



## **Majandus- ja Kommunikatsiooniministeerium Riigi infosüsteemide osakond**

### **Riigi infosüsteemide semantilise koosvõime raamistik Versioon 0.7 (01.08.2007)**

Käesolev raamistiku dokument töötati välja ekspertide poolt koosöös riigi- ja kohaliku omavalitsuse ja erasektori asutustega. Raamistik on läbinud avaliku konsultatsiooniperioodi ja kooskõlastusringi avaliku sektori asutustega (Majandus- ja Kommunikatsiooniministri kiri 19.06.2007 nr 22:-1/5888 ministeeriumitele, põhiseaduslikele institutsioonidele, Eesti Maaomavalitsuste Liidule, Eesti Linnade Liidule, Riigi Infosüsteemide Arendukeskusele).

Raamistiku sisu vaadatakse läbi ja vajadusel uuendatakse igal aastal. Uutes versioonides arvestatakse eelmisel perioodil laekunud ettepanekuid. „Riigi IT koosvõime raamistik“ ja temast tulenevad dokumendid on allalaaditavad veebiaadressilt <http://www.riso.ee/et/koosvoime/>. Käesolev dokument on avatud ettepanekuteks avaliku, era ja kolmanda sektori asutustele ning kõigile asjast huvitatud isikutele. Ettepanekud palume saata e-kirjana aadressile [koosvoime@riso.ee](mailto:koosvoime@riso.ee).

# SISUKORD

Kokkuvõte.....	3
1 Sissejuhatus.....	4
1.1 Raamistiku mõiste ja eesmärk.....	4
1.2 Semantilise koosvõime võtmepõhimõtted.....	5
1.3 Semantiliste varade rakendamine.....	6
2 Raamistikku mõjutavad regulatsioonid.....	7
2.1 Seadusandlus.....	7
2.1.1 Andmekogude seadus.....	7
2.1.2 Avaliku teabe seadus.....	7
2.1.3 Andmekogude riikliku registri asutamine.....	7
2.1.4 Arhiiviseadus.....	8
2.1.5 Infosüsteemide andmevahetuskihi rakendamine.....	8
2.1.6 Klassifikaatorite süsteemi kehtestamine.....	8
2.1.7 Aadressandmete süsteemi kehtestamine.....	8
2.1.8 Geodeetilise süsteemi kehtestamine.....	8
2.1.9 Infosüsteemide turvameetmete süsteemi kehtestamine.....	8
2.1.10 Arhiivieeskiri.....	8
2.1.11 Vabariigi Valitsuse seadus.....	8
2.2 Riiklikud prioriteedid.....	9
2.2.1 Eesti infoühiskonna arengukava 2013.....	9
2.3 Puuduolev regulatsioon.....	9
2.3.1 XML-i kasutamist puudutav regulatsioon.....	9
2.3.2 Semantilise koosvõime organisatsiooni regulatsioon.....	10
2.3.3 Riigi infosüsteemi haldussüsteem ja selle haldamine.....	10
3 Semantiline koosvõime.....	10
3.1 Semantilise koosvõime mõiste.....	10
3.2 Semantikavarad.....	11
3.2.1 Valdkonnasõnastikud.....	11
3.2.2 Tesaarused.....	11
3.2.3 Klassifikaatorid.....	11
3.2.4 Taksonoomiad.....	12
3.2.5 Vastavustabelid.....	12
3.2.6 Ontoloogiad.....	12
3.2.7 Teenuste registrid.....	12
3.3 Semantikavarad kui intellektuaalne omand.....	13
3.4 Semantilise koosvõime arhitektuur.....	13
4 Semantilist koosvõimet tagav organisatsioon.....	13
4.1 Osalised.....	14
4.1.1 Tsentraalne organ.....	14
4.1.2 Tugikeskus.....	14
4.1.3 Valdkondade esindusorganid.....	15
4.1.4 Andmeteenuste haldajad.....	15
4.1.5 Valdkondadevahelised töögrupid.....	15
4.2 Semantilise koosvõime tagamise protsessid.....	15
4.2.1 Valdkondade määramine.....	16
4.2.2 Semantikavarade loomine valdkonniti.....	16

<a href="#">4.2.3 Semantiliste varade haldamine valdkonniti.....</a>	<a href="#">18</a>
<a href="#">4.2.4 Valdonnaülene semantiliste varade mestimine.....</a>	<a href="#">19</a>
<a href="#">4.2.5 Semantiliste varade riskasutus ja taaskasutus.....</a>	<a href="#">20</a>
<a href="#">4.2.6 Tööriistade ning keskkondade loomine.....</a>	<a href="#">21</a>
<a href="#">4.2.7 Semantiliste varade publitseerimine.....</a>	<a href="#">21</a>
<a href="#">4.2.8 Nimeruumide haldamine.....</a>	<a href="#">21</a>
<a href="#">5 Toetavad tegevused.....</a>	<a href="#">21</a>
<a href="#">5.1 Semantikavarade haldamise keskkond.....</a>	<a href="#">22</a>
<a href="#">5.2 XML repositoorium.....</a>	<a href="#">22</a>
<a href="#">5.3 XML levikeskus.....</a>	<a href="#">23</a>
<a href="#">Mõistete ja lühendite loetelu.....</a>	<a href="#">23</a>
<a href="#">Kasutatud kirjandus.....</a>	<a href="#">25</a>

## Kokkuvõte

Semantikatehnoloogiatega põhiülesandeks on lähimate aastate jooksul vähendada riigi infosüsteemide integreerimisele kuluvat aega ning luua seeläbi võimalus info kiireks ja asjakohaseks edastamiseks, suurendades sealhulgas valmisolekut kriisisituatsioonide lahendamiseks. Nimetatud eesmärkide täitmine nõuab riigisiseste selgete reeglite ja kokkulepete fikseerimist ja ühiste standardite kasutamist infosüsteemide kirjeldamisel.

Eestis on viimase 4 aasta jooksul rajatud andmevahetuskiht X-tee, millega oli liidestatud 2007. aasta alguse seisuga 69 teenust osutavat infosüsteemi ja ligi 400 riigi inforessurssi kasutatavat infosüsteemi. X-tee teenuste arv ületab 700, millele lisanduvad nn KIT teenused, teavitusteenused, blanketiteenused, infoteenused jms. Seega oli 2007. aasta alguseks suur osa riigi infosüsteemist teenustepõhine. Infosüsteemide teenustepõhisus on oluliseks eeltingimuseks semantilise koosvõime saavutamisel ja tänu selle tingimuse täitmisele on riik nüüd võimeline astuma konkreetseid samme ka infosüsteemide *semantilise koosvõime* saavutamiseks. Esimeste sammudena on tänaseks välja töötatud semantilise koosvõime strateegia ning riigi infosüsteemide semantilise kirjeldamise juhised ja meetodika. Käesolev raamistik on 2005. aastal publitseeritud semantilise koosvõime strateegia dokumendi edasiarendus ning täiendus 2004. aastal avaldatud riigi IT koosvõime raamistiku dokumendile. Lisaks on võetud arvesse Eesti infoühiskonna arengukavas (2007-2013) esitatud põhimõtteid, visiooni ja meetmeid.

Raamistik eeldab riigi kodanikukesksust ning infosüsteemide teenusepõhisust. Eeldatakse, et Eesti Euroopa Liidu ja NATO liikmesriigina tagab semantilise koosvõime teiste liikmesriikidega. Raamistik hõlbustab ettevõtete, riigi- ning kohaliku omavalitsuse asutuste olemasolevate infosüsteemide integreerimist, kaardistamist ja mõjusat info vahetamist. Raamistiku järgimine on kohustuslik riigiasutuste ja kohaliku omavalitsuse asutuste infosüsteemide omavahelisel suhtlemisel, ühisprojektides ning ühiselt kasutatavates infosüsteemides.

Käesolev dokument koostati IT ekspertide poolt koostöös riigi ja kohalike omavalitsuste, erasektori ja kolmanda sektori asutustega. Ekspertide töögrupi korraldas Majandus- ja Kommunikatsiooniministeeriumi riigi infosüsteemide osakond koostöös IT ja konsultatsioonifirmaga OÜ SOA Trader.

# 1 Sissejuhatus

## 1.1 Raamistiku mõiste ja eesmärk

Kui IT koosvõime on infosüsteemide ja tema poolt toetavate tegevusprotsesside võime vahetada andmeid ja ühiselt kasutada informatsiooni ja teadmisi, siis *semantiline* koosvõime on nii infosüsteemide kui arendajate võime üheselt aru saada milliseid andmeid vahetatakse ning millist informatsiooni kasutatakse. Ehk siis teisisõnu on semantiline koosvõime erinevate organisatsioonide võime mõista vahetatud informatsiooni tähendust ühte moodi.

Infosüsteemide semantilise koosvõime raamistik on e-riigi raames kehtestatud standardite ja juhendite kogum, mis tagavad ühese arusaamise infosüsteemidest ning seeläbi avaliku halduse asutuste, eraettevõtete ja kodanike parema teenindamise nii riigi, üleeuroopalises kui ka laiemas rahvusvahelises kontekstis. Infosüsteemide semantilise koosvõime raamistik ja sellest lähtuvad dokumendid on kohustuslikud riigi ja kohalike omavalitsuste asutuste infosüsteemide arendamisel ning haldamisel. Raamistiku dokumente ei saa siiski vaadelda õigusaktidena. Nende kohustuslikkus väljendub järgnevates asjaoludes:

- Raamistik ja temast lähtuvad dokumendid läbivad konsultatsiooniperioodi, mille vältel riigi- ja kohaliku omavalitsuse asutused, erasektor, kolmanda sektori asutused ja eraisikud saavad esitada omi ettepanekuid. Seega on raamistiku dokumendid kohustuslikud kui eri osapoolte vaheline kokkulepe.
- Vastavalt „Vabariigi Valitsuse seadusele“, „Andmekogude seadusele“ ja „Eesti infoühiskonna arengukava 2013“ dokumendile koordineerib Majandus- ja Kommunikatsiooniministeerium riigi infosüsteemi arendamist. Raamistik ja temast tulenevad dokumendid on riigi infosüsteemi põhidokumendid.

Infosüsteemide semantilise koosvõime raamistik arvestab alljärgnevaid dokumente:

- Euroopa Liidu koosvõime raamistik ja sellest tulenevad dokumendid;
- riigi IT koosvõime raamistik;
- riigi IT semantilise koosvõime strateegia;
- Eesti infoühiskonna arengukava 2013.

