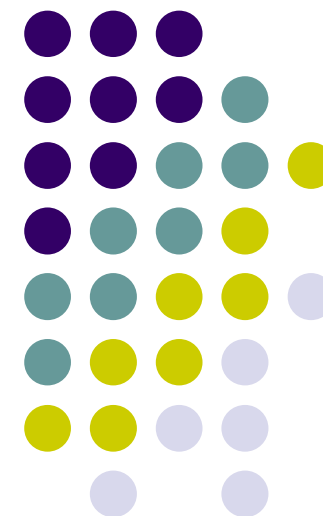
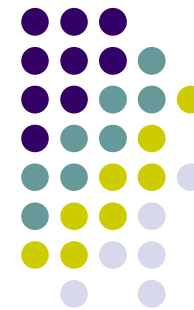


Standardizace otevřených archivů: popis a výměna agregovaných webových zdrojů prostřednictvím OAI-ORE

Petr Novák

Ústav informačních studií a knihovnictví FF UK
a
Ústav výpočetní techniky UK





- OAI-ORE: další projekt
Iniciativy pro otevřené archivy
- okolí: vazby na
 - sémantický web a jeho technologie vč.architektury budoucího webu
 - infrastrukturu repozitářů



Open Archives Initiative
Object Reuse and Exchange



Cesta k ORE

http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.14.2836

CiteSeer^x beta

Documents

Authors

Search

Include Citations | [Advanced Search](#) | [Help](#)

Summary

Related Documents

Version History

Modeling Web Site Structure in UriGraph

by Hrvoje Simic

<http://bib.irb.hr/datosite/uri-graph/www2003simic.pdf>

[Add To MetaCart](#)

[Add to Collection](#)

Co

vazby různých typů

persistentní
identifikátory

příbuzné
dokumenty

historie
verzí

DOWNLOAD:
PDF PS LATEX

formáty ke
stažení

Abstract:

This paper introduces a Web site structure model called UriGraph and, using the model, describes several important patterns of site structure. Web site structure is defined as the collective information about the identity, ...

citace

Citations

238 [Web Modeling Language \(WebML\): a modeling language for designing Web sites](#), presented at - Ceri, Fraternali, et al. - 2000

174 [Designing Web Usability: The practice of Simplicity](#) - Nielsen - 2000

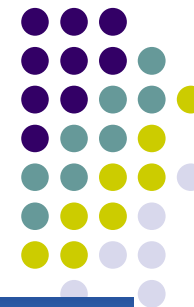
POPULAR TAGS

No tags have been applied to this document.

Add a tag:

BIBTEX | ADD TO METACART

```
@MISC{Simic_modelingweb,  
  author = {Hrvoje Simic},  
  title = {Modeling Web Site Structure in UriGraph},  
  year = {}  
}
```



Cesta k ORE

http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.14.2836

CiteSeer^x beta

Documents

Authors

Search

Include Citations | [Advanced Search](#) | [Help](#)

Summary

Related Documents

Version History

Modeling Web Site Structure in UriGraph

by Hrvoje Simic

<http://bib.irb.hr/datosite/uri-graph/www2003simic.pdf>

[Add To MetaCart](#)

příbuzné dokumenty

historie verzí

persistentní identifikátory

DOWNLOAD:
PDF PS LATEX

formáty ke stažení

>>> agregace <<<

citace

Abstract:

This paper introduces a Web site structure model called UriGraph and, using the model, describes several important patterns of site structure. Web site structure is defined as the collective information about the identity, ...

POPULAR TAGS

No tags have been applied to this document.

Add a tag:

Citations

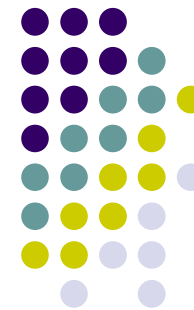
238 [Web Modeling Language \(WebML\): a modeling language for designing Web sites](#), presented at - Ceri, Fraternali, et al. - 2000

174 [Designing Web Usability: The practice of Simplicity](#) - Nielsen - 2000

BIBTEX | ADD TO METACART

```
@MISC{Simic_modelingweb,
  author = {Hrvoje Simic},
  title = {Modeling Web Site Structure in UriGraph},
  year = {}
}
```

Terminologie ORE



agregace - aggregation

zdrojová mapa - resource map

agregovaný zdroj - aggregated resource

serializace - serialization

ATOM - standard pro syndikaci

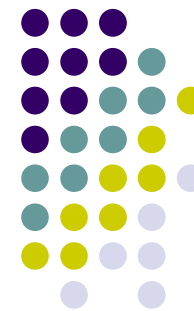
RDF - standard pro metadatový popis webových zdrojů

