

**Atminties institucijų bendradarbiavimas:
bendras istorinių asmenvardžių,
vietovardžių ir chronologijos tezauras**

Rimvydas Laužikas
Vilniaus universitetas
Komunikacijos fakultetas
rimvydas.lauzikas@kf.vu.lt

Atminties institucijos

- Metafora (Ellis G. Burcaw, 1975; R. Hjerppe, 1994)
- Institucijos, kuris saugo paveldą, konkrečios visuomenės atminčiai svarbius dokumentus, objektus
- Lietuvoje – muziejai, archyvai, dalis bibliotekų, Kultūros paveldo departamentas... religinės bendruomenės....

Keli terminai

- **Tezaurus**
 - bendrasis ar kurios nors srities žodynas, nurodantis prasminius žodžių ryšius
 - žodžių (terminų) sąrašas, kuriame žodžiai sugrupuoti į tam tikras grupes ir yra nurodyti jų tarpusavio ryšiai
 - Garshol L.M. Metadata? Thesauri? Taxonomies? Topic? Maps! Making sense of all it. Prieiga per internetą: <<http://www.ontopia.net/topicmaps/materials/tm-vs-thesauri.html>>
- **Ontologija**
 - formali objektų ir jų tarpusavio ryšių aprašymo žinių sistema
 - metaterminų žodynas, kurio tikslai yra susieti skirtingų mokslų žinias į vieną sistemą, palengvinti žmogaus ir kompiuterio bendravimą
 - Gruber T. What is ontology? Prieiga per internetą: <<http://www.miskatonic.org/library/facet-web-howto.html>>
- **Semantiniai ryšiai**
 - pagal prasmę: Pasvalys - Vytė Nemunėlis
 - reikalingas aukštas sistemos žinių lygis
 - W3C Semantic Web Frequently Asked Questions. Prieiga per internetą: <<http://www.w3.org/2001/sw/SW-FAQ>>

Kodėl reikia BAVIC ?

- **Menkas paveldo ir humanitarinių mokslų duomenų IS struktūravimo ir standartizavimo lygis**
- **Nepakankamas paveldo ir humanitarinių mokslų mokslinių duomenų IS sąveikumas**
- **Fakto ir nuomonės problema paveldo tyrimuose**
- **Trijų lygių dokumentų problema paveldo skaitmeninime**
- **Kitų šalių tezaurų riboto taikymo problema**
- **Istorinės geografijos problemos**
- **Istorinės chronologijos problemos**
- **Istorinės biografistikos problemos**

Istorinės geografijos problemos

- dažnai keitėsi administracinis padalijimas
- keitėsi teritorijos administracinė priklausomybė
- keitėsi vietovardžiai
- lietuviški vietovardžiai buvo užrašinėjami įvairiomis kalbomis ir rašto sistemomis
- daug vienodų vietovardžių
- daug išnykusių vietovardžių
- vėlyva lietuvių kalbos rašytinė tradicija
- vėlyva Lietuvos kartografijos tradicija
- šiuolaikinė daugiakalbiškumo problema
- tarmybių (mikroregioniškumo) išsaugojimas

Istorinės chronologijos problemos

- vartojamos skirtingos chronologijos schemos
- vartojamos skirtingos chronologijos tradicijos
- periodizacijos netikslumas
- Lietuvos istorinės chronologijos sąsajos su regiono ir bendraeuropiniais chronologijos kontekstais
- skirtingi datų užrašymo ir laiko atskaitos metodai
- istorinis datavimas yra kaip interpretacija

Istorinės biografistikos problemos

- iš dalies panašios į istorinės geografijos problemas
- to paties asmens įvardinimas skirtingais vardais (slapyvardžiais)
- Lietuvoje dirbę skirtingų tautybių asmenys, kurių asmenvardžiai turi skirtingas užrašymo tradicijas gimtąja ir kitomis kalbomis
- kaip paveldas aktualios asmenvardžių tarminės bei istorinės formos

Faktinė situacija Lietuvoje, 2010 m.

