

CC71X : La Web de Datos

# Grandes Volúmenes de Datos en Astronomía

Guillermo Cabrera  
[gcabrera@dcc.uchile.cl](mailto:gcabrera@dcc.uchile.cl)

2 de Mayo, 2011



# Introducción

- Astrónomos se enfrentan al problema del manejo de grandes volúmenes de datos.
- “Astroinformática”: disciplina enfocada en resolver problemas informáticos para la astronomía.
- Observatorios Virtuales



# Contenido

- **Astroinformatics: a 21st Century Approach (2009)**
- **Stellaris: An RDF-based Information Service for AstroGrid-D (2007)**
- **A standard transformation from XML to RDF via XSLT (2009)**



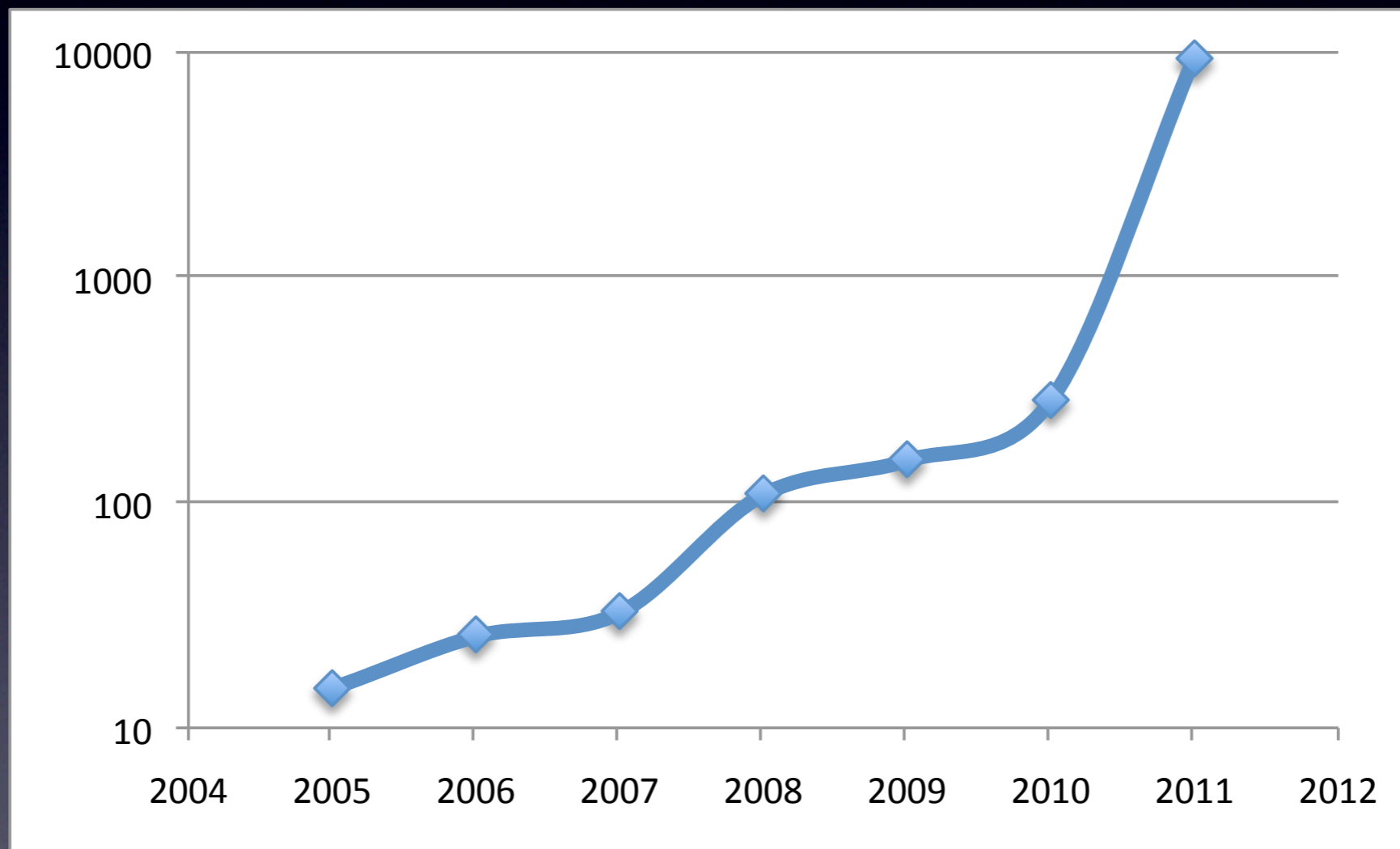
# Astroinformatics: a 21st Century Approach

Astro2010 State of the Profession Position Paper (March 2009)

Kirk D. Borne and 81 coauthors

- Crecimiento de datos astronómicos de GB a TB durante la década pasada y a PB durante la próxima década.
  - Organización de datos.
  - Descripción de datos.
  - Taxonomías para clasificar datos astronómicos.
  - Ontologías conceptuales.
  - Data mining.
  - Machine learning.
  - Visualización.
  - Astroestadística.

