osalta keskusteltiin aineistolähtöisyydestä eli lähtökohtana ovat aineistot ja käyttäjätarpeet. Aineistosta on olemassa oleva dokumentaatio ja tästä saatavilla olevasta dokumentaatiosta luodaan skeemat. Kohdetta tarkastellaan suoraan eli tässä tapauksessa kohteena ovat Astuvansalmen kallio ja sen kalliomaalaukset.

Astuvansalmen kallio muodostui aineistoksi, koska siitä olemassa oleva dokumentaatio on sopivasti rajallinen ja aihe on mielenkiintoinen maailmanlaajuisesti. Työtä tehdessä tuli huomioida tuotteistamisen mahdollisuudet, vaikka tarkoituksena ei ainakaan tässä vaiheessa ollut luoda valmista tuotetta ontologisoinnin ympärille. Projektipäällikkö Juhani Grönhagenilla merkittävä asiantuntemusta koskien kalliota ja sen dokumentaatiota tuli hyödyntää mahdollisuuksien mukaan.

Museoviraston näkökulmana on ollut, että heidän käytössään olevat tietokannat ja niiden nykyiset metatietoskeemat eivät ole riittäviä. Tämä oli yksi syy, miksi lähdettiin etenemään tapahtumapohjaisuutta kohti. Tapahtumapohjaisuus ilmenee CIDOC CRM:ssä, joka valittiin ontologisoinnin pohjalle.

Alkuvaiheessa etenemisjärjestykseksi eli prosessiksi sovittiin:

1. *Aineistojen kartoitus.* Sisältää tarvittaessa aineistojen skannauksen ja kuvaamisen. Aineistot sisältävät sekä bibliografisia tietoja, piirroksia, kuvia että käsiteltyjä kuvia. Vastuu Henna Mölsä.
2. *Pienimuotoisen tietokannan mallinnus.* Aineistojen syöttämistä varten. Vastuu jäi epäselväksi.
3. *Tietokannan luominen Protégélla.* Tietokantaa ja ontologiaa käytetään apuna kysymyksenasettelussa koskien aineistoja. Vastuu Mika Nyman.
4. *Ontologisten suhteiden miettiminen.* Mietitään prosessi ja kontekstit, jotka liittyvät esineisiin ja kuviin. Tarvittavia tietoja ovat mm. aika, paikka, nimi, tunniste, fyysiset objektit, toimijat, käsite, symboli ja tapahtuma. Protégéä käytettäessä on mietittävä miten relaatiot ja ominaisuudet saadaan ilmentymään, miten moniperiytyminen on mahdollista ja miten tiedot ovat siirrettävissä muualle. Bibliografisten aineistojen kuvailussa käytettäisiin FRBR-käsittemallia. Kalliomaalauksia analysoidaan käyttäen hyväksi CIDOC-CRM:ää, joka on malli kulttuurihistoriallisen sisällön kuvailuun. Vastuu Henna Mölsä ja Miia Herrala.
5. *Prosessikuvauksen tekeminen dokumentoinnista.* Vastuu Henna.

Viva3- ja Ephemera –hankkeissa kilpailutettiin asiantuntijapalvelut. Asiantuntijaksi valittiin Mika Nyman Synapse Computing Oy:stä. Nymanin osuus ontologisoinnissa oli toimia ohjaajana. Tarkoituksena oli saada tuntuma erilaisiin mahdollisuuksiin

ilman loputonta harhailua. Mikalta kaivattiin näkemystä ontologisointiprosessin sisältämistä työkuluista ja käytetyistä apuvälineistä.

Nyman (2009b) selvensi omaa näkökulmaansa ja toimintatapaansa tätä projektia laajemmin seuraavasti. Hänen tavoitteenaan on luoda kulttuuriperinnön hajautettuja, kansainvälisiä ja yhteistoiminnallisia infrastruktuureja. Tällainen hajautettu palveluinfrastruktuuri vaatii tuekseen standardeja, mutta toisaalta on pystyttävä yhdistämään standardipohjaisuus mukautuvuuteen eri käyttötilanteissa ja käyttäjien erilaisiin vaatimuksiin. Järjestelmän aito hallinta ja ymmärtäminen ovat välttämättömiä tekniikan lisäksi, että välitettävä tieto säilyttää oikean merkityksensä ja että tieto rikastuu. Aito ymmärrys syntyy ajan kanssa ja menetelmänä voidaan käyttää tekemiseen pohjautuvaa reflektoivaa oppimista. Toiminnalliset prototyypit ovat tästä syystä hyödyllisiä. Käsitelmalleista, ontologioista, metatietostandardeista, sisällöistä ja prosesseista muodostuu prosessuaalinen kokonaisuus. Ontologiat ovat formalisoituja käsitelmalleja. Käsitelmallit kuvaavat sisältöä ja prosesseja (deskriptiivisyys) tai määrittävät jonkin tietojärjestelmän vaatimat tietosisällöt (preskriptiivisyys). Metatietostandardien tulisi perustua taustalla olevaan käsitelmalliin tai ontologiaan. Usein näin ei ole, vaan taustalla on vain implisiittinen tieto ja kokemus jonkin toimialueen luonteesta. Aito ymmärrys vaatii metatietostandardin yhdistämisen kohdeprojektin taustalla oleviin käsitelmalleihin ja ontologioihin. Inhimillisen ymmärryksen lisäksi järjestelmien on ymmärrettävä toisiaan, mikä vaatii jaettuja käsitelmalleja ja ontologioita. Jaetut käsitelmallit vaativat vuorovaikutusta, mistä seuraa yhteisten skeemojen eli tietorakenteiden kehittyminen ja tietojen nivoutuminen toisiinsa. Kansainvälinen yhteistyö tarjoaa pohjaa jaettujen käsitelmallien syntymiselle paikallisesti. Tästä esimerkkinä ovat ylätasen käsitteistöt CIDOC CRM ja RDA, joiden erikoistunut käyttö paikallisesti on mahdollisesta niiden tyyppi-käsitteen vuoksi.

Asiantuntijapalvelun kilpailutusvaiheen tarjouksessa Mika Nyman (2009c) tarkensi ontologisoinnin ohjaamista koskevia lähtökohtia. Yhtenä lähtökohtana on Miia Herralan jo tekemä metatietoformaatteja koskeva metatietoselvitysluonnos, jota projektin kuluessa on päivitetty, sekä XML- ja RDF-muotoiset hahmotelmat. Lähdeaineistona käytetään XML-muotoisia metatietoskeemoja kuten SPECTRUM ja Midas Heritage. Pällekkäisyydet tunnistetaan ja tulevan skeeman tulee ilmentää luokkia, alaluokkia, ominaisuuksia ja suhteita siten että ne ovat selvästi erotettavissa

toisistaan. Tästä alustavasta skeemasta tuotetaan luettelomuotoinen esitys alustavaa palautetta varten. Skeema integroidaan CIDOC CRM:n OWL-muotoiseen esitykseen. Työskentelymenetelmänä on dialogi eri asiantuntijatahojen kanssa. Aineistojen (skeemojen, datan ja työvälineiden) kanssa työskennellessä syntyy oppimista, oivalluksia ja uutta tietoa. Emootiot syventävät tätä oppimista. Oppiminen eli uuden tiedon integroituminen aiempaan tietoon tapahtuu pääsääntöisesti alitajuisesti ajan kuluessa. Tämä antaa pohjaa tiedon edelleen opettamiselle. Evaluaatiovaiheessa CRM-skeema arvioidaan aineistolähtöisesti. Tärkeitä arviointikohteita ovat: mihin tapahtumiin skeeman elementit kytkeytyvät, kuinka hyvin ominaisuudet ovat periytyneet ja että suhteet vastaavat aineistoa. Työn toisena erillisenä tavoitteena on skeeman pohjalta laatia relaatiomalliin perustuva tietokanta, jossa toteutetaan fasettipohjainen navigaatorakenne.

Viva3- ja Ephemera hankkeiden kokouksessa joulukuussa 2009 Mika Nyman esitteli tarkemmin ontologioita ja niiden prosesseja. Ontologioita on eritasoisia – sekä sanastolähtöisiä ontologioita että ylätasoon ontologioita. Näillä molemmilla tavoilla on mahdollista koota eri metatiedot yhteiseen ylätasoon malliin. Tällä hetkellä on olemassa tarve tällaiselle muistiorganisaatioiden yhteiselle core-ontologialle. FinnOnto on päätynyt sanastolähtöisten ontologioiden käyttöön, kun taas CIDOC CRM:n perustuva ontologia edustaisi ylätasoon ontologioita. Tämä työ aloitetaan CIDOCin kokouksessa tammikuussa 2010 ja lähtökohtana on kokemusperäinen näkökulma aineistoihin. Tämä vaatii tietoa muistiorganisaatioiden prosesseista ja niiden perusteella voidaan luoda yhteinen kuvauskieli. KDK:n prosessikuvaustasoa olisi mahdollista käyttää yhtenäisen tarkkuustason saavuttamiseksi. Nyman oli myös keskustellut Ruotsin yhteistyökontaktien kanssa. Svenskt hällristningsforsknings arkiv:issa (SHFA) on tehty työtä, jota mahdollisesti olisi hyvä hyödyntää Astuvansalmen kallioon liittyvään ontologisointiin liittyen. (Metatiedon keruu digitointiprosesseissa ja keskustelua 2010.)

Kuten aiemmin on todettu Staabin ym. (2000) aloitusvaiheeseen kuuluu ontologiavaatimusten määrittelydokumentti. Koska prosessin aloitusvaiheessa ei ollut ehtinyt tulla tietoon Staabin prosessi, tein dokumentin jälkikäteen aloitusvaiheen selkeyttämiseksi (liite 2).

Kokonaisuudessaan aloitus- ja määrittelyvaihe kostui siis hankkeen tavoitteen määrittelystä, esimerkkikohteen valinnasta, prosessin alustavasta hahmottelusta, asiantuntijapalvelujen kilpailutuksesta sekä tavoitteiden ja prosessin tarkentamisesta asiantuntijan avulla (liite 3). Käyttäjätarpeiden selvityksen tekeminen olisi parantanut lopputulosta, joten se olisi ollut hyvä lisä prosessiin.

6.2 Toteutusvaihe

Varsinaisessa toteutusvaiheessa tehtiin alustavaan prosessiin muutoksia tarpeen mukaan. Mm. aineiston kartoitus ei sisältänyt ainoastaan aineistojen kartoitusta vaan myös kuvien ja kuva-aiheiden kartoituksen. Prosessissa merkittävä muutos oli lopulta toiminnallisen prototyyppitietokannan luomisesta luopuminen. Ennen tätä päätöstä kuitenkin mallintaminen ja ontologisten suhteiden miettiminen jalostuivat tarkemmiksi prosesseiksi koostuen käsitellinnuksesta ja metatietoskeemasta ja näihin liittyvästä ontologisten suhteiden miettimisestä. Lähtökohtana ei ollut enää aloittaa Protégella vaan käsitellinnuksen teko pikemminkin ajatustasolla ja aineistolähtöisesti.

6.2.1 Aineiston kartoitus

Lähtökohtana aineiston keruulle oli kerätä erityyppisiä aineistoja eri aloilta eli arkisto-, kirjasto- ja museoaloilta bibliografisia viitteitä, valokuvia, piirroksia, kaivauskertomuksia, www-sivustoja jne. Aineistojen monipuolisuustavoite liittyi esimerkkiaineiston muodossa valtakunnallisella tasolla käynnistyneelle tavoitteelle luoda yhteinen kirjastojen, arkistojen ja museoiden yhteinen ylätasoinen ontologia. Erilaisten aineistojen käyttöön liittyy aina tekijänoikeuskysymykset. Joko aineistojen on oltava tekijänoikeudesta vapaita tai käyttöoikeudet täytyy olla selvitetty. Tässä tapauksessa on hyödynnetty projektin omia tuotoksia sekä muutoin kysytty lupa aineistojen käyttöön.

Koska Astuvansalmen kalliota käsittelevää ”virallista” aineistoa ei ole loputtomasti, päädyttiin lähes kattavaan aineistojen keräämiseen. Neuvottelut Museoviraston kanssa eivät kuitenkaan tuottaneet tulosta eikä näitä aineistoja saatu käyttöön, vaikka useammat ihmiset yrittivät saada tietoja eri kautta. Samalla eri alojen eli arkisto-, kirjasto- ja museoalan aineistojen erilaisuutta ei täysin saatu toteutettua.

Museoviraston kuvat koostuivat pääsääntöisesti Ismo Luukkosen vuonna 2001 Museovirastolle tekemästä valokuvausdokumentoinnista. Koska näitä aineistoja ei saatu käyttöön, Luukkosen itsensä kanssa neuvoteltiin hänen vuonna 1996 ottamien kuvien käyttöoikeuksista.

Käsiteltävänä aineistona oli Viva3-hankkeessa vuonna 2010 otettuja valokuvia, Luukkosen vuonna 1996 ottamia valokuvia, Pekka Kivikkään piirroksia ja hänen kirjojensa valokuvia, lehtiartikkeleita, bibliografista aineistoa sekä www-sivustoja. Kivikkään piirroksia kerättiin hänen kotoaan digitoitavaksi Mikkelin ammattikorkeakoulussa. Pekka Kivikkäältä ja hänen kustantajaltaan saatiin myös lupa kirja-aineistojen käyttöön. Kivikkään kirjoista tärkeimpänä oli uusin teos Suomen kalliomaalausten merkit: kalliot, kuvakentät ja merkitykset vuodelta 2009.

Tekstiaineiston kartoitus aloitettiin käsillä olevista kirjoista, lehtiartikkeleista ja www-sivustoista tutustumalla aineistoihin ja tekemällä hakuja sekä yliopistokirjastojen yhteisluettelo Lindaan, Mikkelin kaupunginkirjaston kokoelmatietokantaan että Mikkelin korkeakoulujen kirjastoverkko Mikkiin. Hakusanoina käytettiin Yleisen suomalaisen asiasanaston (YSA) mukaisia asiasanoja kuten kalliotaide, Astuvansalmi ja Ristiina. Aineistojen viitetiedot tallennettiin exceliin. Museoviraston tietokantojen Muinaisjäännösten hoitorekisterin, Muinaiskalupäiväkirjan ja Suomen museot onlineen tiedot tallennettiin erilliseen exceliin. Tiedot luokiteltiin jo valmiiksi nimi-, kuvaus- ja arvosarakkeisiin. Myöhemmin nämä kaksi exceliä yhdistettiin yhdeksi tiedostoksi Viva3ONTOaineistot_v1.0. (sis. nimikuvausarvo010710 ja Monografia-aineisto_Astuvansalmi (29.3.2010). Elektronisista lähteistä oleellisia olivat viralliset linkitykset viitetietoihin sekä metatietojen tarkkuus.

Aineistojen kartoitus nosti esiin seuraavia kysymyksiä. Ensinnäkin millä perusteella aineistot valitaan ja mihin vedetään raja valittavilla aineistoilla. Riittääkö, että Astuvansalmi on mainittu aineistossa vai tulisiko sitä olla käsitelty laajemmin esim. artikkelin verran. Otetaanko yleiskirjallisuutta Suomen kalliomaalauksista tai esim. kivikaudesta ja Saimaasta. Saimaasta puhuttaessa esim. vedenpinnan vaihtelut vaikuttavat oleellisesti Astuvansalmen kalliomaalauksiin, mutta onko tällainen pelkkää Saimaata koskeva aineisto riittävää. Rajauksena aineistovalinnassa käytettiin, että kirjassa tulisi olla isompi kokonaisuus koskien Astuvansalmen kalliota. Lisäksi usein nousi esille aineiston luotettavuus ja millä perusteella se määritellään.

Vaadimme ko käsittelemältämme aineistolta ”tiedepiirin” hyväksyntää vai ei. Aineistoa pyrittiin ottamaan mukaan kuitenkin edustavasti kaikilta tasoilta, koska laadukasta ja muutoin arvokasta työtä on tehty, vaikkei sitä olisikaan tieteellisesti tutkittu. Kalliomaalauksia koskevat tulkinnat erityisesti ovat niin moninaisia kuin ovat tulkitsijatkin.

Alkuun tiedostoon tallennettiin ainoastaan kirjojen ja lehtiartikkelien viitteet. Mitä tietoja sitten näistä aineistoista tulisi koota tiedostoon? Tähän ei lopulta saatu kunnan vastausta vaan aineistoista otettiin sen verran tietoja, että ne olivat tunnistettavissa ja löydettävissä myöhemmin. Kirjoista otettiin tiedot ylös kuitenkin artikkelitasolla niiden artikkelien osalta, jotka koskivat Astuvansalmen kalliota. Mietinnässä oli myös miten suhtaudutaan aineistoissa käsiteltyihin henkilöihin ja tulisiko heidän tietonsa näkyä tiedostossa. Osaltaan tässä vaiheessa ei vielä tiedetty miten aineistot tullaan tallentamaan ja missä tallennusmuodossa, joten lopullinen ratkaisu tallennettavista tiedoista jäi odottamaan myöhempää ajankohtaa.

Tässä vaiheessa en itse vielä ymmärtänyt miten tärkeää on tutustua ja perehtyä Astuvansalmea koskevaan aineistoon melko tarkasti. Vaikka aineistolähtöisyydestä oli puhuttu ja se oli tiedossa, niin sen todellinen merkitys selvisi tarkemmin vasta myöhemmin.

Aineiston kartoitus koostui siis tiedonlähteiden alustavasta selvityksestä, tallennusformaatin ja -tavan valinnasta, kartoitettavan aineiston rajauksesta, tiedonhankinnasta, lähdeviitteiden tallennuksesta, käyttöoikeuksien selvityksestä ja ongelmakohtien kirjaamisesta (liite 3).

