Mapping existing medical terminologies to SNOMED CT: An investigation of the novice user's experience

D M Vickers, M J Lawley
The Australian e-Health Research Centre,
CSIRO ICT Centre, Brisbane, Australia

THE AUSTRALIAN E-HEALTH RESEARCH CENTRE

CSIRO

Objective

- Investigate the novice user's experience mapping an existing terminology to SNOMED CT using a dedicated mapping tool
- Gain an understanding of the difficulties
- Inform future software tool development efforts

SNOMED CT

- Systematized Nomenclature of Medicine – Clinical Terms
- Is an international standard maintained by IHTSDO (International Health Terminology Standards Development Organisation)
  - Australia is a member country
- A terminology designed for clinical purposes
- ~300,000 active concepts
- Despite its size, its coverage can be patchy – particularly in specialist domains
- Some benefits of SNOMED CT include
  - No more data translation issues
  - Sophisticated querying such as subsumption queries

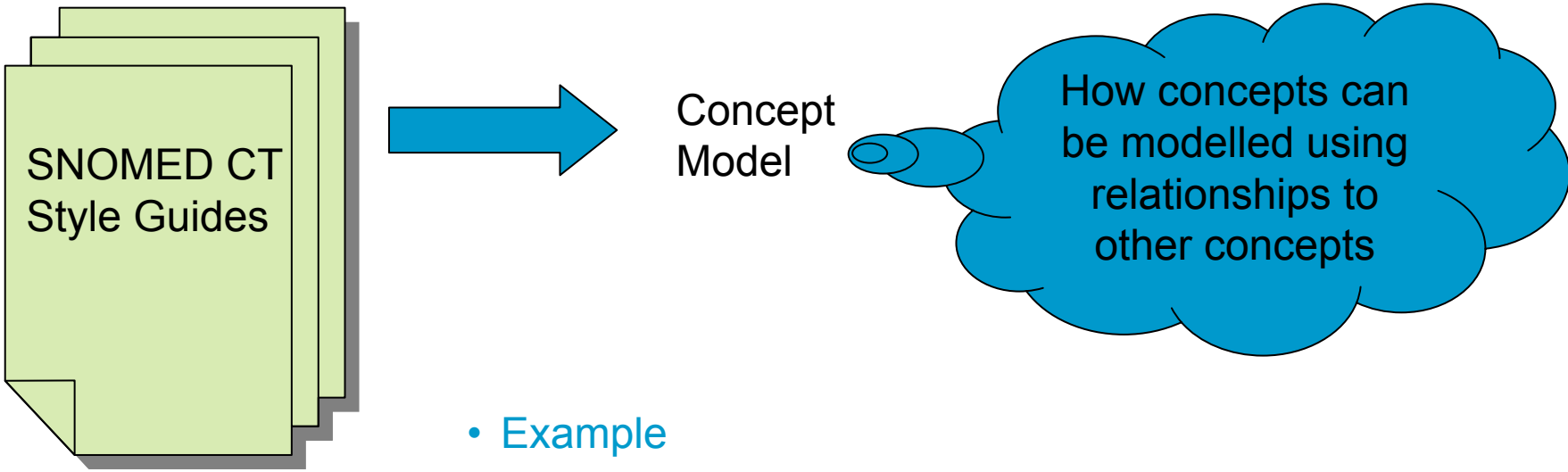
Mapping to SNOMED CT

- To switch over to SNOMED CT a mapping of the existing terms to SNOMED CT will need to be done
- Sometimes this will be “simple” – a direct mapping to a single SNOMED CT concept
  - E.g. “craniotomy” → “**25353009** | Craniotomy (procedure) |”
- Sometimes there will be no single SNOMED CT concept that encapsulates the existing term and a post-coordinated expression will need to be created
  - E.g. “Cut right calf” → “**283437002** | Cut of calf (disorder) | **: 272741003** | Laterality (attribute) | **= 24028007** | Right (qualifier value) |”

