Exchanging observations and measurements: a generic model and encoding

Simon Cox
Research Scientist
22 May 2007

www.csiro.au

CSIRO

Science relies on observations

- **Provides evidence & validation**
- **Involves sampling**

- **This paper is about a domain-independent terminology and information-model**

- **Fast overview, much more detail available …**

Examples

- **The 7th banana weighed 270gm on the kitchen scales this morning**
- **The attitude of the foliation at outcrop 321 of the Leederville Formation was 63/085, measured using a Brunton on 2006-08-08**
- **Specimen H69 was identified on 1999-01-14 by Amy Bachrach as Eucalyptus Caesia**
- **The image of Camp Iota was obtained by Aster in 2003**
- **Sample WMC997t collected at Empire Dam on 1996-03-30 was found to have 5.6 g/T Au as measured by ICPMS at ABC Labs on 1996-05-31**
- **The X-Z Geobarometer determined that the ore-body was at depth 3.5 km at 1.75 Ga**
- **The simulation run on 2004-09-09 indicated that the pressure in the hanging-wall at 618 Ma was reduced 4 MPa**

What is “an Observation”

- **Observation act involves a procedure applied at a specific time *(Fowler & Odell, 1997ish)***
- **The result of an observation is an estimate of some property**
- **The observation domain is a feature of interest at some time**

- **[0..*] locations may be of interest, associated with the procedure and feature of interest**

Observed property

- **Observed property**
  - Length, mass, temperature, shape
  - location, event-time, orientation
  - colour, chemical concentration
  - count/frequency, presence
  - species or kind (classification)
- **Expressed using a reference system or scale**
  - Scale may also be ordinal or categorical
  - May require a complex structure
- **“Sensible”, but not necessarily physical …**

Feature-of-interest

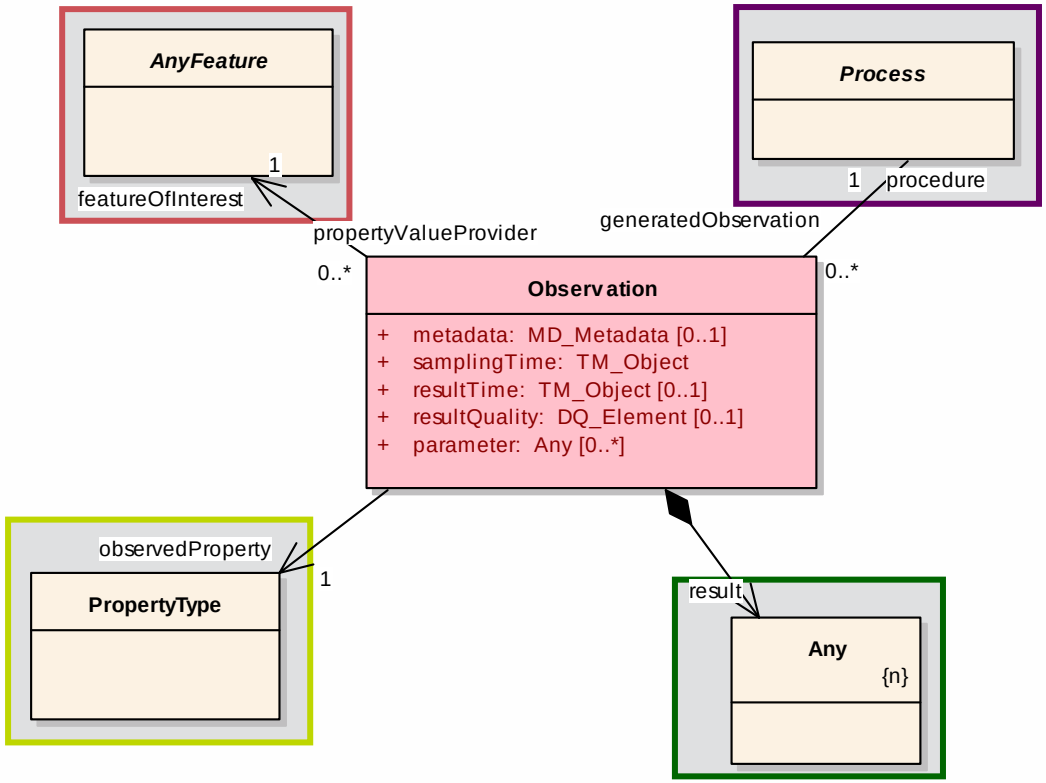
- **The observed property is associated with something**
  - Location does not have properties,
    the *substance* or *object* at a location does
  - Observed property must be logically consistent with the feature-of-interest
    - E.g. rock-*density*, pixel-*colour*, city-*population*, ocean-*surface-temperature*

- **… i.e. the Observation "target"**

Procedure

- **Instruments & Sensors**
  - Respond to a stimulus from local physics or chemistry
  - Intention may concern local or remote source
  - Sample may be in situ or re-located

