Registry Relational Schema at IA2
Implementation Notes

Marco Molinaro
INAF-OATs IA2

[Registry Session]

IVOA Interoperability Meeting, Heidelberg 12-17 May 2013

Outline

- Available
  - where we started from
- Missing
  - what needed to be done
    - implementation steps
- RDB schema in MySQL
  - focus on implementation notes
- Validation on ingestion
  - focus on resources' requirements
- Conclusions

Starting from …

- RDB schema description
  - Registry Relational Schema document
- VO Resources' records
  - GAVO endpoint
    - already exposing "RegTAP"
- IA2TAP web application
  - TAP service generator
    - MySQL compatible

… we needed to:

- Generate the RDB schema in MySQL
  - Implementation Notes appendix
- Prepare the TAP_SCHEMA records for it
  - and configure the IA2TAP webapp
- “harvest” the GAVO registry
- Ingest records into the RDB schema
  - resource validation issues

MySQL RDB schema

- Straightforward … but
  - MySQL limitations on VARCHAR (e.g.)
    - “Implementation Notes” appendix to the document
- Implementation issues led to
  - more general considerations on resource integrity preservation
    - “Implementations MUST NOT truncate strings of length equal or smaller than…”
  - tips for user interface development

Implementation Notes appendix

- Mainly devoted to preserve minimum length of resource's elements
  - Record integrity
- Categorized string field types
  - Setting minimums based upon resource records as of February-March 2013

TAP_SCHEMA and service config

- Registry Relational schema doc
  - Direct translation into inserts for TAP_SCHEMA
    - Utype dependency
- IA2TAP webapp
  - few configs
  - some issues (identified during test phase)
    - not directly Registry Relational Schema related

Resource retrieval

- Many ways, choosen one
- GAVO registry TAP endpoint
  - Retrieval by ADQL queries
  - We used no dedicated coding
    - STILTS scripting
      - including record filtering
- Fast dump (~2 min. in XML format)
  - Various attempts, also due to schema changes
- Only 1 major issue
  - NULL vs. EmptyString

Resource ingestion

- From full dump
  - table-by-table ingestion
  - ~13000 resources
    - time consuming (first ingestion)
- Direct resource validation by schema definition and constraints
  - Led to a set of iterations consisting of
    - RDB schema definition and generation
    - records ingestion
    - Identify issue (this shouldn’t be null, shortName has max length 16, does it make sense to leave this empty, …)
    - schema changes and/or ingestion modifications

Conclusions

- Implementation notes: to be kept in mind rather than be considered requirements
- NULL allowed and "empty" allowed resource fields (and other records requirements): DM to work on it?
- ...some bug-fix left to do, not RegTAP dependent

Thank you!

THE IA2 TEAM