6.2.2 Kuvien ja kuva-aiheiden kartoitus

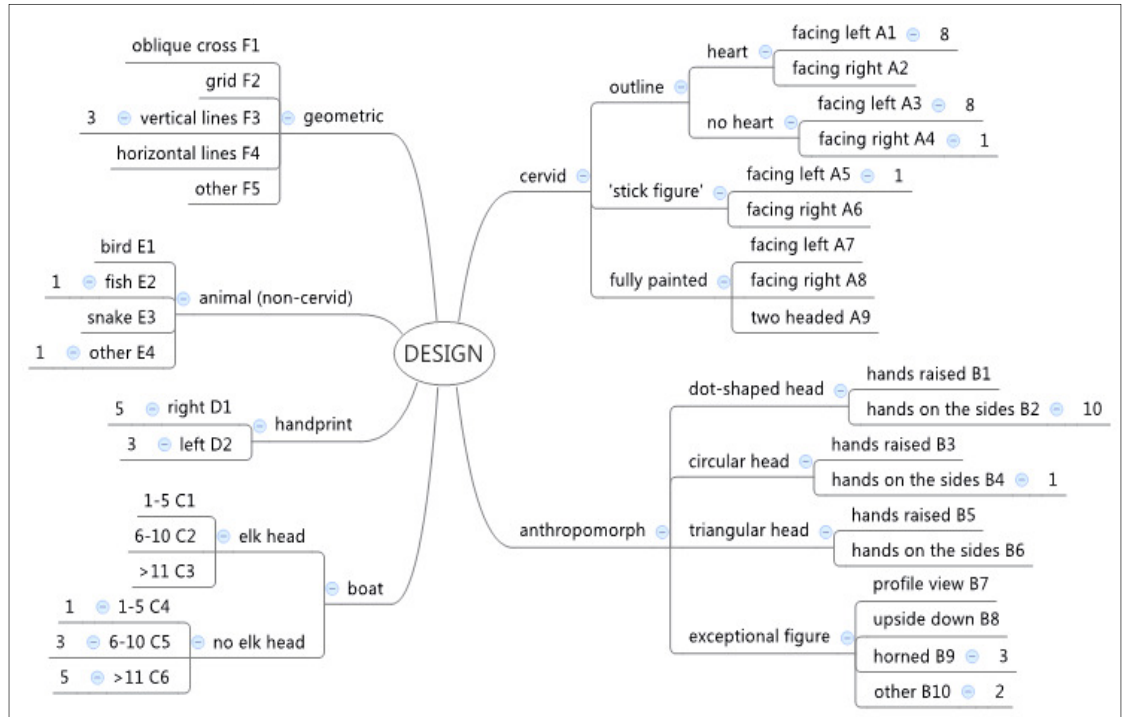
Alkuun prosessiin omalta osaltani kuului aineiston kartoitus. Tavallaan yllätyksenä tuli tieto, että kuvat ja kuva-aiheet itsessään vaativat kartoituksen ja että teen tämän työn. Toki alkuperäisen prosessin sisältämät tehtävät olivat laajoja kokonaisuuksia ja suuntaa antavia, joten kaikkea ei ollut mahdollisuus alusta asti tietää tarkasti.

Oma osuutensa aineiston keruuseen liittyen oli saavuttaa ymmärrys Astuvansalmen kallon kokonaisuudesta, sen kuvaryhmistä, kuvista ja kuva-aiheista. Tiedot

tallennettiin Viva3ONTOKuvaAiheet v3.1 –exceeliin. Tiedosto koostuu kuva-alueen tunnuksesta, sen mahdollisesta nimestä, kohteista kuva-alueilla määriteltynä sekä Pekka Kivikkään että Antti Lahelman (<Pekka Sarvas) mukaan, Lahelman kuva-aihejaotteluskemasta, kuva-aiheen eheys- ja tunnistettavuustiedoista, kuva-alueiden ja kuva-aiheiden mahdollisesta sijainnista korkeussuunnassa sekä kuva-alueen että kuva-aiheen korkeudesta.

Vaikka voi kuulostaa helpolta tutkia kirjallisuutta ja kuvia ja kerätä sen pohjalta kallion sisältämät kuvat, tämä kartoitus on haastavaa. Kalliosta on toki löydettävissä joitain selkeitä kuvia, jotka lähes kaikki pystyvät tunnistamaan, kuten esim. jousikäisen naisen. Näiden selkeiden kuvien määrä on vain melko vähäinen. Epäselvät kuvat voivat olla eritavoin epäselviä kuten päällekkäisiä, katkelmia, muuteltuja, jäänteitä, tuhoutuneita tai huonosti havaittavissa olevia. Kalliosta havaittavat asiat riippuvat pitkälti vuodenajasta ja muista luonnonolosuhteista kuten kasvillisuudesta, jäkälästä ja ilmansaasteista, jotka vaihtelevat ja muuttuvat vuosien myötä. Sadevesi puolestaan on irrotanut kalliosta piioksidia, joka sekä suojaa että peittää kalliomaalauksia. Nämä olosuhteet vaikuttavat erityisesti valokuvien ottamiseen ja siten myös tutkimiseen. Kuvaaminenkin voi olla haastavaa, koska useimmiten kalliomaalaukset sijaitsevat vesistön äärellä, joten tarvitaan kestävä jääpeite kuvausten suorittamista varten. Nykyaikaisella kuvankäsittelyllä on kuitenkin puolensa, koska kuvakentistä on saatu uutta tietoa.

Oman osansa vaikeusasteeseen tuottaa, että eri ihmiset ovat tulkinneet erityisesti epäselvät kuva-aiheet eri tavoin ja mitä ne esittävät. Hyvä esimerkki tulkinnallisuudesta ovat hirven sarvet ja veneet. Joidenkin mielestä veneet ovat osa hirven sarvia, kun toisten mielestä ne ovat pelkästään veneitä. Toisaalta samankin ihmisen tulkinnat voivat muuttua ajan kanssa ja uusien menetelmien tuoma lisäinformaatio voi muuttaa tulkinnan toiseksi. Esim. Kivikäs (2010) on ennen tulkinnut yhden kuvan lahnaksi ja nykyään hän ajattelee sen olevan osittainen karhu. Lukumäärien laskemisessa oli myös epäselvyyttä. Onko mm. seitsemän vierekkäistä pystyviivaa yksi kuva vai seitsemän kuvaa. Keskeisenä kysymyksenä ontologisointia tehdessä olikin kuinka nämä kuva-aiheiden eri tulkinnat on mahdollista yhteen sovittaa vai onko tämä edes tarpeen. Kuva-aiheita exceeliin kerättyä päätettiin käyttää sekä Pekka Kivikkään uusimman kirjan (2009) tietoja että Antti Lahelman väitöskirjan (2008) sisältämiä tietoja.

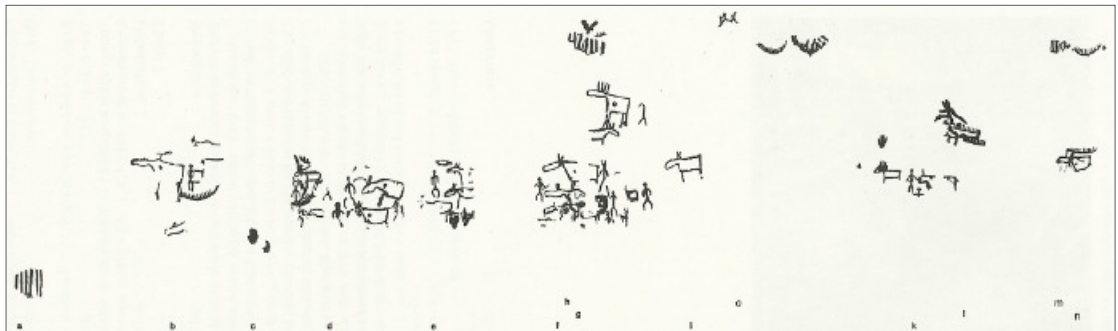


KUVA 16. Lahelman skeema sisältäen kalliomaalausten lukumäärät (Lahelma 2008, 194)

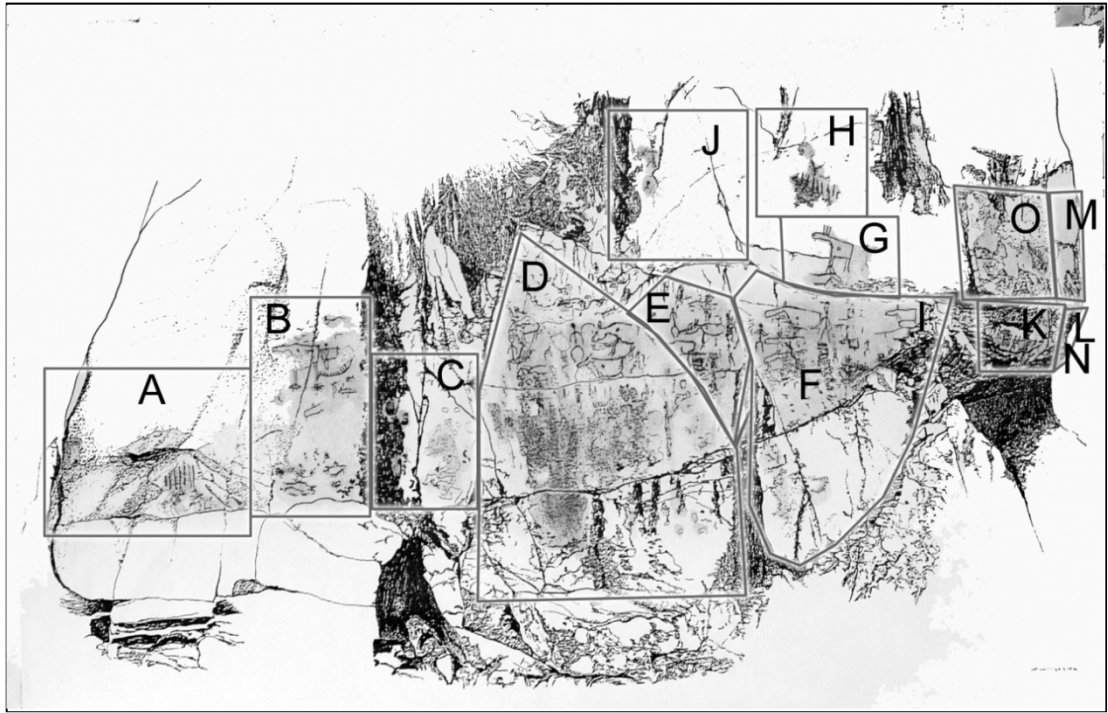
Lahelma (Lahelma 2008, 194) on väitöskirjassaan esitellyt kuva-aiheiden mukaisesti järjestetyn skeeman (kuva 16) Suomen kalliomaalauksista. Hän on ottanut huomioon ainoastaan selkeät kuva-aiheet. Itse olen lisännyt Lahelman mukaiset selkeät kuva-aihemäärät kuvan skeemaan Astuvansalmen kallion osalta. Tarkoituksena oli käyttää mahdollisimman paljon hyödyksi tällaisia valmiiksi tehtyjä asioita kuten tässä jaottelua eri luokkiin. Ongelmallisen skeemasta teki, ettei kuvia oltu yksilöity väitöskirjassa kuva-alueelta erikseen eli ainoastaan lukumäärät eri kuvista oli eroteltu Astuvansalmen kallion osalta. Lahelma jaottelee skeemassaan kuvat hirvieläimiin, ihmismäisiin hahmoihin, veneisiin (venekuviin), kämmenpainamiin, ei hirvieläimiin ja geometrisiin kuvioihin. Skeema jaottelee kuvat hyvinkin tarkkoihin ja pitkälti tulkittamista vaativiin luokkiin, mikä tekee luokittelun käytön vaikeaksi. Esim. veneet on jaoteltu sekä hirven päähän että ei hirven päähän kuuluviin. Tämän tulkitseminen voi olla hyvin haastavaa. Luokittelu menee lisäksi tätä syvemmälle, koska ryhmittely vaatii veneiden sisältämien ihmisten/sarven piikkien lukumäärän laskemista. Osittain tätä luokittelua ei ollut mahdollista tehdä näin tarkasti. Ajateltiin, että riittävää on, että kuva tunnistetaan veneeksi, mikäli sen ihmisten/sarven piikkien lukumäärää ei voi varmasti tunnistaa. Samantyyppisiä ongelmia oli erityisesti hirvieläinten sydänpisteiden tunnistamisessa eli ovatko hirvieläimet sydänpisteellisiä vai eivät.

Kirjoissa ei myöskään välttämättä ollut mainintaa kämmenpainamien oikea tai vasenkätisyydestä. Kysyin Pekka Kivikkäältä (2010) näihin ongelmiin liittyen pystyykö hän kertomaan tarkasti lukumääriä, suuntia tai sydänpisteellisyyksiä. Hän ei ole järjestelmällisesti tehnyt tällaista työtä, joten vahvistus tarkkoihin luokituksiin jäi saamatta.

Lopullinen kuva-aiheiden tunnistaminen oli riippuvainen omista tulkinnoista, vaikkakin Lahelma (2008) kirjassaan käyttää Pekka Sarvaksen (1969) tekemää aiempaa tutkimusta hyväkseen. Koska Sarvaksen tutkimuksesta on aikaa, nykytekniikalla on saatu kalliomaalauksista huomattavasti enemmän tietoa tuon aikaiseen tutkimukseen verrattuna. Tästä johtuen on vaikea sanoa mitä asioita aiemmin on pidetty selkeinä. Yhtenä tärkeänä apuvälineenä olen käyttänyt Sarvaksen tekemää piirrosta Astuvansalmen kallion kuvakentästä (kuva 17). Sarvas on nimennyt kuvaryhmät aakkosin a-o. Ryhmittely eri alueille helpotti huomattavasti selkeiden kuvien tunnistamista. Yksilöivien tunnisteiden tarve kävi ilmeiseksi tässä vaiheessa. Toisena merkittävänä tulkintojen apuvälineenä on ollut Kivikkään (2009) piirustus (kuva 18). Piirustuksessa on huomioitu kallion muodot ja selkeämmin kuva-alueiden rajaukset.



KUVA 17. Astuvansalmen kalliomaalaukset Pekka Sarvaksen (1969) mukaan. Kuvaryhmä o on lisätty alkuperäiseen piirroksen (Lahelma 2008, 258)



KUVA 18. Astuvansalmen kalliomaalausalueet Pekka Kivikkään (2009) mukaan

Miia Herrala lisäsi tiedostoon kuva-aiheen eheys- ja tunnistettavuustiedot, kuva-alueiden ja kuva-aiheiden mahdollisen sijainnin korkeussuunnassa sekä kuva-alueen että kuva-aiheen mahdollisen korkeuden. Koordinaatti- ja korkeustietojen ajateltiin hyödyntävän Esa Hannuksen työtä sijoitellessaan valokuvia 3D-mallin pinnalle. Lisäksi ne ovat tietysti metatietoja koskien kuvia.

Kuva-aiheita koskevia tiedostoja päivitettiin tarpeen mukaan ja projektin kuluessa. Esim. alun perin kuva-alueet oli eritelty aakkosin, mutta tätä tarkennettiin, koska joillain kuva-alueilla jaottelua oli tehty tarkemmin.

Kuvien ja kuva-aiheiden kartoitus alkoi kohdeaineistoon tutustumisella. Prosessissa seuraavana olivat tallennusformaatin ja –tavan valinta, kuvaryhmien, kuvien ja kuva-aiheiden erittely ja tallennus sekä eheyden ja tunnistettavuustietojen tallennus (liite 3). Lopuksi tallennettiin koordinaattitiedot ja kirjattiin ongelmakohtat.

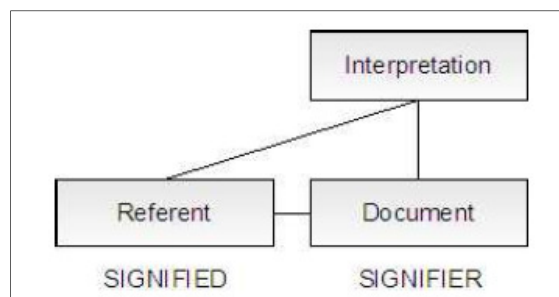
6.2.3 Kuvailu ja käsittemallinnus

Kulttuuriperinnön käsittemallinnuksen teon taustalla ovat metatieto, sen merkitys ja siitä tehtävät tulkinat. Käsitellen tässä kappaleessa ensin teoreettisesti metatietoa ja

sen yhteyttä tulkintoihin. Lisäksi käsittelen tehtäväkuvaukseen liittyen kartoitettujen aineistojen kuvailun kautta tehtävää käsittemallinnusta.

Metatiedon merkitys voidaan jaotella formaaliin, viittaavaan ja kognitiiviseen merkitykseen. Merkitykset ovat aina kulttuurisidonnaisia ja riippuvaisia kustakin yksilöstä ja asiayhteydestä, eikä näitä erilaisia merkityksen tasoja välttämättä voi erottaa toisistaan, vaan ne ovat toisiinsa sidoksissa ja toisiaan täydentäviä tasoja. (Hannus ym. 2010.)

Formaalit merkitykset muotoutuvat tiettyjen määritelmien tai sääntöjen mukaisesti. Näistä säännöistä muodostuu käsitteen merkitys. Joitain tällaisia merkityksiä pystytään tulkitsemaan koneella, mutta osa säännöistä ei ole yksiselitteisiä ja siten on mahdollisuus useammille tulkinnoille. Viittauksellinen merkitys viittaa tunnistettavaan objektiin. Paikannimiä voi olla samoja, mutta tarvitaan esim. koordinaattijärjestelmä tai maan nimi viittaamaan haluttuun paikannimeen. Kuvassa 19 on esitetty Astuvansalmen tapauksessa ajateltu tulkintojen suhde aineistoihin ja niiden tarkoitteeseen eli viittaajaan. (Hannus ym. 2010.)