Infosüsteemide semantilise koosvõime raamistikust lähtuvad alljärgnevaid dokumendid:

- andmekogude ja andmekogude poolt osutatavate teenuste semantilise koosvõime meetoodika;
- andmekogude ja andmekogude poolt osutatavate teenuste semantilise kirjeldamise juhise.

Kui meetoodika ülesandeks on formuleerida juhise kirjeldamise reeglid ning luua selgitav taust, siis juhise on suunatud arendajale ning semantikavarade loojatele.

Infosüsteemide semantilise koosvõime raamistik on:

- juhendmaterjal teenustepõhiste infosüsteemide semantiliseks kirjeldamiseks;
- juhendmaterjal teenustepõhiste infosüsteemide semantiliseks kirjeldamiseks vajalike tööriistade ja keskkondade loomiseks;
- juhendmaterjal semantilise koosvõime tagamiseks oluliste organisatsiooniliste protsesside disainimiseks ning käivitamiseks.

Infosüsteemide semantilise koosvõime raamistiku eesmärk on muuta Eesti avaliku sektori toimimine efektiivsemaks, mõjusamaks ning parendada info kättesaadavust Eesti ja Euroopa Liidu elanikele ja äri sektorile pakutavate teenuste abil. Raamistiku rakendamise konkreetsed eesmärgid on järgnevad:

- lihtsustada Eesti ning Euroopa Liidu kodanikele ja ettevõtetele suunatud informatsiooni leidmist ning rakendamist;
- kiirendada infosüsteemide integreerimise protsessi ning tõsta seeläbi muuhulgas ka valmisolekut kriisisituatsioonide operatiivseks lahendamiseks;
- kaasa aidata ja läbi viia asutuspõhise maailma muutmise/tõlkimise kodaniku jaoks teenusepõhiseks maailmaks, kus kõik isikud saavad oma asjaajamised avaliku sektoriga toimetada ilma, et nad peaksid teadma midagi avaliku sektori hierarhisest struktuurist ja rollide jaotusest selles;
- vähendada tervikuna avaliku sektori IT kulusid keskselt realiseeritud lahenduste laialdase kasutuse kaudu;
- parandada riigi IT süsteemide koordineerimist, juhtimist ja kiirendada uute lahenduste väljatöötamist;
- kaardistada riigi infosüsteemide teenuste ning potentsiaalsete andmevoogude tasemel;
- luua vahendid teadmiste talletamiseks, seostamiseks haldamiseks ning rakendamiseks riiklikes institutsioonides;
- luua mehhanismid semantikavarade ning andmete korduvkasutamiseks;
- tagada riigi infosüsteemide toimimine ühtse tervikuna läbi avatud standardite ning masinloetava semantika;
- koostada avaliku sektori põhiliste dokumendiliikide XML kirjeldused ning täiendada need semantiliste kirjeldustega;
- toetada ekspordipotentsiaali omavate semantika kirjeldamise ja haldamise lahenduste loomist osaledes vastavates projektides pilootkliendina;
- teha riigi infosüsteemide sisu ning ülesanne üheselt arusaadavaks nii tarkvara arendajatele kui vastava valdkonna ekspertidele;
- semantikavarade, kaasaarvatud haldustesauruse, loomine;
- XML repositooriumi ja levikeskuse (*ClearingHouse*) loomine ning selle kaudu (ka semantiliste) andmekirjelduste standardimise ja korduvkasutuse tagamine.

Käesolev dokument ei käsitle semantilise koosvõime tagamiseks vajalikku metoodikat ega juhiseid, mis on esitatud eraldi dokumentidena. Raamistiku eelnev versioon avaldati 2005. aastal riigi IT semantilise koosvõime strateegia nime all. Käesolev versioon on eelmise ümbertöötlus ning täiendus.

## **1.2 Semantilise koosvõime võtmepõhimõtted**

- Riigi infosüsteemide semantilisel kirjeldamisel rakendatakse järgnevaid keeli ning nende laiendusi: XSD, OWL, WSDL, SA-WSDL, WS-BPEL.
- Riigi infosüsteemide nii inim- kui masinloetavad semantilised kirjeldused on tasuta kättesaadavad nii avaliku ja erasektori esindajatele kui kõigile EL-i kodanikele.
- Iga ministeerium on kohustatud looma ning seejärel haldama oma valitsemisalaga seotud ontoloogiaid, valdkonnasõnastikke, nendega seotud veebiteenuste semantilisi kirjeldusi ning muid semantilisi varasid.

- Ontoloogiatega ja valdkonnasõnastike valdkonniti kattuvate osade loomine ning haldamine toimub meeskonnatöö vormis, kus osalisteks on erinevate valdkondade eksperdid ning meeskonnatöö toetamiseks kasutatakse vastavat Internetipõhist keskkonda.
- Kõik institutsioonid on vabad oma infosüsteemide sisemise arhitektuuri ja koosvõime põhimõtete valikul, kuid nende omavahelisel seostamisel on kohustuslik on järgida raamistiku põhimõtteid (nn subsidiaarsuse põhimõte).
- Infosüsteemide semantiline koosvõime Eestis on strateegiline valik, kusjuures avalik sektor näitab semantilise koosvõime põhimõtete järgimisel eeskuju.
- Semantiliste varade loomiseks ja haldamiseks vajalike tehnoloogiate väljatöötamisele suunatud tegevused seostatakse teadus- ja arendustegevusega Eestis.
- Semantiliste varade loomisel ja haldamisel arvestatakse korduvkasutatavuse põhimõtteid.
- Ka erasektorile tagatakse ligipääs infosüsteemide semantilise kirjeldamise keskkonnale ning võimalus kirjeldada semantiliselt omi infosüsteeme sarnastel alustel mil avaliku sektori esindajad riigi infosüsteeme.
- Semantiline koosvõime toetab e-Äri lahenduste loomist, vajadusel ettevõtete infosüsteemide liidestamist riigi infosüsteemidega ning riigi poolt kodanikele pakutavate teenuste personaliseerimist.
- Riik soosib eraettevõtete ja nende ühenduste kaasatust semantiliste varade loomisel, publitseerimisel ja rakendamisel.

### **1.3 Semantiliste varade rakendamine**

Semantiliste varade otsesteks rakendusteks on peamiselt spetsialiseeritud otsingumootorid. Vastavaid otsingumootoreid rakendavad esialgu peamiselt infosüsteemide arendajad:

- vajalike andmeteenuste ning –baaside otsimisel, integreerimisel ja rakendamisel,
- infosüsteemide ja teenuste arenduse ja halduse toetamisel.

Ametnikud ning kodanikud rakendaksid vastavaid otsimootoreid:

- dokumendivormide otsimisel,
- nutikal otsimisel riigi veebis läbi spetsiaalsete veebirakenduste.

Tulevikku suunatud rakendused võiksid olla järgnevad:

- digitaalse kultuuripärandi säilitamine ja rakendamine,
- otsustustööriistad ametnikele,
- teadmuse haldamine ja kommunikatsiooni juhtimine riigi- ja kohalike omavalitsuste asutustes,
- intelligentne sisuhaldus veebi jaoks,
- masintõlke teostamine,
- riigi infosüsteemide komponentide ning nende vaheliste seoste analüüs,
- andmeteenuste poolautomaatne orkestreerimine ning seeläbi kompleksteenuste kiire loomine,
- uute teenuste loomise võimaldamine ärianalüütiku tasemega spetsialistile.

Semantiliste varade rakendamisel:

- paraneb andmete kvaliteet; eri allikatest pärit andmeid saab omavahel integreerida, vähenevad vead ja ebatäpsused andmete kasutamisel ja otsuste tegemisel,
- vähenevad investeeringud andmete tootmises, sest andmeid on hõlpsam korduvkasutada,
- väheneb organisatsioonide infosüsteemide integreerimise aeg,
- paraneb infosüsteemideätkusuutlikus ja lihtsustub edasiarendus. Infosüsteeme puudutav teadmiste baas säilib organisatsioonis ka peale võtmeisikute lahkumist.

## 2 Raamistikku mõjutavad regulatsioonid

### 2.1 Seadusandlus

Kuigi enamike semantilise koosvõime aspektide reguleerimine toimub mitmesuguste kokkulepete, standardite või soovitustega, siis semantilise koosvõime organisatsiooni on mõistlik protsessi tõhustamiseks koordineerida vastavate seaduste ja määrustega. Kuna hetkel otseselt semantilist koosvõimet puudutavad ning reguleerivad õigusaktid puuduvad, siis on käesolevas jaotises ära toodud loetelu vaid otseselt digitaalteavet puudutavatest ja reguleerivatest õigusaktidest. Viimane on tingitud asjaolust, et semantilise koosvõime tagamine tähendab endast sisuliselt olemasolevate digitaalsete ressursside rikastamist.

#### 2.1.1 Andmekogude seadus

(RT I 1997, 28, 423; terviktekst <https://www.riigiteataja.ee/ert/act.jsp?id=745339>) Seadus sätestab riigi ja kohaliku omavalitsuse andmekogu valdamise ja kasutamise korra ning riigi, kohaliku omavalitsuse ja eraõigusliku isiku andmekogu pidamise, andmete väljastamise ja andmete kasutamise üldalused. Kehtiv andmekogude seadus asetab rõhu küll andmekogude asutamisele ja sätestab vastava korra, kuid jätab vaatluse alt välja olulised infosüsteemi komponendid – andmekogude kasutamise ja teenused ehk *front-office*'i reguleerimise.

#### 2.1.2 Avaliku teabe seadus

(RT I 2000, 92, 597; terviktekst <https://www.riigiteataja.ee/ert/act.jsp?id=921835>) Seaduse sätted tagavad üldiseks kasutamiseks mõeldud teabele avalikkuse ja igaühe juurdepääsu võimaluse, lähtudes demokraatliku ja sotsiaalse õigusriigi ning avatud ühiskonna põhimõtetest, ning loob võimalused avalikkuse kontrolliks avalike ülesannete täitmise üle.

Alates 01.01.2008 hakkab kehtima avaliku teabe seaduse redaktsioon (<https://www.riigiteataja.ee/ert/act.jsp?id=12790561>, millesse on lisatud ka andmekogusid reguleerivad sätted. Seejuures muutub alates 2008. aastast kehtetuks andmekogude seadus.