výrok - assertion

URI - identifikátor zdroje

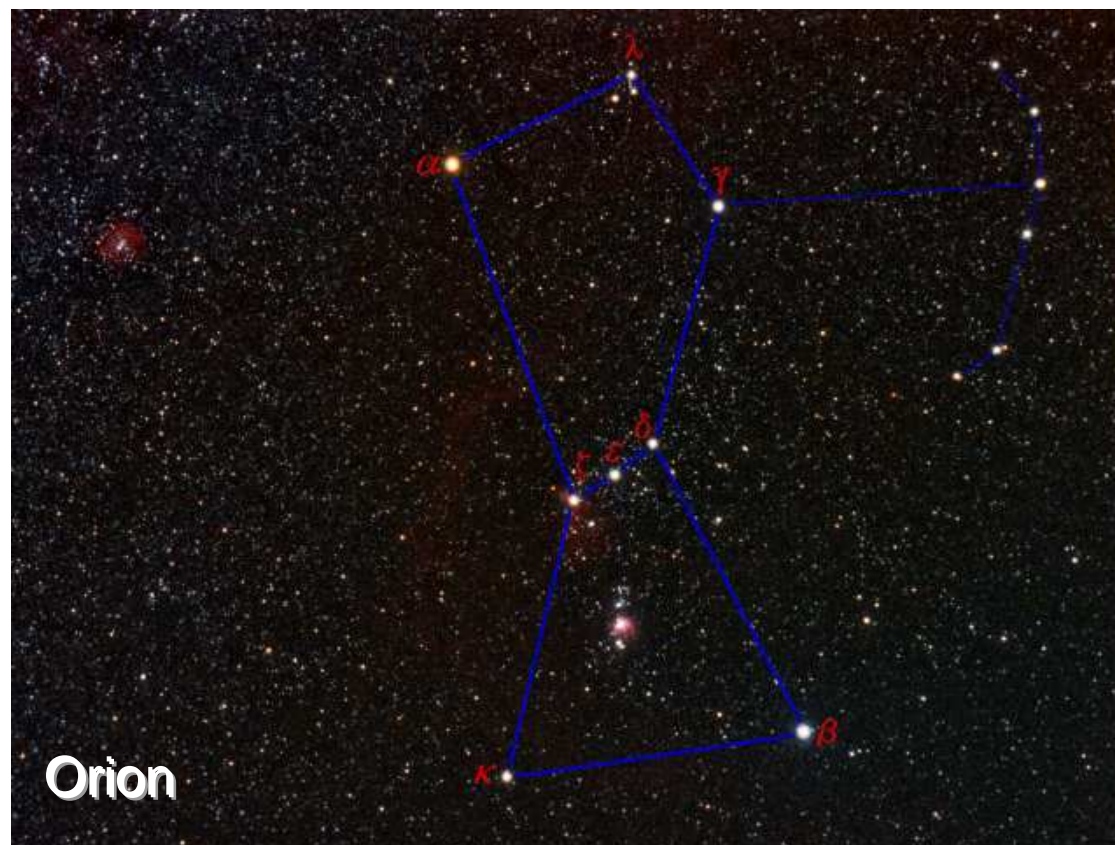
Dublin Core

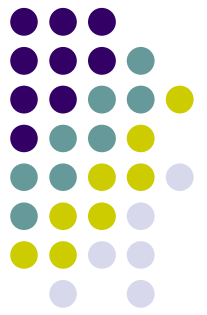
Graf - grafické znázornění vazeb



Agregace – analogie suhvězdí

- Vazby definovány ex post a v různých dobách různě
- Objektům – hvězdám – je fakt, že tvoří souhvězdí, vcelku lhostejný
- Každé souhvězdí
 - je unikátní (jen 1x)
 - je popsáno
 - je identifikovatelné

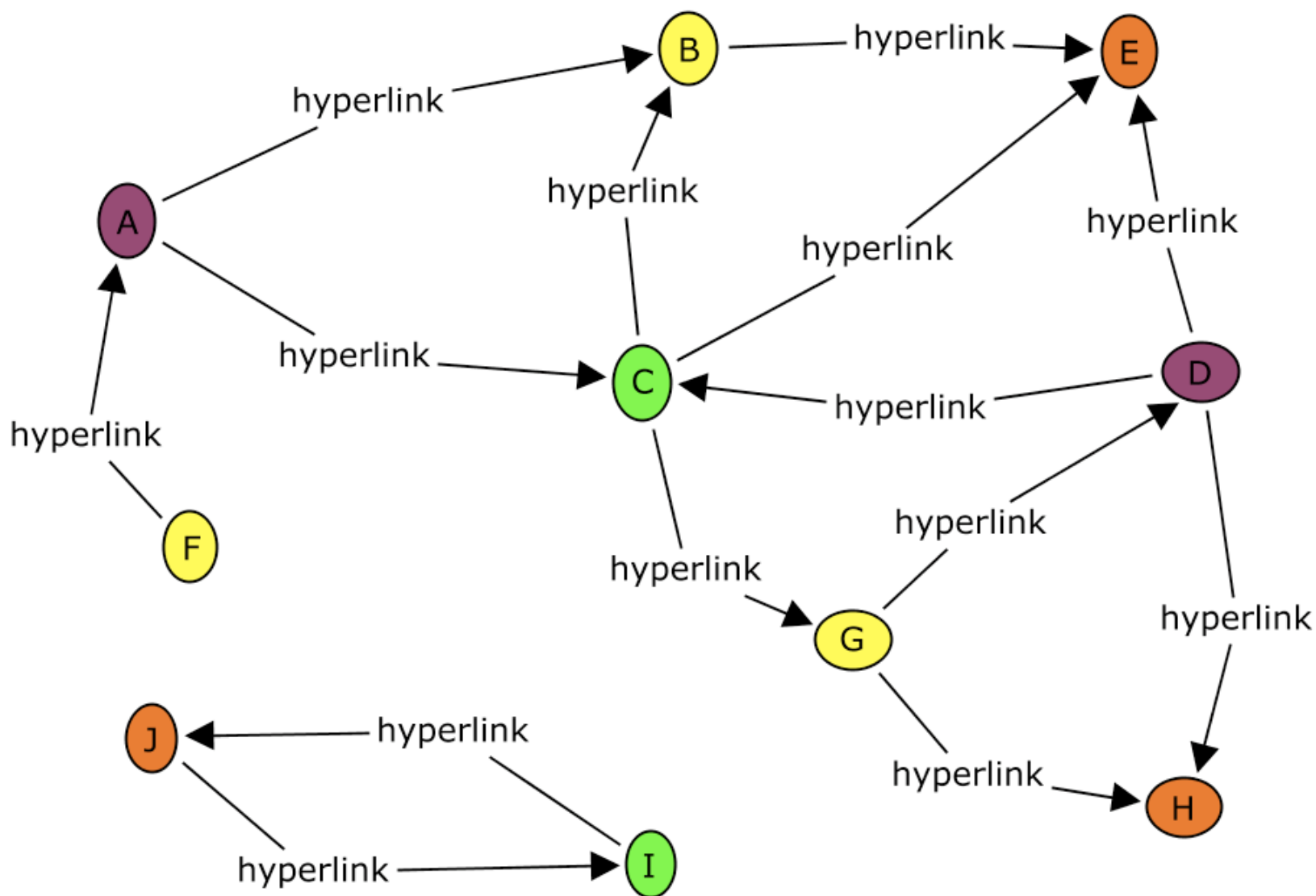


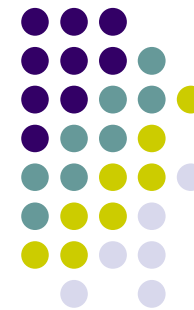


Původní vize (9.2004)