Generic pattern for observation metadata



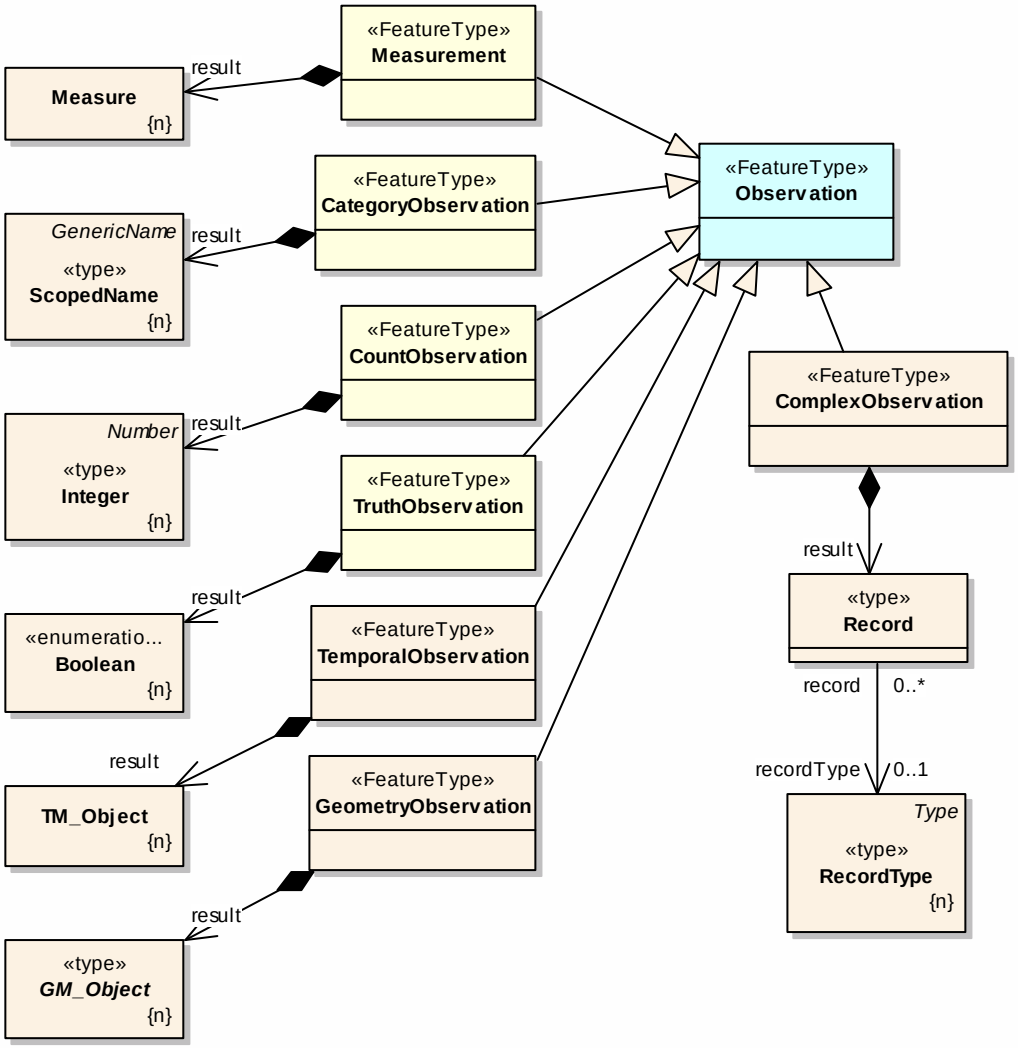
An **Observation** is an action whose **result** is an **estimate** of the **value** of some **Property** of the **Feature-of-interest**, obtained using a specified **Procedure**

Feature-of-interest concept reconciles remote and in-situ observations

When is this viewpoint interesting?

- **Primarily if the data-acquisition metadata is of concern**

Specialize by result-type



Procedures are usually process chains

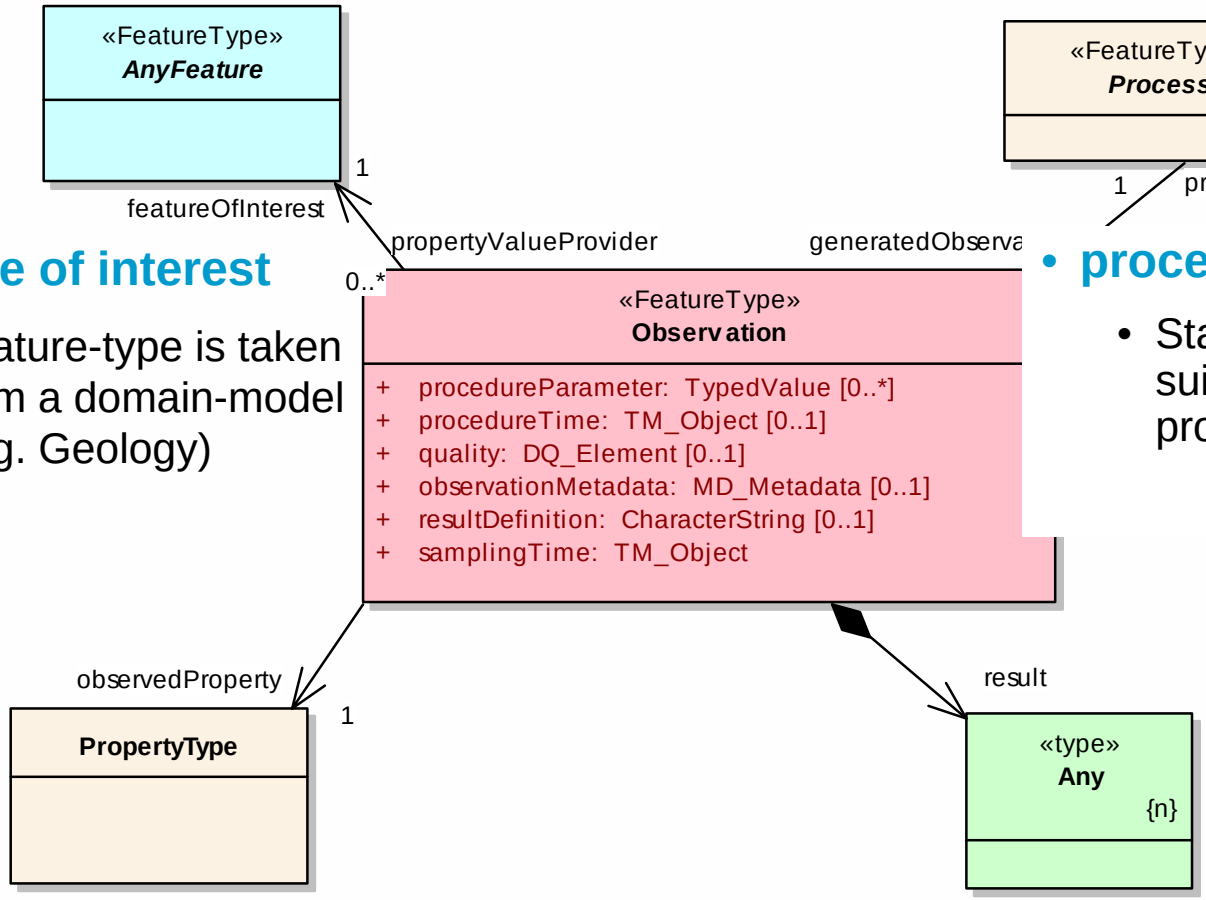
- **Procedure often includes data processing, to transform "raw" data to semantically meaningful values**
  - Voltage ➔ orientation
  - count ➔ radiance ➔ NDVI
  - Position + orientation ➔ scene-location
  - Mercury meniscus level ➔ temperature
  - Shape/colour/behaviour ➔ species assignment
- **This requires consideration of "sensor"-models and calibrations**