#### 2.1.3 Andmekogude riikliku registri asutamine

(RT I 1998, 63, 997; terviktekst <https://www.riigiteataja.ee/ert/act.jsp?id=769739>) Määruses on sätestatud andmekogude riikliku registri pidamine, registrisse kantavate andmete koosseis, registrisse andmete esitamise ja nende kasutamise kord, registri pidamise finantseerimise, registri pidamise üle järelevalve teostamise ning registri likvideerimise kord. Registri pidamise eesmärk on andmekogude üle arvestuse pidamine, Vabariigi Valitsusele ja andmekogude vastutavatele töötlejatele ning riigi infosüsteemidealaseid töid koordineerivale

asutusele ettepanekute tegemine sarnaste ja teineteist sisuliselt kordavate andmekogude pidamise vältimiseks, andmekogu laiendamiseks, ühendamiseks või likvideerimiseks, andmete riskkasutuseks ning andmetöötluse või andmehõive korrastamiseks.

#### **2.1.4 Arhiiviseadus**

(RT I 1998, 36/37, 552; täistekst <https://www.riigiteataja.ee/ert/act.jsp?id=742357> )  
Seadus sätestab arhivaalide kogumise, hindamise, arhiveerimise, säilitamise ja nendele juurdepääsu korraldamise ning arhiivide tegevuse alused.

#### **2.1.5 Infosüsteemide andmevahetuskihi rakendamine**

(VV, RTI, 23.12.2003, 83, 568; täistekst <https://www.riigiteataja.ee/ert/act.jsp?id=688079> )  
Määrusega kehtestatakse ühtsed põhimõtted infosüsteemide andmevahetuskihi haldamiseks ja toimimiseks.

#### **2.1.6 Klassifikaatorite süsteemi kehtestamine**

(VV, RTI, 15.10.2003, 65, 435; <https://www.riigiteataja.ee/ert/act.jsp?id=632688>)  
Määrus sätestab klassifikaatorite haldamise ja kasutamise ühtsed põhimõtted riigi- ja kohaliku omavalitsuse asutuste andmekogudes.

#### **2.1.7 Aadressandmete süsteemi kehtestamine**

(VV, RTI, 17.12.2003, 77, 520; <https://www.riigiteataja.ee/ert/act.jsp?id=681650> )  
Määrus sätestab aadressandmete töötlemise ja kasutamise ning aadressiteenuste ühtsed põhimõtted.

#### **2.1.8 Geodeetilise süsteemi kehtestamine**

(VV, RTI, 18.02.2004, 17, 267; <https://www.riigiteataja.ee/ert/act.jsp?id=707092> )  
Määrusega kehtestatakse geodeetiline süsteem.

#### **2.1.9 Infosüsteemide turvameetmete süsteemi kehtestamine**

<https://www.riigiteataja.ee/ert/act.jsp?id=791875>

Määrusega kehtestatakse riigi ja kohaliku omavalitsuse andmekogude pidamisel kasutatavate infosüsteemide ning nendega seotud infovarade turvameetmete süsteem. Turvameetmete süsteemi rakendamine seisneb infoturbe eesmärkidele vastavate turvaklasside määramises ja nende vastavate turvameetmete valimises vastavalt infosüsteemide kolmeastmelise etalonturbe süsteemi (edaspidi *ISKE*) rakendamisjuhendile ja nende rakendamises.

#### **2.1.10 Arhiivieeskiri**

(RT I 1998, 118/120, 1904 <https://www.riigiteataja.ee/ert/act.jsp?id=544218> ja VV, RTI, 11.03.2003, 26, 162 <https://www.riigiteataja.ee/ert/act.jsp?id=263235>)

Arhiivieeskiri sätestab dokumentide säilitustähtaegade kehtestamise, arhivaalide hindamise, hävitamiseks või arhiivi üleandmiseks eraldamise, arhiveerimise, säilitamise, kaitse, kasutamise ja avalike arhiivide poolt arhivaalide kogumise nõuded ning korra digitaalarhivaalide kohta.

#### **2.1.11 Vabariigi Valitsuse seadus**

<https://www.riigiteataja.ee/ert/act.jsp?id=12788414>



Määrab ministriumide valitsemisalad, mida omakorda rakendatakse semantiliste varade haldamisel valdkondade määramisel.

## **2.2 Riiklikud prioriteedid**

### **2.2.1 Eesti infoühiskonna arengukava 2013**

[http://www.riso.ee/et/files/Infoyhiskonna\\_arengukava\\_2013.pdf](http://www.riso.ee/et/files/Infoyhiskonna_arengukava_2013.pdf)

„Eesti infoühiskonna arengukava 2013” on Vabariigi Valitsuse poolt kinnitatav valdkondlik arengukava, mis sätestab infoühiskonna arendamise üldise raamistiku ja seab eesmärgid ning määratleb neist tulenevad tegevusvaldkonnad IKT laialdaseks kasutamiseks teadmispõhise majanduse ning -ühiskonna arendamisel Eestis aastail 2007-2013. EL liikmesriigina on strateegia väljatöötamisel arvesse võetud 2005. aastal heakskiidetud infoühiskonna raamdokumenti i2010 ja 2006. aastal heakskiidetud EL-i eValitsuse tegevuskava. Infoühiskonna arengukava raames kavandatud tegevused on kooskõlas „Eesti majanduskasvu ja tööhõive tegevuskava 2005-2007” ja „Struktuurivahendite kasutamise strateegia 2007-2013” prioriteetidega.

## **2.3 Puuduolev regulatsioon**

### **2.3.1 XML-i kasutamist puudutav regulatsioon**

Kuna enamik semantilise koosvõime realiseerimiseks vajalikke tehnoloogiaid baseerub XML-il, siis, tagamaks semantilise koosvõime maksimaalne mõjus ja semantiliste varade taaskasutus, vajab esmalt korrastamist XML-i puudutav regulatsioon. Hetkel on XML-i kasutuselevõtuks vajalik õiguslik regulatsioon küll ulatuslik, kuid pole sisuliselt ühtlustatud. Määratlemata on XML-i kasutus dokumentide ja andmekogude osas ning nende seostes. Oluliselt raskendab ühtse regulatsiooni aluse määratlemist see, et XML-i on vaja rakendada nii andmekogudes, andmevahetuses kui dokumentide kasutamisel. Andmekogude valdkonda reguleerib Andmekogude seadus (8. 04. 2004 ( RT I 2004, 30, 204) 28. 04. 2004), milles on muuhulgas kirjas, et andmekogude pidamist kindlustav süsteem on infosüsteemide andmevahetuskiht (§53 lg 1 p 5), milleks on hetkel X-tee.

Kuna tehnoloogilisi protsesside reguleerimise võimalused õigusaktidega on piiratud, siis kehtestatakse vajalikud regulatsioonid semantilise koosvõime raamistikuga ja temast tulenevate dokumentidega.

Olukord on veelgi komplitseeritum asjaajamise ja dokumendihalduse valdkonnas. Aastast 2002 seisab „Asjaajamise ja arhiivinduse arengukavas aastateks 2002–2005“, et „olukord, kus volitusnorm asjaajamiskorra ühtsete aluste kehtestamiseks sisaldub eriseadustes, mis seonduvad ajaajamisega ainult selle seaduse reguleerimise objektiks oleva valdkonna ulatuses, ei soodusta ühtse ja stabiilselt areneva asjaajamissüsteemi kujunemist riigi avalikus sektoris. Asjaajamise õigusliku reguleerimise keerukus raskendab avaliku sektori inforessursi haldamist, sealhulgas asjaajamise arengut ja üleminekut elektroonilisele asjaajamisele.“ See tähendab, et puudub dokumendihalduse üldregulatsioon ning volitusnorm tuleneb mitmest seadusest.

Kuna ka dokumendisüsteemide tehnoloogilisi protsesside reguleerimise võimalused õigusaktidega on piiratud, siis kehtestatakse vajalikud regulatsioonid dokumendisüsteemide koosvõime raamistikuga ja temast tulenevate dokumentidega.

### **2.3.2 Semantilise koosvõime organisatsiooni regulatsioon**

Tagamaks semantilise koosvõime organisatsiooni jätkusuutlikkus ning tõhusus on semantilise koosvõime raamistikus fikseeritud semantilise koosvõime protsessid Raamistikus esitatud regulatsioon määrab konkreetsetes protsessides osalejad, nende rollid, kohustused ning õigused. Raamistiku esmaseks ülesandeks on tagada RIHA-s registreeritud andmeteense semantiline kirjeldamine ning järk-järgult toimuks sama protsessi laiendamine teisele semantilistele varadele.

Praeguses versioonis puuduvad veel nõuded riigi infosüsteemide ja XML skeemide semantiliste kirjelduste ning muude semantiliste varade auditeerimisele. Hetkel on auditeerimise nõuded liialt üldised ning võrdlemisi kasutatud semantilise koosvõime tagamise seisukohalt võttes.

### **2.3.3 Riigi infosüsteemi haldussüsteem ja selle haldamine**

Hetkel on väljatöötamisel RIHA põhimääruse eelnõu, mis asendab andmekogude riikliku registri määruse. Määrusega kehtestatakse riigi infosüsteemi haldussüsteem ning sätestatakse infosüsteemide kooskõlastamise ja registreerimise tingimused ja kord riigi infosüsteemi haldussüsteemis. Kui RIHA saab semantiliste varade publitseerimise keskkonnaks, siis peab vastav määrus käsitlema ka semantiliste varade publitseerimist.

## **3 Semantiline koosvõime**

### **3.1 Semantilise koosvõime mõiste**

Semantilise koosvõime all mõistame infosüsteemide võimet teistelt infosüsteemidelt saadud andmeid adekvaatselt kasutada. Ehk siis teiste sõnadega, semantilise koosvõime tagamine eeldab infosüsteemide inim- ja masinloetava semantika olemasolu. Koosvõime muudab keerukaks asjaolu, et tarkvarasüsteemide kasutusviisid, eesmärgid ja kontekstid on erinevad, seega on erinevad ka andmete esitusviisid, kodeerimine ja tähendusnüansid.