KUVA 19. Tarkoitteen, tulkintojen ja aineiston suhde toisiinsa (Hannus ym. 2010)

Kognitiiviset merkitykset kuten tuntemukset, tunteet, asenteet ja ajatukset ovat ihmisen tulkintoja asiasta, esim. auton valot eivät tarkoita autoa vaan pelastusta kiperästä tilanteesta. Näiden ilmaiseminen konekielisesti voi olla vaikeaa ja yksi vaihtoehto on jättää tällaiset tapaukset käsittelemättä. Joissain tapauksissa näiden asioiden ilmaiseminen on välttämätöntä, kuten taiteessa, jonka pääasiallinen tehtävä on herättää ajattelemaan laajemmin tai välittää herkkiä tai voimakkaita tuntemuksia. Kalliokuvien eri merkitykset voivat olla eri ihmisille aivan erilaisia ja herättää erilaisia tunteita. Erityisesti kognitiiviset merkitykset ovat yksilösidonnaisia. Tietojenkäsittelyn

maailmassa näitä kognitiivisia merkityksiä voidaan saada esiin mahdollistamalla yksilöiden tuntemusten ilmaiseminen esim. tagien avulla tai mahdollistamalla linkitysten teko ulkopuolisiin lähteisiin. Kalliokuvien osalta tällaista aineistoa löytyy mm. kirjoista, artikkeleista ja asiakirjoista. Vaikka ihmisten tulkinnat muuttuisivat ajan myötä, samalla tulee kerättyä kalliokuvan proveniensiin eli objektin historiaan liittyvää dokumentointia. (Hannus ym. 2010.)

Käsitellinnuksen ideana oli ajatella, luokitella, tyypitellä ja kuvailla aiemmin kerättyä aineistoa eri näkökulmista ja huomioida moniulotteisuus mallin luomisessa. Huomattavaa on, että aineiston kuvailu on aina tulkintojen tekemistä. Lähtökohtana kalliomaalauskuviin ja -kuva-aiheiden perusteella tehtävään käsitelmalliin oli Mika Nymanin (2010a) hahmottelema kalliomaalauksen kuvailumalli. Kuvailumalli syntyi projektiryhmän kokouksessa, jossa keskusteltiin yhteistyössä Mikan ehdotuksesta. Viva3-hankkeen tarkoituksena oli tarkastella ontologisointiprosessia aineistolähtöisesti, johon tarpeeseen kuvailumalli antoi vastauksen. Kuvailumalli toimi kuvailun apuvälineenä. Aineistoa eli Astuvansalmen kalliota koskevaa dokumentointia tarkasteltiin kuvailumallin näkökulmasta. (Nyman 2010b.)

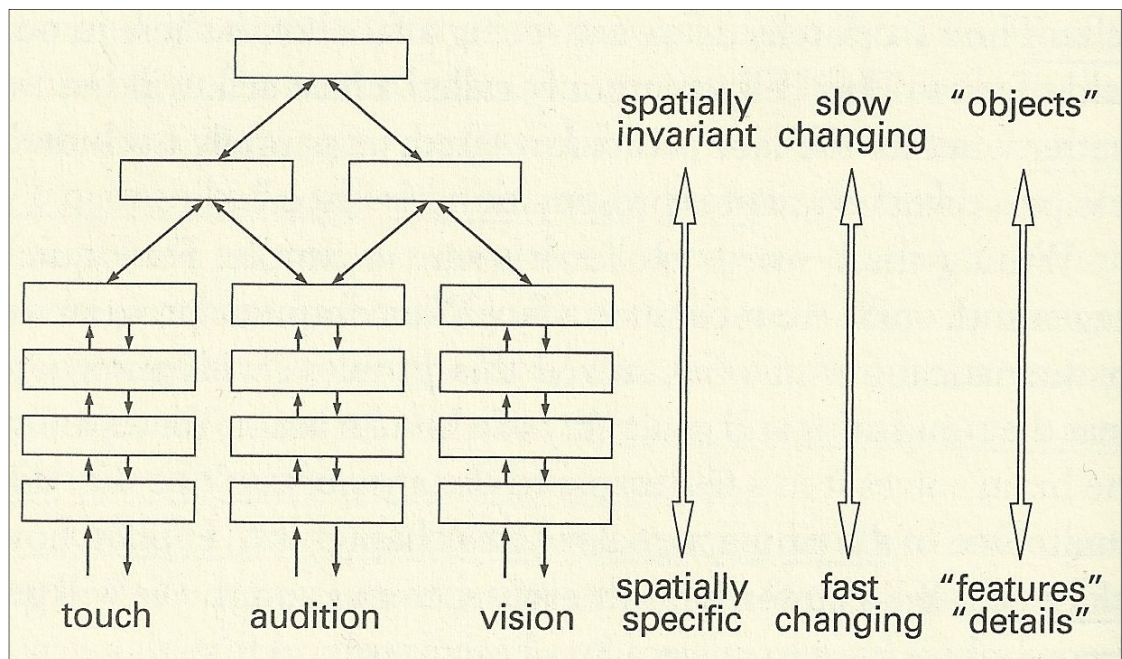
Mika Nymanin (2010c) kuvailumalli:

- ”piirteet (eivät esiinny itsenäisinä)
- fyysiset objektit ja niiden osat
- fyysiset kontekstit (laajemmat kokonaisuudet, struktuurit)
- abstraktit käsitteet ja merkitykset
- intentiot (päämäärät, tavoitteet, tarkoitukset)
- toiminta ja tapahtumat

Lisäksi pitäisi huomioida kieleen liittyvät tunnisteet ja termit.”

Kuvailumallin viitekehystenä Nyman on käyttänyt perception - action cyclea, josta mm. neurobiologi Joaquín M. Fuster (2003, 106-109) on kirjoittanut. Kuvailtaessa tulee siis huomioida havaitsemiseen liittyvät asiat, koska havaitseminen vaikuttaa ihmisen tulkintaan ja siten kuvailuun. Eri aikakausina kyseinen ajanjakso ja siihen liittyvät kulttuuriset tapahtumat vaikuttavat kunkin yksilön tulkintoihin. Se, kuinka ihminen reagoi havaintoon, riippuu Fusterin mukaan ihmisen kullakin hetkellä olevasta sisäisestä tilasta eli mm. tavoitteista ja motiiveista. Toisekseen toimintaan vaikuttavat käyttäytymiseen liittyvät, tiedolliset ja emotionaaliset yhteydet. Tämä

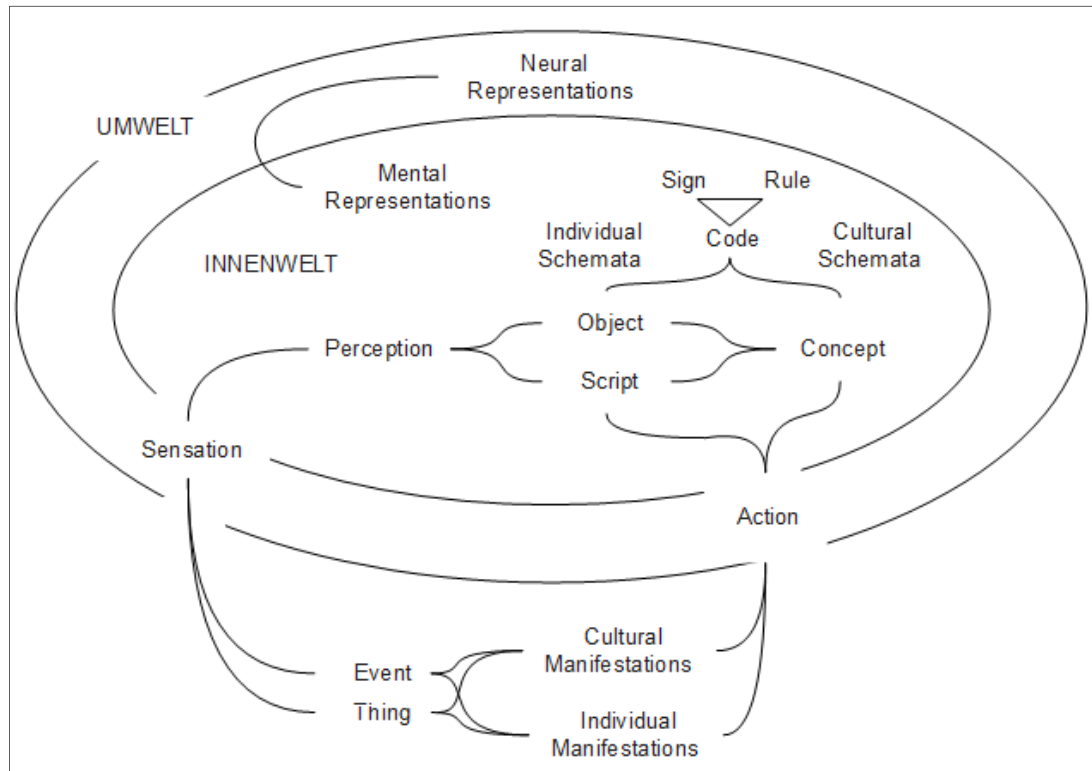
aiheuttaa sen, että ihminen ennakoii tilanteita ja asioita sekä sisäisten että ulkoisten käsitystensä sekä tiedollisten ja tiedostamattomien asioiden perusteella. Havainto sisältää vuoropuhelua entisen muistissa olevan tiedon kanssa, jonka perusteella havainto tehdään. Joissain tapauksissa tämä aiheuttaa ennakkoasenteita, joiden valossa tulkitaan ehkä vääristyneestikin asioita. Toisaalta sama asia voi aiheuttaa ihmisissä hyvinkin erilaisia tunnereaktioita tai tulkintoja. Oikeat ennakko-oletukset johtavat ymmärrykseen muutoin ennakko-oletuksista seuraa hämmennystä ja toisaalta kyseinen asia herättää huomiota. Hawkinsin (2004, 86 – 89) memory-prediction malli edustaa tätä ennakkointia. Huomattavaa on, että ihmisen havainnot eivät tule pelkästään aistien kautta, vaan pääasiassa ihmisen sisäisestä muistimallista. Kuvasta 20 käy ilmi kuinka tietoa kulkee molempiin suuntiin eli ennakkointia tapahtuu. Ihminen muodostaa piirteistä ja yksityiskohdista objekteja aistiensa avulla eli tekee tulkinnan objektista sen piirteiden perusteella. Tämä on suoraan sovellettavissa CIDOC CRM –malliin ja kuvailumalliin.



KUVA 20. Kuulon, näön ja kosketuksen muuttumattomien kuvausten muodostuminen memory-prediction mallin mukaan (Hawkins 2004, 118)

Kuva 21 selventää kuvailumallin teoreettista taustaa – asioita, jotka luovat ihmisen tietoisuuden ja käsitteitä, jotka ilmentävät tietoteknistä todellisuutta. Pääideana on, että ihmisen todellisuus on subjektiivista ja liittyy havaintoihin. Havainnot puolestaan riippuvat ihmisen aiemmista kokemuksista ja kulttuurisesta ympäristöstä, jossa hän elää. Havainnon merkitys voi olla ihmiselle kaksitahoinen riippuen siitä tulkitaanko

havainto yksilöllisestä vai kulttuurisesta näkökulmasta. Havaintojen kautta signaalit tunnistetaan joksikin jo ennestään mielessä olevaksi kohteeksi eli identifioidaan aistimukset. Jokaiseen kohteeseen liittyy toimintatapoja ja tunnistettavia käsitteitä. Manifestaatiot ilmentävät ihmisen vallitsevia kulttuurisia skeemoja tai yksiköllisiä skeemoja tuottaen käyttäytymistä ja asioita. (Nyman 2010c; 2011a)



KUVA 21. Luomisen ja tuottamisen kierto (Nyman 2011b)

Kuvailumallissa yhdistyvät Nymanin mukaan myös 1900-luvun tieteenalat kuten strukturalismi, funktionalismi, systeemitheoria sekä semiotiikka. Myöhemmässä kokouksessa Nyman toi esille myös fraktaalisuuden huomioimisen käsitelmän kokoamisessa. (Nyman 2010c.)

Nymanin kuvailumallin pohjalta teimme Miia Herralan kanssa yhteistyössä esimerkivalokuvan kuvailun kuvailumallin mukaisesti. Liitteessä 4 on otos tästä kuvailusta. Esimerkkikuvailuun pohjautuen lähdin rakentamaan exceliin käsitelmää, joka käsittää kallion ja kalliomaalaukset sekä näiden eri tulkinnat. Tästä käsitelmästä on otos liitteessä 4, josta käy ilmi kallion piirteet ja fyysinen konteksti. Tarkoituksena oli tarkastella erityisesti aineistojen sisältöjä. Tässä vaiheessa keskityttiin aineistojen sisältämään informaatioon. Tiedon käsitteen mukaisesti informaatio pitää sisällään aineiston sisällön eli viestittävän sisällön, kun taas tietämys on omaksuttua sisältöä.

Tässä vaiheessa ei myöskään keskitytty itse aineistoihin ja niiden muodollisiin ja sisällöllisiin kuvailuihin. (Vakkari 2000, 19.)

Heti alkuun törmäsimme kuvailun kohteen ongelmaan. Kuvailemme valokuvassa olevaan maalausta vai myös kuvassa näkyvää kalliota eli kuvan kontekstia. Päädyimme ottamaan myös kallion huomioon, koska se palvelee joka tapauksessa aiheen käsittelyä. Toisaalta heräsi kysymyksiä koskien kuvailumallia. Mitä eroa on esim. piirteillä ja fyysisillä objekteilla. Onko kallion ihmiskasvoisuus piirre vai osa fyysistä objektia. Toisaalta kuvailumallin tarkoitus ei ollutkaan saada asioita suoraan sen jaotteluun vaan pääasiassa herättää kysymyksiä ja varmistaa, että eri näkökulmat asioista tulevat esille. Mietimme myös tiettyjen asioiden tärkeyttä. Onko tarvetta miettiä ja selvittää mm. kalliomaalausten viivojen leveyttä. Onko tätä ylipäättään mahdollista saada selvitettyä. Halusimme mallista jotenkin ilmentyvän sellaisten asioiden jotka puuttuvat kuvakentästä kuten Astuvansalmen kohdalla palvonta-asentoisten hahmojen (adorantit) puute.

Käsittemallinnusta tehdessä erityisesti erilaiset tulkinnat tuli ottaa huomioon. Tulkintoja oli erilaisia erilaisista lähteistä. Kysymyksenä oli kehen uskotaan. Lopputuloksena oli kuitenkin, ettemme välitä lähteen laadusta vaan tulkintaan sidotaan aina lähde, jolloin vastuu on lukijalla ja hänen arvostuksellaan kyseisestä lähteestä.

Luokittelin käsitemallia lisäksi pakollisiin ja suositeltaviin/vaihtoehtoisiiin asioihin, koska ajattelin sen hyödyntävän sekä omaa asioiden käsittelyäni että tulevaa tietokantamallinnusta.

Käsitemalliehdotusta käsiteltiin kokouksessa 2.8.2010, jonka lopputuloksena sovittiin, että tietokannan käyttöliittymä otetaan käyttöön ennen käsitemallin jatkoa. Käsitemallista tuli kuitenkin vielä etsiä ensisijaisesti fyysisiä objekteja eli objekteja, joihin piirteet tai fyysiset kokonaisuudet kuuluvat. Painopiste tulisi olla osaan ja kokonaisuuteen. Tätä työtä voitaisiin hyödyntää tietokannan mallinnuksessa. Käsitemallinnuksessa tulisi myös keskittyä tunteisiin ja asenteisiin/arvoihin. Lopulta käsitemallista ei tehty ns. viimeistä versiota, vaan metatietoskeemaan lähdetiin tekemään ehdotuksen pohjalta, koska tietokannan käyttöliittymää ei ollutkaan käytettävissä.

Käsittemallia tehdessä olisi hyvä ollut paremmin dokumentoida mm. eri tulkintojen lähteet ja sanatarkasti kirjata lähteissä olleet tekstit. Alkuvaiheessa ei vain ollut tarpeeksi hyvää käsitystä jatkokäytöstä ja miten tarkasti asiat tulisi olla dokumentoitu vaan annettiin vapaat kädet käsittemallinnuksen tekemiseen.

Käsittemallinnuksen osalta prosessi koostui siis kuvailumallin luomisesta, kuvailun kohteen valinnasta, itse kuvailusta, kuvailutietojen tallennuksesta, aiheiden priorisoinnista sekä objektien kartoituksesta (liite 3).

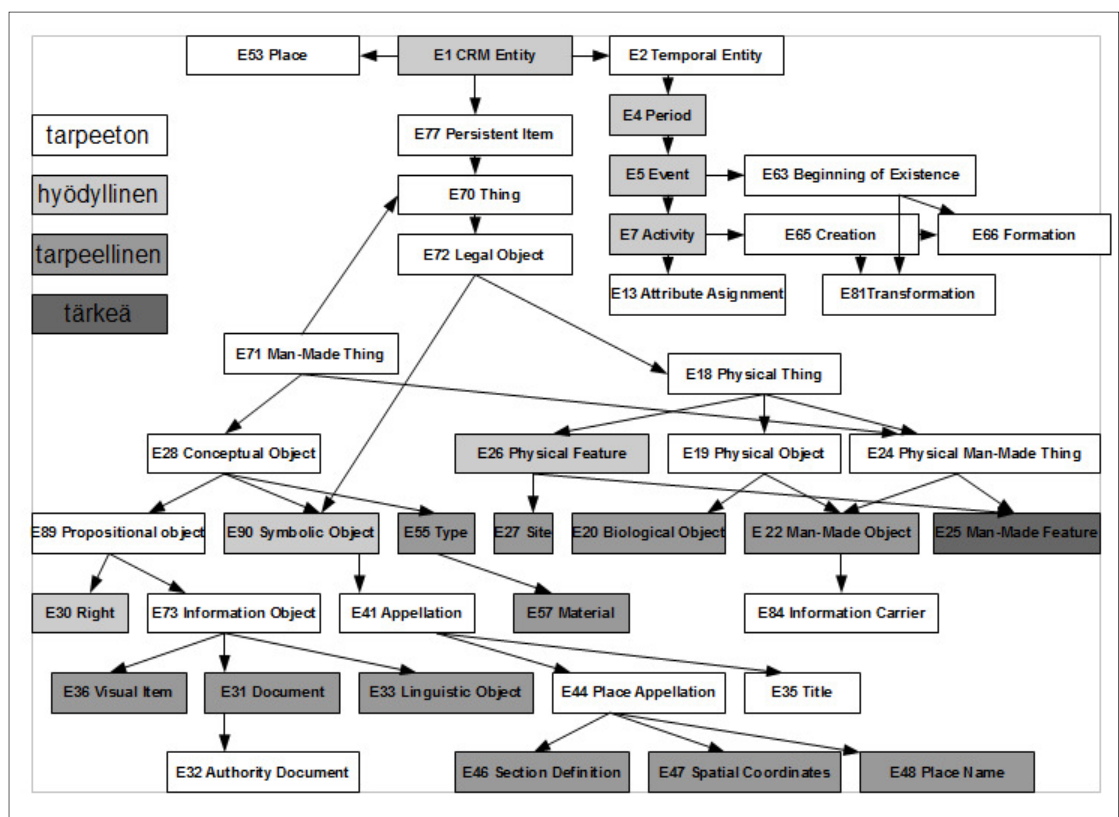
6.2.4 Ontologisten suhteiden miettiminen: CIDOC CRM

Kulttuuriperinnön dokumentoinnin ongelmana on erilaisten standardien ja suositusten määrä. Osittain nämä ohjeet ovat päällekkäisiä, mutta osa on tarkoitettu hieman eri tarkoituksiin. Astuvansalmen kallion ontologisoinnin osalta metadataskeeman ilmaisemistavaksi valikoitui formaali ontologia CIDOC CRM. Jo aiemmin on käsitelty käsittemallia (conceptual model), jonka pohjalta luodaan metatietoskeema (metadata schema). Metatietoskeeman muuttamiseksi ontologiaksi voidaan käyttää valmista ontologiaa eli tässä tapauksessa CIDOC CRM:ää. CIDOC CRM:n avulla on mahdollista parantaa metatiedon laatua, koska se on itsessään käsitemalli, joka määrittelee tarkasti käsiteltävien asioiden kontekstin ja niiden merkityksen. (Hannus ym. 2010.)