- + Turime ne mažiau kaip 50 faktiškai veikiančių ir paveldo ir mokslinių duomenų skaitmeninimui skirtų įvairaus lygmens duomenų bazių ir informacinių sistemų
- + + Daugumas jų nenaudoja jokių standartų
- + + + Daugumas jų nepasirengę sąveikumui ir nėra jo suplanavę

Geroji patirtis, 2010 metai

- BAVIC poreikių erdvėje:
 - LNB Autoritetinių įrašų DB – aukštas standartizavimo lygis (UNIMARC), išsamus turinys
 - IS “Aruodai” - gera vietovardžių ir asmenvardžių pateikimo struktūra, išsamus turinys, sąsajos su žemėlapiu, iš dalies sprendžiamas istorinės chronologijos klausimas
 - Lietuvių kalbos instituto “Vietovardžių duomenų bazė” – išsamus turinys, geras lingvistinis aprūpinimas

BAVIC pateikimo formatai

- **Ontologinis aprašas**
 - **Tekstinė dalis**
 - **Duomenų elementų sąrašas**
 - **Duomenų elementų hierarchija**
 - **Duomenų elementų aprašymas**
 - **Savybių sąrašas**
 - **Savybių hierarchija**
 - **Savybių aprašymas**
 - **Sąsajos (crosswalk) su BAVIC tematikos standartais ir DB**
 - **Sąsajos su atminties institucijose naudotinais standartais ir esamomis bei kuriamomis IS**
- **Protege byla**
- **E-R diagrama**

BAVIC ontologinė struktūra

- **Duomenų elementas (klasė)**
 - vieną ar daugiau bendrų savybių turinčių elementų visuma,
 - daiktavardis (kas?)
 - **BAVIC DE - 71**
- **Savybė**
 - požymis, apibūdinantis specifinės prigimties ryšį tarp dviejų klasių
 - veiksmazodis (ką veikia?)
 - **BAVIC S - 60**
- **Savybių simetriškumas – BAVIC - asimetriškos**
- **sritis - klasė, kurios savybė yra formaliai apibrėžta – veiksnys, kas?**
- **kategorija - klasė, kuri apima visas potencialias savybės vertes – papildinys, ko? kam? ką? kuo?**

Pagrindiniai duomenų elementai

- DE57 Objektas
- - DE23 Asmens vardas
- - DE28 Grupės pavadinimas
- - DE36 Laikotarpio pavadinimas
- - DE42 Vietovės vardas