# Número de Resultados en Google de “Astroinformatics” por Año





# Observatorios Virtuales

- Gran número de imágenes y catálogos.
- Distintos telescopios.
- Distintos data centres con diferentes datos, software y hardware.





# Observatorios Virtuales

- Gran número de imágenes y catálogos.
- Distintos telescopios.
- Distintos data centres con diferentes datos, software y hardware.





# GAVO: AstroGrid-D

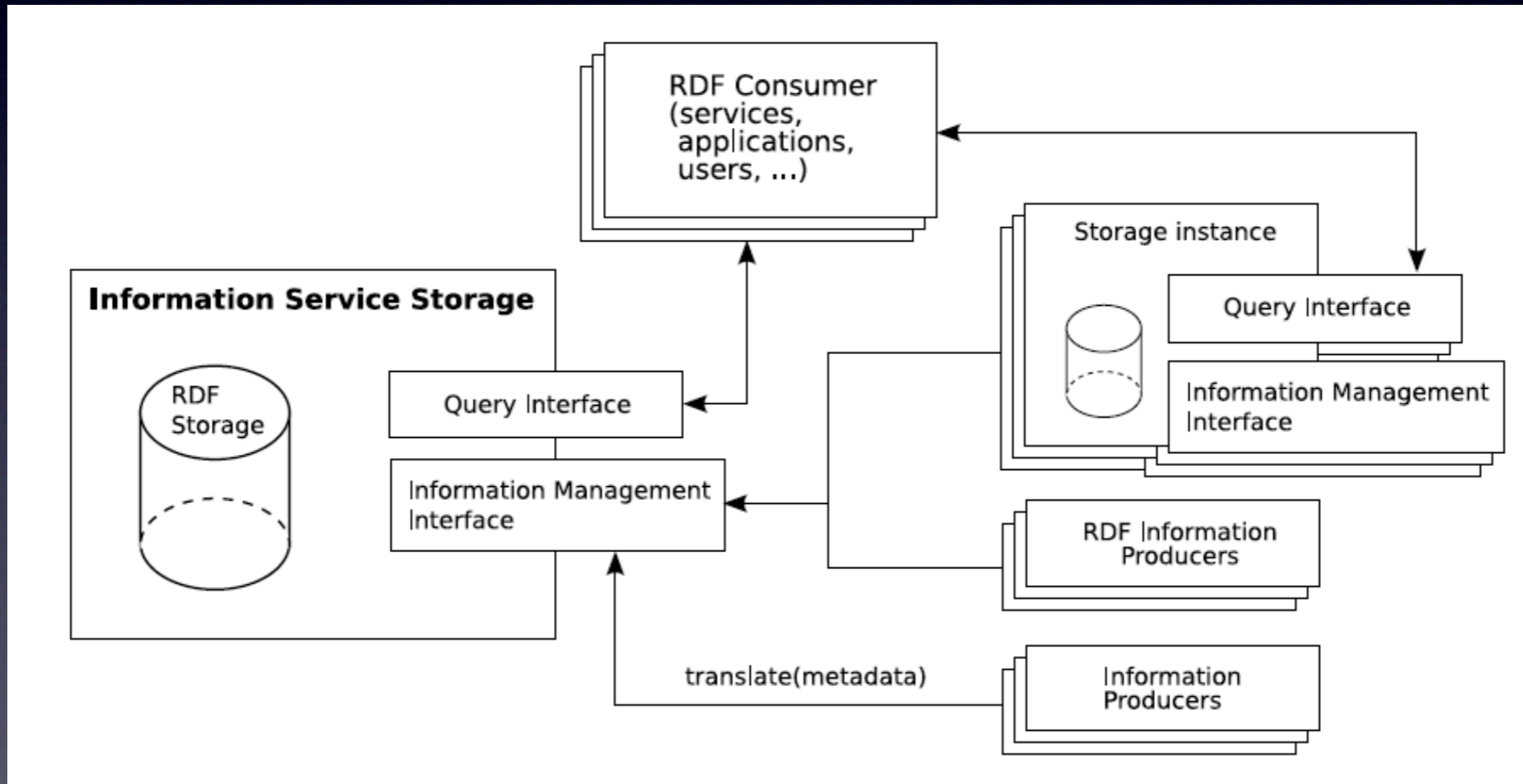
## Tecnología Grid para Aplicaciones científicas