Koska itselläni ei ollut aikaisempaa kokemusta CIDOC CRM:stä, lähdin tutustumaan CIDOC CRM:ään sen luokkahierarkian pohjalta. Tein alustavan mindmap-tyylisen hahmotelman keskustelun pohjaksi. Mahdollisia käytettäviä luokkia CIDOC CRM:ssä on 90. Kaikkia näitä ei missään tapauksessa ollut tarve käyttää ja malli tuli pitää yksinkertaisena. Kalliotaide on esimerkinomaisesti mainittu kahdessa CIDOC CRM luokassa: E24 Physical Man-Made Thing ja E25 Man-Made Feature. E25 on E24-luokan alaluokka eli alaluokka E25 on varsinainen kalliotaiteen sijoitusluokka. Luokka sisältää fyysiset piirteet, jotka ihminen on tietoisesti luonut. (Definition of the CIDOC Conceptual Reference Model 2010.)

Keskustelun pohjalta mindmappia muokattiin hieman ja todettiin, että parempi tapa kuvata on esim. piirrosohjelma, joka ei pakota luomaan hierarkisia suhteita.

Keskustelun pohjalta muokattiin Astuvansalmen kallion ontologisoinnissa mahdollisesti hyödylliset CIDOC CRM luokat. Näitä luokiteltiin tärkeisiin ja vähemmän tärkeisiin. Kuvassa 22 näkyvät eri luokkien väliset periytyvyudet eli kaikkia välivaiheita ei ollut tarkoitus käyttää vaan hahmottaa mitä kautta tiedot periytyvät. Muitakin hahmotelmia luokista ja ominaisuuksista tehtiin. Tämä alustava tutustuminen CIDOC CRM:än mahdollisuuksiin tapahtui prosessuaalisesti jo ennen aikojaan, mutta toisaalta siitä oli hyötyä myöhemmässä vaiheessa. Alkuperäisenä tavoitteenahan oli aineistolähtöisyys, josta harhauduimme välillä CIDOC CRM:n puolelle eli ns. toteutuspuolelle.



KUVA 22. Astuvansalmen kallion ontologisoinnissa mahdollisesti hyödylliset CIDOC CRM luokat

Aiemmin on esitelty havaitsemista ja kuinka havainto tapahtuu piirteiden kautta. Piirteiden perusteella luodaan objekteja ja objekteista tehdään yleistyksiä. CIDOC CRM:stä löytyy itsestään piirteen käsite, E25 Man-Made Feature ja E26 Physical Feature. CIDOC CRM ymmärtää piirteen käsitteen rajallisena, kun taas havaintopsykologisesti piirteeksi riittää vähäisempikin. Havaintopsykologiassa piirteenä voidaan ajatella mm. ääriviivat tai liikkeen suunta. Yhteys näiden käsitteiden välillä on kuitenkin olemassa ja mahdollisuus tulkita piirteen käsite

moniulotteisemmin. Tulevassa ratkaisussa olisi hyvä saada esille havaintokohteen esim. eri ajalliset kerrostumat. (Nyman 2011c.)

Tässä vaiheessa päädyttiin ratkaisuun, joka eroaa hieman CIDOC CRM:n perinteisestä terminologisesta käytöstä. Dokumentti sana koski aineistoa ja ymmärrettiin laajasti ja sillä tarkoitettiin kaikkea kommunikointiin liittyviä asioita ei vain kommunikointitarkoitukseen alunperin luotuja eli sekä museoesineet, kirjalliset aineistot että kalliopiirroksiset ja –uurokset ymmärrettiin aineistoksi. Ratkaisu perustui semioottiseen teoriaan. (Hannus ym. 2010.)

Perustelut aineisto -määritelmän käytölle ovat viisitahoiset. Ensinnäkin haluttiin liittää kuhunkin fyysiseen aineistoon yksilöivä tunniste ja tutkia kunkin ilmentymän sisäisiä suhteita. Toisekseen oli tarve keskittyä objektin fyysisiin piirteisiin, koska objektit saattoivat olla eri tavoin epäselviä. Kolmanneksi tärkeää oli ilmentää aineiston fyysistä kontekstia. Neljänneksi oli tarpeen selkeästi erottaa eri tulkinnat, jotka saattavat liittyä joko objektin fyysisiin ominaisuuksiin (jotka ovat tiedossa) tai objektiin itseensä (jotka voivat olla tiedossa tai ovat epävarmoja). Viidenneksi pyrittiin selventämään kuvan 19 tarkoitteen, tulkintojen ja aineiston suhdetta toisiinsa. (Hannus ym. 2010.)

Aineistot ajateltiin siis fyysisinä kappaleina (E18), joilla on fyysinen muoto, jolla on (P128 carries) älyllistä sisältöä (E73 Information Object). Aineisto sanaa ei siis tullut sekoittaa perinteisesti CIDOC CRM:ssä käytettyihin E84 Information Carrier eikä E31 Document –luokkiin, koska niiden määritelmät eivät ole näin laajoja. E84 ymmärretään perinteisesti CIDOC CRM:n mukaisesti ainoastaan tiedon välittäjäksi ja E31 ilmaisee aineetonta kappaletta. Käytännössä tämä tarkoitti CIDOC CRM:n osalta, että kalliotaiteen luokkana ei käytetty suoraan E25 Man-Made Feature vaan yläluokkaa E18 Physical Thing, joka ei ole riippuvainen onko kalliotaide tehty tietoisesti vai ei. (Hannus ym. 2010.)

CIDOC CRM työskentelyyn haluttiin tuoda esille lisäksi fraktaalit ja fraktaalisuus kokoavana kehyksenä. Olio-ohjelmoinnissa fraktaalimenetelmä esiintyy kompositionalisuuden nimellä. Fraktaalisuudella tarkoitetaan kokoelmaa luokkia, jotka pilkkoutuvat vielä pienemmiksi alaluokiksi. CIDOC CRM:stä tämä eroaa ottamalla huomioon periyttämällä myös komponentteja ylä- ja alaluokkien lisäksi eli

luokan mukana tulevat myös sen komponentit. CIDOC CRM ominaisuudet oli mahdollista ajatella luokkina, jolloin eri malleja oli ideatasolla helppo ilmaista. (Nyman 2010d.)

6.2.5 Tunnisteet

Hankkeen yhtenä osa-alueena oli luoda Astuvansalmen kalliosta 3D-malli, mikä oli Esa Hannuksen vastuulla. Tämä osa-alue ei kuulu tämän opinnäytetyön rajauksen sisälle, mutta näille osaprojekteille yhteisiä ovat käytetyt tunnisteet. Mika Nymanin ja Esa Hannuksen yhteistyön tarkoituksena oli luoda yksinkertainen tunnistepalvelu, mikä vaatii tyyppien nimeämisen ja tunnistamisen selvittämisen. Yksilöiviä, pysyviä ja jaettuja id:tä eli tunnisteita työskentelyssä tarvittiin datan semanttiseen rikastamiseen. Toisaalta ne mahdollistavat datan jakamisen maailmanlaajuisesti. Ennalta määritellyn tunnisteskeeman lisäksi tarvitaan lisätietoa tunnistamiseen. Lisätietona voi toimia linkki temaattiseen portaaliin, joka sisältää kuvia tai muuta kuvailevaa dataa sivusta. Astuvansalmen kohdalla päädyttiin koordinaattipohjaiseen ratkaisuun. (Hannus ym. 2010.)

3D-mallissa kallion pinta peittyy valokuvilla. Tähän pintaan pystytään linkittämään muita valokuvia ja aineistoja. Valokuvien linkittäminen vaatii koordinaattien määrittämistä kullekin kohdalle. Koordinaattien määrittäminen ei ole yksinkertaista, koska se ei onnistu helposti yhden koordinaattipisteen avulla. Koordinaatteja voidaan luoda myös sekä yksittäiselle kuvalle että kuvaryhmälle. Vaihtoehtoisia tapoja koordinaattien luomiselle on useita, mutta tärkeintä on tehdä koordinaatit julkisiksi maailmanlaajuisesti URI-tunnisteen avulla, johon liittyy kuvailevaa dataa. Tunnisteiden jakaminen tekee mahdolliseksi luoda yhtymäkohdan, johon uutta dataa on mahdollista linkittää. Tämä mahdollistaa hankkeessa kartoitettujen fyysisten ja elektronisten aineistojen tuomisen yhteen museoista, kirjastoista ja arkistoista. Myös myöhemmin aihepiiriin liittyvät aineistot on mahdollista linkittää, kun tunnisteet ovat tiedossa ja jaettuja muille tahoille. (Hannus ym. 2010.)

Koordinaattiratkaisuksi valittiin ”bounding box” eli neliö, joka rajoittuu kuvan tai kuvaryhmän ympärille. Koska on kyse 3D-mallista, on tietenkin huomioitava myös kolmas ulottuvuus eli syvyyskoordinaatti. Neliön keskipiste on yksilöity kohta. Vaihtoehtoisia tapoja määrittää yksilöity kohta ovat mm. käyttää kaikkia neliön

kulmien pisteitä tai käyttää sääntöjä, jotka määrittävät kohdan esim. neliön oikea yläkulma. Todennäköisyys, että kahdella kuvalla olisi sama koordinaattipiste, on hyvin minimaalinen. Ongelmallista menetelmässä on neliön rajojen määrittäminen, joka vie aikaa ja jota ei ole mahdollista automatisoida. Kuvien määrä ja laatu vaikuttavat merkittävästi työmäärään. (Hannus ym. 2010.)

Hankkeessa ei ole pyritty nopeaan ratkaisuun vaan miettimään ja kartoittamaan ongelmia. Toiveissa on löytää tulevaisuudessa ratkaisuja näihin ongelmiin. Tulevaisuudenratkaisujen toivotaan kannustavan rikkaan ja erilaisen datan julkaisemista, jotka liittyvät läheisesti yksilöiviin, jaettuihin ja kestäviin/pysyviin tunnisteisiin. (Hannus ym. 2010.)

6.2.6 Metatietoskeeman mallinnus

Metatietoskeema nimensä mukaisesti sisältää metatietoa eli ”tiedon kontekstia, sisältöä ja rakennetta sekä niiden hallintaa ja käsittelyä koko elinkaaren ajan kuvaavaa tietoa” (KDK 2010b). Yksinkertaisesti sanottuna metatietoskeema koostuu kentän nimistä. Tähän läheisessä yhteydessä varsinkin aineistolähtöisessä tavassa kuvata asioita ovat kentän arvot, jotka voivat olla mm. auktorisoituja lähteitä eli sovittuja käsitteistöjä.

Kansalliskirjaston standardisalkussa on mainittu mm. sekä Dublin Core, Lido, VRA Core, MARC21 että CIDOC CRM kuvailevan metatiedon mahdollisena formaattina. Näistä kaikista formaateista keskusteltiin projektin edetessä. KDK:n tiedot on tarkoitus siirtää Europeanaan, jossa käytetään siirtoformaattina ESEä (European Semantic Elements). ESE ollaan korvaamassa EDM:llä (European Data Model). Lähtökohtana Astuvansalmen metatietoskeeman teolle oli ESE:n elementit sisältävä skeema. (KDD 2010a)

Metatietoskeeman ja tietokannan mallinnuksen pohjalla oli käsittemallista löydettävät objektit eli kentät. Käsittemallia lähdettiin luokittelemaan ja teemoittamaan Listbaseen, jonka Mika Nyman oli pystyttänyt Djangoa hyväksikäyttäen. Listbaseen koottiin objekteja ja otsakkeita ja niiden sisältöjä. Sisällöt pystyivät olemaan siis ihan mitä tahansa kalliomaalauksiin liittyviä asioita niputettuina eli teemoitettuna. Tavoitteena tässä ryhmittelyssä oli päästä käsiksi eri aineistoihin ja niiden merkityksiin ja saada

uusii ideoiia. Listbase loi siis kontrolloiiuua sanastoja tietokantaa ja sen käyttöliittymää varten.

Varsinaista metatietoskeemaa lähdeiiin rakentamaan käsiitemallin pohjalta valiten tärkeimpiä kentiiä. Kenttien valitsemista tein yhteistyössä projektipäällikkö Juhani Grönhagenin kanssa. Otos tästä alustavasta metatietoskeemasta on liitteessä 6. Tässä vaiheessa työskentelyyn mukaan tuli Jyväskylän yliopiston taiteiden ja kulttuurin tutkimuksen laitos, jonka tavoitteena oli hierarkkisen kalliomaalauskiisiteistöön luominen. Työstä vastuussa on projektitutkija Karoliina Järvinen. Itse kokosin käsiitemallin pohjalta Viva3-projektin näkökulmasta mallissa kuvattavat asiat ja Järvinen oli omalla tahollaan tehnyt alustavaa kiisiteistöä. Tavoitteena ei ollut varsinaisesti skeemojen yhdistäminen, mutta ideana oli hyödyntää Viva3-hankkeessa molempien tekemää työtä ilman suurempaa päällekkäisyyttä. Näiden kahden raakilemallin perusteella hahmottelin metatietoskeemaa siten että se sisältää sekä semanttisesti vahvat kentät että hierarkkisen kiisiteistöön (liite 7). Lähtökohtana oli päättää aluksi objektit eli kenttien nimet, jonka jälkeen kenttien arvojen määrittäminen eli kalliomaalauskiisiteistöön luominen olisi Jyväskylän yliopiston tehtävänä. Viva3-hankkeen osuus ei varsinaisesti kata kalliomaalauskiisiteistöä vaan tarkoitus on ollut keskittyä semanttisesti vahvan ontologisoinnin kokeiluun, joka päätettiin tehdä CIDOC CRM –luokkatasolla.

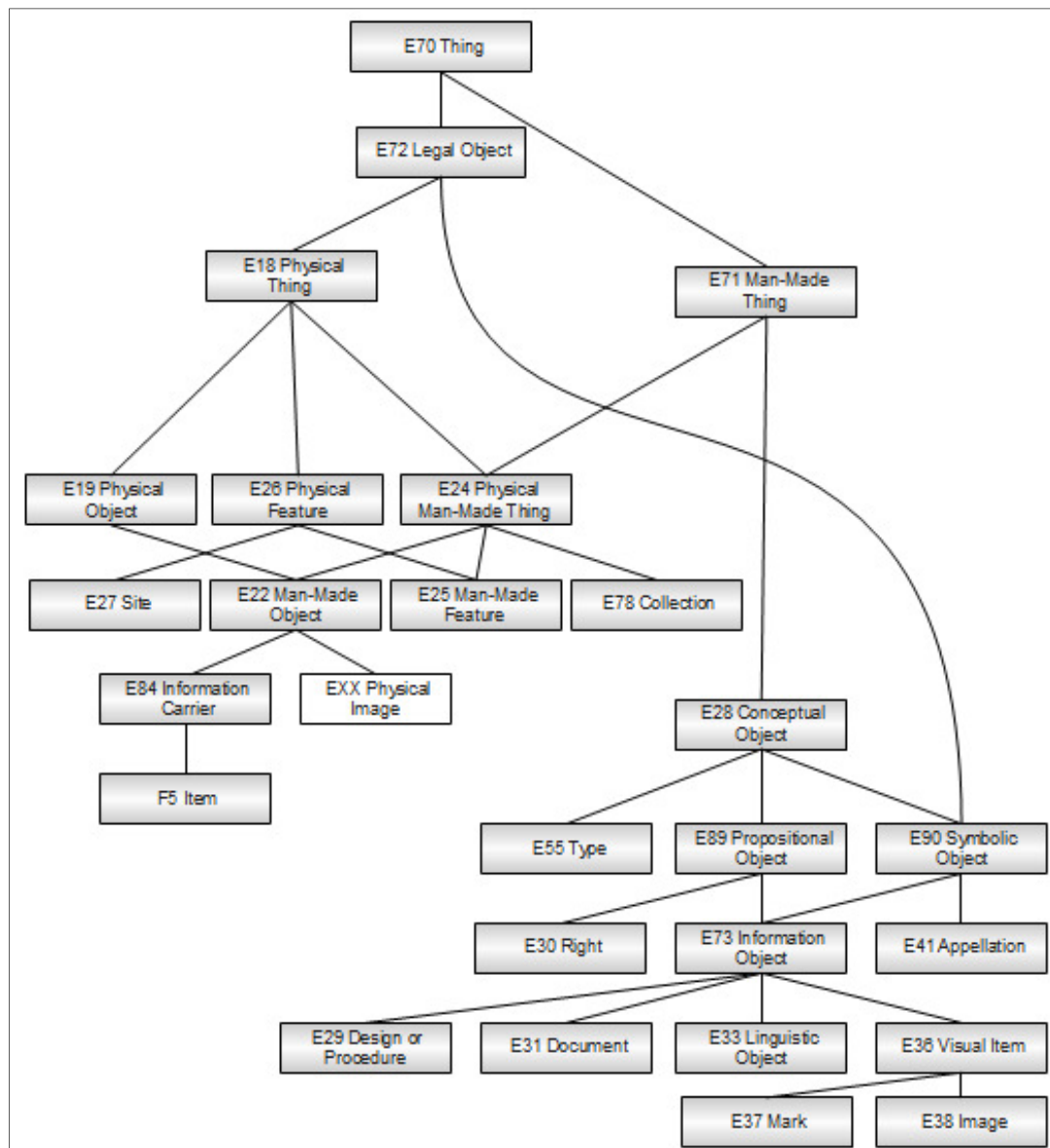
Metatietoskeemaa tehdessä tuli eteen monia haasteita. Yksi suurimmista ongelmista tässä vaiheessa oli projektin rajaus koskemaan testausta. Toisaalta tämä tarkoittaa ongelmien kartoittamista ilman varsinaisia lopullisia ratkaisuja näihin ongelmiin ja toisaalta, ettei oltu määritelty varsinaista kohderyhmää, jota varten ontologisointi tehdään. Totta kai voidaan sanoa, että ontologisoinnin lopputulokset koskettavat kaikkia kalliotaiteesta kiinnostuneita, mutta jos työtä olisi tehty todellista kohderyhmää varten, olisi ollut hyödyllistä tehdä selvitys käyttö- ja käyttäjätarpeista. Tähän käyttäjätarpeiden selvittämiseen käytettiin kuitenkin Grönhagenin asiantuntemusta valikoitaessa tärkeitä kentiiä. Toisena haasteena oli ajanpuutteen vuoksi luopuminen tietokannan käyttöliittymän tekemisestä projektissa. Projekti keskittyi siis pääasiassa mallintamiseen. Alkuvaiheen projektin määrittely koskemaan aineistolähtöistä kiisittelytapaa osaltaan teki käyttäjätarpeista toissijaisen lähtökohdan.