Post-coordinated expressions

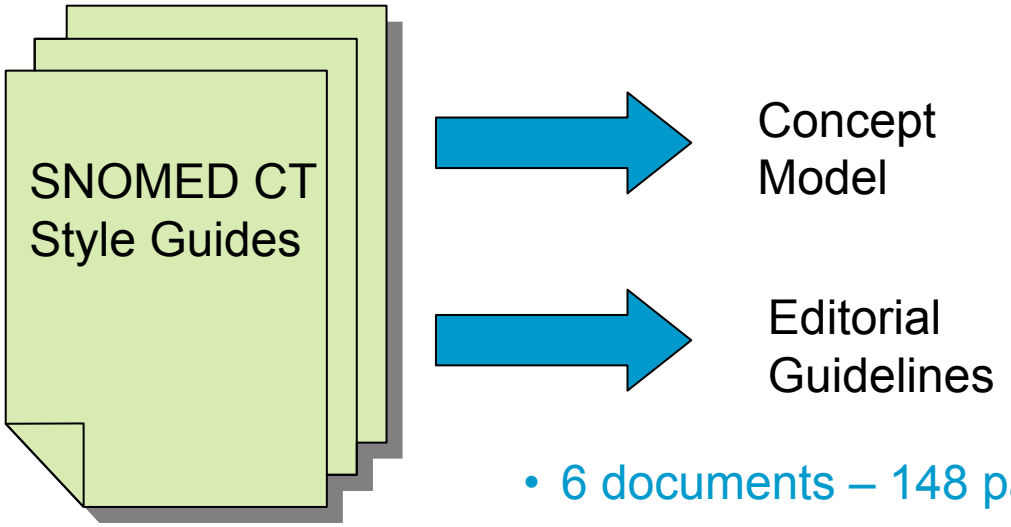
- Composing post-coordinated expressions is no easy task ..
  - SYNTAX Rules
    - SNOMED CT Compositional Grammar
  - SEMANTICS (Meaning) Rules
    - SNOMED CT Style Guides
    - Expressions violating semantic rules ..
      - **238516007** | Suntan (disorder) | **: 32911000** | Homeless (finding) | **= 387207008** | Ibuprofen (substance) |
      - **16331000** | Heartburn | **: 246454002** | Occurrence (attribute) | **= 70232002** | Frequent |

SNOMED CT Style Guides



- Example
  - Procedures may specify an intent and valid values for intent are palliative, guidance, diagnostic, therapeutic, preventative or screening