Semantilise koosvõime saavutamiseks praktikas tuleb selgelt määratleda milliseid infosüsteemi komponente on üldse mõistlik semantiliselt kirjeldada. Arvestades, et infosüsteemide semantika loomine ja haldamine on võrdlemisi kallis protsess, orienteerutakse raamistikus kõige laialdasemalt kasutatavatele komponentidele, milleks hetkel on X-teega liitunud infosüsteemid. Semantilise koosvõime saavutamise küsimusele lähenetakse kui ülesandele lihtsustada tarkvarasüsteemide arhitektide ja arendajate tööd, kes peavad realselt ehitama liideseid teiste tarkvarasüsteemidega.

Koosvõime saavutamise teed on suuresti organisatoorse, sotsiaalse ja haridusliku laadi. Esmajoones kavandatakse abistada eri valdkondade süsteemispetsialiste üksteise valdkondade vastastikusel mõistmises, andmestruktuuride ja protokollide mõistlikus dokumenteerimises ning nende dokumentatsioonide kerges leidmises. Infosüsteemid kasutavad salvestatud teadmiste esitamiseks mitmesuguseid vahendeid alates keelest, sõnastikest, klassifikatsioonidest, reeglitest kuni keeruliste ontoloogiateni. Analoogselt infosüsteemi riist- ja tarkvaraga võime me rääkida infosüsteemi semantikavaradest.

## 3.2 Semantikavarad

Semantilises koosvõime varades eristatakse süntaksivarasid ja semantikavarasid. Kahe infosüsteemi semantilise koosvõime tagamiseks on vajalik nendevahelise semantikalüüsi (semantic gateway) olemasolu. Semantikalüüs peaks tagama semantilised teisendused, mille tulemusena infosüsteemid kasutaksid üksteise andmeid adekvaatselt. Riigi infosüsteemi semantikalüüs on mitmepoolsete kokkulepete ja reeglite kogum, mis lihtsustaks süsteemide omavahelist seostamist semantilisel tasemel.

**Süntaksivaradeks** on XML skeemid, metaandmete skeemid, mudelid. Selles lõigus on lähiaasta ülesandeks fikseerida põhimõtted, kuidas avalikustatakse andmeskeeme ja metaandmete definitsioone. Koosvõime süntaksi tase on esmane etapp semantilise koosvõime saavutamiseks. Koosvõime süntaksi tasemel on saavutatav XML skeemide repositooriumide loomisega.

Semantilise koosvõime **semantikavarad** on infosüsteemide koosvõime tagamiseks loodud inforessursid. Eristame järgnevaid semantilise koosvõime varade tüüpe (jaotus on võetud IDABC töödokumendist „IDABC Content Interoperability strategy”):

- sõnastikud,
- tesaurused,
- klassifikaatorid (nomenklatuurid),
- taksonoomiad,
- vastavustabelid (*mapping tables*),
- ontoloogiad,
- teenuste registrid.

### 3.2.1 Valdkonnasõnastikud

Valdkonnaspetsiifilise sõnavara selgitamiseks valdkonnavälistele osapooltele rakendatakse (valdkonna)sõnastikke. Traditsiooniline sõnastik on alfabeetiliselt järjestatud mõistete loetelu koos viimaste juurde käivate selgitustega. Alfabeetiline järjestatus tagab otsitavate definitsioonide kiire leidmise kui mõiste on teada. Kuigi valdkonnasõnastike peamine otsene väärtus inimeste seisukohast on vastava valdkonna terminoloogia lahtiselgitus kasutades üldisemat sõnavara, omab masintöödeldavuse seisukohast suurimat väärtust sõnastik, mis viitab lisaks sama tähendusega mõistetele (sünonüümidele) teiste valdkondade sõnastikes. Taolised sõnastikud kombineerivad inimloetavad selgitused tesauruse funktsionaalsusega.

### 3.2.2 Tesaurused

Tesaurused on tavatähenduses liik mõistelist sõnaraamatut, kus sõnavaraüksused ei ole organiseeritud mitte alfabeetiliselt, vaid sisuseoseid pidi. Tesaurusele on omane mõistetevaheline hierarhiline struktuur ja alluvussuhted. Suhteliike võib defineerida erinevalt (nt. soo- ja liigiseosed, osa ja terviku seosed jne.) Süsteemi üldise arhitektuuri ja seosevõrgustiku järjepidevuse huvides võib tesaurus üksustena sisaldada mitte üksnes loomuliku keele sõnu ja väljendeid, vaid ka kokkuleppelisi abstraktse sisuga deskriptoreid. Seda eriti organisatsiooni kõrgematel, suurema üldistusastmega tasanditel.

### 3.2.3 Klassifikaatorid

Klassifikaatorid täidavad analoogseid funktsioone tesaurustega. Mõistete asemel opereeritakse siin liiginimedega. Klassifikaator on üht tüüpi objekte teatud tunnuse või tunnuste hulga alusel liigitav ja üheselt identifitseeriv tähistuseeskiri. Struktuuri järgi võib klassifikaatoreid liigitada

hierarhilisteks ja lineaarseteks. Hierarhiline klassifikaator on mitme tasemega klassifikaator, mida on võimalik kujutada puustruktuurina ehk tasemelise struktuurina, kus on võimalik eri tasemete vahel üles-alla liikuda. Lineaarne klassifikaator on ühe tasemega klassifikaator, mis kujutab koodide ja nende vastavate nimetuste loetelu. Rahvusvahelisel tasemel liigitatakse klassifikaatoreid vastavalt kasutusotstarbele (tegevusalad, tooted ja teenused, territooriumid jne) ning vastavalt kasutuse ulatusele (rahvusvahelised, regionaalsed, riigisisised).

### **3.2.4 Taksonoomiad**

Taksonoomiad on puukujulised struktuurid objektide kategoriseerimiseks. Seega saab klassifikaatoreid vaadelda taksonoomiate erijuhuna. Peamine erinevus taksonoomiate ning klassifikaatorite vahel tuleneb taksonoomiate kõrgemast abstraktsusetasemest võrreldes klassifikaatoritega.

### **3.2.5 Vastavustabelid**

Vastavustabelite peamiseks ülesandeks hetkel on nii erinevate klassifikaatorite vaheliste vastavuste kui ka klassifikaatorite eri versioonide vaheliste üleminekute kirjeldamine. Taolised üleminekud on vajalikud koosvõime tagamiseks nii tehnilisel, organisatoorsel kui ka semantilisel tasandil. Seega põhimõtteliselt võiks vastavustabeleid kasutada ka ontoloogiate ja teiste semantiliste varade vaheliste seoste hoidmiseks.

Infosüsteemi semantikavarade väljatöötamisel on sageli vajalik arvestada/kasutada teise infosüsteemi semantikavarasid. On võimalik korduvkasutada teise infosüsteemide semantikat või töötada välja vastavustabelid suhtlemiseks teiste infosüsteemidega. Avalikus sektoris on praeguseks fikseeritud reeglid üht tüüpi semantikavarade (klassifikaatorite) korduvkasutuse korraldamiseks. Vastavustabelite loomine on lihtne sõnastike jaoks ja juhul kui mõlemad infosüsteemid kasutavad sama hierarhiat. Keerulisematel juhtudel pole semantikalüüsi loomine vastavustabelite abil võimalik ja korrektsete vastavustabelite loomiseks on vajalik muuta suhtlevate infosüsteemide semantikat.

### **3.2.6 Ontoloogiad**

Filosoofiateadustes käsitletakse ontoloogiat kui olemise õpetust. Mõistel „ontoloogia” on infotehnoloogias konkreetsem tähendus kui filosoofiateadustes. Laiemalt võetult on ontoloogia mõistete kogum koos nendevaheliste seostega. Selles mõttes on tesaurused, taksonoomiad ning klassifikaatorid ontoloogiate erijuhud. Kitsamalt võttes on valdkonna ontoloogia vastava valdkonna mõistete süsteem.

### **3.2.7 Teenuste registrid**

Teenuste registrid on olemasolevate teenuste kirjeldamist ja otsimist hõlbustavad keskkonnad. Vastavad registrid võivad eksisteerida näiteks lihtsa veebilehena (kui tegemist on inimloetava registriga). Masinloetavad teenusteregistrid võimaldavad infosüsteemidel endil otsida vajadusel konkreetseid teenuseid ning tõsta seeläbi vastavate infosüsteemide dünaamilisust. Tuntuimad masinloetavad teenusteregistrid baseeruvad UDDI või eb-XML registrite standardeil.

Riigi infosüsteemide nimistu kontekstis peetakse silmas eelkõige WSDL dokumentidega seotud teenuste kirjeldusi ning kasutatakse UDDI registreid (RIHA-s)..

### **3.3 Semantikavarad kui intellektuaalne omand**

Et vältida potentsiaalseid konflikte semantiliste varade rakendamisel ning haldamisel peab varakult sätestama kellele kuulub nende intellektuaalne omand. Rakendama peaks järgnevaid põhimõtteid:

- semantiliste varade intellektuaalne omand kuulub nende loojatele ning haldajatele;
- avalikule sektorile kuuluvad semantilised varad on inimloetaval kujul tasuta kättesaadavad kõigile;
- avalikule sektorile kuuluvad semantilised varad masinloetaval kujul on tasuta kättesaadavad avaliku sektori asutustele ja teadus- ning õppetöoks. Ülejäänud osapooltele on ligipääs tagatud vastavate lepingute alusel, kusjuures piisab ühest lepingust ligipääsu omandamiseks kõigi avaliku sektori semantilistele varadele.

### **3.4 Semantilise koosvõime arhitektuur**

Infosüsteemide arhitektuuri planeerimisel on peamised suunised semantilise koosvõime hõlbustamiseks järgmised:

- kasutada andmevahetuseks XML formaati üle HTTP või HTTPS protokollid;
- kasutatav XML formaat peab olema varustatud nii inim- kui masinloetava semantikaga;
- nii liht- kui kompleksandmete süntaktilisel kirjeldamisel kasutada tööstuse poolt tunnustatud standardeid;
- kasutada semantiliste varade loomisel ja edasise elutsükli jooksul meeskonnatööd võimaldavat keskkonda;
- tagada semantiliste varade maksimaalne taaskasutus ning kättesaadavus masinloetaval kujul kolmandatele osapooltele ka teistest EL riikidest.