Tulkinnoista on puhuttu tässä työssä aiemminkin ja myös tässä metatietoskeeman teon alkuvaiheessa sen sijoittaminen osoittautui hankalaksi, koska tavallaan tulkinnat koskevat kaikkea esitettyä tietoa. Mikä siis on ns. auktorisoitujen kontrolloitujen käsitteistöjen rooli suhteessa lähteisiin. Aiemminhan oli päädytty, että tulkintaan liittyy aina lähde. Mutta jos tietojen syötössä käytetään lomaketta, jossa on arvoina pääsääntöisesti käytetty sovittua käsitteistöä, niin miten tulkinnat saadaan tähän mukaan. Esimerkiksi kuva-aiheita voidaan ilmaista valmiin käsitteistön avulla esim. Lahelman skeema, mutta entä jos kuva-aiheita on tulkittu eri tavoin, mihin tämä tieto sijoitetaan metatietoskeemassa. Eli onko lopputulos, että on olemassa useita eri tahojen tekemiä kontrolloituja sanastoja ja toisaalta miten saadaan sellaiset tulkinnat näkymään, jotka eivät ole auktorisoidun eli ns. hyväksytyt tahon tuottamia. Tähän yhteyteen liittyvät myös käyttäjien tulkinnat kuva-aiheista ja käyttäjien lisäämän kuvailun saaminen näkymään metatietoskeemassa.

Metatietoskeeman tekovaiheessa todettiin, etteivät CIDOC CRM:n kentät ole riittäviä kuvaamaan todellisuutta, vaan päätettiin laajentaa CIDOC CRM:ää. CIDOC CRM on luonteeltaan luotu laajennettavaksi, joten tässä tapauksessa tämä vaihtoehto koettiin parhaimmaksi. Tarve laajennukselle tuli, koska CIDOC CRM:n kehittämisessä on painotettu käsitteellistä puolta eli erilaiset dokumentit on mallissa sijoitettu E28 eli käsitteellinen objektin (Conceptual Object) alaisuuteen. Astuvansalmen tapauksessa on kuitenkin kyse fyysisestä kokonaisuudesta eli kalliosta, fyysisistä objekteista eli kalliomaalauksista sekä fyysisistä kuvista – ei ainoastaan kuvien käsitteellisestä sisällöstä. Luokkien laajennuksen tuli siis ulottua E18 Physical Thing alaisuuteen. Laajennuksen tavoitteena on mahdollistaa polkujen syntymisen kalliosta eli originaaliaineistosta kirjallisiin ja kuvallisiin aineistoihin. Laajennetun CIDOC CRM:n voi testata vakiintuneella skeemalla kuten ESEllä (Europeana Semantic Elements). (Nyman 2011a.)

CIDOC CRM laajennuksen teosta vastasi Mika Nyman. Kuvasta 23 käy ilmi ehdotus laajennuksesta. Ehdotukseen on lisätty EXX Physical Image, joka sijaitsee E22 Man-Made Object alaluokkana. CIDOC CRM:n mukaan fyysiset kuvat sijaitsevat E25 Man-Made Featuren alaisuudessa. Tässä tapauksessa koettiin tarpeelliseksi, että kuva on itsenäinen objekti, kuten havaintopsykologisesti se havaitaan. Kuvaa tulisi pystyä käsittelemään objektina, jolla on piirteitä. Piirteitä ovat mm. objektin osat, osan muodot jne. Kalliomaalauksien osalta piirteiden avulla voidaan kuvata esim. käsien

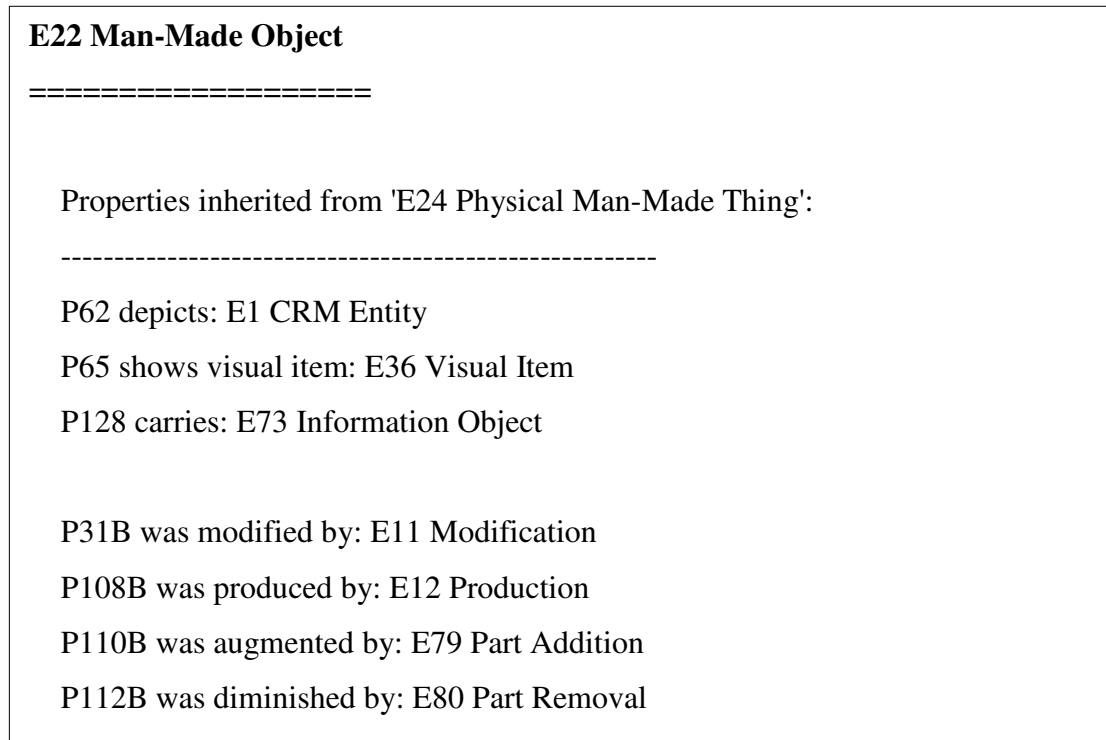
asentoa. Tarve soveltaa CIDOC CRM:ää varsinaisen käyttötarkoituksen ulkopuolelle johtuu pääosin sen käytöstä pääasiassa museoiden tarpeisiin. (Nyman 2011d.)



KUVA 23. CIDOC CRM –laajennusehdotus (Nyman 2011d)

Laajennusta tehdessä keskustelua herätti mm. E27 Collection. Museopuolella tätä käytetään tietyn kokoelman esittämiseen, mutta sitä olisi mahdollista soveltaa kuvakentän kuvaamiseen. Periaatteessahan kuvakenttään voidaan lisätä ja poistaa kohteita kuten museoiden kokoelmistakin. Toinen vaihtoehto olisi luoda E78 – luokalle rinnakkaisluokka E XY Aggregation (ryhmäytymä), jonka avulla olisi paremmin mahdollista kuvata originaaliaineiston kuvakenttää. Tämä rinnakkaisluokka koostuisi joko E25 Man-Made Featureista tai sovellettuna E22 Man-Made Objecteista. (Nyman 2011d.)

Laajennusehdotuksen lisäksi Mika Nyman teki luettelon CIDOC CRM:n ja FRBRoo:n periytyvistä ominaisuuksista, että asioiden moniperiytymisen selvittäminen on helpompaa. Luettelo ottaa huomioon käännteissuhteet sekä yläluokilta periytyvät suhteet. Kuvassa 24 on otos E22 Man-Made Objectille periytyvistä ominaisuuksista E24 Physical Man-Made Thingin kautta. Luettelosta on myös mahdollista helposti tarkistaa tiettyjä kohtia. (Nyman 2011d.)



KUVA 24. Otos E 22 Man-Made Object -luokalle periytyvistä ominaisuuksista (Nyman 2011e)

Astuvansalmen kalliosta ja sen yksittäisistä kuvista koostuvaa CIDOC CRM:n mallia käytetään siten, että lähdetään liikkeelle yhdestä erilliskohteesta ja selvitetään kohteen CIDOC CRM -luokka. Tämän jälkeen hyödynnetään CIDOC CRM:n luettelodokumentaatiota, että saadaan selville luokkaan liittyvät suhteet eli periytyvät ominaisuudet. Kaikki suhteet eivät yleensä ole kuvailun kannalta relevantteja eli käytetään ainoastaan niitä luokkia ja ominaisuuksia, jotka ovat kuvailun kohteen kannalta hyödyllisiä. (Nyman 2011f.)

Astuvansalmen kalliokuvaan pohjautuvaa ontologiaa Nyman tulee hyödyntämään Pohjoisten ja arktisten alueiden kalliomaalausportaalia luotaessa (Nyman 2011f). Ontologiadokumentti on löydettävissä verkosta (Nyman 2011g).

Metatietoskeemaa ja laajennusta tehdessä prosessi (liite 3) alkoi käsitelmällin luokittelulla objekteihin ja niiden sisältöihin, jatkui kontrolloitujen sanastojen luomisella ja ontologisten suhteiden miettimisellä CIDOC CRM:n osalta sekä loppui laajennusehdotuksen ja moniperiytyvyysluettelon tekemiseen.

6.3 Arviointivaihe

Arviointia voi tapahtua etukäteen, jatkuvasti ja jälkikäteen. Arvioitava kohde on riippuvainen hankkeen päämääristä ja tavoitteista (Jalava 1998, 132). Astuvansalmen osalta tavoitteiksi määriteltiin oppiminen ja vahvaan semantiikkaan perustuvan ontologisoinnin testaus.

Arviointivaiheessa oli alun perin tarkoitus arvioida CIDOC CRM:n pohjalta tehtyä metatietoskeemaa aineistolähtöisesti. Tärkeitä arviointikohteita olivat: mihin tapahtumiin skeeman elementit kytkeytyvät, kuinka hyvin ominaisuudet ovat periytyneet ja että suhteet vastaavat aineistoa. (Nyman 2009c.)

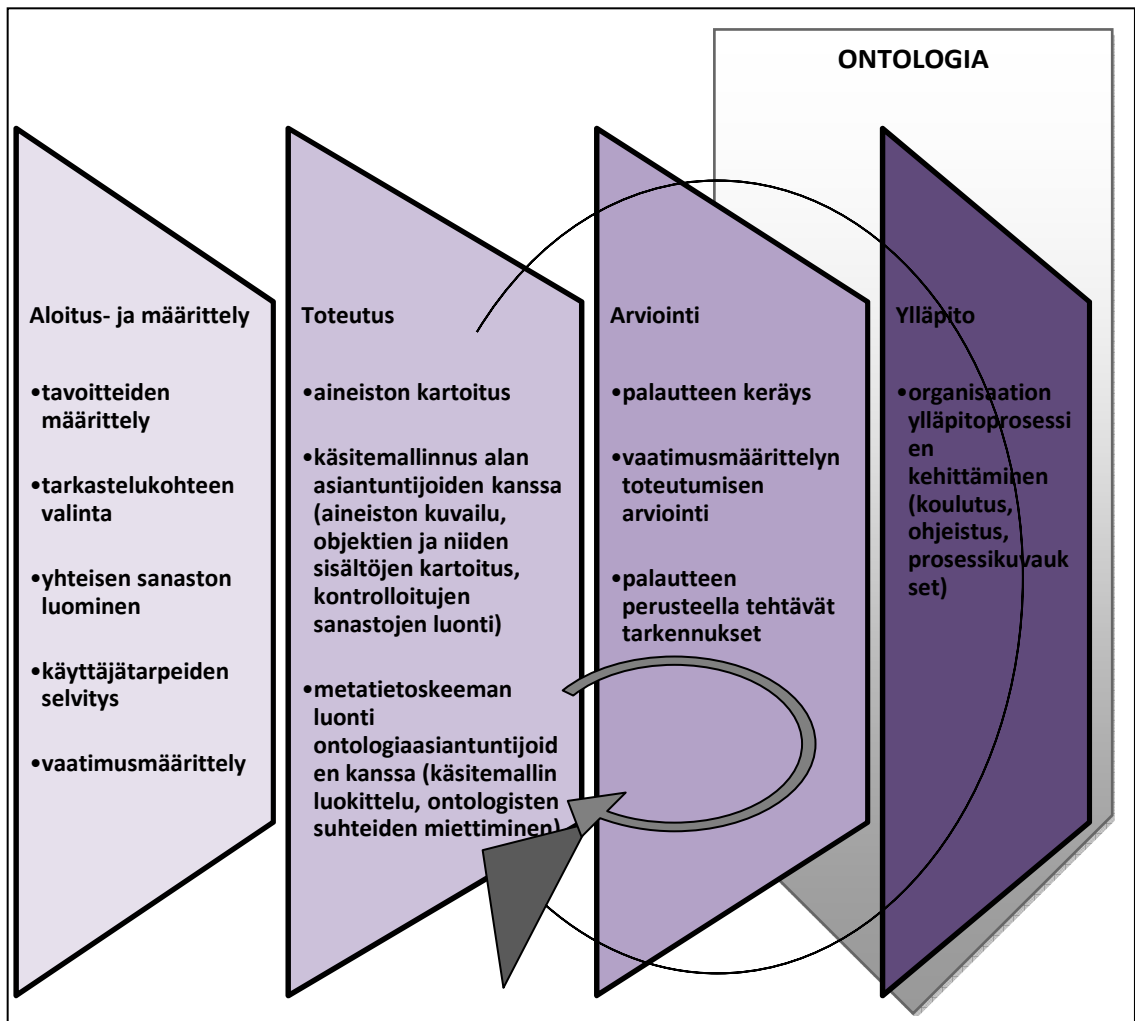
Varsinaista ontologiaa ei arvioitu, koska lopullista versiota CIDOC CRM:n laajennuksesta ei ollut aikataulullisista syistä käytettävissä. Aiemmin tehtyä vaatimusmäärittelyn toteutumista ei myöskään ollut mahdollista täysipainoisesti arvioida tästä syystä. Toisaalta prosessissa pidettiin tärkeänä avoimuutta uusille vaihtoehdoille ja päätyminen metatietoskeeman loppuun viemisen sijasta CIDOC CRM:n laajentamiseen on todiste tästä avoimuudesta. CIDOC CRM –laajennus antaa joka tapauksessa yhden ratkaisun ontologisoinnin tekemiseen aineisto huomioiden. Tämä uusi CIDOC CRM:n tulkintatapa ja lähestymiskulma ovat merkittäviä saavutuksia prosessissa. CIDOC CRM –laajennusta tullaan arvioimaan sen perusteella miten hyvin polku originaaliaineistoista syntyy erilaisiin kirjallisiin ja kuvallisiin tuotoksiin. Mikäli olisi kyse tuotantokäyttöön tuotetusta ontologiasta, olisi tarpeen kerätä palautetta käyttäjiltä ja myös tehdä tarvittavia korjauksia ja hienosäätöä annetun palautteen perusteella.

Ontologian lisäksi arvioitavana kohteena on ollut itse ontologisointiprosessi. Prosessi on kuvattu sekä sanallisesti että toimintamallien että kaavion avulla. Tällaisessa prosessissa mukana oleminen väistämättä tuottaa oppimista. Toisaalta selkeät prosessikuvaukset mahdollistavat oppimisen muille asiasta kiinnostuneille. Prosessista on tehty lisäksi muistilista Astuvansalmen kallion ja sen kalliokuvien aineistolähtöisessä ontologisoinnissa huomioitavista asioista (liite 8), joka omalta osaltaan tuo esiin sekä toteutukseen liittyviä että ongelmallisia kohtia. Tarkemmin tutkimusprosessia arvioidaan seuraavassa luvussa Tutkimusprosessin arviointi ja kehittämisideat.

6.4 Yhteenveto työvaiheiden kuvauksesta ja ontologisointiprosessikaavio

Tässä kappaleessa esittelen vielä lopuksi ontologisointiprosessikaavion ja vedän yhteen toteutuneet työn vaiheet. Ontologisointiprosessi voidaan kuvata kuvan 25 mukaisesti, joka yleistää visuaalisesti aineistolähtöisen ontologisointiprosessin vaiheet. Kaaviossa on otettu huomioon kehitettävät asiat.