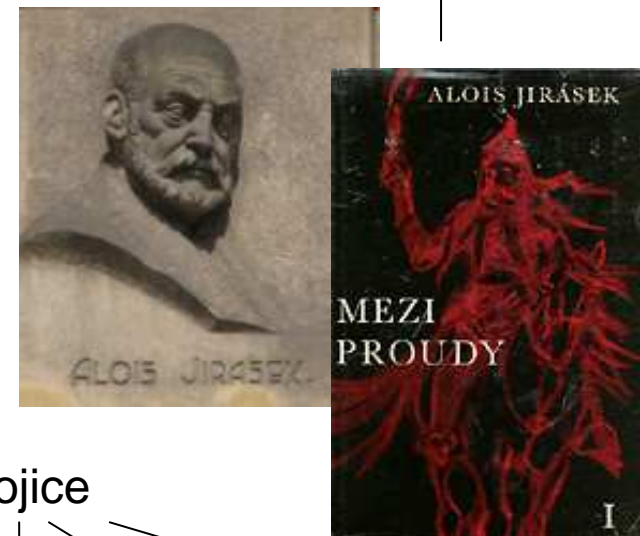
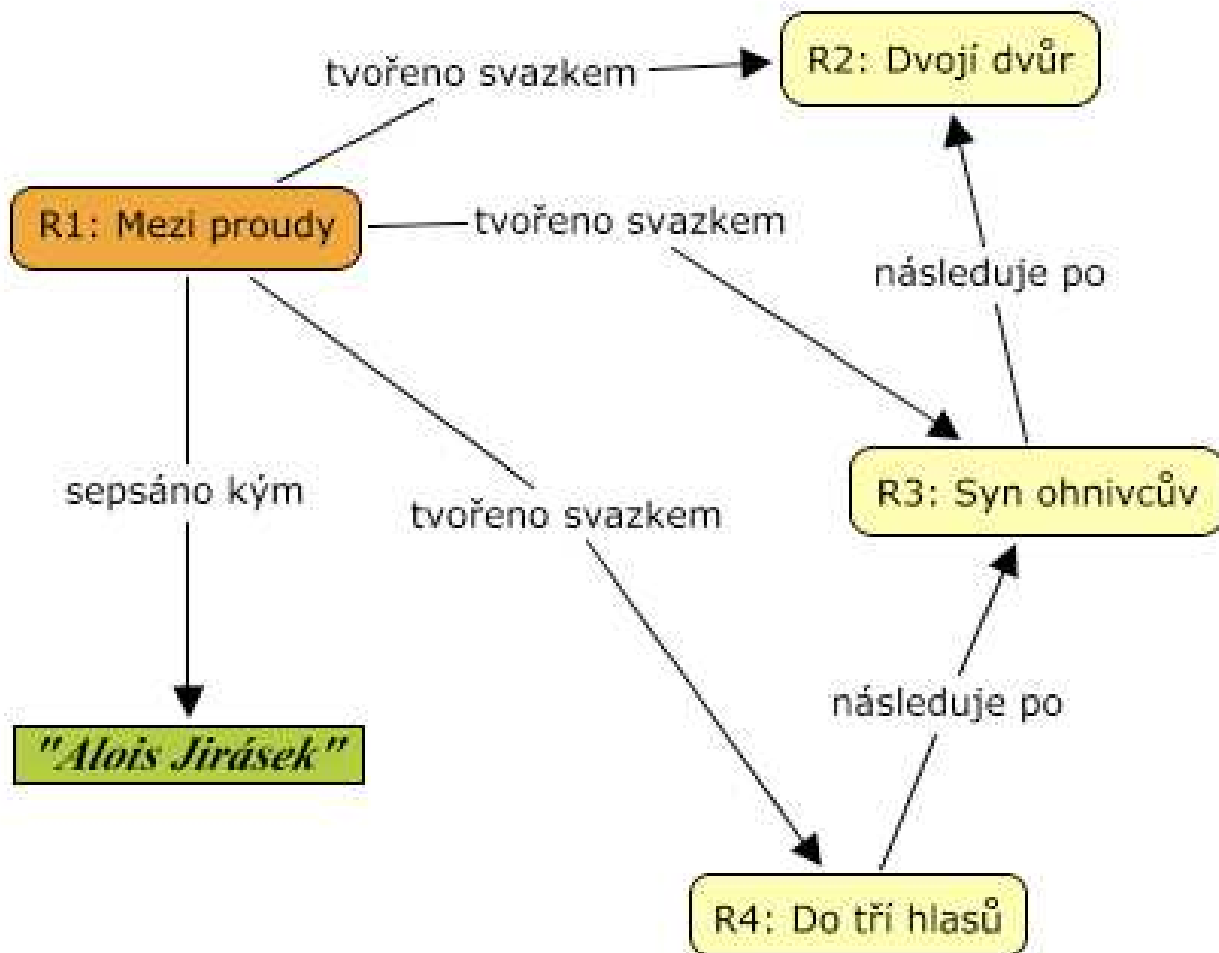
- umocnit **informační hodnotu** materiálů skrytých v repozitářích
- překonat bariéru a informace z repozitářů **nabídnout jinde**
- učinit z repozitářů **aktivní uzly** v globálním kontextu
- materiály (digitální objekty) **začínají** svůj virtuální život **v repozitářích => nekončí tam**
- tyto materiály musí být **znovuvyužitelné** v různých (dalších) kontextech