Advanced procedures

- **Modelling, simulation, classification are procedures**
  - “raw” data == modeling constraints
    (sensor-outputs=process-inputs)
  - “processed” data == simulation results (outputs)
  - “interpreted” data == classification results (outputs)

- **SensorML provides a model and syntax for describing process-chains**

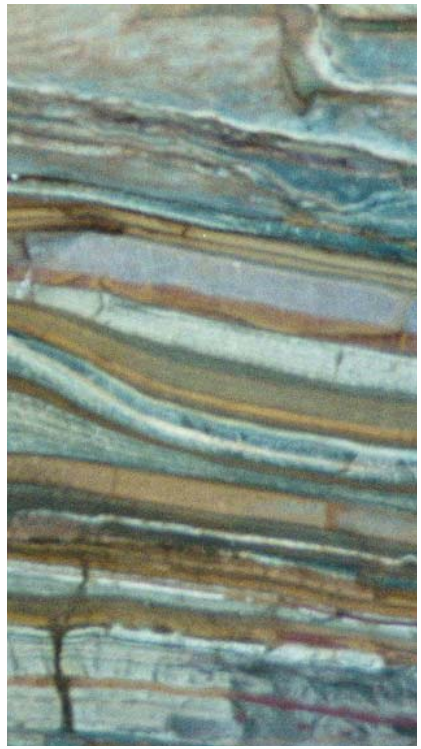
Domain profile



- **feature of interest**
  - Feature-type is taken from a domain-model (e.g. Geology)
- **procedure**
  - Standard procedures, suitable for the property-type
- **observed property**
  - Belongs to feature-of-interest-type

Conceptual object model: features

- **Digital object corresponding with identifiable, typed, object in the real world**
  - mountain, road, specimen, event, tract, catchment, wetland, farm, bore, reach, property, license-area, station
- ***Feature-type* is characterised by a specific set of *properties***



- **Specimen**
  - ID (name)
  - description
  - mass
  - processing details
  - sampling location
  - sampling time
  - related observation
  - material
  - …

Geology domain model – (e.g. GeoSciML)

- **type(featureOfInterest) = any of these classes**
- **observedProperty = any of these properties**

Borehole

- **collar location**
- **shape**
- collar diameter
- length
- operator
- logs

Fault

- **shape**
- **surface trace**
- **displacement**
- age
- ...

License area

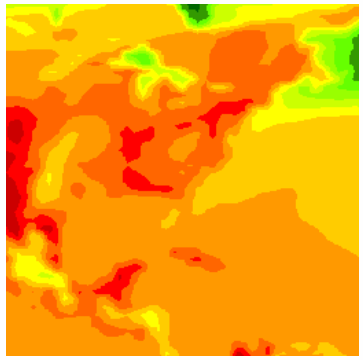
- issuer
- holder
- interestedParty
- **shape(t)**
- right(t)
- ...

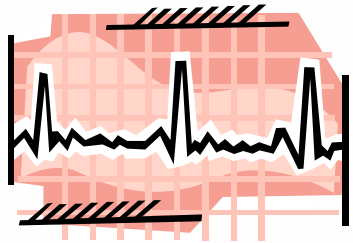
Ore-body

- commodity
- deposit type
- host formation
- **shape**
- resource estimate
- ...

- **shape**
- sampling frame
- age
- dominant lithology
- ...

Some properties vary within a feature





- colour of a Scene or Swath varies with position
- shape of a Glacier varies with time
- temperature at a Station varies with time
- rock density varies along a Borehole

- **Variable values may be described as a *Function* on some axis of the feature**

- **Corresponding Observation/result is a Function**
  - if spatio-temporal also known as *coverage* or *map*

Proximate vs ultimate feature-of-interest

The ultimate (project) thing of interest may not be directly or fully accessible

1.Sensed property is a proxy

- e.g. want *land-cover,* observe *colour*
- Post-processing required
- FoI may change during processing – e.g. “scene” ➔ “tract”

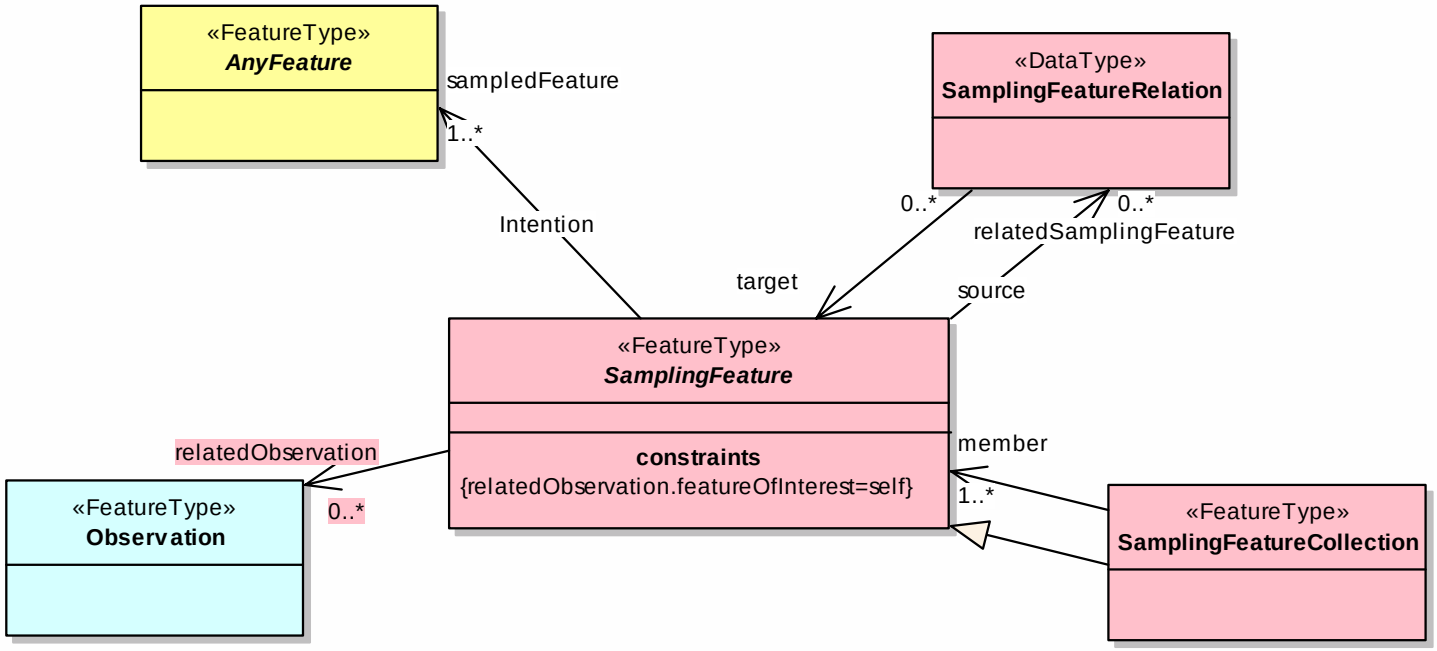
2.Proximate feature of interest embodies a sample design