Aloitus- ja määrittelyvaiheessa olisi ollut hyvä alustavasti määritellä yhteisesti käytettävä sanasto, selvittää käyttäjätarpeet ja tehdä yksityiskohtaisempi vaatimusmäärittely. Vaikka sanaston merkityksen tiesi prosessiin lähtiessä, se olisi kuitenkin tarvinnut kirjata heti alussa, koska se olisi yksinkertaistanut ja helpottanut tulevaa työtä. Käyttäjätarpeiden selvitystä ei tässä tutkimuksessa koettu välttämättömäksi, koska oli kyse testauksesta, mutta se olisi auttanut lopputuloksen arvioinnissa. Todellista ontologiaa kehitettäessä käyttäjätarpeet tulee kuitenkin selvittää ja huomioida.



KUVA 25. Ontologisointiprosessikaavio

Toteutusvaiheessa ilmeni uusia näkökulmia ja huomioitavia asioita. Vaiheiden tarkempi sisältö avautui paremmin työtä tehtäessä. Mm. kuvien ja kuva-aiheiden kartoitus on loogista näin jälkikäteen ajatellen, mutta sen erottelu alun perin aiotusta aineiston kartoituksesta on oleellista. Myös CIDOC CRM –laajennus tuli aiottuun suunnitelmaan uutena näkökulmana. Lopulta toteutusvaihe noudatti seuraavaa järjestystä:

1. Aineiston kartoitus (vastuu Henna Mölsä)
2. Kuvien ja kuva-aiheiden kartoitus (vastuu Henna Mölsä)
3. Käsitellinnus ja kuvailu (vastuu Henna Mölsä)
4. Metatietoskeeman mallinnus, välivaihe (vastuu Henna Mölsä)
5. Ontologisten suhteiden miettiminen: CIDOC CRM (vastuu Miia Herrala, Mika Nyman)
6. CIDOC CRM:n laajentaminen projektin tarpeisiin (vastuu Mika Nyman)

7. Yksinkertaisen prosessikuvauksen tekeminen työvaiheista (vastuu Henna Mölsä)

Pääasiassa aikataulullisista syistä varsinaista ontologisoinnin arviointia ei tehty. CIDOC CRM –laajennus antaa kuitenkin yhden ratkaisuvaihtoehdon kalliomaalausten ontologisoinnille. Todellisen ontologian teon arvioinnissa tulisi huomioida laajennuksen lisäksi ainakin palautteen keräys, vaatimusmäärittelyn toteutumisen arviointi ja palautteen perusteella tehtävät tarkennukset.

Ylläpitovaiheeseen asti ei testauksessa päästy, minkä takia toimintamalleissa ei ole kuvattu ylläpitoprosesseja. Mikäli ylläpitovaiheeseen asti olisi päästy, ylläpidon tulisi palvella ontologisoinnin jatkuvaa kehittämistä. Kehittäminen vaatii mm. ontologisointiprosessien kuvaamista, ohjeistusta ja koulutusta.

Yksi merkittävä Viva3-hankkeen työntekoa ja ontologisointiprosessia jäsentävä asia oli raportin kokoaminen CIDOC 2010 konferenssiin 7. – 12.11.2010 Shanghaihin Kiinaan. Tiivistelmän tuli olla valmis huhtikuussa 2010 ja valmiin raportin elokuussa 2010. Raportin tavoitteina oli mm. tuoda esille Mikkelin 3D-osaamista sekä mahdollistaa kontakteja ja yhteistyösuhteita. Raportin ideana oli esittää vaihtoehtoinen tapa toimia eli ”näinkin voisi tehdä” ja muistaa, että kyseessä on nimenomaan menetelmäkehityshanke. Tätä artikkelin kirjoittamista ei mietitty alustavassa prosessissa, mutta se toi yhteen Viva3-hankkeen eri aiheita kootusti ja jäsentäen niitä ja niiden välivaiheita samalla asettaen selkeän aikataulun tekemiselle. Konferenssiraportti on löydettävissä CIDOC 2010 konferenssin nettisivustolta. (Hannus ym. 2010.)

7 TUTKIMUSPROSESSIN ARVIOINTI JA KEHITTÄMISIDEAT

Mikkelin ammattikorkeakoulun sähköiseen aineiston hallintaan liittyvissä hankkeissa on pohdittu metatietojen käsittelyä myös ontologioiden kannalta alkaen ElkaD-hankkeesta. Koulutuksen puolella sähköisen asioinnin ja arkistoinnin koulutusohjelmassa ontologisointia käsiteltiin ensimmäisen kerran tammikuussa 2010. Aineistolähtöiset ja vahvaan semantiikkaan perustuvat lähtökohdat prosessin

etenemiselle poikkesivat valtakunnallisesta ontologioiden kehittämisestä, jotka perustuvat pääasiassa heikon semantiikan sanastopohjaisille ontologioille.

Alussa tavoitteiksi ontologisointiprosessille määriteltiin oppiminen sekä prototyypin testaus. Ontologisointiprosessi tuli dokumentoida ja kuvata helposti ymmärrettävästi. Ontologisoinnin osalta oma asiantuntijuus kehittyi prosessin edetessä ja koulutukseen osallistuessa. Oppimistavoite saavutettiin sekä henkilökohtaisella että organisaation oppimista palvelevalla tavalla. Ontologisoinnin testaus antaa Mikkelin ammattikorkeakoululle tietoa aineistolähtöisestä lähestymistavasta, kuinka vahvan semantiikan CIDOC CRM –mallia on mahdollista hyödyntää sekä prosessiin ja tutkimusaiheeseen liittyvistä haasteista. Prosessin kuvaaminen niin sanallisesti työnkuvauksen avulla kuin toimintamallikaavioiden avulla sekä yleistettävissä olevan kaavion avulla, avaa prosessin osa-alueita ja kokonaisuutta myös ulkopuolisille asiasta kiinnostuneille palvelleen näin oppimista. Erityisesti ontologisointiprosessikaavio vie ontologisoinnin yleistettävälle tasolle, mikä mahdollisesti helpottaa tulevien ontologisointiprosessien suunnittelua. Prosessissa kerättyjä kokemuksia ja prosessikuvauksia on mahdollista soveltaa sekä kansallisesti että kansainvälisesti kulttuuriaineistojen ontologisointiprosesseissa. Prosessien kuvaamisen osalta päästiin tavoitteisiin.

Prosessit on kuvattu sekä kirjallisesti että kuvallisesti, mikä palvelee erilaisten oppimistyylien tapaa hahmottaa asioita. Toki prosesseja olisi voinut kuvata hyvin eri tavoin. Kuvauksessa olisi voinut hyödyntää esim. liikkuvaa kuvaa ja Internettiin sijoitettuja materiaaleja. Yhtenä kehittämisideana olisikin luoda ontologisointiprosessista oppimateriaali verkkoon. Englannin kielen käyttäminen laajentaisi huomattavasti aineiston hyödyntämismahdollisuuksia kansainvälisesti.

Tutkimusmenetelmäksi valittu refleктоiva oppiminen ja dialogi eri asiantuntijatahojen kanssa osoittautui ennakko-oletuksia vaikeammaksi. Prosessiin ja sen lopputulokseen vaikuttavat oleellisesti käytettävissä olevat yhteistyö- ja asiantuntijatahojen asiantuntemus kohdealueesta ja työryhmän laajuus. Yhteistyö eri asiantuntijatahojen kanssa oli mielenkiintoista, mutta samalla myös haasteellista. Vaikka ydintyöryhmän koko ei ollut kovin suuri, monia asioita olisi helpottanut, jos työryhmän jäsenet olisivat olleet samalla paikkakunnalla ja tehneet töitä yhdessä saman katon alla, jolloin keskustelujen käyminen olisi lisääntynyt ja ideointi helpottunut. Vaikka sähköisellä

aikakaudella elämmekin, vuorovaikutus on kuitenkin helpointa kasvotusten erityisesti kun on kyse itselle ennestään lähes tuntemattomasta aihepiiristä.

Vuorovaikutus eri asiantuntijatahojen kanssa on kuitenkin välttämätöntä, mikäli pyritään kehittämään laadullisesti hyvä ontologia. Ontologisointiprosessissa korostuvat ontologisoinnin asiantuntijuus ja alan asiantuntijuus – tekniikka ja sisältö. Tietoteknisten henkilöiden ja tutkimuskohteen alan asiantuntijoiden välistä vuorovaikutusta olisi helpottanut yhteinen sanaston luominen jo alkuvaiheessa. Nyt skeema, käsitelmä tai formaatti saattoi eri tapauksissa tarkoittaa eri asiaa. Monet kerrat termejä joutui tarkentamaan ymmärtääkseen puhutaanko samasta vai eri asiasta kuin mitä itse ajattelee.

Toisaalta oppimisprosessissa korostui oma ajattelu ja asiantuntijuuden kunnioittaminen, jotka ovat mielestäni hyviä lähtökohtia työskentelyssä. Valmius muuttaa alkuperäisiä suunnitelmia keskustelun ja ajatusten pohjalta tarpeen mukaan kertoo osaltaan itseään kehittävästä työryhmästä. Työskentelyssä korostettiin erityisesti kaikkien työskentelyssä mukana olevien asiantuntemusta, olipa asiantuntemus miltä alalta tahansa. Myös avoimuus korostui, ei ollut välttämättä väärää tapaa tehdä vaan pikemminkin vaihtoehtoisia lähtökohtia, joista valittiin työskentelysuuntia. Tällainen työskentelytapa palvelee oppimista, oivalluksia ja uuden tiedon luomista, tasavertaisuutta sekä aitoa ymmärrystä aiheesta. Toisaalta se, ettei kukaan ole sanomassa ”oikeaa” tapaa toimia, luo epävarmuutta valinnoista.

Ontologisointiprosessin tarkoituksena ei ollut varsinaisesti luoda uutta tietoa, vaan testata vahvan semantiikan ontologisointia. Prosessissa päästiin välivaiheeseen metatietoskeeman osalta ja CIDOC CRM –laajennuksesta tehtiin ehdotus. Valitettavasti prosessissa ei päästy prototyypin tekemiseen ja testaukseen asti eli metatietoskeeman toimivuutta ei pystytty todentamaan. Hankkeen tavoitteiden toteutuminen jäi jonkin verran vaillinaiseksi.

Prosessi olisi toki voinut toteutua hyvin eri tavallakin. Aineistolähtöisyyttä olisi voitu lähestyä mm. alussa suunnitellun Protégé-ohjelman avulla. Toisaalta pääpaino olisi voinut olla aineistolähtöisyyden sijasta teknisessä toteutuksessa mallintamisen sijasta.

Ontologioita on mahdollista hyödyntää ja käyttää nykyistä enemmän. Nykyiset sähköiset arkistot on yleensä rakennettu yleisen suomalaisen asiasanaston (YSA) hyödyntämistä ajatellen, joten loogisesti tästä edetään YSON hyödyntämiseen. Toisaalta vahvan semantiikan ontologisointi vaatii vahvaa ammattitaitoa ontologisoinnista ja toisaalta aineistolähtöinen vahvaan semantiikkaan pohjautuva lähestymistapa on hyvin raskas pienelle organisaatiolle.

Koska itsellä ei ollut ennestään kokemusta ontologisoinnista, prosessin suunnittelu ja toteuttaminen järkevästi tuotti ongelmia. Yritin saada käyttöön mm. wikin, joka olisi helpottanut tiedonkulkua. Idea ei kuitenkaan ottanut tulta alleen, joten tämä mahdollisuus jäi hyödyntämättä. Opiskeltujen asioiden soveltaminen ei aina ollut mahdollista, kun ei tiennyt tarpeeksi hyvin tulevia työvaiheita. Ontologisointiprosessikaavio on tehty yleistettäväksi malliksi, että se voi helpottaa muita, jotka lähtevät miettimään ontologisointia.

Ontologiat koetaan vaikeasti lähestyttävänä. Tämä tutkimus osaltaan valottaa työtä, joka liittyy ontologisointiprosessiin. Haasteena on löytää kuhunkin prosessiin sopiva työryhmä, jolta löytyy asiantuntemusta kohteeksi valitusta aiheesta ja ontologioista. Toisaalta aloittelijan mukana oleminen ryhmässä voi tarjota uusia näkökulmia ja toisaalta näin tietoa voidaan jakaa käytännön oppimisen kautta. Aihe on hyvä rajata selkeästi, vastuuttaa ja aikatauluttaa. Projektinhallintataidot ovat ensiarvoisen tärkeitä kuitenkin unohtamatta vapautta poiketa alustavista suunnitelmista. Näistä lähtökohdista on hyvä lähestyä ontologioiden maailmaa.

8 LOPUKSI

Astuvansalmen kallio tutkimuskohteena oli erittäin mielenkiintoinen. Se herättää mielenkiintoa sekä paikallisesti, kansallisesti että kansainvälisesti. Toisaalta opinnäytetyön tutkimusaiheena kivikautisen ja nykyaikaisen tiedonkulun yhdistäminen on itsessään ajatuksia herättelevä ja hedelmällinen lähtökohta. Kalliokuvat kertovat kullekin meille omaa tarinaansa elämästä ja ontologiat pyrkivät tätä ympäröivää todellisuutta hahmottamaan koneille (Hietala 1993, 81).

Ontologioiden epämääräisyys on muuttunut selkeämmäksi näkemykseksi eri ontologioiden eroista, kuinka ontologioita voidaan lähteä rakentamaan ja mitä asioita on huomioitava ontologisointiprosessia suunnitellessa. Prosessissa on tuotettu muiden hyödynnettävissä olevia kuvauksia prosessista ja näin tehty ontologiat myös muille helpommin lähestyttäviksi.

Maaailma kehittyy edelleen tietoteknisempään suuntaan. Informaatio- ja tietomäärät kasvavat kasvamistaan. Ontologioiden toimivuudesta kiistellään, mutta tosiasia on, että ne ovat yksi vaihtoehto tehdä konemaailmasta ihmisille ymmärrettävämpää ja parantaa tiedonhakujen tuloksia. Ontologiat antavat mahdollisuuden sellaisten asiayhteyksien tietoon saattamiseen, joista tiedonhakijalla ei ole tietoa tiedonhakutilanteessa. Tiedonhankintataitojen merkitys muuttuu, kun palvelut pystyvät suosittelemaan ja visuaalisesti esittämään informaatiota totuttua paremmin. Käyttäjien tarpeiden ennakointi on hyvä suuntaus. Kirjastoihmisenä koen, että tällainen lähtökohta lupaa paljon.

Ontologioiden kehittäminen tulevaisuudessa vaatii yhteistyötä. Olisikin toivottavaa, että tulevaisuudessa kirjastot, arkistot ja museot voisivat tehdä enemmän yhteistyötä laadukkaiden ontologioiden rakentamisessa. KDK-hanke ja yleinen suomalainen ontologia ovat askel tähän suuntaan, mutta tulevaisuuden tavoitteena tulisi olla vahvemman semantiikan ontologiat, jolloin asioiden suhteet säilyisivät huomattavasti paremmin ja toisivat aivan uusia oivalluksia tiedonhakijoille. Haasteita riittää, mutta mahdollisuudet ovat jo nyt olemassa.

Astuvansalmen kallion ontologisointiprosessissa mukana oleminen oli monellakin tapaa haastavaa. Oppiminen ei rajoittunut ainoastaan opiskeltavaan alaan vaan myös kohdealueen eli kalliokuvien tuntemukseen. Tietoa tuli omaksua niin ontologioista kuin kalliokuvistakin samanaikaisesti. Itselleni tärkeää prosessissa oli kokea ja oppia kuinka verkostoitumista voi hoitaa tehokkaasti ja toisaalta kuinka luodaan ilmapiiri, jossa kaikki työskentelyssä olevat kokevat olevansa asiantuntijoita omalla allaan. Haluankin kiittää kaikkia työskentelyyn osallistuneita, jotka ovat omalta osaltaan tehneet mahdolliseksi ontologisointiprosessin toteutumisen.

9 LÄHTEET

Apilo, Tiina, Taskinen, Tapani & Salkari, Iiro 2007. Johda innovaatioita. Talentum: Helsinki.

Definition of the CIDOC Conceptual Reference Model 2010. Version 5.0.2. ICOM/CIDOC CRM Special Interest Group. http://www.cidoc-crm.org/docs/cidoc_crm_version_5.0.2.pdf. Päivitetty 27.8.2010. Luettu 4.10.2010.

CIDOC CRM 2010. The CIDOC CRM version 5.0.1 in a graphical format. http://www.cidoc-crm.org/cidoc_graphical_representation_v_5_1/graphical_representaion_5_0_1.html WWW-dokumentti. Päivitetty 15.7.2010. Luettu 16.1.2011.

Denny, Michael 2002. Ontology Building: A Survey of Editing Tools, Table 1. Ontology editor survey results. WWW-dokumentti. http://www.xml.com/2002/11/06/Ontology_Editor_Survey.html. Päivitetty 5.5.2003. Luettu 7.11.2010.

Doerr, Martin & Stead, Steve 2009. The CIDOC CRM. A standard for the integration of cultural information. The CRM tutorial at Imperial College, UK, May 22, 2009. CIDOC Conceptual Reference Model Special Interest Group. PPT-esitys. Luettu 16.1.2011.

Fernández-López, M. 1999. Overview of the methodologies for building ontologies. Teoksessa V.R. Benjamins (toim.) Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI 99), Workshop on Ontologies and Problem-Solving Methods, Stockholm, Sweden, July 31 - August 6. Morgan Kaufmann, 33-46.

Fuster, Joaquín M. 2003. Cortex and mind. Unifying cognition. Oxford University Press: Oxford.

Gruber, Thomas R. 1993. A translation approach to portable ontology specifications. Knowledge systems laboratory. Technical report. KSL 92-71. Stanford University: California. PDF-dokumentti. <http://tomgruber.org/writing/ontolingua-kaj-1993.pdf>. Päivitetty 27.6.2006. Luettu 25.10.2010.

Guarino, Nicola & Giaretta, Pierdaniele 1995. Ontologies and knowledge bases. Towards a terminal clarification. National research council, LADSEB-CNR: Italy. Institute of history of philosophy, University of Padova: Italy. PDF-dokumentti. <http://www.loa-cnr.it/Papers/KBKS95.pdf>. Päivitetty 1.4.2004. Luettu 25.10.2010.