Webový graf - grafické znázornění webových vazeb





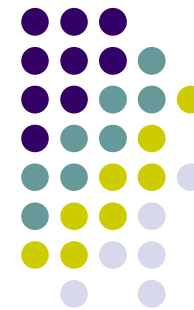
Knihy graficky znázorněná pomocí RDF grafu



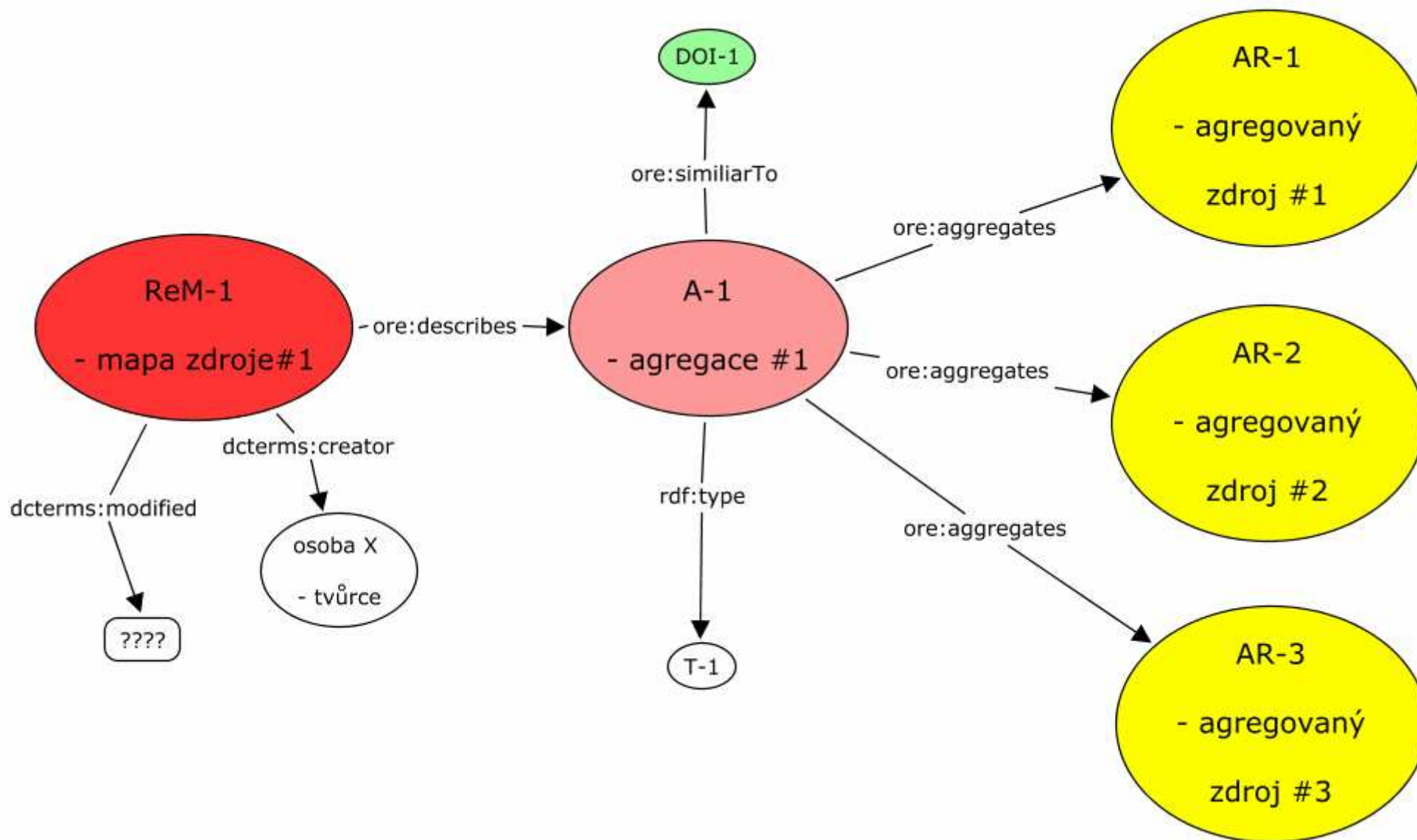
trojice

zdroj	vlastnost	objekt
R1	tvořeno svazkem	R2
R3	následuje po	R2
R1	sepsáno kým	"Alois Jirásek"
...