Papildomi duomenų elementai

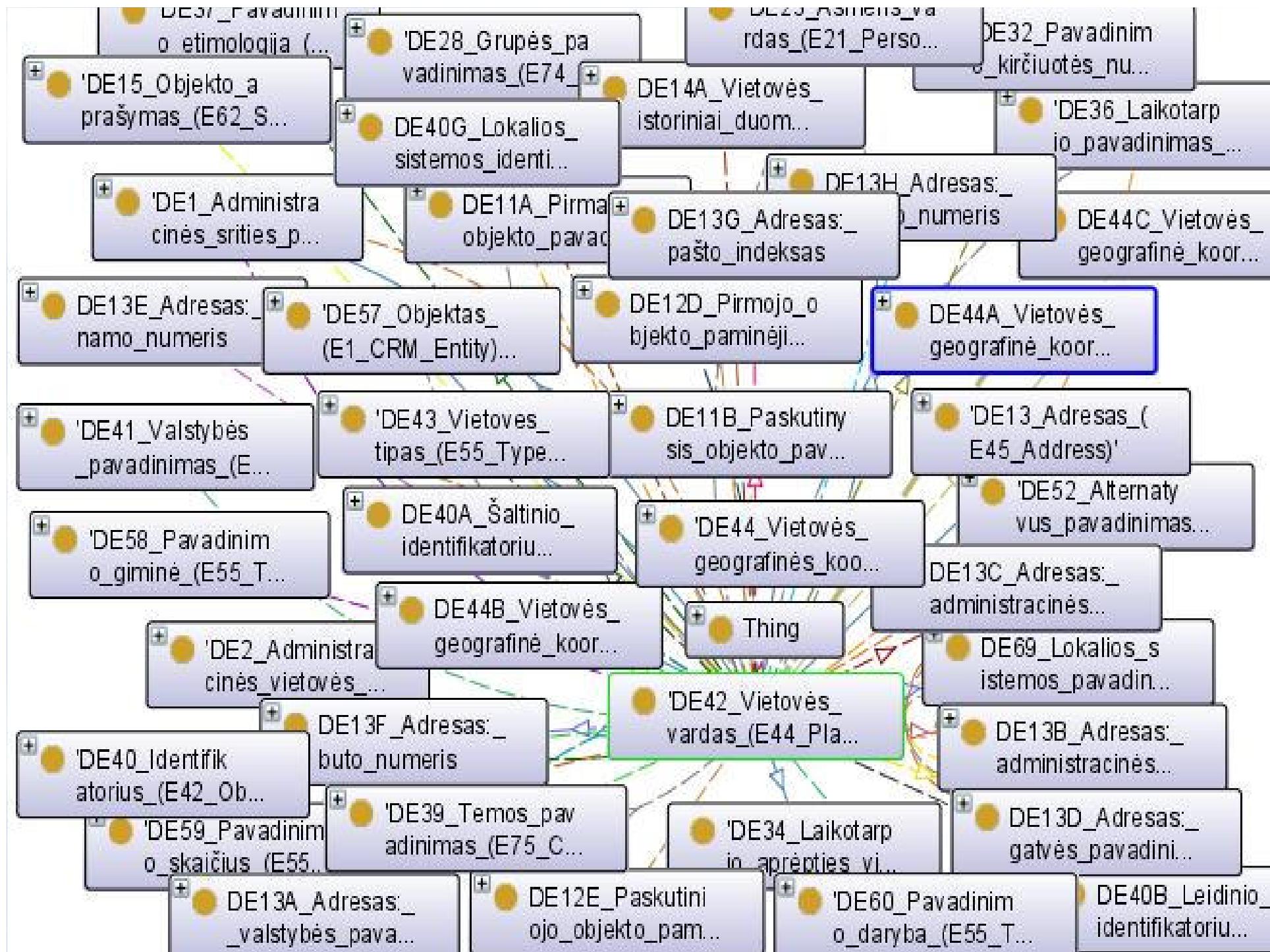
- - DE39 Temos pavadinimas
- - DE52 Alternatyvus pavadinimas

Pagalbiniai duomenų elementai

- DE68 Pirminė reikšmė
- - išreiškiami absoliučiais skaičiais (matmenys, koeficientai),
- - absoliučiais skaičiais išreikštas datas,
- - laisvo teksto eilutėmis išreiškiamus duomenų elementus

+ +

- - identifikatoriai
- - tipai
- - kalbos, tarmės, rašto sistemos
- - matavimo metodai



Projektuojamos BAVIC sąsajos

- Sąsaja standartas – standartas
(konceptcija – konceptcija)
- Sąsaja standartas – sistema
(konceptcija – realizacija)
- Sąsaja sistema – sistema
(realizacija – realizacija)

Sąsajos (crosswalk) su BAVIC tematikos DB

- Getty Thesaurus of Geographic Names (TGN)
- Union List of Artist Names (ULAN)
- LNB Autoritetinių įrašų DB
- IS “Aruodai” geografijos, personalijų bankai ir terminų žodynas
- Kultūros vertybių registro geografijos dalis
- Lietuvių kalbos instituto “Vietovardžių duomenų bazė”