Grönhagen, Juhani 2010. Sähköpostiviesti 20.9.2010. Projektipäällikkö. Viva3-hanke.

Hannus, Esa, Herrala, Miia, Grönhagen, Juhani, Mölsä, Henna & Nyman, Mika 2010. Global sharing of information about prehistoric rock art. CIDOC 2010. ICOM General conference. Shanghai, China. 8. – 10.11.2010. http://cidoc.meta.se/2010/full_papers/herrala.pdf

Hawkings, Jeff & Blakeslee, Sandra 2004. On intelligence. Times books: New York.

Heinemann, Laila 2007. Tiedonorganisoinnin uudet haasteet : miten globalisoitua yhteisö hakee tietoa. Tietolinja. Kansalliskirjasto. <http://urn.fi/URN:NBN:fi-fe20072115>. Ilmestynyt 12.12.2007. Luettu 20.2.2010.

Heikkilä, Jorma & Heikkilä, Kristiina 2000. Dialogi. Avain innovatiivisuuteen. WSOY: Helsinki.

Herrala, Miia 2009. Sähköpostiviesti 20.7.2009. Projektityöntekijä. Viva3-hanke.

Herrala, Miia 2009. Hankkeen metatietoselvitys 10.9.2009. Sähköpostiviesti 10.9.2009.

Herrala, Miia 2010. Sähköpostiviesti 22.9.2010. Projektityöntekijä. Viva3-hanke.

Hietala, Veijo 1993. Kuvien todellisuus. Johdatusta kuvallisen kulttuurin ymmärtämiseen ja tulkintaan. Kirjastopalvelu: Helsinki.

Hitzler, Pascal, Krötzsch, Markus & Rudolph, Sebastian 2009. Foundations of semantic web technologies. CRC Press: London.

Hotti, Virpi, Huttunen, Riikka, Kajander, Ani, Lehmuskoski, Antero, Ojala, Matti, Taskinen, Teppo & Tiihonen, Timo 2008. Tietämyksenhallinta ja ontologiat sosiaalihuollon näkökulmasta. Stakesin raportteja 11/2008. Sosiaali- ja terveystietoyhteiskuntayksikön julkaisuja. PDF-dokumentti. <http://www.stakes.fi/verkkojulkaisut/raportit/R11-2008-VERKKO.pdf>. Luettu 11.2.2010.

Hovi, Ari, Huotari, Jouni & Lahdenmäki, Tapio 2005. Tietokantojen suunnittelu & indeksointi. Docendo: Jyväskylä.

Hyvönen, Eero 2001. Semantic web. Kohti uutta merkitysten Internetiä. Esitelmä Semantic Kick-Off in Finland -tilaisuudessa. Helsingin yliopisto. Porthania 2.11.2001. <http://www.cs.helsinki.fi/u/eahyvone/stes/semanticweb/SemanticWebVisio>. Päivitetty 1.11.2001. Luettu 25.10.2010.

Hyvönen, Eero, Seppälä, Katri, Viljanen, Kim & Frosterus, Matias 2007. Yleinen suomalainen ontologia YSO. Kohti suomalaista semanttista webiä. Semanttisen laskennan tutkimusryhmä (SeCo), TKK Viestintätekniikan laboratorio, Helsingin yliopisto, tietojenkäsittelytieteen laitos: Helsinki. PDF-dokumentti. <http://www.seco.tkk.fi/publications/2007/hyvonen-et-al-yso-2007.pdf> Päivitetty 2.9.2010. Luettu 25.10.2010.

Hyvönen, Eero 2008a. FinnONTO-malli kansallisen semanttisen webin sisältöinfrastruktuurin perustaksi - visio ja sen toteutus. TKK viestintätekniikan laitos ja Helsingin yliopisto, tietojenkäsittelytieteen laitos: Helsinki. PDF-dokumentti. <http://www.seco.tkk.fi/publications/2008/hyvonen-ONKI-yleisesitys-2008.pdf>. Päivitetty 2.9.2010. Luettu 24.10.2010.

Hyvönen, Eero 2008b. Kulttuurisampo - suomalainen kulttuuri semanttisessa webissä. Muistiorganisaatioiden ja kansalaisten yhteisöllinen kansallinen julkaisujärjestelmä.

PDF-dokumentti. <http://www.seco.tkk.fi/publications/2008/hyvonen-Kulttuurisampo-2008.pdf>. Päivitetty 2.9.2010. Luettu 30.11.2010.

Jalava, Urpo & Virtanen, Petri 1998. Tietoa luova projekti. Polku oppivaan organisaatioon. Kirjayhtymä: Helsinki.

Jussila, Timo 1999. Saimaan kalliomaalausten ajoitus rannansiirtymiskronologian perusteella. Teoksessa Kivikäs, Pekka 1999. Saimaan ja Päijänteen alueen kalliomaalausten sijainti ja syntyaika. Kivikäs, Muinaistaidekeskus: Jyväskylä.

Jussila, Timo 2004. Suomen esihistorian kronologiataulukko. <http://www.mikroliitti.fi/aikakaa/aikakaaf.htm>. Päivitetty 4.8.2005. Luettu 24.11.2010.

JHS 152 2002. Prosessien kuvaaminen. JUHTA Julkisen hallinnon tietohallinnon neuvottelukunta. PDF-dokumentti. <http://www.jhs-suositukset.fi/suomi/jhs152>. Julkaistu 13.12.2002. Versio 6.6.2008. Päivitetty 20.6.2010. Luettu 4.9.2010.

Kaario, Kimmo & Peltola, Tuomo 2008. Tiedonhallinta. Avain tietotyön tuottavuuteen. Docendo: Jyväskylä.

Karhulampi, Ellen 2008. Prosessikuvauksesta vauhtia taideteollisuusalan ontologia TAO:n ensiaskeleihin. Signum 3, 10-14. <http://pro.tsv.fi/stks/signum/> Verkkolehti. Luettu 19.10.2010.

Kauppa- ja teollisuusministeriö 2007. Luovien alojen yrittäjyyden kehittämisstrategia 2015. KTM Julkaisuja 10/2007. PDF-dokumentti. http://www.tem.fi/files/19795/Luovat_alat.pdf. Päivitetty 26.6.2008. Luettu 2.9.2010.

KDK 2010a. Kokonaisarkkitehtuuri. Liite B Standardisalkku. 23.3.2010. Kansallinen digitaalinen kirjasto. <http://www.kdk2011.fi/fi/kokonaisarkkitehtuuri/sanasto>. PDF-dokumentti. Luettu 27.11.2010.

KDK 2010b. Kokonaisarkkitehtuuri. Liite C Sanasto. 17.3.2010. Kansallinen digitaalinen kirjasto. <http://www.kdk2011.fi/fi/kokonaisarkkitehtuuri/sanasto>. WWW-dokumentti. Luettu 18.10.2010.

Kettula, Suvi 2009. Semanttisen webin ontologisen tekstiilikäsitteistön kehittäminen ja liittäminen museoiden luettelointitietoihin. Helsingin yliopisto. Kotitalous- ja käsityötieteiden julkaisuja 21. Väitöskirja. PDF-dokumentti. <http://urn.fi/URN:ISBN:978-952-10-5845-5> Päivitetty 26.5.2010. Luettu 25.10.2010.

Kivikäs, Pekka 1995. Kalliomaalaukset. Muinainen kuva-arkisto. Paintings on rock. An ancient picture archive. Atena: Jyväskylä.

Kivikäs, Pekka 1997. Kalliomaalausten kuvamaailma. Atena: Jyväskylä.

Kivikäs, Pekka 2009. Suomen kalliomaalausten merkit. Kalliot, kuvakentät ja kuvamerkit. Atena: Jyväskylä.

Kivikäs, Pekka 2010. Haastattelu 25.8.2010.

Koivunen, Hannele 1995. Kirjasto merkityksen tuottajana tietoyhteiskunnassa. Kirjastopalvelu: Helsinki.

Kotkavirta, Jussi & Nyyssönen, Seppo 2005. Cognito. Johdatus filosofiaan. WSOY: Helsinki.

Koukkunen, Kalevi 1990. Nykysuomen sanakirja. Vierassanojen etymologinen sanakirja. WSOY: Helsinki.

Kuronen, Timo 1995. Virtuaalikirjasto. Tulevaisuuden kirjastojen kehittämiskonsepti. Teoksessa Kuronen, Timo (toim.) Tiikerinloikka. Kirjastot tiedon viidakossa. Kirjastopalvelu: Helsinki, 174-242.

Lahelma, Antti 2008. A touch of red. Archaeological and ethnographic approaches to interpreting Finnish rock paintings. The Finnish antiquarian society: Helsinki. ISKOS 15.

Lassila, Ora & McGuinness, Deborah 2001. The role of frame-based representation on the semantic web. Software Technology Laboratory, Nokia Research Center. Knowledge Systems Laboratory, Stanford University. PDF-dokumentti. <http://www.ep.liu.se/ea/cis/2001/005/cis01005.pdf>. Luettu 21.8.2010.

Metatiedon keruu digitointiprosesseissa ja keskustelua 2010. Muistio 4.1.2010. Viva3-Ephemera metatietoryhmän kokous 15.12.2009.

Muinaisjäännösrekisteri. Astuvansalmi 696010003. Meripihkariipukset 1955. <http://kulttuuriymparisto.nba.fi/netsovellus/rekisteriportaali/portti/default.aspx>. Päivitetty 24.11.2010. Luettu 24.11.2010.

Nordman, Tarja 2002. Viestintä terveydenhuollon organisaatiossa. Teoksessa Torkkola, Sinikka (toim.) Terveysviestintä. Vammala: Tammi, 121 - 149.

Nyman, Mika 2009a. Kansainvälisiä hankkeita ja verkottumismahdollisuuksia. Muistio 12.6.2009. Synapse Computing Oy.

Nyman, Mika 2009b. Sähköpostiviesti 7.7.2009. Asiantuntija. Synapse Computing Oy.

Nyman, Mika 2009c. Tarjous Viva3 ja Ephemera –hankkeiden asiantuntijapalvelusta. Tarjous 18.8.2009. Synapse Computing Oy.

Nyman, Mika 2009d. Sähköpostiviesti 6.10.2009. Asiantuntija. Synapse Computing Oy.

Nyman, Mika 2010a. Sähköpostiviesti 9.6.2010. Asiantuntija. Synapse Computing Oy.

Nyman, Mika 2010b. Sähköpostiviesti 22.6.2010. Asiantuntija. Synapse Computing Oy.

Nyman, Mika 2010c. Sähköpostiviesti Miia Herralalle 28.6.2010. Asiantuntija. Synapse Computing Oy.

- Nyman, Mika 2010d. Tapaaminen 2.8.2010. Asiantuntija. Synapse Computing Oy.
- Nyman, Mika 2011a. Tapaaminen 4.2.2011. Asiantuntija. Synapse Computing Oy.
- Nyman, Mika 2011b. Sähköpostiviesti 21.2.2011. Asiantuntija. Synapse Computing Oy.
- Nyman, Mika 2011c. Tapaaminen 24.2.2011. Asiantuntija. Synapse Computing Oy.
- Nyman, Mika 2011d. Tapaaminen 1.3.2011. Asiantuntija. Synapse Computing Oy.
- Nyman, Mika 2011e. Tuloste CIDOC CRM:n luokista ja perityistä ominaisuuksista. Synapse Computing Oy. WWW-dokumentti. <http://www.synapse-computing.com/resource/bc215100-4621-11e0-a536-70f395cfa66e>
- Nyman, Mika 2011f. Sähköpostiviesti 4.3.2011. Asiantuntija. Synapse Computing Oy.
- Nyman, Mika 2011g. Ontologiadokumentti. Synapse Computing Oy. WWW-dokumentti. <http://www.synapse-computing.com/resource/f1b0680f-4621-11e0-b5ad-70f395cfa66e>
- Opetusministeriö 2006. Yksitoista askelta luovaan Suomeen. Luovuusstrategian loppuraportti. Opetusministeriön julkaisuja 2006:43. PDF-dokumentti. http://www.minedu.fi/OPM/Julkaisut/2006/Yksitoista_askelta_luovaan_Suomeen. Päivitetty 16.2.2010. Luettu 2.9.2010.
- Poli, Roberto 1996. Ontology for knowledge organization. Teoksessa Green, R. (toim.). Knowledge organization and change. Indeks, Frankfurt: 113–319. PDF-dokumentti. <http://robertopoli.co.cc/wp-content/uploads/papers/Poli%201996%20Ontology%20for%20Knowledge%20Organization.pdf>. Päivitetty 20.3.2010. Luettu 25.10.2010.
- Poli, Roberto 2003. Ontology, step by step. Teoksessa Fujita, Hamido & Johannesson, Paul (toim.). New trends in software methodologies, tools and techniques. Proceedings of Lye_W03, IOS Press, 60-70. PDF-dokumentti. <http://robertopoli.co.cc/wp-content/uploads/papers/Poli%202003%20Step%20by%20Step.pdf>. Päivitetty 20.3.2010. Luettu 25.10.2010.
- Ruohotie, Pekka 2000. Oppiminen ja ammatillinen kasvu. WSOY: Helsinki.
- Ruotsalainen, Niko 2008. Ontologioiden oppiminen tekstistä. Jyväskylän yliopisto. Tietojenkäsittelytieteiden laitos. Pro-gradu. PDF-dokumentti. https://jyx.jyu.fi/dspace/bitstream/handle/123456789/18853/URN_NBN_fi_jyu-200808175659.pdf?sequence=1. Päivitetty 21.6.2010. Luettu 30.11.2010.
- Sanastokeskus 2006. Terminologian sanasto. TSK 36. Sanastokeskus TSK: Helsinki. <http://www.tsk.fi/fi/info/TerminologianSanasto.pdf>. Päivitetty 23.8.2007. Luettu 6.11.2010.

Semantic Web Stack 2010. Wikipedia. http://en.wikipedia.org/wiki/Semantic_Web_Stack. Päivitetty 27.8.2010. Luettu 18.10.2010.

Simovaara, Jyrki 2004. Metadata ja tietovarantojen yhteiskäyttö. Teoksessa Muistiorganisaatioiden tietovarannot yhteiskäyttöön. Toteuttamismahdollisuudet ja toimenpidesuositus. KAMUT 2 -yhteishankkeen loppuraportti 26.1.2004. Museovirasto: Helsinki, 49-97.

Staab, Steffen, Schnurr, Hans-Peter, Studern Rudi & Sure York 2000. Knowledge processes and ontologies. Institute AIFB, University of Karlsruhe. <http://www.uni-koblenz.de/~staab/Research/Publications/isystems-knowledgeprocess.pdf>. PDF-dokumentti. Päivitetty 14.10.2004. Luettu 18.12.2010.

Strauss, Anselm & Corbin, Juliet 1998. Basics of qualitative research. Techniques and procedures for developing grounded theory. Sage Publications: London.

Suonuuti, Heidi 2006. Sanastotyön opas. Sanastokeskus TSK ry: Helsinki.

Taskinen, Helena 2007. Suomen kalliomaalausten ja -hakkausten dokumentointimenetelmät. Teoksessa Saarikoski, Helena (toim.). Aurinkopeura 3. Suomen muinaistaideseura: Helsinki, 123-143.

Uotila, Pekka 2009. Luovuus ja projektit. Osaaja.net 2. Verkkolehti. <http://ojs.seamk.fi/index.php/osaaja/article/view/1076/943>. Luettu 2.9.2010.

Vakkari, Mikael 2004. Metadata – soveltamismahdollisuudet kirjasto-, arkisto- ja museosalalla. Teoksessa Vakkari, Mikael, Simovaara, Jyrki & Valanto, Sirkka. Kamut 2 – Muistiorganisaatioiden tietovarannot yhteiskäyttöön. Toteuttamismahdollisuudet ja toimenpidesuositus. KAMUT 2 -yhteishankkeen loppuraportti 26.1.2004. Museovirasto: Helsinki. PDF-dokumentti. <http://www.nba.fi/tiedostot/8c050dc5.pdf>. Päivitetty 26.1.2004. Luettu 30.11.2010.

Vakkari, Pertti 2000. Tiedonhankinnan tukeminen ja informaatiotutkimus. Teoksessa Mäkinen, Ilkka (toim.) Tiedon tie. Johdatus informaatiotutkimukseen. BTJ Kirjastopalvelu: Helsinki, 9-31.

Viva3-Ephemera metatietoryhmän kokous 2009. Muistio 11.6.2009. Kansalliskirjasto. Mikkelin ammattikorkeakoulu. Elinkeinoelämän keskusarkisto. Synapse Computing oy.

Viva3 ja Ephemera –projektien yhteiset työryhmät 2009. Muistio 15.4.2009. Kansalliskirjasto. Mikkelin ammattikorkeakoulu. Elinkeinoelämän keskusarkisto. Synapse Computing oy.

Viva3 projektisuunnitelma 2008. Versio 3.1. Mikkelin ammattikorkeakoulu, informaatio- ja mediatekniikan laitos.

W3C 2004. Resource Description Framework (RDF). Concepts and Abstract Syntax. W3C Recommendation 10 February 2004. WWW-sivusto. <http://www.w3.org/TR/rdf-concepts/> Päivitetty 10.2.2004. Luettu 20.10.2010.

W3C 2009. OWL 2. Web Ontology Language. Document Overview. W3C Recommendation 27 October 2009. <http://www.w3.org/TR/owl2-overview/> Päivitetty 27.10.2009. Luettu 20.10.2010.