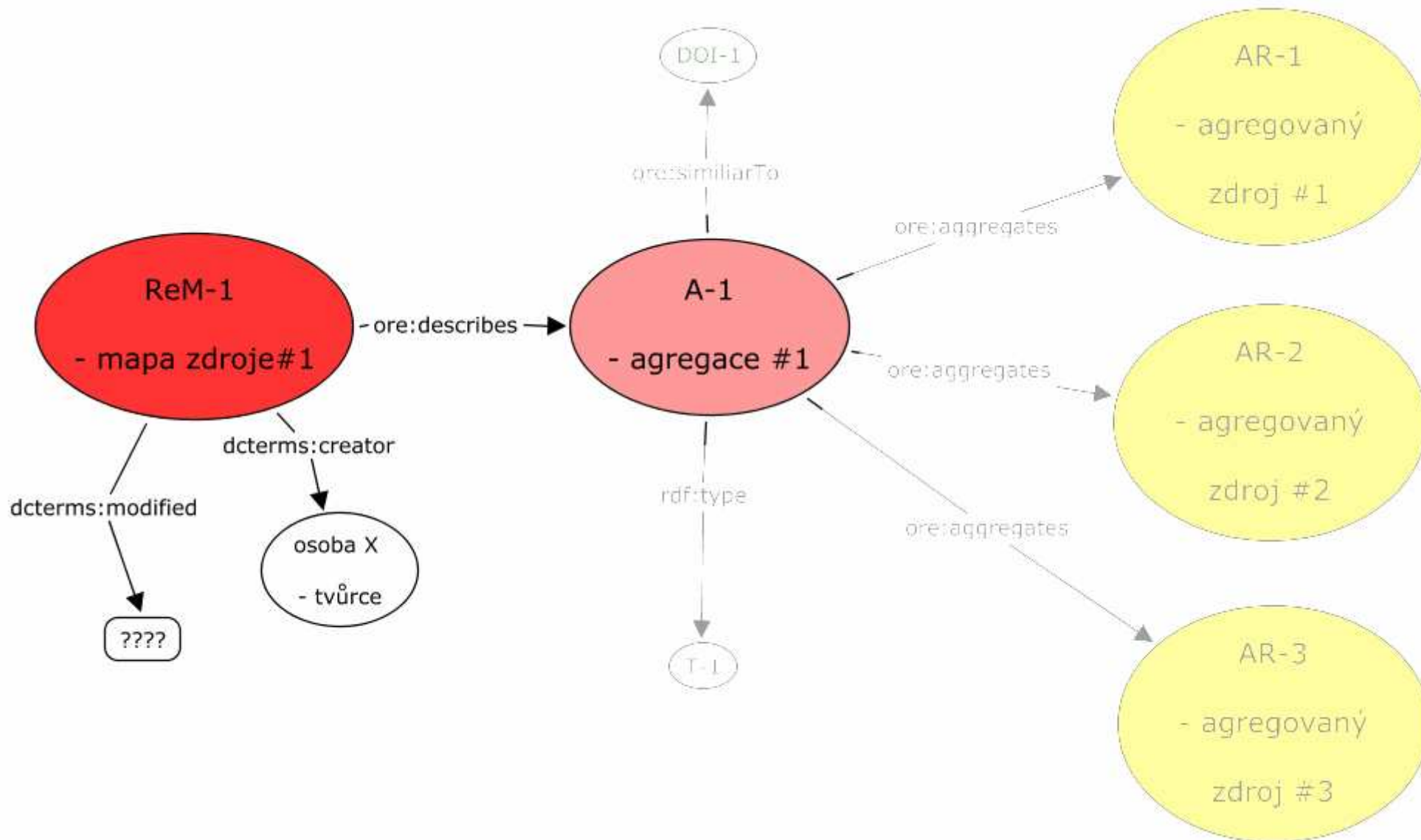
... lze si představit analogickou strukturu např. pro pokračující zdroj (seriál), VŠKP, e-learningový materiál, článek s obrázky a tabulkami ...



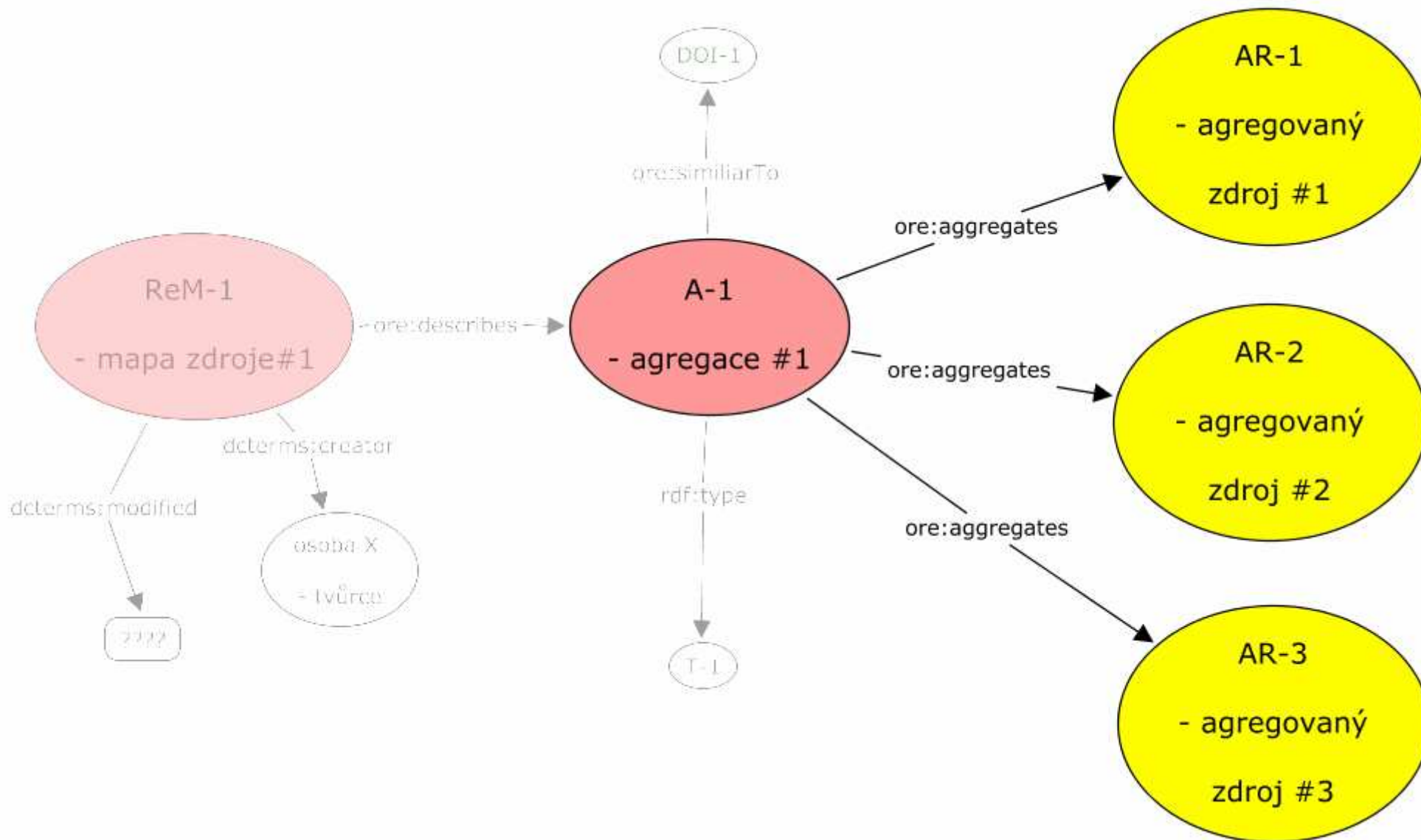
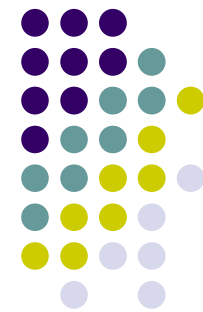
Podstata ORE: Mapa zdroje