- Astrónomos y científicos en computación alemanes aplicando tecnologías grid para aplicaciones científicas.
- Stellaris: An RDF-based Information Service for AstroGrid-D (2007)
- A standard transformation from XML to RDF via XSLT (2009)



# Stellaris: An RDF-based Information Service for AstroGrid-D

M. Hogqvist, T. Roblitz, A. Reinefeld, German e-Science 2007





# A standard transformation from XML to RDF via XSLT

F. Breitling, *Astronomical Notes* 2009

- Gran parte de información astronómica en XML.
- Necesidad de transformar de XML a RDF.
- Soluciones actuales complejas y requieren de software no disponible o no funciona para información astronómica.
- Virtuoso: XML a RDF, necesita de esquema específico para los datos, por lo que no puede ser aplicado directamente.



# A standard transformation from XML to RDF via XSLT

F. Breitling, Astronomical Notes 2009

- **Evitar nodos blancos:**
  - No se puede acceder a ellos, por lo que algunas operaciones (reemplazos) no se pueden realizar.
  - Desempeño de SPARQL disminuye.
- **Mapeo uno a uno de XML a RDF:**
  - Para pasar de RDF a XML también.
- **Independencia de XML schemas:**
  - Poder aplicar la transformación directamente.

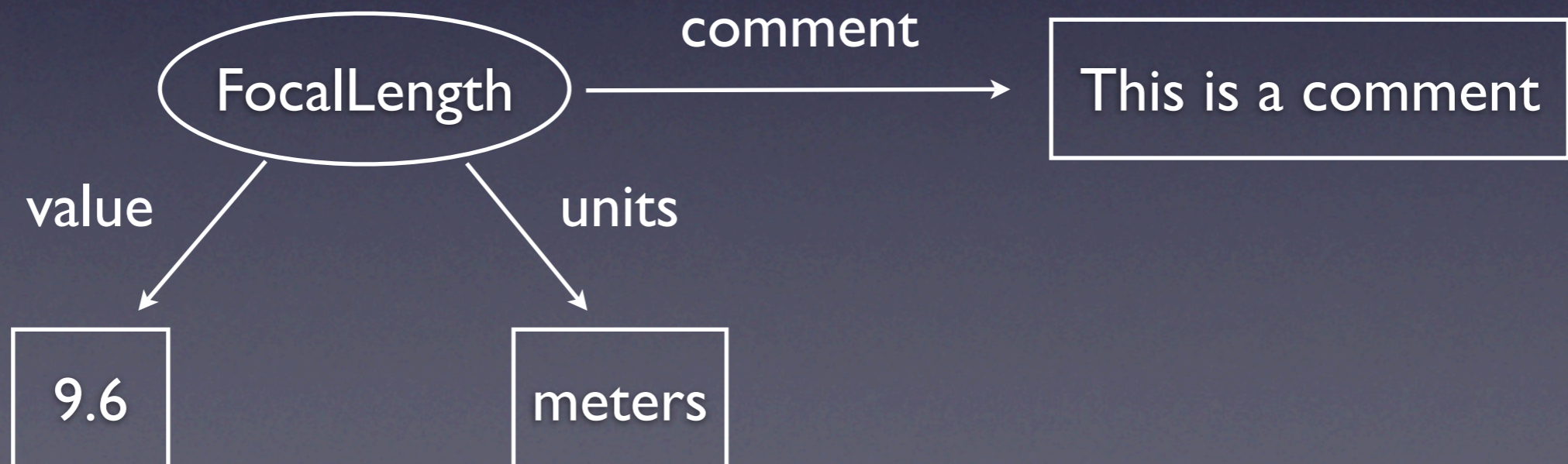


# A standard transformation from XML to RDF via XSLT

F. Breitling, Astronomical Notes 2009

- **Atributos, Texto y Comentarios**

```
<FocalLength units="meters">  
  <!-- This is a comment-->  
  9.6  
</FocalLength>
```



