the IVOA Standard Vocabulary
for events,
object types,
instruments

A. Preite Martinez - inaf/it
F. Hessman – mo.ne.t./de
S. Lesteven – cds/fr
R. Williams - caltech/usa
...

the IVOA Standard Vocabulary Status – recent docs & events

- VOEvent I - Caltech — 2004
- UCDs, Doc (Rec)
- VOConcepts
- SV draft (v0.3) — 2005
- VOEvent II - Tucson
- UCDs, List (Rec)
- SV draft (v0.8) — 2006

the IVOA Standard Vocabulary
why, what, how

- Standardise the vocabulary used to describe the data the IVOA community is working on
- Categories (of immediate IVOA usage):
  - Processes
  - Object types
  - Intrumentation (imagers, spectrographs)   + theory, sim., ...
- Use UCD-like syntax:

  root-concept [[[.subordinate-concept(s)] ; concept-specification(s)]..]
- Based upon:
  - A&A keywords (324)
  - Simbad object types (155),
  - UCD1+ word-list (379)
  - Literature (6 Journals, last 6 years)

the IVOA Standard Vocabulary Syntax rules (FH + APM)

- Use of normal vocabulary, abbreviations when generally accepted (ISM), hyphens and capital letters when normally used (X-ray, HII) or to bridge words (dwarfNova, absLineSystem)
- Use UCD-like dots to generate hyerarchies and semicolons to concatenate concepts:
  - stars.binary.low-mass;em.X-ray (LMXB)
- object names assumed as archetypal of an object class (RRLyrae) or discoverer's name describing an object class (Seyfert) will be sub-words of concept "class"

the IVOA Standard Vocabulary Syntax rules (FH + APM)

- if an object or a class of objects are commonly designated by an abbreviated form or acronym, the standard word can be replaced by its abbreviated form or acronym
  - *process.variation.burst;em.gamma* == ***alias.GRB***
  - *stars.variable.cataclysmic;class.AMHer* == ***alias.AMHer***
- to indicate that an object (objectX) is a member of a multiple/composite object (objectY), the qualifier "stat.member" is used:
  - *objectX;objectY;[objectZ;...]stat.member*
- If an object is the subject of a simulation, the qualifier "*stat.sim*" can be used.

the IVOA Standard Vocabulary general problems (1)

- **If the object is simple, the SV representation is simple (and, contrary to the title, this is not a problem!!):**

  "Star of spectral type B" => "*stars.spType.B*"

  "Cluster of stars" => "*stars.cluster*"

  "Cluster of galaxies" => *"galaxies.cluster"*

  "Active galaxy" => *"galaxies.active"*

  "White dwarf" => *"stars.whiteDwarf"*

  "Elliptical galaxy" => *"galaxies.elliptical"*

  "HII region" => *"ISM.region.HII"*

the IVOA Standard Vocabulary .. to general problems (2)

- **If the object is complex (the majority?) the SV representation (i) is long/complex, (ii) could be not unique:**

  “Star in cluster of stars” => “*stars;stars.cluster;stat.member*”

  “Semi-regular pulsating star” =>
  “*stars.variable;process.variation.quasi-periodic”,*

  but why not:

  “stars.variable.quasi-periodic” or:

  “stars; process.variation.quasi-periodic”

  “cataclysmic variable” =>

  stars.binary.close || *stars.variable.cataclysmic* ||
  stars.variable..dwarfNova* || *stars.*;class.UGem*

the IVOA Standard Vocabulary general problems (3)

“Eclipsing binary of Algol type” that could be declined as:

“stars.binary;process.eclipse;class.Algol”

“stars.binary.eclipsing.class-Algol” ....

“alias.Algol”

“Galactic nebula” => *“ISM.nebula.galactic” or “ISM.nebula;Galaxy;stat.member”*

“Supenova of type Ia” => *“stars.superNova.Ia” or “stars.superNova.Ia;process.explosion”, or “alias.SNIa”*

the IVOA Standard Vocabulary general problems (4)

- **the SV representation is sometimes implicit as our language:**
  - a "Low-mass X-ray binary" is a complex system made of (i) an evolved low-mass star, (ii) the remnant of a high-mass star, (iii) an accretion disk, heated by loss of gravitational energy of matter from the low-mass star, origin of the X-ray emission.

    In SV => "*stars.binary.low-mass;em.X-ray*", forgetting about half a dozen physical processes (last but not least: "*process.emission*")
  - In a similar way, we say "Pulsar" and the SV suggests *"stars.pulsar",* but an equally effective way of using words of the SV would be to write *"stars.neutron;process.rotation;process.emission;em.radio"* and we still miss the "very fast" qualifier!

the IVOA Standard Vocabulary general problems (5)

- **So in most cases the SV representation is the result of :**
  (i) a choice between “object” or “hierarchy” and “process”
  (ii) a choice of the process most relevant to the observer

- **How to deal with “common-language” qualifiers:**
  Late/early, high/medium/low, hot/cold, bright, fast/slow, *etc.*
  My suggestion is to put them in the hierarchy rather than in “stat.*” to avoid ambiguities in the concatenation of concepts.

different views of the same object:

variable

cataclysmic | explosive | nova-like | dwarfNova

close binary (eclipsing binary)

dwarf | late-type K-M giant filled Roche lobe

white dwarf, accretion disk, hot spot

X-ray source

low-state, emission lines

high-state, absorption lines

class UGem (SSCyg, SUUMa, ZCam)

the IVOA Standard Vocabulary .. is that all?