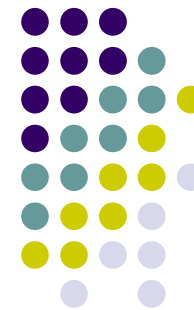


Zvýrazněná metadata zdrojové mapy

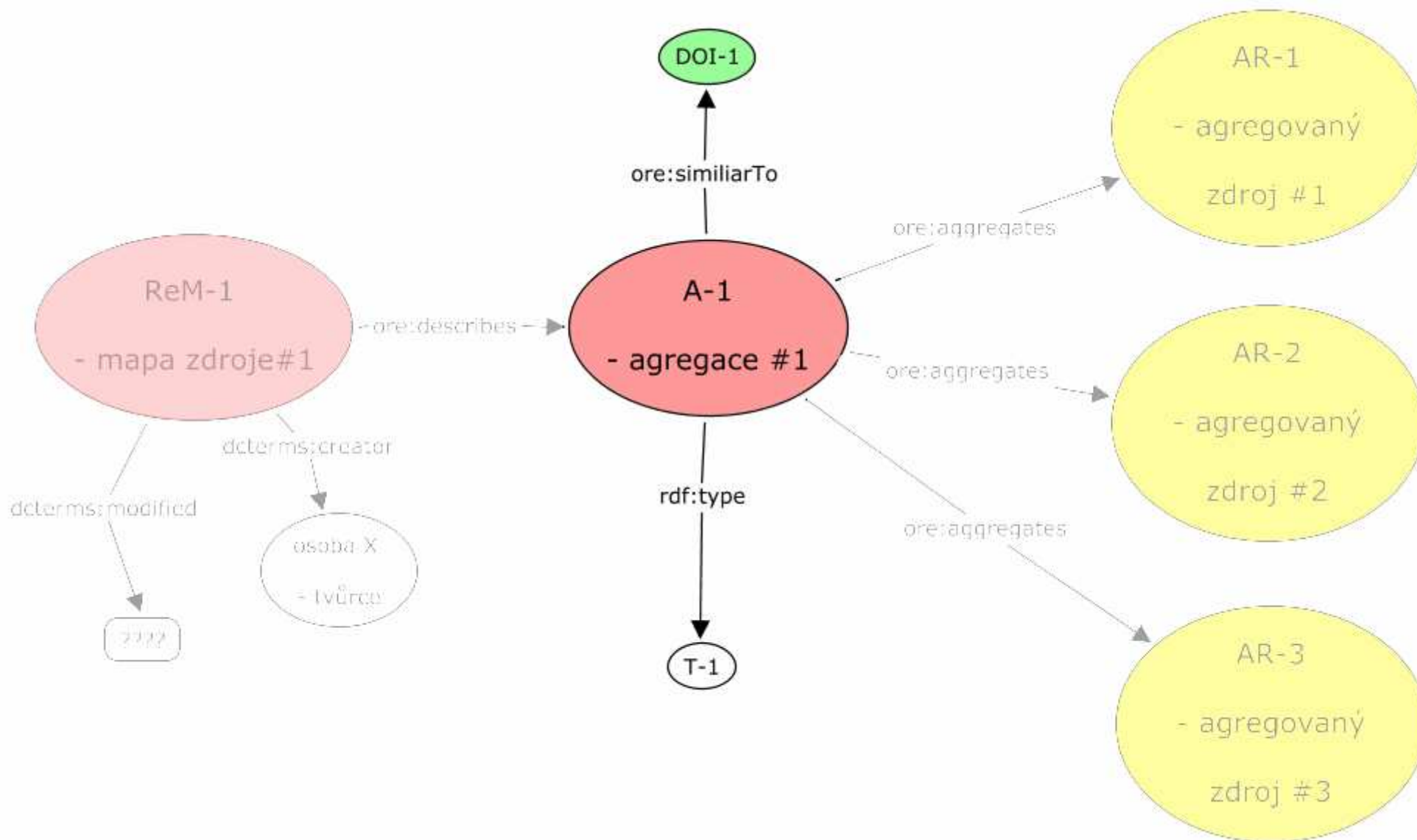


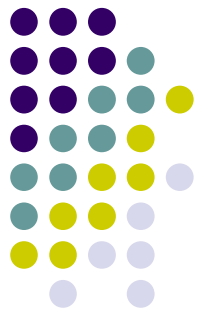
Zvýrazněné vazby mezi agregací a jejími zdroji





Zvýrazněná metadata agregace

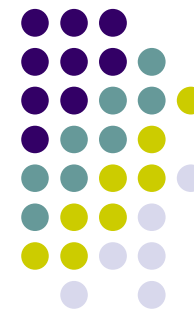




Reuse = znovuvyužití

- klíčový cíl celého úsilí
- komunikace hotových zdrojových map popisujících agregace skrz **proxy** mechanismy
- platí:
 - **jedna mapa** zdroje popisuje **jednu agregaci**
 - **jedna agregace** může být popsána **více zdrojovými mapami**
 - pokud je agregace popsána **vícero zdrojovými** mapami, jedna ze zdrojových map je **autoritativní**
 - agregace může být chápána jako zdroj a jako taková agregovaným zdrojem v jiné agregaci