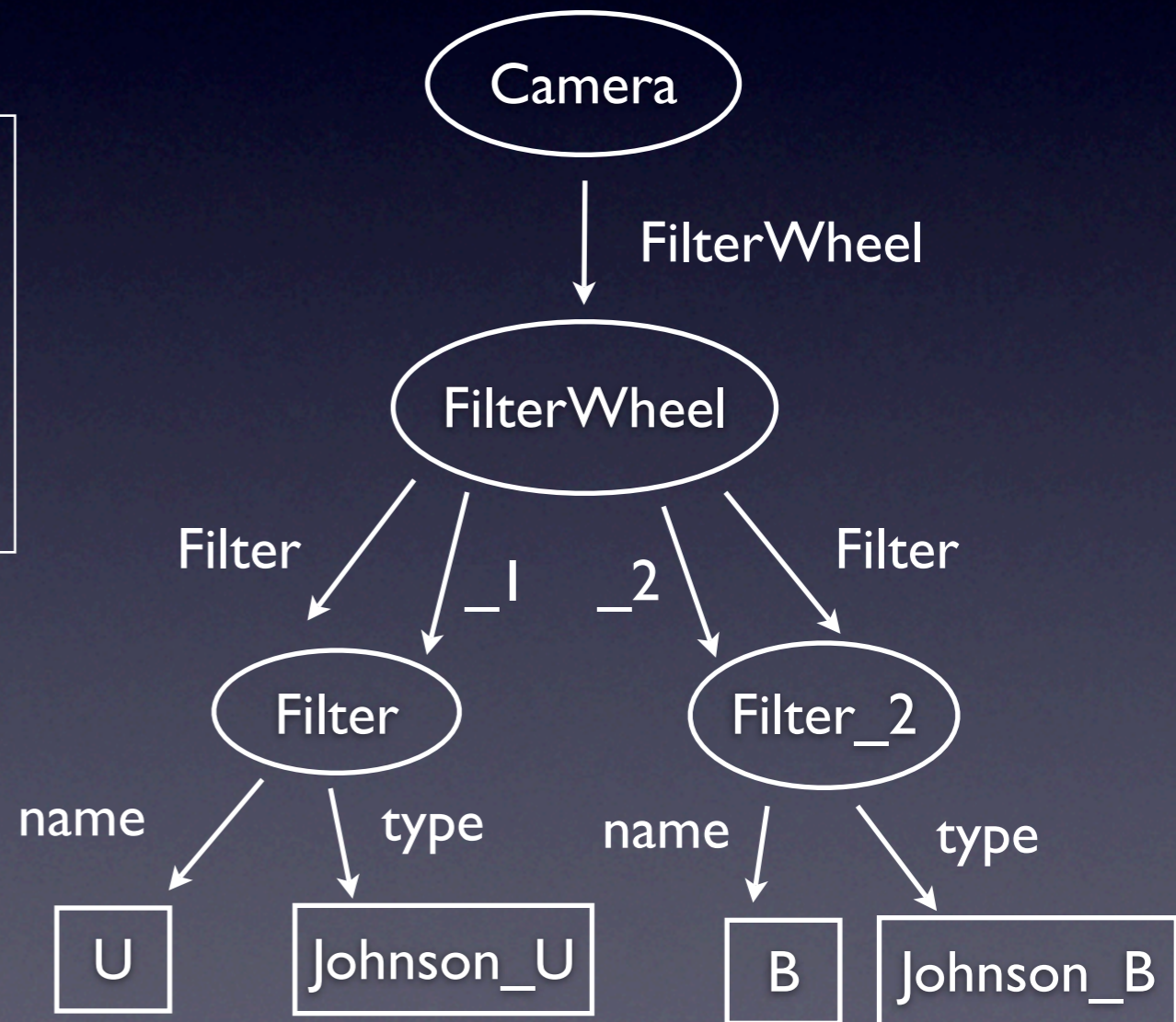


# A standard transformation from XML to RDF via XSLT

F. Breitling, Astronomical Notes 2009

## ● Elementos

```
<Camera>
  <FilterWheel>
    <Filter type="Johnson_U" name="U"/>
    <Filter type="Johnson_B" name="B"/>
  </FilterWheel>
</Camera>
```