- Rock-specimen samples an ore-body or geologic unit
- Well samples an aquifer
- Profile samples an ocean/atmosphere column
- Cross-section samples a rock-unit

Some standard designs are common

Sampling Features model

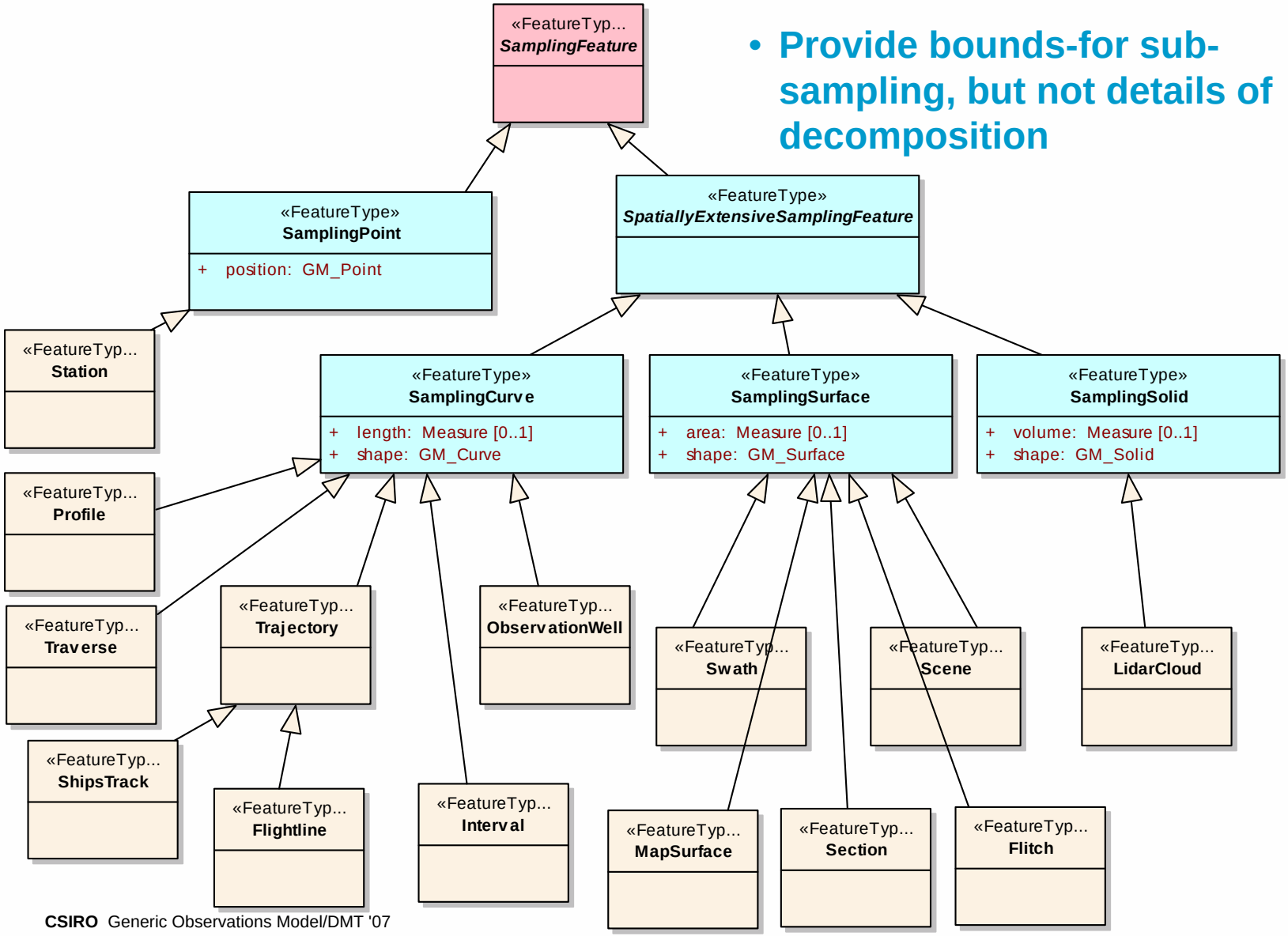
Name: SamplingBase
Package: pedagogy
Version: 1.0
Author: Simon Cox