Příklad syndikace ve formátu ATOM



```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">
  <title>Nové diplomky</title>
  <link href="http://vskp.cz/prirustky/" />
  <updated>2010-10-08T12:30:02Z</updated>
  <author>
    <name>František Vomáčka</name>
    <email>vomacka@vskp.cz</email>
  </author>
  <id>urn:uuid:80d71c70-c494-21d3-b91c-0095932f1de1</id>
  <entry>
    <author>Maruška Vopičková</author>
    <title>Ekonomické přínosy otevřených archivů na národní hospodářství</title>
    <link href="http://vskp.cz/dp/hdl:787.10/etd-aba-mgr-2009363" />
    <id>urn:uuid:8754a231-aca8-7baa-8eaa-98ca231baca6</id>
    <updated>2009-01-02T00:31:02Z</updated>
    <content>Popis obsahu...abstrakt...</content>
    <rights>Creative Commons Attribution Non-Commercial 3.0 License</rights>
  </entry>
  <entry>... atd ...</entry>
</feed>
```

Příklad serializace ORE ve formátu ATOM – zdroj (feed)



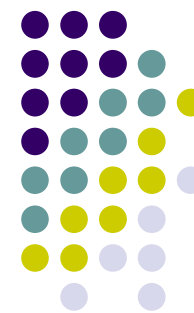
- „kostra“ zdrojové mapy

```
<?xml version="1.0" encoding="utf-8"?>
<feed xmlns="http://www.w3.org/2005/Atom">

  <id>http://arxiv.org/rem/astro-ph/0601007#aggregation</id>
  <link href="http://arxiv.org/rem/astro-ph/0601007" rel="self" type="application/atom+xml"/>
  <generator uri="http://arxiv.org/">arXiv.org e-Print Repository</generator>
  <updated>2007-10-10T18:30:02Z</updated>
  <category scheme="http://www.openarchives.org/ore/terms/"
term="http://www.openarchives.org/ore/terms/Aggregation" label="Aggregation" />

</feed>
```


Příklad serializace ORE ve formátu ATOM – agregované zdroje - záznamy



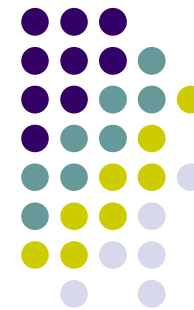
- jednotlivé záznamy tvořící agregaci

```
<entry>
  <id>http://oreproxy.org/r?what=http://arxiv.org/ps/astro-ph/0601007&where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
  <link href="http://arxiv.org/ps/astro-ph/0601007" rel="alternate" type="application/postscript"/>
</entry>

<entry>
  <id>http://oreproxy.org/r?what=http://arxiv.org/pdf/astro-ph/0601007&where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
  <link href="http://arxiv.org/pdf/astro-ph/0601007" rel="alternate" type="application/pdf"/>
</entry>

<entry>
  <id>http://oreproxy.org/r?what=http://arxiv.org/e-print/astro-ph/0601007&where=http://arxiv.org/rem/astro-ph/0601007%23aggregation</id>
  <link href="http://arxiv.org/e-print/astro-ph/0601007" rel="alternate"/>
</entry>
```

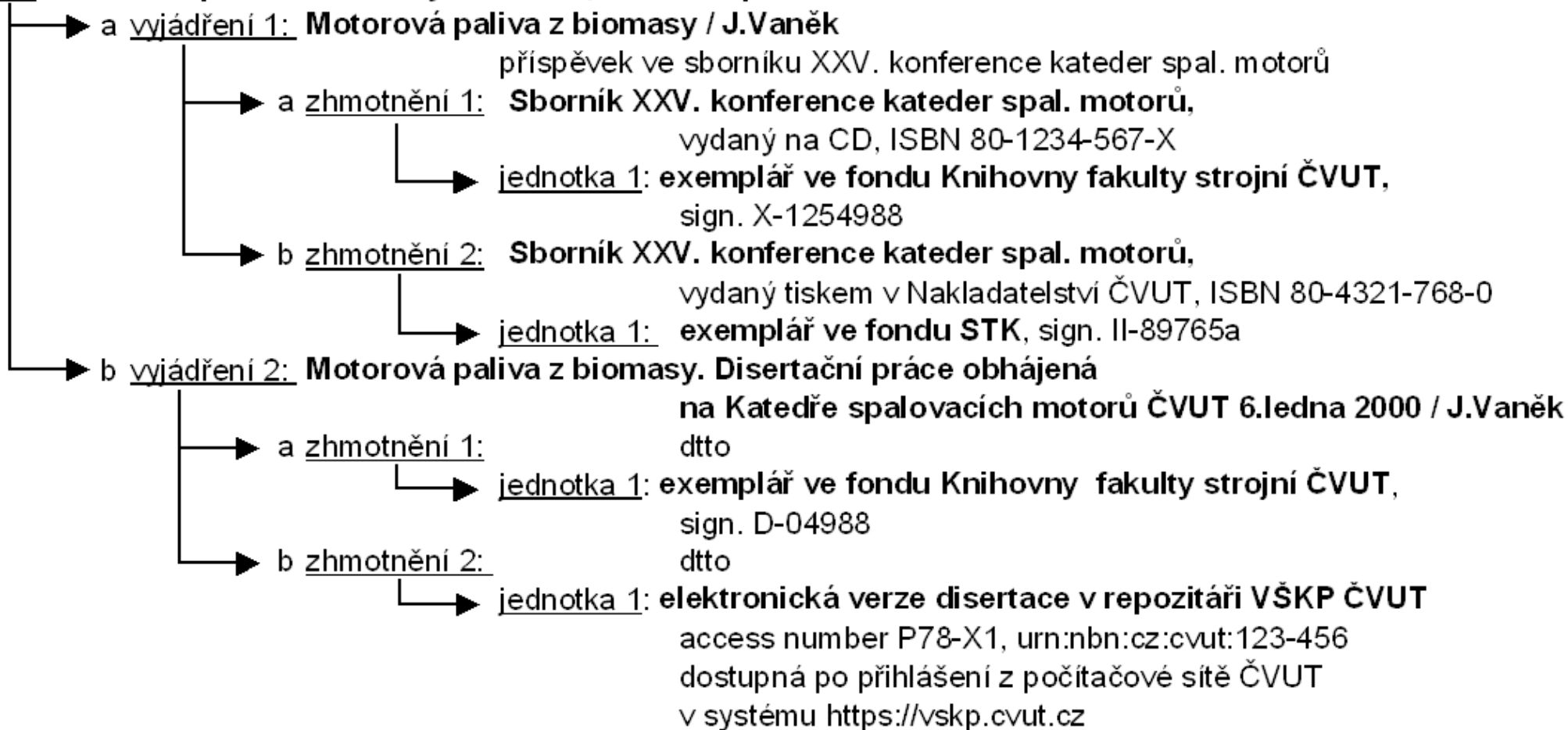
zdroj: <http://www.openarchives.org/ore/0.9/atom-implementation>