## **4 Semantilist koosvõimet tagav organisatsioon**

Semantiline koosvõime sõltub eelkõige andmekogude, teenuste, rakenduste ja valdkondade kvaliteetsest dokumenteerimisest. Niisuguse dokumentatsiooni kooskõlastatud arendamine ja regulaarne täiustamine ongi seda tagava organisatsiooni peamine eesmärk. Vastav dokumentatsioon peab olema nii inim- kui masinloetav. Kui inimloetav dokumentatsioon on vajalik arendajatele ning aluseks masinloetavale, siis masinloetav on vajalik semantikavarade taaskasutuse toetamiseks ning riigiinfosüsteemide analüüsimise, monitoorimise, ümberseadistamise ja haldamise automatiseerimiseks.

Semantilise koosvõime taset saab riigis tõsta (valdkonna)sõnastike, tesauruste, klassifikaatorite ja muude semantikavarade väljatöötamisega. Sealjuures saab neile semantikavaradele viidata õigusaktides ja muuta vajadusel nende kasutamine kohustuslikuks.

Riigi infosüsteemi semantilist koosvõimet tagava organisatsiooni väljakujundamisel lähtutakse järgmistest põhimõtetest:

- keskse koordineerimise rolli võtab kanda Majandus- ja Kommunikatsiooniministeeriumi riigi infosüsteemide osakond, mille koosseisu on otstarbekas täiendada semantilise koosvõime arhitektiga;
- igas suuremas valdkonnas luuakse ekspertide grupp, kes loob ja haldab vastava valdkonna semantilisi varasid. Üldiselt määratakse valdkonnad ministeeriumite

valitsemisaladega ning prioriseeritakse andmeteenuste semantiliseks kirjeldamiseks vajalike semantikavarade väljatöötamist;

- semantiliste varade loomisel välditakse ühelt poolt võimalikku varade dubleerimist ning teiselt poolt tagatakse dublikaatide eemaldamine;
- valdkondadevaheliste semantikavarade loomiseks rakendatakse Internetipõhist meeskonnatöö keskkonda;
- juhul kui semantilise koosvõime saavutamiseks ei piisa mitmepoolsetest kokkulepetest, luuakse valdkondadevahelised töögrupid. Valdkondadevaheliste töögruppide eesmärgiks on luua ja hallata juhendeid selle kohta, kuidas ühe valdkonna semantikavarad teisendada teise valdkonna semantikavaradeks;
- rahvusvahelises valdkonnas on Eesti huvides osaleda IDABC semantilise koosvõime gruppides, mille eesmärgiks on välja töötada mitmepoolsed kokkulepped ja semantikalüüsid eri riikide infosüsteemide semantilise koosvõime tagamiseks. Olulisemate liidestamisprojektide korral välismaiste süsteemidega luuakse kahepoolne töögrupp kummagi poole ekspertidest;
- semantikavarad publitseeritakse RIHA-s.

#### **4.1 Osalised**

Semantiliste varade loomise, haldamise ning nende protsesside delegeerimise jaoks pakume välja järgneva hierarhilise organisatsiooni loomise. Organisatsioon koosneb tsentraalsest toetavast organist, tugikeskusest ning valdkonnaeksperte ühendavatest organitest.

##### **4.1.1 Tsentraalne organ**

Tsentraalse organi ülesandeks on organiseerida või delegeerida:

- andmekogude ja -teenuste registri ülalpidamine ja selles leiduva info veebis publitseerimine;
- semantiliste varade loomist, haldamist ning veebis publitseerimist toetava meeskonnatööks sobiva veebikeskkonna haldamine;
- koolituste ja juhendite loomise organiseerimine;
- koolituste regulaarse läbiviimise organiseerimine;
- valdkondade/ministeeriumite töögruppide organiseerimine;
- valdkondadevaheliste töögruppide töö toetamine;
- rahvusvahelise koostöö organiseerimine,
- semantiliste varade loomise ning haldamise järelvalve teostamine.

Tsentraalse organi igapäevast tööd sobib tegema inimene, kes peab vastutama andmekogude/teenuste registri ülalpidamise eest. Lisaks taolisele inimesele on mõistlik semantilise koosvõime töögrupi loomine erinevatest ekspertidest, kes vastavalt vajadusele, kuid mitte harvemalt kui kord kvartalis, peaksid töögrupi koosoleku.

##### **4.1.2 Tugikeskus**

Tugikeskuse ülesandeks on toetada tsentraalset organit organiseeritud tegevuste täideviimisel.

### **4.1.3 Valdkondade esindusorganid**

Arvestades, et semantilise koosvõime saavutamiseks on mõistlik läheneda semantikavarade loomisele ning haldamisele valdkonniti, siis on igas suuremas valdkonnas tarvis ekspertide gruppi, kes koostab ja haldab vastava valdkonna semantikavarasid. Kuna peamised valdkonnad kattuvad enam-vähem ministeeriumidega, siis on otstarbekas moodustada iga ministeeriumi juurde ekspertgrupp, kelle ülesanne oleks koostada ja hallata vastava valdkonna semantikavarasid. Taoline ekspertgrupp saaks üldjuhendeid oma tegevuseks eelmainitud tsentraalselt organilt, kes omakorda publitseerib ekspertgrupi loodavaid/hallatavaid semantilisi varasid.

On oluline, et ekspertgruppi kuuluksid nii antud valdkonda reguleerivad ametnikud ja valdkonnaekspertid kui vastava valdkonna semantiliste varadega otseselt seotud arendajad ja vastava koolituse saanud ontoloog. Semantiliste varade loomisel arvestatakse nii puhtalt infotehnoloogilisi kui sisulisi, korralduslikke ja igapäevasest tegevusest lähtuvaid aspekte. Semantiliste varade loomise algusjärgus on mõttekas lähtuda antud ministeeriumi valitsemisalas olevatest andmeteenustest.

### **4.1.4 Andmeteenuste haldajad**

Iga avalikus halduses kasutatava andmeteenuste komplekti jaoks peab olema määratud isik, kes vastutab antud objektide semantilise kirjeldamise eest ning nende publitseerimise eest RIHAs. Sama isik peab sisuliselt valdama andmeteenuste kodeerimispõhimõtteid ja andmete semantikat. Samuti peab ta olema valmis nõu andma vastava ministeeriumi semantikatöögrupile ja konsulteerima arendajaid, kes tegelevad muude rakenduste liidestamisega antud andmeteenustega.

### **4.1.5 Valdkondadevahelised töögrupid**

Valdkondadevaheliste töögruppide eesmärgiks on luua ja hallata juhendeid selle kohta, kuidas ühe valdkonna semantilisi varasid teisendada teise valdkonna semantilisteks varadeks. Taolised teisendused ei ole enamasti triviaalsed – tihtipeale kaob nende käigus informatsiooni või vastupidi, saadakse uus info erinevate semantiliste varade mestimise teel.

Valdkondadevahelisi töögrupe on mõistlik moodustada mahukamat sorti liidestamistöode jaoks, kus ühe valdkonna rakendus peab hakkama regulaarselt andmeid saama teise valdkonna käest. Niisuguse projekti käigus on mõttekas kokku panna ajutine töögrupp, kuhu kuuluvad liikmed kummagi valdkonna semantika-ekspertgrupist. Loodav teisendusjuhend tuleb muuhulgas anda ka tsentraalsele organile, kes selle omakorda publitseerib.

## **4.2 Semantilise koosvõime tagamise protsessid**

Semantilise koosvõime tagamise protsessid on järgnevad:

- valdkondade määramine,
- semantikavarade loomine valdkonniti,
- semantikavarade haldamine valdkonniti,
- semantikavarade valdkonnaülene mestimine,
- tööriistade ning keskkondade loomine semantikavarade haldamiseks ja rakendamiseks,
- meetodikate ning juhendmaterjalide väljatöötamine ja koolituste korraldamine,

- avalikkuse teadlikkuse tõstmine semantilise koosvõime olemusest,
- semantikavarade publitseerimine.

Ressursside eraldamisel semantiliste varade loomise ning haldamise jaoks peaks lähtuma 20:80 printsiibist – kui varade loomiseks kulub 20 ühikut ressursse, siis varade haldamiseks 80 ühikut. Semantikavarade haldamine toimub kahes suunas: valdkonniti ning valdkondade üleselt. Esimesel juhul tagatakse semantikavarade maksimaalne taaskasutus ning liiasuse minimeerimine läbi konkreetse valdkonna semantikavarade ristkasutuse. Teisel juhul luuakse vastavustabelid erinevate valdkondade raames loodud semantikavarade teisendamiseks ning semantikavarade üldkogud, mis on valdkonniti ühised.

#### **4.2.1 Valdkondade määramine**

Valdkonnasõnastike koostamisel, nagu enamike semantikavarade puhul, on üheks oluliseks probleemiks piiri määramine erinevate valdkondade vahele. Valdkonnad määratakse ministeeriumite valitsemisalade järgi, kusjuures arvestatakse olemasolevaid semantikavaraid – sõnaraamatuid, klassifikaatoreid, andmeteenuste süntaktilisi kirjeldusi jms, mida vastava ministeeriumi alluvuses hallatakse. Juba loodud varad määravad esialgsed piirid valdkondadele. Taolist praktikat rakendades juhtub paratamatult, et osa erinevate valdkondade raames loodud semantilistest varadest kattuvad. Lahenduseks on semantiliste varade valdkonnaülene mestimine ning meeskonnatööd toetavad keskkonnad, mis aitavad avastada sarnaseid või kattuvaid semantikavaraid.

#### **4.2.2 Semantikavarade loomine valdkonniti**

##### **4.2.2.1 Terminoloogia loomine**

Erinevad inimesed, isegi tegutsedes samas valdkonnas, võivad saada erinevalt aru mõistetest, mida nad kasutavad oma igapäevases töös. Ühise valdkonnaspetsiifilise terminoloogia loomine tagab semantiliste varade definitsioonide ühese mõistetavuse valdkonniti. Terminoloogia loomisel on aluseks valdkonnas kasutatavad dokumendid ja neis määratletud terminid, mõisted ning nende tähendused. Terminite defineerimisel järgitakse ISO/IEC 11179 standardit.