- **Are we sure we are getting all the necessary semantic richness with the few hundred keywords of A&A?**

  Some work was done to dig into the most recent literature searching for the way the "users" describe events, processes, astronomical objects, instruments

  The approach was:

  (1) derive statistics on the use of all "A&A keywords" in recent years

  (2) select keywords with high frequencies (>100 citations/year)

  (3) find all the sentences (in the abstract) containing the selected keywords

  (4) extract all the different meaningful semantic content

Different expressions with different semantic meaning of selected objects/keywords in the literature (years: 2000-2005):

Detail of:	journals	issues	sentences	expressions	DR(%):
abundances	2	235	2792	45	1.6
accretion	2	248	2672	130	4.9
circumstellar	2	204	592	129	21.8
galaxies active	6	445	854	41	4.8
galaxies cluster	6	781	4481	120	2.7
GRB	6	396	2205	37	1.7
hydrodynamics	2	115	192	8	4.2
mass-loss	2	122	261	30	11.5
neutron stars	2	214	1486	52	3.5
numerical	2	229	1170	76	6.5
stars binary close	6	277	409	37	9.0
stars formation	2	267	1769	129	7.3
supernova	2	248	1969	34	1.7
X-ray	6	856	14995	740	4.9

supernova core-collapse	**stars.superNova;process.collapse**
supernova explosion	**stars.superNova;process.explosion**
supernova explosion thermonuclear	**stars.superNova;process.explosion.thermonuclear**
supernova fallback	**stars.superNova.fallback**
supernova blast-wave	**stars.superNova;process.shock**
supernova shock	**stars.superNova;process.shock**
supernova ejection pencil-beam	**stars.superNova;process.mass-loss.jet**
supernova kick	**stars.superNova;process.angMomentumTransfer**
supernova enrichment	**stars.superNova;process.enrichment**

neutron star accreting	**stars.neutron;process.accretion**
neutron star bursting	**stars.neutron;process.variation.burst**
neutron star coalescing	**stars.neutron;process.merging**
neutron star cooling	**stars.neutron;process.cooling**
neutron star evolution	**stars.neutron;process.evolution**
neutron star glitches	**stars.neutron;process.variation.glitch**
neutron star magnetar/magnetized	**stars.neutron.magnetar**
neutron star mergers	**stars.neutron;process.merging**
neutron star millisecond	**stars.neutron;process.rotation**
neutron star non-rotating	**stars.neutron;process.rotation;stat.not ?**
neutron star oscillations	**stars.neutron;process.pulsation**
neutron star pulsations	**stars.neutron;process.pulsation**
neutron star rotating	**stars.neutron;process.rotation**
neutron star rotating differentially	**stars.neutron;process.rotation.differential**
neutron star rotating fast	**stars.neutron;process.rotation.fast**
neutron star rotating slow	**stars.neutron;process.rotation.slow**
neutron star spin/spinning	**stars.neutron;process.rotation**
neutron star spinning rapidly	**stars.neutron;process.rotation.fast**

(4-c/4-w): bright soft X-ray loops
(4-c/4-w): modeled X-ray burst oscillations
(4-c/5-w): center-filled X-ray properties of SNRs
(4-c/5-w): high-mass X-ray binary pulsar
(4-c/5-w): transient low-mass X-ray binaries
(4-c/5-w): X-shaped soft X-ray morphology
(4-c/6-w): high-luminosity accretion-powered X-ray pulsars
(4-c/6-w): intermediate-mass black hole X-ray binaries
(4-c/6-w): redshift-limited, X-ray-selected cluster sample
(4-c/6-w): stellar-mass black hole X-ray binaries
(4-c/7-w): X-ray-selected high-redshift radio-quiet quasar
(5-c/5-w): intrinsically absorbed X-ray emission regions
(5-c/5-w): inverse S-shaped X-ray sigmoid structures
(5-c/5-w): transient binary supersoft X-ray source
(5-c/6-w): accretion-powered, transient, millisecond X-ray pulsars
(5-c/6-w): eclipsing halo low-mass X-ray binary
(5-c/6-w): extended electron-scattered hard X-ray emission
(5-c/6-w): low-luminosity globular cluster X-ray sources
(5-c/6-w): X-ray-bright, eclipsing magnetic cataclysmic variable
(5-c/7-w): type I X-ray-bursting low-mass X-ray binary
(5-c/7-w): X-ray-heated Roche lobe-filling secondary star