SNOMED CT Style Guides

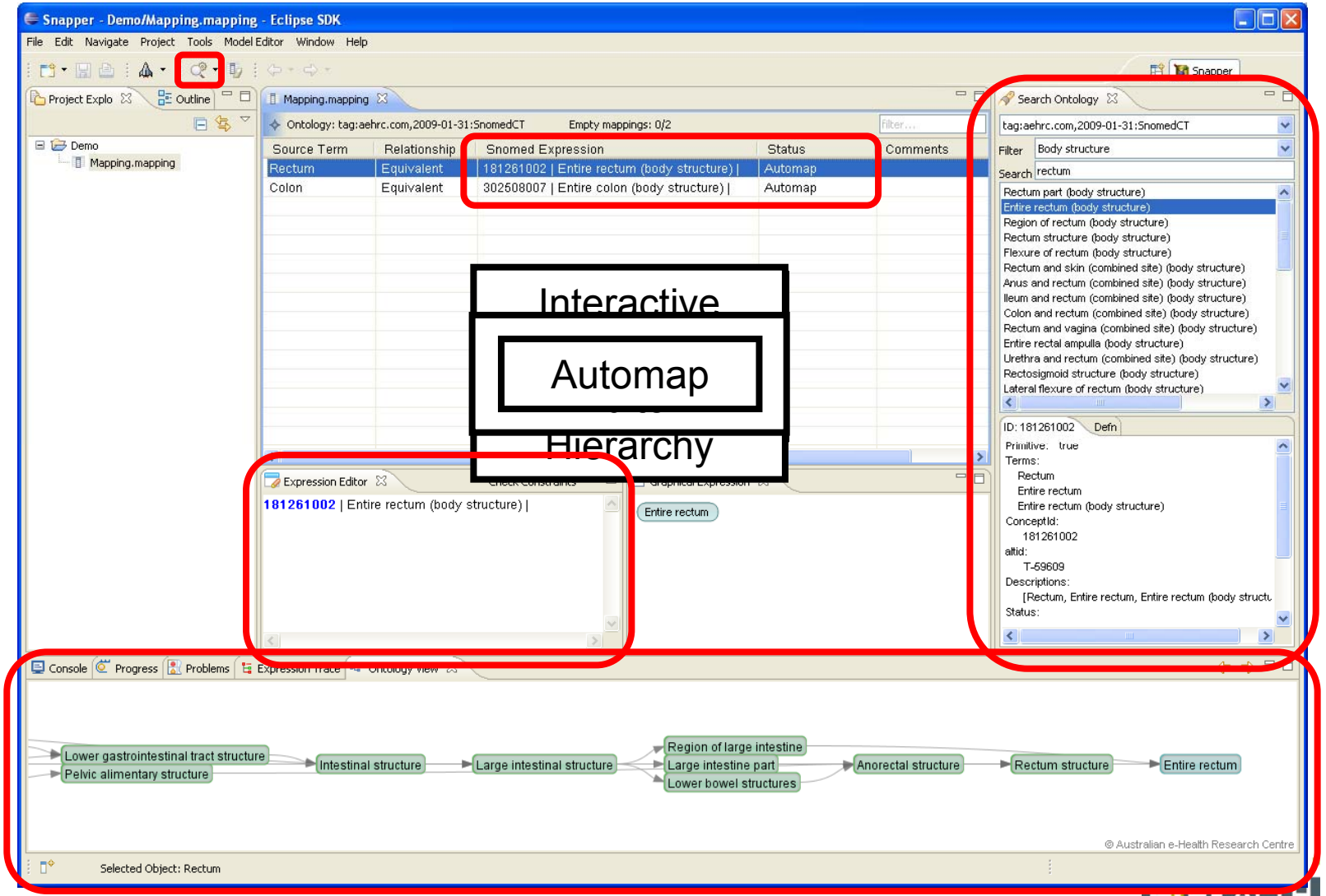


- 6 documents – 148 pages in total
  - Introduction and Overview
  - Body Structures - Anatomy
  - Clinical Findings
  - Morphologic Abnormalities
  - Procedures
  - Situations with Explicit Context
- http://www.ihtsdo.org/publications/inside-ihtsdo/