Valdkonna terminoloogia luuakse üldjuhul vastava valdkonna ekspertide poolt. Terminoloogia loomisel võetakse aluseks terminoloogia teaduse printsiipe ja seda tegevust vaadeldakse kui terminoloogilist tööd. Terminoloogiline töö on mõistete ja nende tähistuste süstemaatiline kogumine, kirjeldamine, töötlemine ja esitamine. Tulemusena tekib reguleeritud sõnastik, mis koosneb võimalike terminite ametlikuks kasutamiseks mõeldud alamhulgast. Vastavat terminoloogiat kasutatakse nii valdkonnasõnastike kui (haldus)tesaurus(t)e väljatöötamisel.

##### **4.2.2.2 Valdkonnasõnastike ehk ontoloogiate loomine**

Valdkondade sõnastikke ehk ontoloogiad luuakse valdkondade ekspertgruppide poolt, kusjuures mõistetena ning viimaste kirjeldamiseks kasutatakse olemasolevat terminoloogiat. Juhul kui tuvastatakse mõni puuduv termin kasutatavast terminoloogiast, lisatakse vastav termin sinna jooksvalt. Arvestades, et ontoloogiad võivad kirjeldada infosüsteemide erinevaid vaateid, siis tuleb enne ontoloogiate loomist täpsustada konkreetne vaatepunkt. Antud juhul määratakse vaatepunktina andmeteenuste semantilise koosvõime tagamise ülesanne ja vajadused.



Ontoloogiate loomine ja arendusprotsess on keerulised ning nende protsesside läbiviimiseks on loodud mitmeid ontoloogiate loomise üldisi metodoloogiaid. Olukorra teeb keerulisemaks asjaolu, et pole olemas ühte standardset ontoloogiate loomise ja arenduse meetodikat. Vaatamata eelnevale, esitame allpool siiski mõned suuremad tegevused, mida ontoloogiate loomise protsess peaks kindlasti hõlmama pärast seda kui konkreetsed vaatepunktid on kindlaks määratud:

- ontoloogiasse lülitavate terminite tuvastamine;
- mõistete klassifitseerimine taksonoomiatesse;
- loodava ontoloogia mõistete omavaheliste seoste määramine;
- mõistete klasside, eksemplaride, konstantide ning nende atribuutide (omaduste) määramine ja kirjeldamine;
- valdkonnas kehtivate kitsenduste formaalne esitamine;
- valdkonnas kehtivate reeglite formaalne esitamine;
- ontoloogia verifitseerimine, selle mittevastuolulisuse kontroll jms.

Nimetatud punktide täitmisel on valdkonnaekspertidel kohustus esitada nõutud info inimkeeles ning ontoloogil masinloetavas ontoloogiate esitamise formaadis. Kitsenduste all peetakse silmas tingimusi, mis on antud valdkonna piires alati tõesed. Reegleid kasutatakse valdkonniti seaduspärasuste kirjeldamiseks.

*Nõue 1. Ministeeriumitel luua oma haldusala valdkondade sõnastikud 31. detsembriks 2008.*

#### **4.2.2.3 Tesauruste loomine**

Tesauruste loomisel kasutatakse olemasolevaid masinloetavaid tesauruseid ning vastavate valdkondade terminoloogiaid. Kindlasti tuleb arvestada, et sama tesaurust rakendatakse ka valdkonnasõnastike ning valdkonnaterminoloogiate täiendamisel. Üheks tesauruste vormiks on haldustesaurus, mis käsitleb kogu avaliku sektori terminite seoseid, näiteks sünonüüme ja lühendeid. Tesauruste väljatöötamisel järgitakse standardit ISO 2788-1986.

*Nõue 2. Riigikantseleil koostada 1. jaanuariks 2008 haldustesauruse loomise tegevuskava. Haldustesaurus peab tagama veebide, dokumentide ja muude avaliku sektori inforessursside indekseerimise.*

#### **4.2.2.4 Andmekogude ja –teenuste kirjeldamine**

Konkreetsed eeskirjad andmekogude ning –teenuste kirjeldamiseks on esitatud dokumendis "Andmekogude ja andmekogude poolt osutavate teenuste semantilise kirjeldamise juhised". Siin kirjeldatakse vaid üldist protsessi ning prioriteete. **Andmeteenuste masinloetavad semantilised kirjeldused** on seotud otseselt riigi infosüsteemide semantilise koosvõime tagamiseks ning seetõttu **kõige olulisemad**.

Andmekogude ning –teenuste kirjeldamisel kasutatakse olemasolevaid semantikavarasid, eelkõige valdkonnasõnastikke ehk ontoloogiaid, milles leiduvaile mõistetele vajadusel viidatakse. Nimelt on andmebaaside väljad ja andmeteenuse sisendid ning väljundid seotud ontoloogiate mõistete kaudu. Inimloetav semantiline kirjeldus ning masinloetav süntaktiline kirjeldus luuakse koostöös andmekogude ning –teenuste haldajate ning arendajate poolt.

Baseerudes sellele informatsioonile lisavad vastava valdkonna eest vastutavad ontoloogid masinloetava semantilise kirjelduse konsulteerides vajadusel inimloetava semantika loojatega. Kui peaks selguma, et konkreetsete andmekogude ning –teenuste masinloetavalt semantiliseks kirjeldamiseks puuduvad vajalikud semantilised varad, millele viidata, luuakse jooksvalt vastavad semantilised varad või nende kohatäitjad.

Andmekogude ning -teenuste kirjeldamiseks tuleb kõigepealt määratleda, milliseid andmekogusid ja –teenuseid ning millises järjekorras semantilisel kirjeldada. Andmeteente puhul tuleb semantilisel kirjeldada kõik RIHA-s või selle järeltulijais registreeritud teenused. Andmekogude semantilisel kirjeldamisel eelistatakse kogusid, mille baasil on kavas arendada uusi andmeteenteid ehk siis andmekogusid, mille andmeid soovitakse jagada või kasutada koos teiste rakenduste ja organisatsioonidega.

*Nõue 3. X-teega ühinenud andmekogudel koostada andmekogude ja teenuste semantilised kirjeldused ja publitseerida need RIHAs 1. juuliks 2008.*

#### **4.2.3 Semantiliste varade haldamine valdkonniti**

Semantiliste varade haldamisel valdkonniti elimineeritakse liiasus erinevate varade lõikes ning tagatakse kõikide semantiliste varade maksimaalne korduvkasutus. Viimane tähendab vastavate valdkondade piires erinevat liiki semantiliste varade seostamist omavahel. Semantiliste varade haldajad lähtuvad järgnevatest reeglitest:

- viidata saab muuhulgas ka semantilistele varadele, mida veel pole loodud. Sel juhul teavitatakse vastavate varade loomise eest vastutajaid vajadusest vastava semantilise vara järele. Vastavate varade loojail on õigus suunata viide mõnele teisele varale kui nad on veendunud, et sama semantikaga vara juba eksisteerib;
- konkreetse semantilise vara kustutamisel tuleb ümber suunata ka viited semantilisel sobivama sisuga varale;
- vara ümbernimetamisel, üleviimisel teise valdkonna alla või üldiste semantiliste varade nimekirja suunatakse vastavale varale osutavad viited ümbernimetatud varale;
- varade mestimisel ehk üldisema semantikaga vara loomisel suunatakse esialgsetele varadele osutavad viited üldisemale varale;
- varade poolitamisel ehk üldisemast varast konkretiseerimise teel spetsiifilisema semantikaga varade loomisel säilitatakse viide kas üldisemale varale või luuakse uus semantiline vara, mis ei välista viidet ühelegi konkretiseerimise tulemusena tekkinud varale;
- enne uute varade lisamist ollakse veendunud, et see ei dublitseeri olemasolevaid vastava valdkonna piires;
- semantilistesse varadesse tehtud muudatused kajastuvad vastavustabelites, mis talletavad semantiliste varade haldamisel tehtud operatsioone ning lisaks kajastavad semantiliste varade teisendusi ühest valdkonnast teise.

Arvestades semantiliste varade loojate ning haldajate killustatust nii valdkonniti kui geograafiliselt, on nimetatud reeglite täimiseks vajalik eraldi meeskonnatööd toetav keskkond.

##### **4.2.3.1 Terminoloogiad**

Kuna terminoloogiatele viidatakse valdkonnasõnastikest ning tesaurustest, siis saab viimatinimetatud semantikavarade poolt viidatud termineid terminoloogias muuta, lisada või

kustutada vaid juhul kui seotud valdkonnasõnastike ning tesauruste haldajad taolise muudatusega nõustuvad.

#### **4.2.3.2 Ontoloogiad**

Ontoloogiade haldamine on seotud nii terminoloogia kui andmekogude ning –teenuste haldamisega. Seega, kuna ontoloogiade mõistetele viidatakse andmeteenuste ning -kogude semantilistest kirjeldustest, siis muudatused ontoloogiates peavad olema heaks kiidetud ka andmeteenuste ja –kogude kirjeldajate poolt kui nendele viidatakse andmekogude kirjeldustest. Lisaks on ontoloogiade koostajail kohustus teha koostööd terminoloogia haldajatega.

#### **4.2.3.3 Sõnastikud**

Sõnastike väljatöötamine on seotud terminoloogiade ning tesauruste loomise ja haldamisega. Sõnastikke luuakse kahel tasemel – sõnastikud tõlkimiseks ühest keelest teise ning sõnastikud terminoloogiade lahtiselgitamiseks. Mõlemad tasemed toetavad lisaks masintõlkele ka Eestikeelsete semantiliste varade mestimist teiste EL riikide semantiliste varadega. Arvestades, et käesolevaks hetkeks on ka erasektori poolt loodud juba piisavalt suur hulk sõnaraamatuid, tuleks mõelda nende kaasamisele semantiliste varade loomise protsessi. Vastavate sõnastike sisseostmise asemel võiks kaaluda nendele sõnastikele ligipääsuõiguse omandamist – sel moel delegeeritakse sõnastike haldamine erasektorile.

#### **4.2.3.4 Klassifikaatorid**

Klassifikaatorite väljatöötamisel ja haldamisel arvestatakse hetkel kasutusel olevaid norme ning juhendeid. Klassifikaatorite muutumisel väljastatakse vastavad vastavustabelid, mis summeerivad muudatused klassifikaatorite eelnevate ning viimaste versioonide vahel. Klassifikaatorid viitavad ka teistele semantilistele varadele - terminoloogiatele ning sõnastikele.

#### **4.2.3.5 Tesaurused**

Tesauruste haldajail on kohustus teha koostööd terminoloogia haldajatega ning praktikas oleks mõistlik, kui nii terminoloogiade kui ka tesauruste väljatöötamine toimiks samade ekspertide poolt.