TYÖRYHMÄN ESITTELY

Juhani Grönhagen, projektipäällikkö, Mikkelin ammattikorkeakoulu

Viva3-hankkeen projektipäällikkö Juhani Grönhagen on työskennellyt laajasti koko keskus- ja maakunnallisten museoiden tehtäväkentällä. Astuvansalmen osalta hänen vedenalaiset tutkimuksensa ovat merkittäviä. Astuvansalmen kallion edustalla hän on tehnyt ensimmäiset alustavat tarkistussukellukset vuodesta 1985 lähtien. Tällöin Grönhagen toimi Museovirastossa Merihistorian toimistossa sukeltavana tutkijana. vuosina 1988, 1990 -1993 hän on tehnyt Savonlinnan maakuntamuseon ja Ristiinan kunnan rahoittamia vedenalaisia kaivauksia Astuvansalmella. Näiden kaivauksien tuloksena löytyivät ainutlaatuiset meripihkariipukset Astuvan Ukko, Akka, Poika ja Karhu. Grönhagen on toiminut lisäksi mm. Suomen Merihistorian seuran hallituksessa, Suomen meriarkeologisessa seurassa hallituksen puheenjohtaja sekä ollut perustamassa Suomen Muinaistaideseuraa, jossa hän on toiminut hallituksen puheenjohtajana. Hän on kirjoittanut useita artikkeleita koskien Astuvansalmen vedenalaisia löytöjä. Grönhagen on toiminut ontologisointityössä määrittelyvaiheessa ja kalliotaiteen asiantuntijana. (Grönhagen 2010.)

Mika Nyman, asiantuntija, Synapse Computing Oy

Mika Nyman on hankkeessa ulkopuolisena asiantuntijana ja suunnannut työskentelyä ja sen vaiheita mm. ontologisointiin ja CIDOC CRM –standardiin liittyen. Nyman on osallistunut Viva3-hankkeessa Astuvansalmen kallion 3D-mallinnuksen tunnistepalvelun luomiseen, jota tämä opinnäytetyö tukee käsitteellisesti. Nymanin osuutena on ollut myös käyttöliittymän tekeminen ontologisoinnin perusteella tehdyn metatietoskeeman pohjalta. Nyman on jäsenenä Unescon alaisen kansainvälisen museoneuvoston ICOM:in dokumentaatiokomitean CIDOC:in työryhmässä, joka kehittää ja ylläpitää museoalan CIDOC CRM -standardia. Nyman on lisäksi toiminut IFLAn (International Federation of Library Associations and Institutions) työryhmässä (FRBR Review Group, Working Group on FRBR/CRM Dialogue), jonka tehtävänä on kirjastoalan tietojärjestelmien toiminnallisen määrittelyn FRBR:n (Functional Requirements for Bibliographic Records) uudistaminen CIDOC CRM:n esikuvan mukaisesti. Nyman toimii puheenjohtajana CIDOC:in Co-reference-työryhmässä. Nymanin yritys Synapse Computing Oy on erikoistunut kulttuuriperintöalan

tietojärjestelmiin. Nymanin kansainvälinen ja kansallinen yhteistyötausta ja aktiivinen kontaktien luonti on tuonut Viva3-hankkeelle uusia yhteistyömahdollisuuksia –ja kumppanuuksia. (Nyman 2009a.)

Miia Herrala, projektityöntekijä, Mikkelin ammattikorkeakoulu

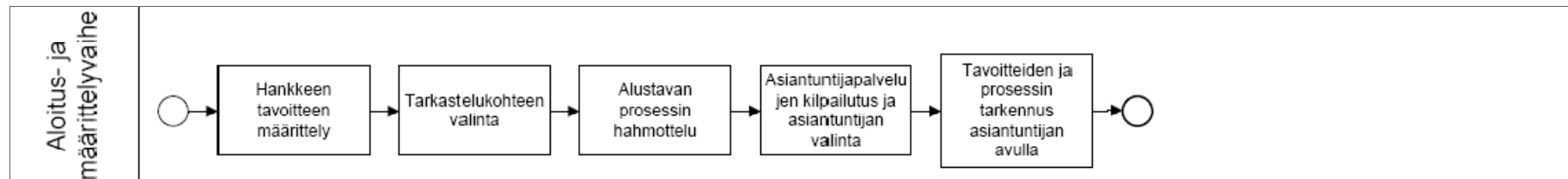
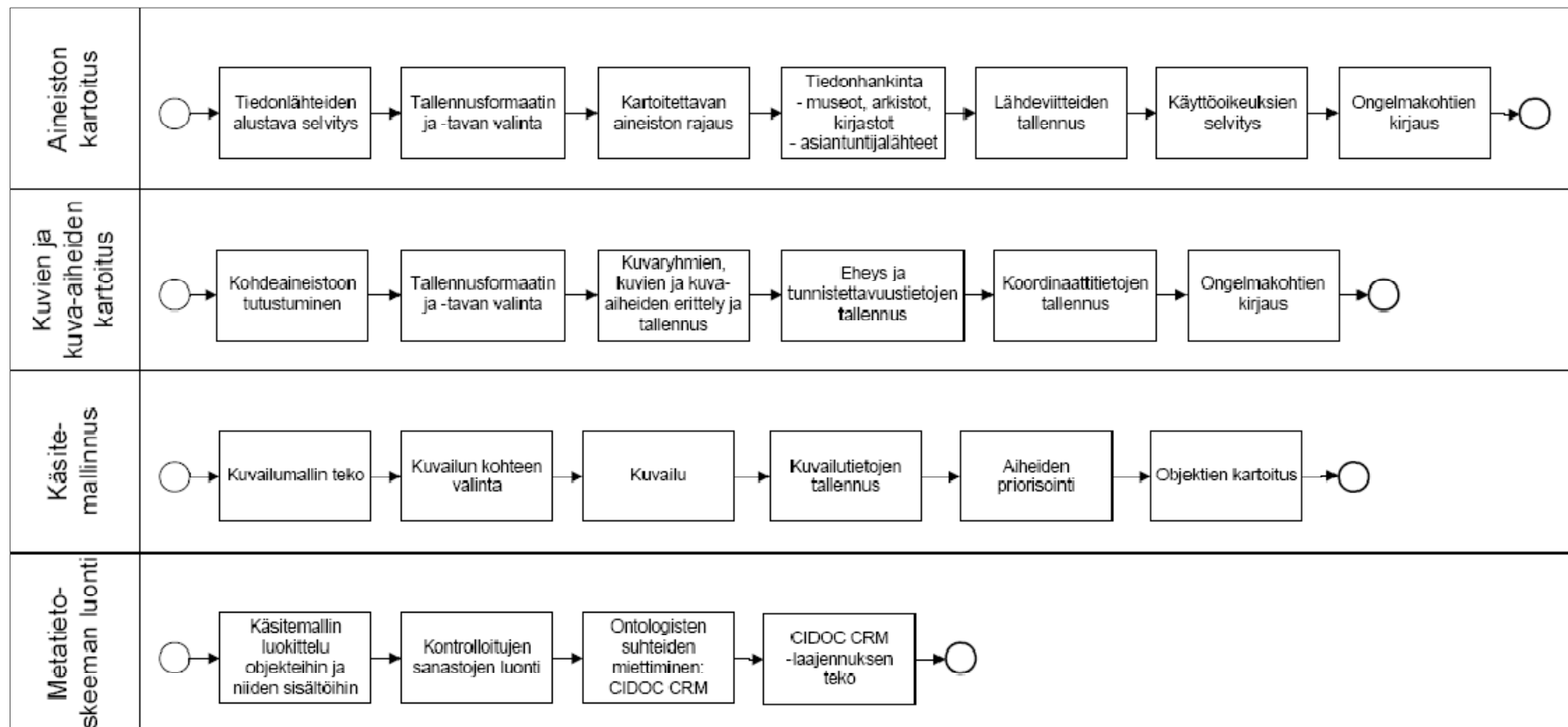
Viva3 –hankkeen projektityöntekijä Miia Herrala on vastannut metatietoselvityksen tekemisestä sekä pitkälti käytännön työstä mm. CIDOC CRM:n osalta. Herralan tavoitteena hankkeessa on ollut tutustua ja oppia käyttämään CIDOC CRM:ää ja ylipäätään ymmärtää kuinka käsitelmalle tehdään. Hänellä on humanististen tieteiden kandidaatin tutkinto kulttuurintutkimuksen alalta. (Herrala 2010.)

Henna Mölsä, opiskelija, Mikkelin ammattikorkeakoulu

Itse olen työskennellyt pääsääntöisesti ammattikorkeakoulukirjastoissa. Olen valmistunut tradenomiksi Turun ammattikorkeakoulusta Kirjasto- ja tietopalvelun koulutusohjelmasta sekä opiskellut Sähköisen aineiston hallinnan ja arkistoinnin erikoistumisopinnot Mikkelin ammattikorkeakoulussa. Erikoistumisopinnot sisälsivät myös arkistonhoitajan tutkinnon. Valokuvausta olen harrastanut vuodesta 2002 lähtien. Erikoistumisopintojen ja ylemmän amk-tutkinnon opintojen lisäksi sekä valokuvaus- että taideharrastukset ovat olleet hyödyllisiä tätä työtä tehdessäni. Ontologisointiprosessissa olen vastannut aineiston keruusta, kuva-aiheiden määrittelystä sekä käsitelmän että metatietoskeeman tekemisestä.

VAATIMUSMÄÄRITTELY: Astuvansalmen kallion ontologisointi

Astuvansalmen kallion ontologisointi
2010
Ontologian tavoite Kalliomaalausten esittäminen ontologian avulla huomioiden ympäristö
Kohde ja ala Lähtökohtana aineistolähtöisyys, tapahtumalähtöisyys ja kalliomaalausten piirteiden erottelu toisistaan ontologian avulla. Tapauskohteena Astuvansalmen kallio ja sen kalliokuvat.
Ontologisointimalli CIDOC CRM
Tietolähteet Museovirasto Juhani Grönhagen, Antti Lahelma, Pekka Kivikäs, Ismo Luukkonen, Mika Nyman www-sivustot aihetta koskien (esim. Muinaistaideseura, Antti Lahelman kotisivut jne.) artikkelit, kirjat (tietokantoina mm. Arto, Aleks, Linda, Mikki, Savotta) Astuvansalmen kallio, kalliomaalaukset ja esineet
Käyttäjät ja käyttötapaukset Käyttäjät: kaikki kalliomaalauksista kiinnostuneet Käyttötapaukset: [ei määritelty]

TOIMINTAMALLI: Aloitus- ja määrittelyvaihe**TOIMINTAMALLI: Toteutusvaihe**

**KUVAILUMALLIN MUKAINEN ESIMERKKIKUVAILU:
D-alueen yläosa (Viva3 kuva nro 8489)
29.6.2010, versio 30.1.2011**

Piirteet:

Kallion piirteet:

- usean värisiä kivimateriaaleja [onko materiaali fyysistä kontekstia?]
- kiviaineksen nimi
- luonnon tekemä piirre: kallion halkeamat, suojaava kalliolippa ja maalauksia suojaava kerrostuma
- onko pintakasvillisuutta poistettu jossain vaiheessa esim. sammaleet, jäkälät? [toisaalta nämä ovat fyysisiä objekteja]
- ihmiskasvoisuus [toisaalta fyysisten objektien ja fyysisten kontekstien piiriin]

Maalauksien piirteet:

- ihmisen tekemät piirteet:
 - väriaine
 - yksittäisen maalauksen koko
 - viivojen paksuus [on toisaalta fyysisten maalausten piirre tai ominaisuus]
- maalauksien sijainti suhteessa kallioon [on toisaalta maalaus objektina ja osana kalliota/kallion erityispiirteitä]
- pohjamateriaali
- muut erityispiirteet

Fyysiset objektit ja niiden osat

- kallio fyysisenä objektina
- maalaus fyysisenä objektina ja osana kalliota
- fyysinen objekti esittää mm. 2x hirvi/peura, kaari, ”ihminen, joka ei näy kuvassa nro 8489”

KÄSITEMALLIESIMERKKI:**Kallion piirteet ja fyysinen konteksti**

<i>metatietoskeemaan otettavat kentät</i>
<i>metatietoskeemaan mahdollisesti otettavat kentät</i>
<i>ei metatietoskeemaan mukaan</i>

KALLIO	CIDOC-CRM
kallio	E19 Physical Object
Kallion piirteet	
kivilaji	E57 Material
materiaalin (kiven) väri	
suojaavan kerroksen materiaali (piioksidi)	E57 Material
suojaavan kerroksen olemassaolo	
pintakasvillisuus ja ympäröivä kasvillisuus	
Kallion fyysinen konteksti (laajemmat kokonaisuudet, struktuurit)	
sijainti:	E27 Site, E53 Place
paikannimi	E48 Place Name
koordinaatit	E47 Spatial Coordinates
käytettävä koordinaatisto	
ilmansuunta, johon päin kallioseinä	
sijainti (mitä kallion ympärillä, vesireitit, metsästysmaiden ja asuinsijojen läheisyys)	
luonnonmuutokset (mm. vedennousu)	E5 Event
veden korkeusarvo	
puiden kasvaminen kallion/maalauksien eteen	
ulkonäkö (halkeamat, kalliolippa, suojaava kerros)	
kaiku	
kunta	
syntymäaika	E5 Event

METATIETOSKEEMAESIMERKKI excel versio 2.0

<i>metatietoskeemaan otettavat kentät</i>
<i>metatietoskeemassa ratkaistavat ongelmat</i>

SKEEMA

kentän nimi
KALLION KUVAILU
kalliomaalauspaikan nimi
kunta
kallion luona tapahtuneet luonnonmuutokset
kallion muoto
kulttuurinen toiminta liittyen kallioon???
kallioon liittyvät tutkimukset (?)
KALLIOMAALAUKSEN KUVAILU
kuvaus kalliomaalauksen sijainnista kuva-alueella ja suhteessa muihin kuvan kuviin
kalliomaalauksen kuva-aihe (selkeät ja epäselvät)
kalliomaalauksetapahtuma

METATIETOSKEEMAESIMERKKI word versio 2.0

(TYPE) = kentän arvo, jokin auktorisoitu lähde, ”käsitteistö”, ”pudostusvalikko”, ei kuulu varsinaiseen metatietoskeemaan
poistetaanko?

[KALLION KUVAILU:]

E18 Physical Thing (NIMI)

E27 Site (PAIKKA)

(TYPE: Kunnat 2010 (ONKI))

maa

maakunta

kunta

kylä

kaupunki

kaupunginosa

E47 Spatial Coordinates (KOORDINAATIT)

KOORDINAATTIJÄRJESTELMÄ

(TYPE: SUO (ONKI))

SAAVUTETTAVUUS (esim. tulo-ohjeet)

MUOTO (sanallinen kuvaus vai käsitteistöllä?)

(TYPE:käsitteistö puuttuu)

kallion jyrkkyys

suojaava kalliolippa

kallion pinnan tasaisuus

kallion ihmismäinen tai muutoin epätavallinen muoto

YMPÄRÖIVÄ LUONTO (sanallinen kuvaus???)

E53 Place (PAIKKA)

E48 Place Name (paikan nimi)

E44 Place Appellation (paikan vaihtoehtoinen nimi)

MUISTILISTA ASTUVANSALMEN KALLION JA SEN KALLIOKUVIEN AINEISTOLÄHTÖISESSÄ ONTOLOGISOINNISSA HUOMIOITAVISTA ASIOISTA

Aloitus- ja määrittelyvaihe

- tavoitteiden määrittely
- tarkastelukohteen valinta
- yhteisen sanaston luominen
- käyttäjätarpeiden selvitys
- vaatimusmäärittely

Toteutusvaihe

Aineiston kartoitus

- mitä aineistotyyppettä kartoitetaan (lehtiartikkelit, kirjat, kuvat jne.)
- aineistojen ja mahdollisesti niiden viitetietojen tallennusmuodon valinta
- viitetietojen tarkkuuden määrittely
- aineistonvalintakriteerit
 - pitääkö olla isompia kokonaisuuksia vai riittääkö maininta
 - kelpaako yleiskirjallisuus
 - lähteen luotettavuuden määrittely (onko merkitystä)
- aineistoon perehtyminen

Kuvien ja kuva-aiheiden kartoitus

- millä tarkkuudella kuvia tulkitaan
 - käsien ja jalkojen asento
 - kulkusuunta
 - ääri viivojen paksuus
 - hirven sarvien piikkien määrä / veneiden ihmisten määrä
- koon määrittely: riittääkö isokokoinen vai cm
- kuvataanko ryhmän sisällä olevien kohteiden keskinäistä suhdetta ja/tai monikerroksellisuutta
 - päällekkäisyys
 - järjestys
- epäselvyyden ilmaisutapa ja tarve
- vaihtoehtoisten tulkintojen ilmaiseminen
- onko merkittävää, että jotain puuttuu (olemattoman tiedon merkitys)
 - esim. ihminen ilman päätä
 - esim. kuva-aiheista puuttuu jotain tyypillistä, kuten palvovat ihmishahmot eli adorantit
- onko muutoshistoria merkittävä
 - esim. suorasta jalasta tehty koukistettu jalka

Mallintaminen

- tallennusmuodon valinta
- palvelimen, ohjelmistojen ja tekniikoiden määrittely ja tarvittaessa hankinta
- CIDOC CRM
 - miten moniperiytyvyys toimii
 - alussa ei tarvitse miettiä ominaisuuksia
 - periytyminen näkyy vasta ominaisuuksissa
- CIDOC CRM:n laajennustarpeen määrittely

Arviointivaihe

- ei irrallinen toteutusvaiheesta -> jatkuva arviointi ja kehittäminen toistamalla vaiheita
- projektin päämääristä tavoitteita, jotka mitattavissa joko määrällisesti tai laadullisesti
- vertaaminen alussa tehtyyn vaatimusmäärittelydokumenttiin
- käsitellin arviointi
 - tulkintojen näkyminen
- metatietoskeeman arviointi aineistolähtöisesti
 - mihin tapahtumiin skeeman elementit kytkeytyvät
 - kuinka hyvin ominaisuudet ovat periytyneet ja suhteet vastaavat aineistoa
- testaus halutussa sovelluksessa
 - toteutuuko moniperiytyminen
- esimerkit miten mallia voi käyttää
- käyttökokemusten keräys ja hyödyntäminen
- kolme ontologista kenttää riittävä Astuvansalmen tapauksessa
- jälkikäteisarviointi
- sekä prosessin että lopputuloksen arviointi

[Ylläpitovaihe – ei Astuvansalmen kallion ontologisoinnissa mukana]

- käyttäjiltä tulleen palauttaa perusteella arviointi ja kehittäminen tarvittaessa
- voimassa niin kauan kuin ontologiaa käytetään