Name: SamplingManifold
Package: pedagogy
Version: 1.0
Author: Simon Cox

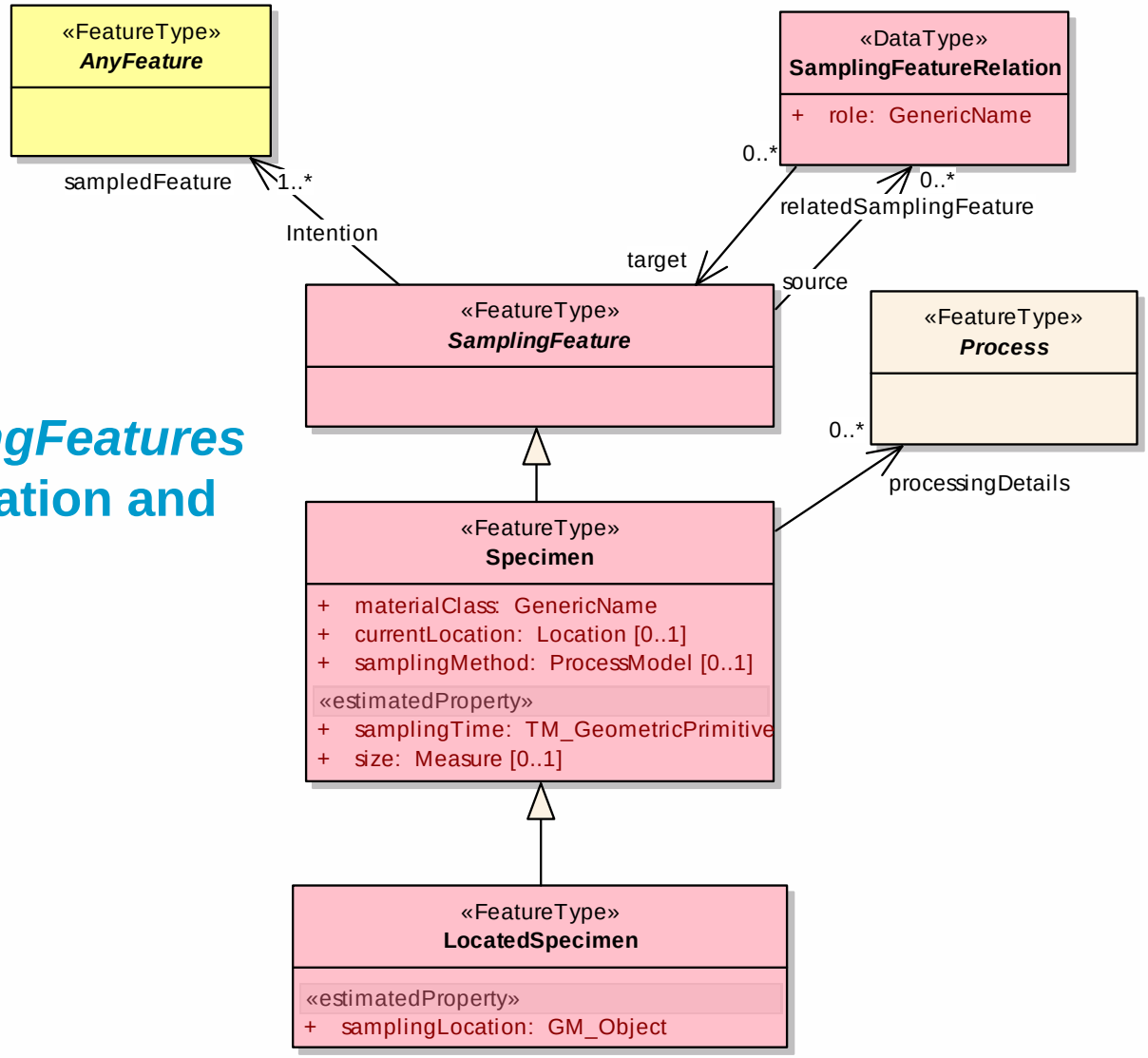
Sampling Manifolds

- **Provide bounds-for sub-sampling, but not details of decomposition**



Specimen

Name: Specimen
Package: pedagogy
Version: 1.0
Author: Simon Cox



- **Specimens are *SamplingFeatures* used for *ex-situ* observation and analysis**

Examples

- **The 7th banana weighed 270gm on the kitchen scales this morning**
- **The attitude of the foliation at outcrop 321 of the Leederville Formation was 63/085, measured using a Brunton on 2006-08-08**
- **Specimen H69 was identified on 1999-01-14 by Amy Bachrach as Eucalyptus Caesia**
- **The image of Camp Iota was obtained by Aster in 2003**
- **Sample WMC997t collected at Empire Dam on 1996-03-30 was found to have 5.6 g/T Au as measured by ICPMS at ABC Labs on 1996-05-31**
- **The X-Z Geobarometer determined that the ore-body was at depth 3.5 km at 1.75 Ga**
- **The simulation run on 2004-09-09 indicated a pressure reduction of 4 MPa at 600 Ma**

Development and validation of O&M

- **Developed in the context of**
  - XMML Geochemistry/Assay data
  - OGC Sensor Web Enablement – environmental and remote sensing
- **Subsequently applied in**
  - Water resources/water quality (WQDP, AWDIP, WRON)
  - Oceans & Atmospheres (UK CLRC, UK Met Office)
  - Natural resources (NRML)
  - Taxonomic data (TDWG)
  - Geology field data (GeoSciML)
- **I could have put dozens of logos down here**

Status

- OGC Best Practice paper, r4 – 2006
- RFC 2007-02-08
- OGC RWG ➔ Adopted Specification – late 2007?

- ISO Standard – 2008-9?

- *Adopted as a key aspect of GeoSciML*

Exploration & Mining
Simon Cox
Research Scientist

Phone: +61 8 6436 8639
Email: Simon.Cox@csiro.au
Web: www.csiro.au/em

Thank you

Contact Us
Phone: 1300 363 400 **or** +61 3 9545 2176
Email: Enquiries@csiro.au **Web:** www.csiro.au

CSIRO

“Cross-sections” through collections

Specimen	Au (ppm)	Cu-a (%)	Cu-b (%)	As (ppm)	Sb (ppm)
ABC-123	1.23	3.45	4.23	0.5	0.34

- **A Row gives properties of one feature**

- **A Column = variation of a single property across a domain (i.e. set of locations)**

Assignment of property values

- **For each property of a feature, the value is either**
  - i. asserted
    - name, owner, price, boundary (cadastral feature types)
  - ii. estimated
    - colour, mass, shape (natural feature types)
    - *i.e. error in the value is of interest*