#### **4.2.3.6 Andmeteenuste ning –kogude kirjeldused**

Andmeteenuste ning –kogude arendajad rakendavad peamiselt juba olemasolevaid semantilisi varasid uute semantiliste varade loomisel. Teenuseid ning andmebaase kirjeldades on võimalik tekitada viiteid semantilistele varadele, mida hetkel veel loodud pole. Sellegipoolest tuleb küsida selleks nõusolekut viidatavate varade haldajailt.

#### **4.2.4 Valdonnaülene semantiliste varade mestimine**

Valdkonnaülest semantiliste varade mestimist võiks vaadelda kahe protsessina:

- vastavustabelite loomine,
- valdkonnaülest semantiliste varade loomine ja haldamine.

Kui esimese protsessi raames luuakse sillad erinevate valdkondade mõistete vahele, siis teisel juhul komplekteeritakse semantilisi varasid, mis on ühised kõikide valdkondade piires.

#### **4.2.4.1 Vastavustabelid**

Vastavustabeleid (*mapping tables*, i.k.) rakendatakse teisenduste ning vastavuste esitamiseks erinevate valdkondade semantiliste varade või sama valdkonna semantiliste varade erinevate versioonide vahel. Vastavustabeleid on kolme liiki:

- ühepoolsed – vastavused on kinnitatud ühepoolselt vastavustabeli haldaja poolt;
- kahepoolsed – vastavused on kinnitatud kahepoolselt kahe valdkonna haldaja poolt;
- mitmepoolsed – vastavused on kinnitatud kahepoolselt kahe valdkonna haldaja poolt, kusjuures vähemalt üks osapooltest on eelnevalt vastava vastavuse kahepoolselt kinnitanud kolmanda osapoolega.

Seega piisab vastavustabelite loomisel vaid kahe osapoole vahelisest nõusolekust, isegi siis, kui osapooli on rohkem kui kaks. Vastavustabeleid hoitakse ja hallatakse neid kinnitanud osapoolte juures. Kahe- või mitmepoolselt kinnitatud vastavustabelitega seostatud semantilisi varasid saab muuta vaid kõikide osapoolte nõusolekul. Samas on igal osapoolel õigus ühepoolselt kinnitus tühistada.

#### **4.2.4.2 Valdcondadeüle sed semantilised varad**

Vastavuste loomisel erinevate valdkondade vahel tekib hulgaliselt vastavustabeleid ning teisendusreegleid, mille haldamine, viimaste suure hulga tõttu, võib aja jooksul hakata koormama semantilise koosvõime tagamise organisatsiooni. Teisendusreeglite ning vastavustabelite hulka aitaks vähendada valdkonnaüleste semantiliste varade loomine. Taoline lähenemisviis eeldab juba kriitilist hulka valdkondadevahelisi kokkuleppeid ühiste semantiliste varade olemuses ja nende kasutamises. Seega oleks mõistlik valdkondadeüleseid semantilisi varasid mitte luua enne kui on tekkinud mõningane ülevaade erinevate valdkondade semantiliste varade ühisosast. Valdcondaüleseid projekte algatab tsentraalne organ.

#### **4.2.5 Semantiliste varade riskasutus ja taaskasutus**

Semantilisi varasid ei looda tühja koha peale. Valdcondasõnastike ja ontoloogiate loomisel kasutatakse vastavate valdkondade terminoloogiaid ning tesauruseid. Tesaurused baseeruvad olemasolevatel terminoloogiatel ning kasutavad loodud klassifikaatoreid. Andmeteenuste kirjeldamisel omakorda viidatakse ontoloogiates defineeritud mõistetele. Taoline mudel tähendab semantiliste varade taaskasutamist ning vajab efektiivseks toimimiseks lisaks organisatsioonilisele toele ka vastavaid semantiliste varade loomise ning haldamise keskkondi.

Keerulisemaks teeb semantiliste varade maksimaalse taaskasutamise asjaolu, et semantiliste varade loomisele ning haldamisele lähenetakse valdkonniti. Seetõttu on liiasuste tekkimine semantiliste varade loomisesse praeguse organisatsiooni puhul „sisseprogrammeeritud“, kuna erinevate valdkondade all võidakse luua semantiliselt sarnaseid varasid. Liiasusel on aga ka positiivne külg – see võimaldab mestida erinevate valdkondade semantilisi varasid ning seeläbi kirjeldada seoseid erinevate valdkondade vahel läbi erinevate vaadete vastavatele semantiliste varadele. See omakorda aga parandab valdkonnaüleste semantilist koosvõimet. Seega tuleb tagada erinevates valdkondades loodud semantiliste varade mestimine. Arvestades aga loodavate semantiliste varade suurt hulka, hajutatust ning nende olemuse dünaamilisust peab semantiliste varade mestimist toetama vastav Internetipõhine keskkond.

#### **4.2.6 Tööriistade ning keskkondade loomine**

Käesolevaks hetkeks on loodud üsna mitmeid nii vabavaralisi kui suletud lähtekoodiga semantiliste varade haldamise tööriistu. Näiteks ontoloogiade loomiseks on olemas järgmised redaktorid: Protégé, OntoEdit, OilEd, SWOOP, SemanticWorks, jt. Kõik eelnevalt nimetatud lihtsustavad vastavate semantiliste varade loomist läbi intuitiivsete kasutajaliidestest. Sarnaseid tööriistu leidub kas vähemal või suuremal määral ka teiste semantiliste varade haldamiseks.

Paraku ei toeta ükski eelpool nimetatuid meeskonnatööd, mis on oluline semantilise koosvõime protsesside optimeerimiseks ning paralleeliseerimiseks. Samuti pole nad veebipõhised, mis oleks ladusa koostöö organiseerimiseks vajalik. Seega peab vastava keskkonna semantiliste varade loomiseks ning haldamiseks arendama.

Kuna tegemist on innovatiivse rakendusega, mille arendamise jaoks pole veel parimad tavad välja kujunenud, siis tuleb luua arenduse jaoks vastav meeskond ning siduda arendus teadustegevusega. Mõistlik oleks algatada jätkusuutlik tootearendus ning teadustegevus kas siis keskse organi juurde või siis luua partnerlussuhe vastavat toodet juurutava ettevõtte ning sellega seotud uurimisasutustega. Arendatavat toodet tuleks käsitleda kui ekspordipotentsiaali omavat tarkvara, mida saaks rakendada ka teistes projektides üle EL-i.

#### **4.2.7 Semantiliste varade publitseerimine**

Semantilised varad publitseeritakse järgnevalt:

- semantiliste varade ning nende vaheliste seoste hetkeseis iga valdkonna haldaja juures,
- publitseeritud ja versioneeritud varad koos vastavustabelitega tsentraalses repositooriumis,
- semantilised varad, millele viidatakse tsentraalselt publitseeritud varadest, publitseeritakse samuti tsentraalselt.

Tsentraalseks publitseerimiseks sobiks RIHA või mõni muu RIA poolt hallatud keskkond. Vastava semantilise vara publitseerimise hetkel fikseeritakse selle versioon, ning genereeritakse kokkuvõttevastavustabel eelmise versiooniga võrreldes. Uute versioonide publitseerimisel teavitatakse eelnevate versioonide kasutajaid uutest versioonidest läbi publitseerimise keskkonna. Publitseerimise keskkond tagab jätkusuutliku versioonihalduse ning muudatuste kättesaadavuse.

#### **4.2.8 Nimeruumide haldamine**

Semantiliste varade loomisel luuakse ka nimeruumid nende haldamiseks ning neile viitamiseks. Nimeruum määratakse konkreetse URL-iga, kusjuures vastavat URL-i veebibrauseris avades ilmub inimloetav kirjeldus vastava nimeruumi kohta koos nimeruumis sisalduvate semantiliste varade nimekirjaga ning nende kirjeldustega. Sama ideoloogiat rakendatakse ka kõikide ülejäänud XML skeemide loomisel.

### **5 Toetavad tegevused**

Semantilise koosvõime tagamine eeldab vastavate repositooriumide, haldus- ja publitseerimiskeskondade loomist. Üldiselt tuleks lähtuda põhimõttest, et semantiliste varade haldamine toimub peamiselt hajutatult ning publitseerimine tsentraalselt.

## 5.1 Semantikavarade haldamise keskkond

Semantiliste varade loomise ja haldamise tööriistad peavad võimaldama järgnevat:

- valdkondade määramine,
- semantiliste varade loomine valdkonniti,
- semantiliste varade haldamine valdkonniti,
- semantiliste varade valdkonnaülene mestimine,
- versioonihalduse teostamine ning teisendustabelite genereerimine,
- semantiliste varade importimine, eksportimine ja publitseerimine,
- otsingumootorid nii süntaktilise kui semantilise otsingu jaoks,
- sarnaste semantiliste varade tuvastamine nii valdkonniti kui valdkonnaüleselt,
- toetab meeskonnatööd ning semantikavarade paralleeltöötlust,
- on veebipõhine, kasutajasõbralik ning internatsionaliseeritav,
- võimaldab ülesannete delegeerimist teistele asutustele või organisatsioonidele,
- toimib nii hajusarhitektuuri kui tsentraalselt koordineeritavana,
- toetab iteratiivset vastavuste loomise protsessi konsensuse saavutamiseks.

*Nõue 4. Semantiliste varade loomise ning haldamise keskkond luuakse 31. detsembriks 2007.*

## 5.2 XML repositoorium

Loodud XML varade publitseerimiseks kasutatakse vastavat XML repositooriumit, kuhu kogutakse kõikide riigi infosüsteemide poolt kasutatavad XML skeemid olenemata sellest, kas need on semantiliselt rikastatud või mitte. Repositooriumi üldiseks ülesandeks on tagada infosüsteemide koosvõime süntaktilisel tasemel. Repositooriumisse paigutatakse:

- dokumendivormid (e-Arve, blanketid, jne) XSD dokumentidena,
- valdkonna- ning valdkonnaülesed sõnastikud OWL dokumentidena,
- andmeteenuste semantiliselt kirjeldused WSDL dokumentidena,
- andmestruktuuride semantiliselt kirjeldused XSD/WSDL dokumentidena,
- semantikavarade valdkondadevaheliste teisendusreeglite/vastavustabelite kogud.