Sąsajos su atminties institucijose naudotinais standartais ir esamomis ar kuriamomis IS

- **CIDOC-CRM**
- **LST ISO 15836:2009 Informacija ir dokumentai. Dublin Core metaduomenų elementų rinkinys (tapatus ISO 15836:2009)**
- **Categories for the Description of Works of Art (CDWA)**
- **SPECTRUM (Standard Procedures for Collections Recording Used in Museums)**
- **The Cultural Objects Name Authority (CONA)**
- **ATHENA LIDO - Lightweight Information Describing Objects**
- **CARARE metadata schema**
- **General International Standard Archival Description (ISAD(G), Second edition)**
- **ISAAR (CPF): International standard archival authority record for corporate bodies, persons and families. Second edition.**
- **Encoded Archival Description (EAD). Tag Library. Version 2002.**
- **Visual Resources Association's Data Standards Committee, VRA Core 4.0**
- **Metadata Mapping & Normalisation Guidelines for the Europeana Prototype (EDM). Version 5.2**
- **Europeana Semantic Elements Specification (ESE). Version 3.3**
- **LST EN 15744:2009 Filmų identifikavimas. Mažiausias kinematografinių kūrinių metaduomenų rinkinys (tapatus ISO 15744:2009)**
- **The European Film Gateway (EFG). Version 3.2.3**
- **DARIAH metaduomenų standartai**
- **UNIMARC (ePaveldas)**
- **IS „Aruodai“ objekto aprašymo laukai**
- **Kultūros vertybių registro objekto aprašymo laukai**
- **LIMIS objekto aprašymo laukai (pagal IRIS ir MERAASI)**

Potencialūs vartotojai

- **lituanistikos mokslų (plačiaja prasme) IS + paveldo IS (atminties institucijoms) + visuomenei**
- **lituanistikos mokslai - nematerialusis ir dokumentinis paveldas yra svarbiausias tyrimų šaltinis**
- **Sektorius:**
 - **apie 10 mokslo ir studijų institucijų**
 - **1 nacionalinė, 65 viešosios ir 4 mokslinės bibliotekos**
 - **4 nacionaliniai, 12 respublikinių ir beveik 100 kitokios priklausomybės muziejų**
 - **5- 8 valstybiniai archyvai**
 - **apie 1200 mokslininkų**
 - **privatus sektorius**

BAVIC nauda

- **IS duomenimis pildančiam specialistui**
- **IS vartotojui, mokslininkui**
- **IS vartotojui – visuomenei**

- **Žinių bazė**
- **Kompleksinė paieška**
- **Aspektinė (fasetinė) paieška**

Perspektyvos

- **BAVIC turi išspręsti šio dokumento pradžioje išdėstytas paveldo skaitmeninimo Lietuvoje problemas.**
- **BAVIC, kaip žinių bazė ir semantinio sąveikumo priemonė daugumoje (visose?) paveldo IS Lietuvoje.**
- **Į šį BAVIC turės būti perkelti dabar esamose ID saugomi geografiniai, chronologiniai ir personalijų duomenys.**
- **BAVIC pildymas duomenimis turėtų būti atliekamas internetu, atviru būdu, paskui papildymai tikrinami ir aprobuojami administratoriaus.**

- **Ateityje:**
 - **BAVIC – LST standartas**
 - **Projektuotina BAVIC GIS dalis ir sąsajos su interaktyviais žemėlapiais.**
 - **Verta dėmesio geoanalitinio, bioanalitinio ir chronoanalitinio modulių kūrimo perspektyva.**
 - **Tezauro kūrimas — nuolatinis procesas, besiremiantis atvirojo pasaulio principu, atvirtąjį pasaulį suprantant kaip: prielaidą, kad informacija, saugoma IS, yra nebaigtinė lyginant ją su pasaulio, kurį sistema siekia aprašyti, visuma**



Ačiū už dėmesį

rimvydas.lauzikas@kf.vu.lt