Observations, features and coverages

Same property on multiple samples is a another kind of coverage

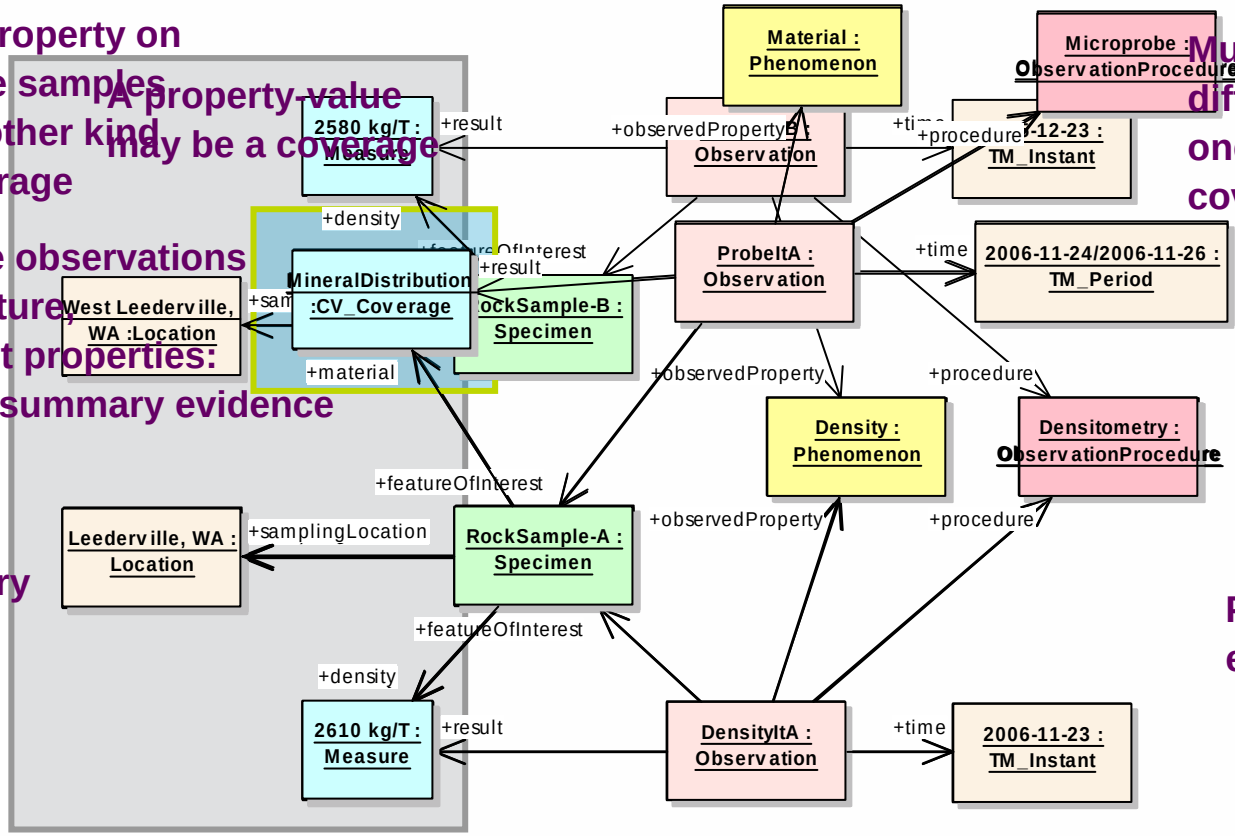
A property-value may be a coverage

Multiple observations one feature, different properties: feature summary evidence