Tsentraalse repositooriumi üheks ülesandeks on hõlbustada infosüsteemide arendajatel, projekteerijatel ja analüütikutel mõista võõra valdkonna andmekogude ja -teenuste andmeid. Repositooriumi teiseks ülesandeks oleks tagada vajalike dokumendivormide kiire leidmine nii ametnikele kui kodanikele. Repositoorium ehitatakse riikliku andmekogude registri (RIHA) laiendus.

Repositooriumi üldpõhimõtted on järgnevad:

- repositooriumis paiknevad vormid on alati aktuaalsed,
- repositoorium talletab avaldatud dokumentide kõik versioonid, kusjuures vaikimisi näidatakse kõige viimast,

- repositooriumis paiknevate XML skeemide aktuaalsuse eest vastutavad nende haldajad,
- repositoorium on avatud infosüsteemidele kas siis info salvestamiseks või allalaadimiseks vastava API või andmeteenuse kaudu,
- repositooriumis paiknevatele XML skeemidele saab viidata veebisaitidelt,
- repositooriumisse sisestatav info valideeritakse sisestamise hetkel,
- repositoorium toetab nii süntaktilist kui semantilist skeemide otsimist.

*Nõue 5. XML repositoorium paigaldatakse 1.veebruariks 2008.*

### **5.3 XML levikeskus**

XML levikeskus (*XML ClearingHouse*, i.k.) on portaal, mis kogub ning levitab infot riigis kasutatavate XML tehnoloogiate, XML skeemide, seotud tehniliste juhendmaterjalide, uudiste, viidete jms. kohta. Eesti XML levikoda vahendab vajadusel infot ka Euroopa XML levikojale ja –kojast. Riigi XML levikoda pakub järgnevat funktsionaalsust:

- keskne suhtlemisplatvorm läbi foorumite,
- XML repositooriumi XML skeemidele viitamine,
- juhendmaterjalidele arhiveerimine,
- lüüs ligipääsuks Euroopa levikojale.

Riigi XML levikoda peaks paiknema RISO kodulehel, kuhu paigutatakse viited XML skeemidele ning muudele materjalidele samas kui skeemid ise paiknevad hajutatult ning XML repositooriumis.

*Nõue 6. XML levikeskus avalikustatase 31. detsembriks 2008.*

## **Mõistete ja lühendite loetelu**

**Andmebaas** – andmete kogu, korraldatud vastavalt kontseptuaalstruktuurile, mis kirjeldab nende andmete omadusi ning neile vastavate olemite vahelisi seoseid, toetades üht või mitut rakendusala.

**Andmekogu** – andmebaasi õiguslik termin. Andmekogu on riigi, kohaliku omavalitsuse, avalik-õigusliku või eraõigusliku isiku peetav korrastatud andmete kogum, mille pidamisel kasutatakse automatiseeritud andmetöötlust või mida peetakse käsitsi ja korrastatud vormidel, mis võimaldavad andmetega lihtsat tutvumist või nende mehhaanilist töötlemist.

**Andmeteenus ehk teenus** – andmekogude teenused on kasutajale lisaväärtust loovad toimingud elektroonilises keskkonnas. Teenuste pakkujateks võivad olla mistahes asutused, ettevõtted, üksikisikud läbi oma infosüsteemi. Vt. ka **veebiteenus**.

**API** – *Application Programming Interface*.

**HTML** – *Hypertext Markup Language*, spetsiaalne keel teksti ja piltide esitamiseks veebis.

---

**HTTP** – *Hypertext Transfer Protocol*, veebis kasutatav andmevahetusprotokoll.

**HTTPS** – *Secure HTTP*, veebis kasutatav turvaline andmevahetusprotokoll.

**Infosüsteemide semantiline koosvõime** – tarkvarasüsteemide võime adekvaatselt kasutada teistelt tarkvarasüsteemidelt saadud andmeid.

**Inimloetav kirjeldus** – semantikavara kirjeldus, mis on oma vormilt sobiv lugemiseks ning arusaamiseks nii arendajale kui valdkonna spetsialistile ning mis esitab vastava semantikavara tähenduse (semantika).

**IT** – *infotehnoloogia*.

**Koosvõime** – infosüsteemide ja nende poolt toetavate tegevusprotsesside võime vahetada andmeid ja ühiselt kasutada informatsiooni ja teadmisi.

**Koosvõime raamistik** – standardite ja juhendite kogum, mida organisatsioonid järgivad üksteisega suhtlemisel.

**Masinloetav kirjeldus** – semantikavara kirjeldus, mis on vormilt arusaadav tarkvarasüsteemile ning mis annab edasi vastava semantikavara tähenduse (semantika).

**Mõiste** – teadmusüksus, mille moodustab ühene tunnuste kombinatsioon.

**Ontoloogia** – valdkonna kirjeldus, nii masintöödeldaval kui ka inimloetavas vormingus. Sisaldab valdkonnas kasutatavate mõistete sõnastikku koos nendevaheliste seostega (n: pärimisseos).

**OWL** – *Web Ontology Language*, masinloetava keel ontoloogiate esitamiseks.

**Portaal** – veebisait, mis koondab enda alla sisu ning funktsionaalsust erinevatest loogiliselt seotud allikatest.

**RDF** – *Resource Description Format*, ressursikirjeldusvorming.

**RDFS** *Resource Description Framework-Schema*, märgendikeel masinloetavate ontoloogiate esitamiseks. Vt ka **OWL**.

**Reegel** – nõue, mille järgimine on kohustuslik.

**Repositoorium** – andmehoidla, võimalik, et koos lisafunktsionaalsusega, mis tüüpilistes andmehoidlates puudub.

**RIA** – *Riigi infosüsteemide arenduskeskus*.

**RIHA** – *Riigi infosüsteemide haldussüsteem*.

**RISO** – *Riigi infosüsteemide osakond* Majandus- ja Kommunikatsiooniministeriumis (MKM).

**SA-WSDL** – *Semantic Annotations for WSDL*, W3C soovitus veebiteenuste ja andmestruktuuride semantiliseks kirjeldamiseks. On seisuga 26. jaanuar 2007 *W3C Candidate Recommendations* statuses.

**Semantiline kirjeldus** – andmeteenuse, andmekogu, valdkonna vm objekti tähendust avav kirjeldus.

**Semantiline koosvõime** – (*semantic interoperability*, i.k.) on erinevate organisatsioonide võime mõista vahetatud informatsiooni tähendust ühte moodi.

**Soovitus** – nõue, mille järgimine ei ole kohustuslik.

**Sõnastik** – suvalise keele kõigi sõnade või teatud kategooria sõnade tähestikuline loend koos samakeelsete seletuste või tõlgetega ühte või mitmesse keelde.

**Teenus** – vt **andmeteenus**, **veebiteenus**.

**Termin** – mingi valdkonna üldmõiste sõnaline tähis.

---



---

**UDDI** – *Universal Description, Discovery and Integration*, standard veebipõhiste teenuste kirjelduste avaldamiseks avatud kataloogide realiseerimiseks.

**URI** – *Uniform Resource Identifier*, üldine ressursilokaator, standard objektide määratlemiseks veebis. Termin URI on W3C üldine nimetus mistahes liiki praeguste või tulevaste ressursi-identifikaatorite kohta, praegu on veebis valdavalt kasutusel selle alamliik **URL**.

**URL** – *Uniform Resource Locator*, globaalne dokumentide või teiste ressursside aadress veebis. URL on üks **URI** alamliik.

**Valdkond** – eriteadmuse aineala.

**Valdkonna spetsialist** – eriteadmuse aineala toimimist tundev isik.

**Valdkonnasõnastik** – konkreetse valdkonna sõnavara koondav sõnastik.

**Veebiteenus** – (*Web service*, i.k.), masinatevaheliseks koosvõimeks arvutivõrgus loodud tarkvarasüsteem. Kitsamas võtmes on veebiteenus WSDL dokumendiga kirjeldatud tarkvarakomponent, mis kasutab andmete edastamiseks Interneti protokolle kusjuures andmed on vormindatud XML põhisteks sõnumiteks.

**W3C** – *World Wide Web Consortium*, rahvusvaheline konsortsium, mille eesmärgiks on veebi arendamine avatud standardite alusel. Peamine WWW standardiseerimisorganisatsioon. W3C standardid on näiteks **XML**, **RDF**, **OWL**.

**WS-BPEL** – *Web Services Business Process Execution Language*, veebiteenustena realiseeritud käivitavate äriprotsesside kirjeldamise keel.

**WSDL** – *Web Services Description Language*, XML-põhine veebiteenuste kirjeldamise keel. WSDL 2.0 on seisuga 26. märts 2007 *W3C Working Draft* staatuses.

**WWW** – *World Wide Web*, omavahel seotud HTML dokumentide kogumik Internetis.

**XML** – *eXtensible Mark-up Language*, tekstiline vorming struktureeritud andmete esitamiseks. Andmed esitatakse siltidega (*tag*) märgendatud elementidele antud väärtuste ja atribuutidena, kus element võib sisaldada skalaarväärtust või teistest elementidest koosnevat struktuuri.

**XML levikeskus** – XML skeemide, juhendmaterjalide ning muu XML temaatika kogumiseks ja levitamiseks loodud portaal.

**XML skeem** – (*XML schema*, i.k.) XML keeles kirjeldatud reeglite kogumik, millele XML dokumendid peavad vastama, et olla 'kehtivad' vastava skeemi alusel.

**XSD** – *XML Schema Definition*, XML skeemide kirjeldamise keel.

**X-tee** – universaalne turvaline andmevahetuskeskkond (vaata <http://x-tee.riik.ee/> ) tagamaks kodanikele, ametiisikutele ja ettevõtjatele andmebaaside kasutamise nende volituste piires.

---

## Kasutatud kirjandus

- Riigi IT koosvõime raamistik. Versioon 2.0 (2005).
- Semantilise koosvõime strateegia. Versioon 0.5 (2005).
- Andmekogude ja andmekogude poolt osutatavate teenuste semantilise koosvõime meetoodika. Versioon 1.1 (2007).
- Andmekogude ja andmekogude poolt osutatavate teenuste semantilise kirjeldamise juhised. Versioon 1.1 (2007).
- European Interoperability Framework.
- Eesti infoühiskonna arengukava 2013.

- IDABC Semantic Interoperability Strategy: The European XML Clearinghouse. Feasibility Study.