

UCD2 vs Ontologies

Sébastien Derriere

CDS, Observatoire astronomique de Strasbourg

derriere@astro.u-strasbg.fr



S. Derriere, Réunion MDA, 12 janvier 2004

UCD: Unified Content Descriptors

Description sémantique de données astronomiques.

UCD2: volontairement simples

- plutôt 20 caractères que 400
- plutôt 1 ou qq mots qu'un paragraphe de XML

Objectifs:

- description de contenu
- localisation de données pertinentes
- manipulations, comparaisons automatisées



S. Derriere, Réunion MDA, 12 janvier 2004

Ontologie:

Philo: Syn. de Métaphysique générale

Investigation sur le sens de l'être distingué des étants

Discours issu de la logique mathématique et de la linguistique, qui traite des termes utilisés pour désigner les êtres constitutifs de la réalité.

En pratique, une ontologie :

- permet la description formelle d'un domaine, compréhensible par les spécialistes du domaine et les programmes;
- rend explicite ce qui est habituellement implicite;
- résulte d'une approche consensuelle;
- facilite l'interopérabilité entre les outils.



S. Derrière, Réunion MDA, 12 janvier 2004

Ontologie:

Définition de:

- **concepts**, comparables aux classes OO (ex: étoile)
- un concept possède des **propriétés** (slots)
 - > distance
 - > type spectral
 - > température
- **facets**, qui sont des contraintes
 - > étoile n'est pas galaxie
- **instances**, qui sont les réalisations (ex: Vega)

Une propriété peut être commune à différents concepts.

ex: température d'un instrument



S. Derrière, Réunion MDA, 12 janvier 2004

Question lors de la construction du vocabulaire utilisé dans les UCD2:

Jusqu'où peut-on conceptualiser sans perdre la simplicité?

Nous avons voulu éviter d'en arriver à se demander:

« What is a quantity ? »

Nous définissons les UCD2 de façon simple. Si l'on veut résoudre des problèmes complexes, on utilise plusieurs UCD, ou des contraintes de la forme UCD=valeur.

**Le mot
primaire
porte
l'essentiel
du sens**

phot .mag; em .opt .R
phot .mag; em .opt .R .Halp
src .veloc
src .veloc .expansion
src .veloc .gc
src .veloc; stat .min



S. Derrière, Réunion MDA, 12 janvier 2004

UCD2 vs Ontologie:

Les UCD2 sont formés à partir d'une liste de mots.
Ceux-ci peuvent être réutilisés.

ex: **phys .temp; src** et **phys .temp; instr**

Certains concepts ont été introduits dans le vocabulaire. Ils ne peuvent cependant pas être utilisés comme mot primaire.

Certains concepts n'ont pas été introduits dans le vocabulaire car ils peuvent être flous. ex: étoile ou galaxie. Mais src.class existe !

Dans certains catalogues, il y a des étoiles et galaxies mélangées et rien pour les distinguer avec certitude.



S. Derrière, Réunion MDA, 12 janvier 2004

UCD2 vs Ontologie:

Comment faire de l'inférence avec des contenus hétérogènes ?



Fruit	Price	Color
item.fruit.type	item.price	item.color
orange	0.50	orange
grape	0.75	violet
apple	0.39	red
...

item.fruit.type='apple'

Price	Color
item.price	item.color
0.19	green
0.25	yellow
0.39	red



S. Derrière, Réunion MDA, 12 janvier 2004

UCD2 vs Ontologie:

Comment décrire les relations mathématiques entre les valeurs?

Ex: B-V indice de couleur

- définition de `phot.colorIndex.B-V` ?
- obtenu à partir de la formule:
 $(\text{phot.mag;em.opt.B}) - (\text{phot.mag;em.opt.V})$
ou bien à partir d'une modélisation.

Cf Stanford Ontology for engineering mathematics.

EngMath (Gruber and Olsen 1994)

<http://www-ksl.stanford.edu/knowledge-sharing/papers/engmath.html>

```
(scalar-quantity k)
(= (physical-dimension k)
  (/ force-dimension length-dimension))
(scalar-quantity d)
(= (physical-dimension d) length-dimension)
(scalar-quantity dn)
(= (physical-dimension Dm) length-dimension)
(scalar-quantity N)
(= (physical-dimension N) identity-dimension)
(scalar-quantity G)
(= (physical-dimension G)
  (* force-dimension
    (expt length-dimension -2)))
(= k (/ (* (expt d 4) G) (* 8 (expt Dm 3) N)))
(= G (* 11.5 (expt 10 6) psi))
```

$$k = \frac{d^4 G}{8 D^3 N}, \quad G = 11,5000 \text{ kpsi}$$



S. Derrière, Réunion MDA, 12 janvier 2004

UCD2 vs Ontologie:

Matching function ?

Trouver les UCD de sens voisins.