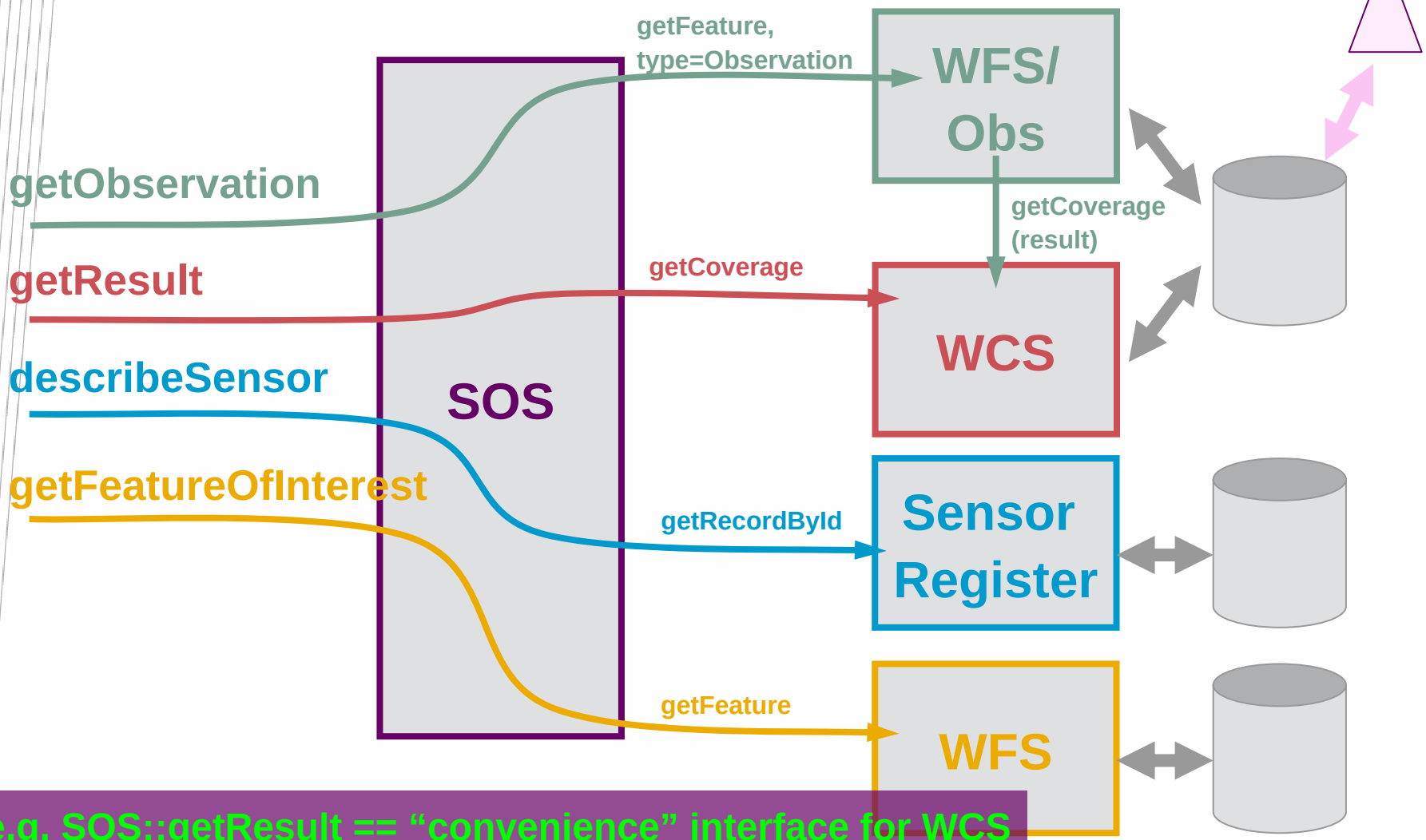
Feature summary

Multiple observations different features, one property: coverage evidence

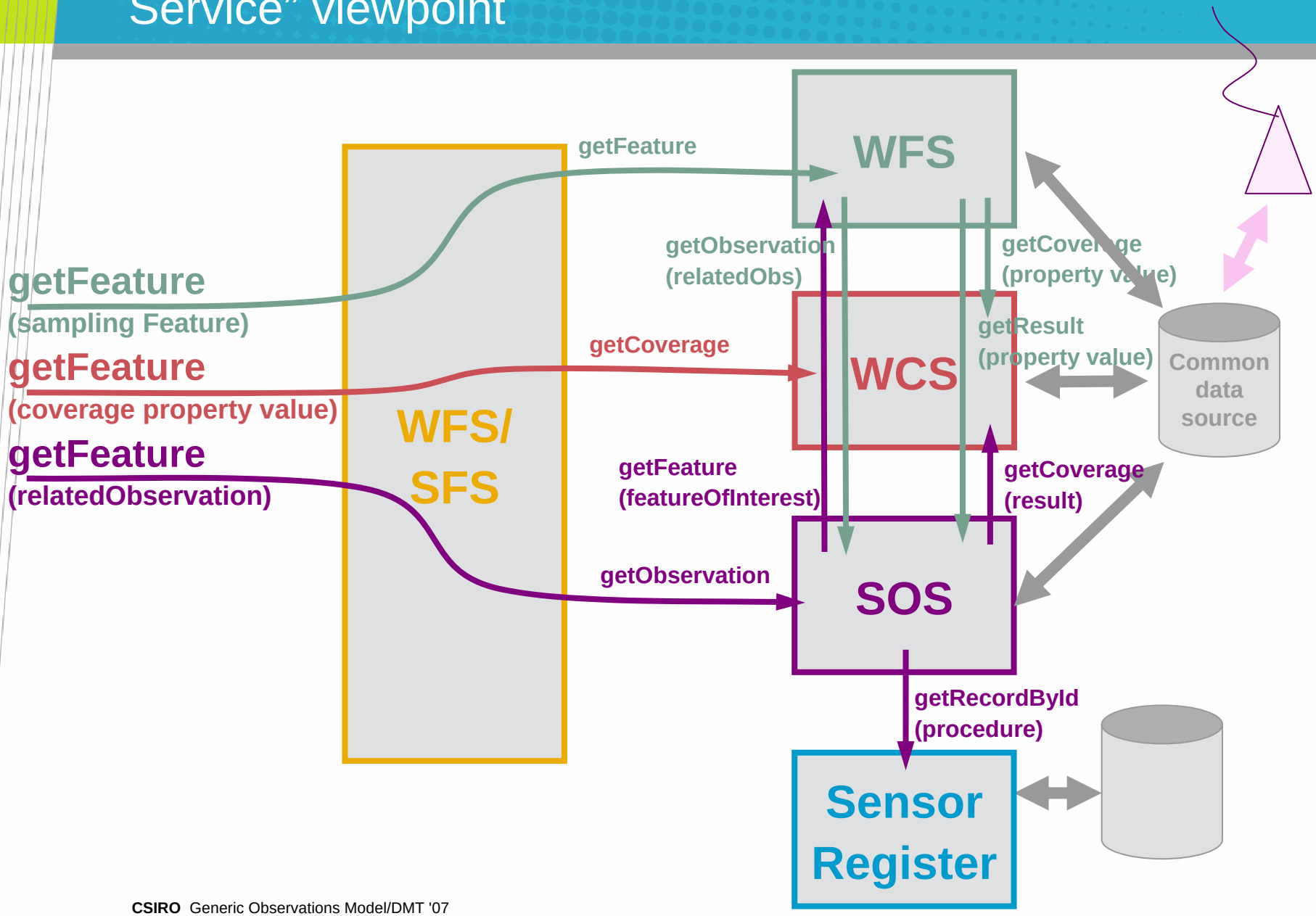
Property-value evidence



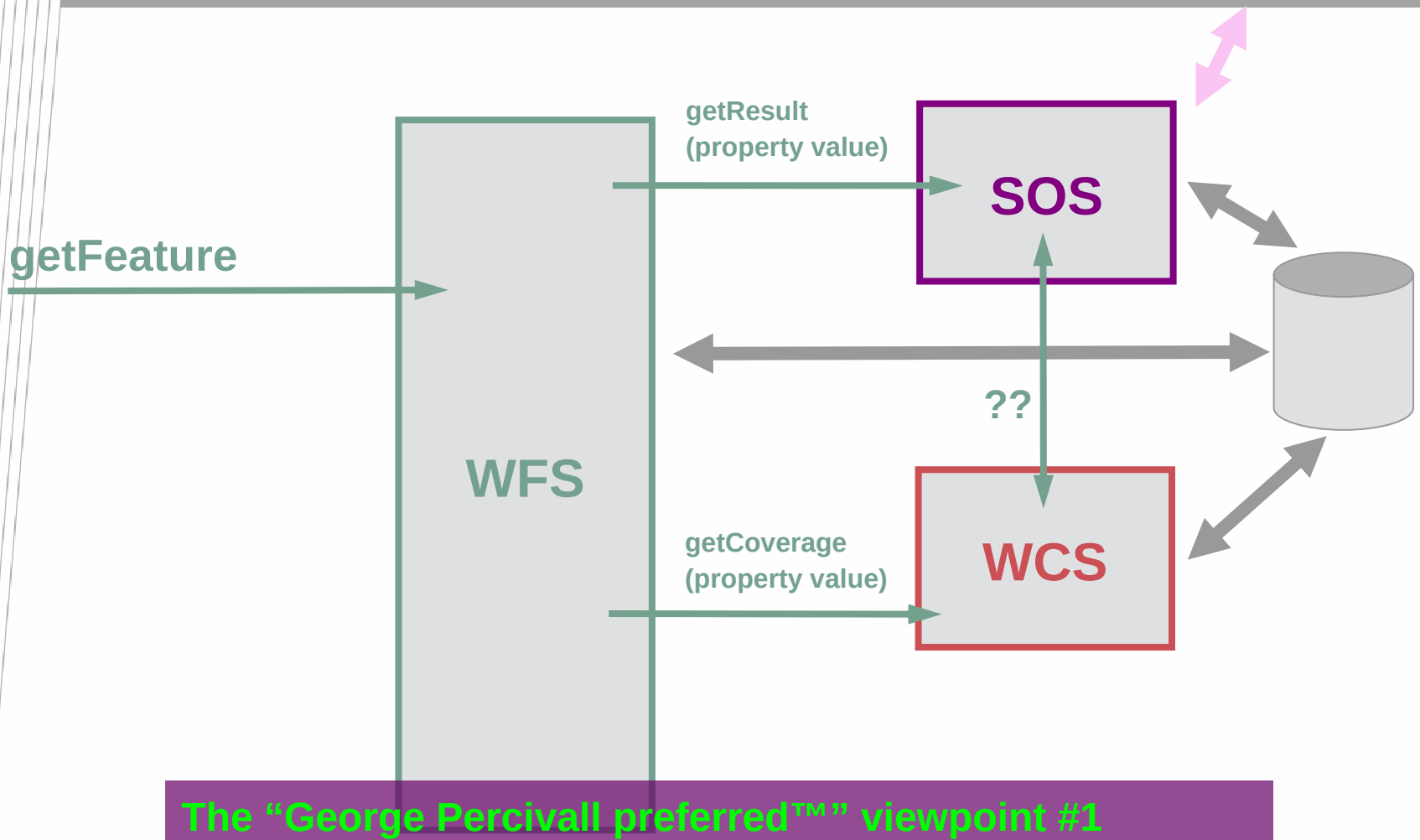
Accessing data using the “Observation” viewpoint