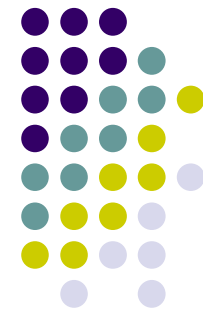


Příklad možného použití v kontextu šedé literatury...

- snímek z prezentace na Knihovnách současnosti 2006, navrhuje použít FRBR pro prezentaci výsledků vyhledávání

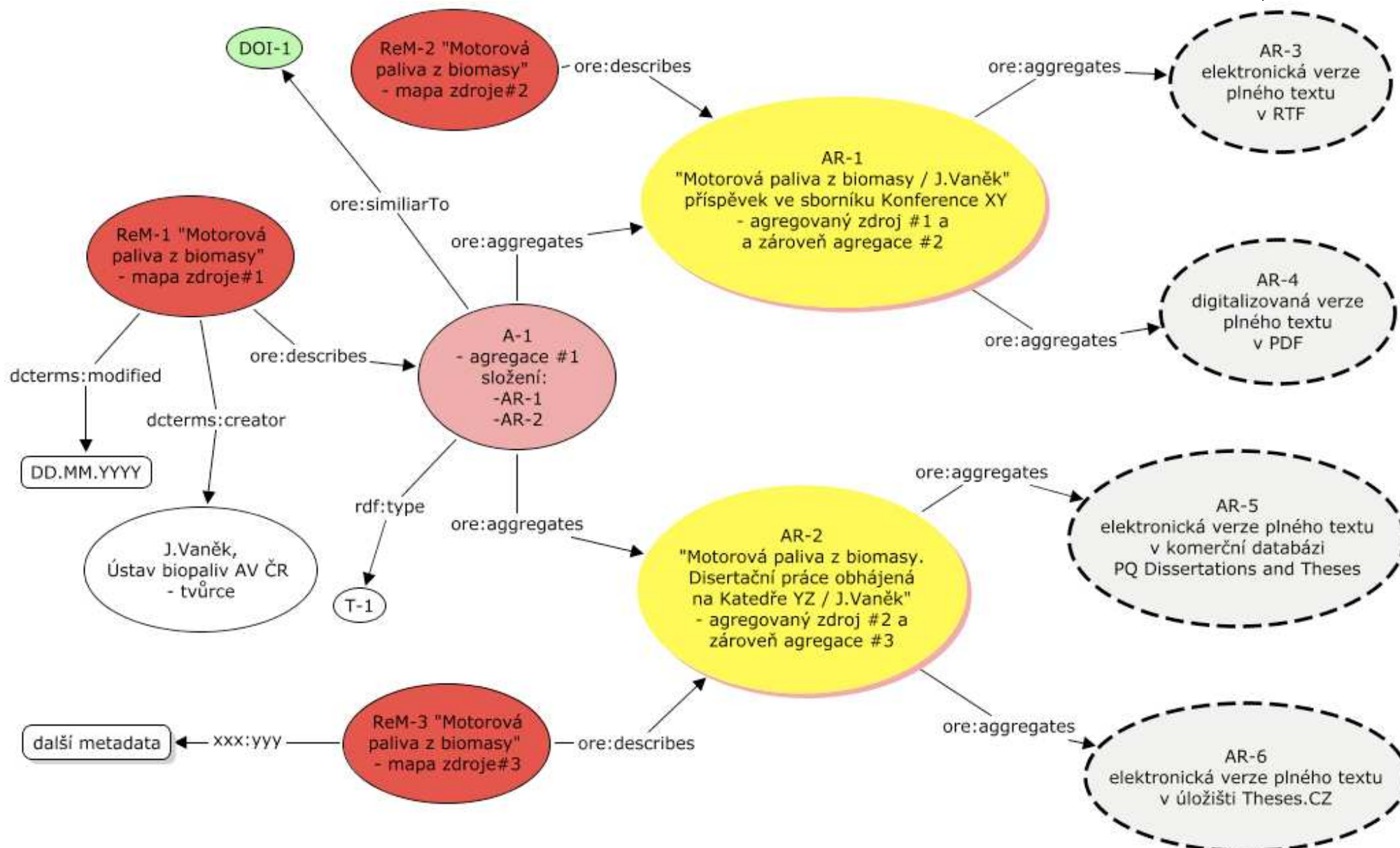
dílo: **Motorová paliva z biomasy / J.Vaněk, Ústav biopaliv AV ČR**

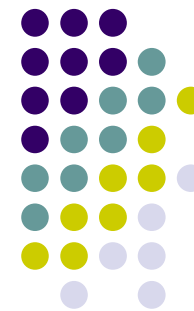




... a totéž pomocí ORE

(uspořádání vazeb mezi metadatovými záznamy 4 verzí 2 dokumentů)





Další zdroje

- OAI-ORE Google Group
<http://groups.google.com/group/oai-ore>
- ORE Wiki
<http://foresite.cheshire3.org/wiki/>
- ORE Challenge – soutěž o nejzajímavější implementaci (9.2008)
<http://www.openarchives.org/ore/RepoCamp2008/>
Vítězem vizualizační nástroj OREsome pro složité složené struktury na úrovni časopis-číslo-článek-stránkování

Logo OAI-ORE:



Děkuji za pozornost

petr.novak@ff.cuni.cz