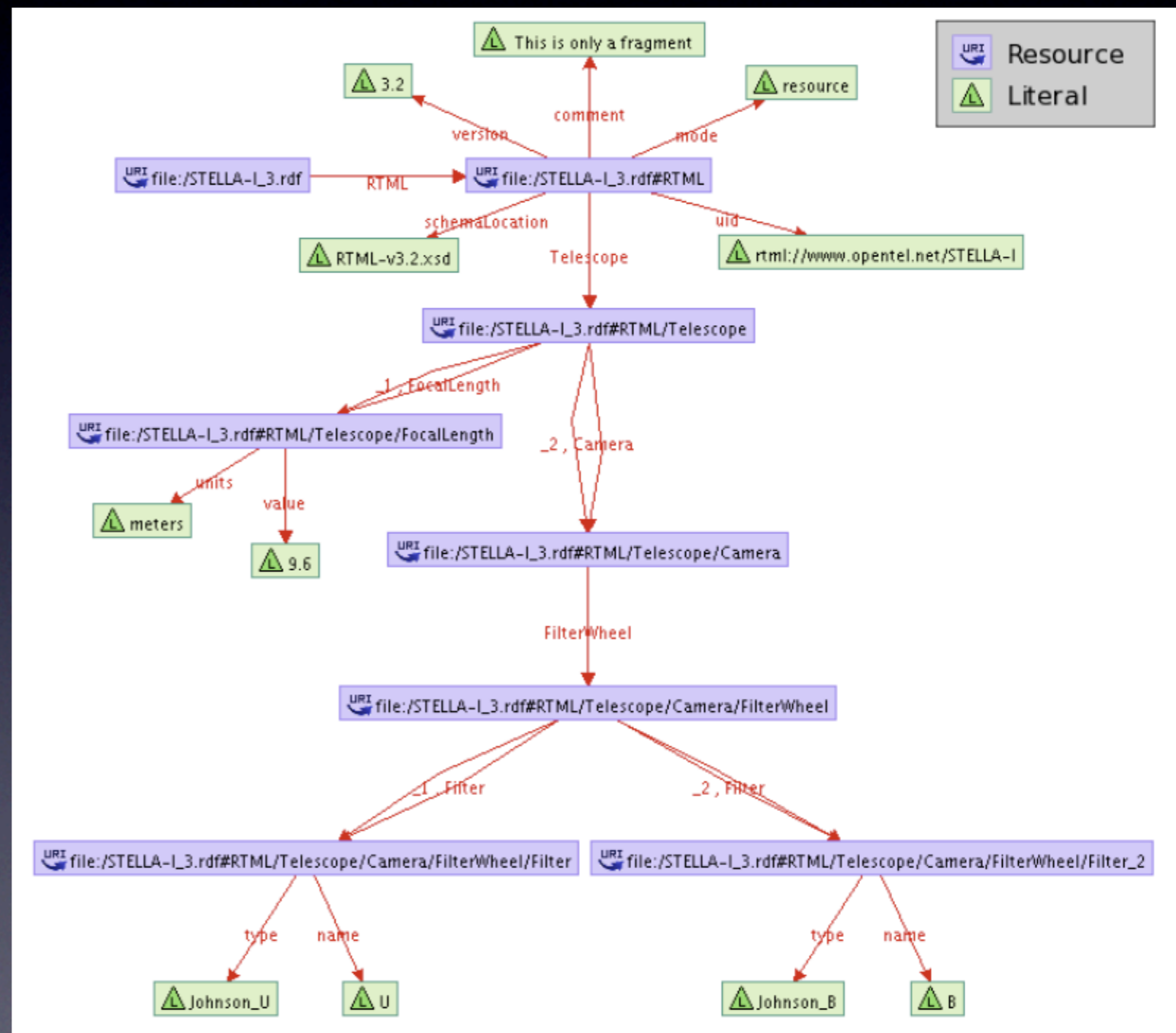
# A standard transformation from XML to RDF via XSLT

F. Breitling, Astronomical Notes 2009

**Listing 1:** Description of the robotic telescope STELLA-I.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <RTML version="3.2" mode="resource" uid="rtml://
  www.opentel.net/STELLA-I"
3   xmlns:xsi="http://www.w3.org/2001/XMLSchema-
  instance"
4   xmlns="http://www.rtml.org/v3.2" xsi:
  schemaLocation="RTML-v3.2.xsd">
5   <!-- This is only a fragment -->
6   <Telescope>
7     <FocalLength units="meters">9.6</FocalLength
8     >
9     <Camera>
10      <FilterWheel>
11        <Filter type="Johnson_U" name="U" />
12        <Filter type="Johnson_B" name="B" />
13      </FilterWheel>
14    </Camera>
15  </Telescope>
16 </RTML>
```

RTML: Remote Telescope Markup Language





# Conclusiones

- Necesidad de ontología para modelo RDF astronómico.
- Necesidad de crear comunidad interdisciplinaria.