Accessing data using the “Sampling Feature Service” viewpoint

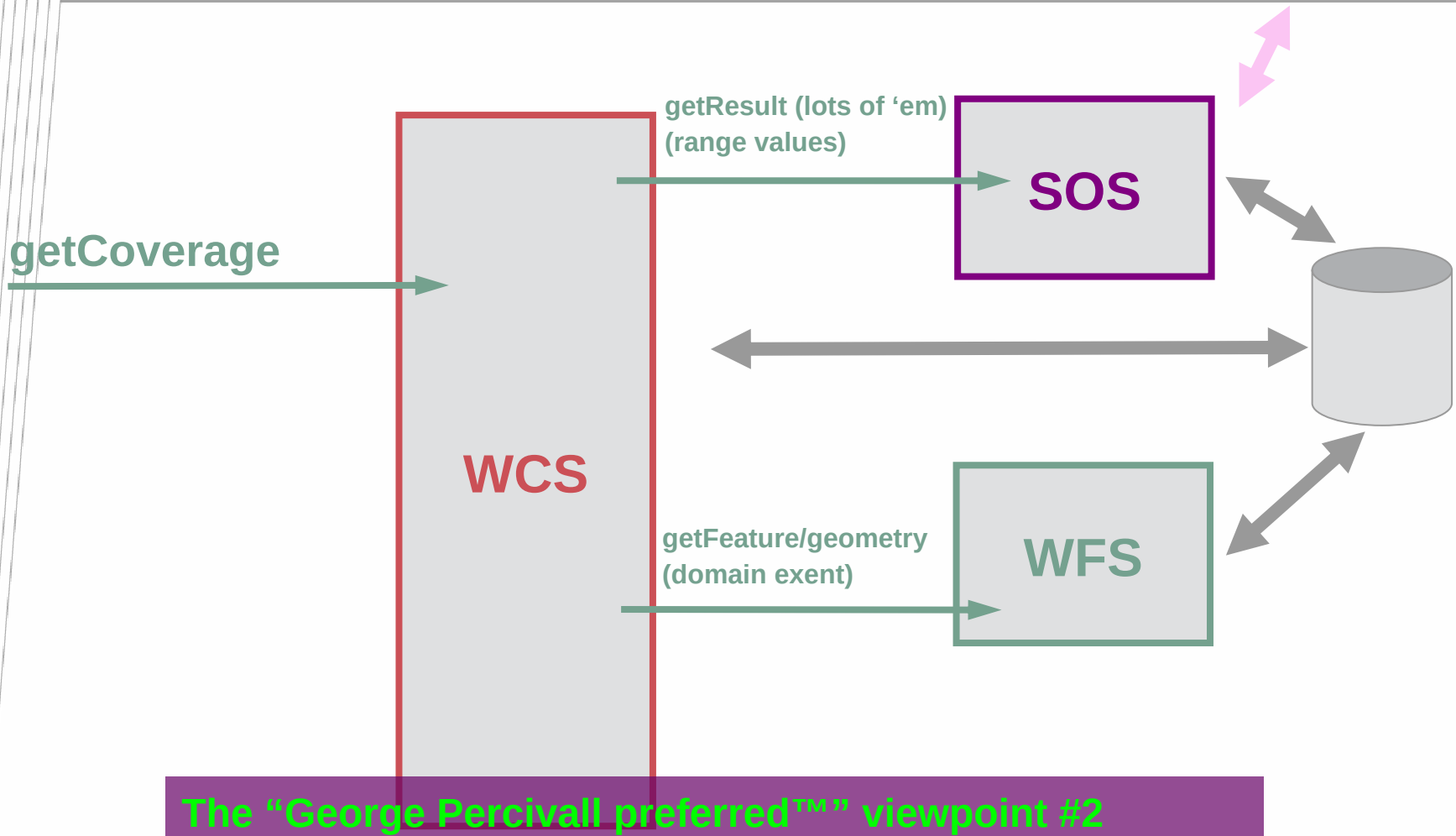


Accessing data using the “Domain Feature” viewpoint



The “George Percivall preferred™” viewpoint #1
– observations are property-value-providers for features

Accessing data using the “just the data” viewpoint



The “George Percivall preferred™” viewpoint #2
– observations are range-value-providers for coverages

Conclusions

- **Different viewpoints of same information for different purposes**
  - Summary vs. analysis
- **Some values are determined by observation**
  - Sometimes the description of the estimation process is necessary
- **Transformation between views important**
- **Management of observation evidence can be integrated**

- **(Bryan Lawrence issues)**
- **For rich data processing, rich data models are needed**
  - Explicit or implicit
- **Data models (types, features) are important constraints on service specification**

Exploration & Mining
Simon Cox
Research Scientist

Phone: +61 8 6436 8639
Email: Simon.Cox@csiro.au
Web: www.csiro.au/em

Thank you

Contact Us
Phone: 1300 363 400 **or** +61 3 9545 2176
Email: Enquiries@csiro.au **Web:** www.csiro.au

CSIRO