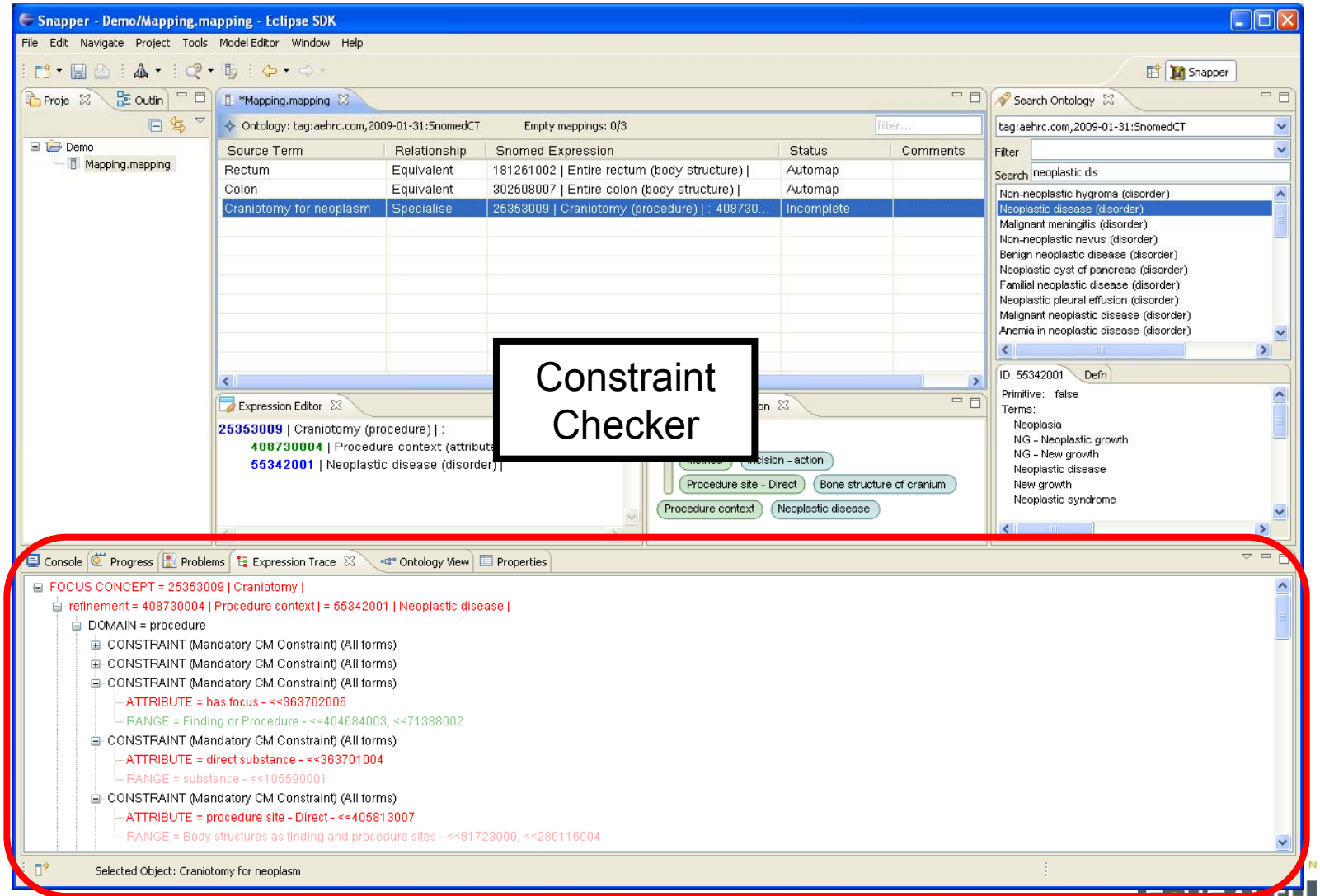
Machine Readable Concept Model (MRCM)

- Software can't assess your expression against a set of documents
  - A machine readable version is needed
  - There is already an awareness of this need
    - IHTSDO Machine Readable Concept Model Working Group
    - Prototype MRCMs have been produced

Snapper



Snapper



An existing medical terminology

- ANZICS (Australian and New Zealand Intensive Care Society) diagnosis codes
  - Real-world candidate codes that could be replaced by SNOMED CT
  - We had access to an ICU clinician

ANZICS Diagnosis Codes

- Consist of 2 levels
  - Diagnosis codes (99)
  - Optional more detailed “sub-codes” grouped under a diagnosis code (396)

Diagnostic Code	Diagnostic Code Description	Diagnostic Sub-code	Diagnostic Sub-code Description
312	GI Cancer	312.01	Cancer of the colon/rectal
		312.02	Cancer of the oesophagous
		312.03	Cancer of the pancreas
		312.04	Cancer of the stomach
		312.05	Cancer of other GI

Data Analyst experienced with health data but no SNOMED CT experience

- ICU Clinician (1 session)
- SNOMED CT Style Guide documents
- SNOMED CT compositional grammar resources
- Informal training session
- ANZICS diagnosis codes loaded into a version of Snapper that did not have the constraint checker

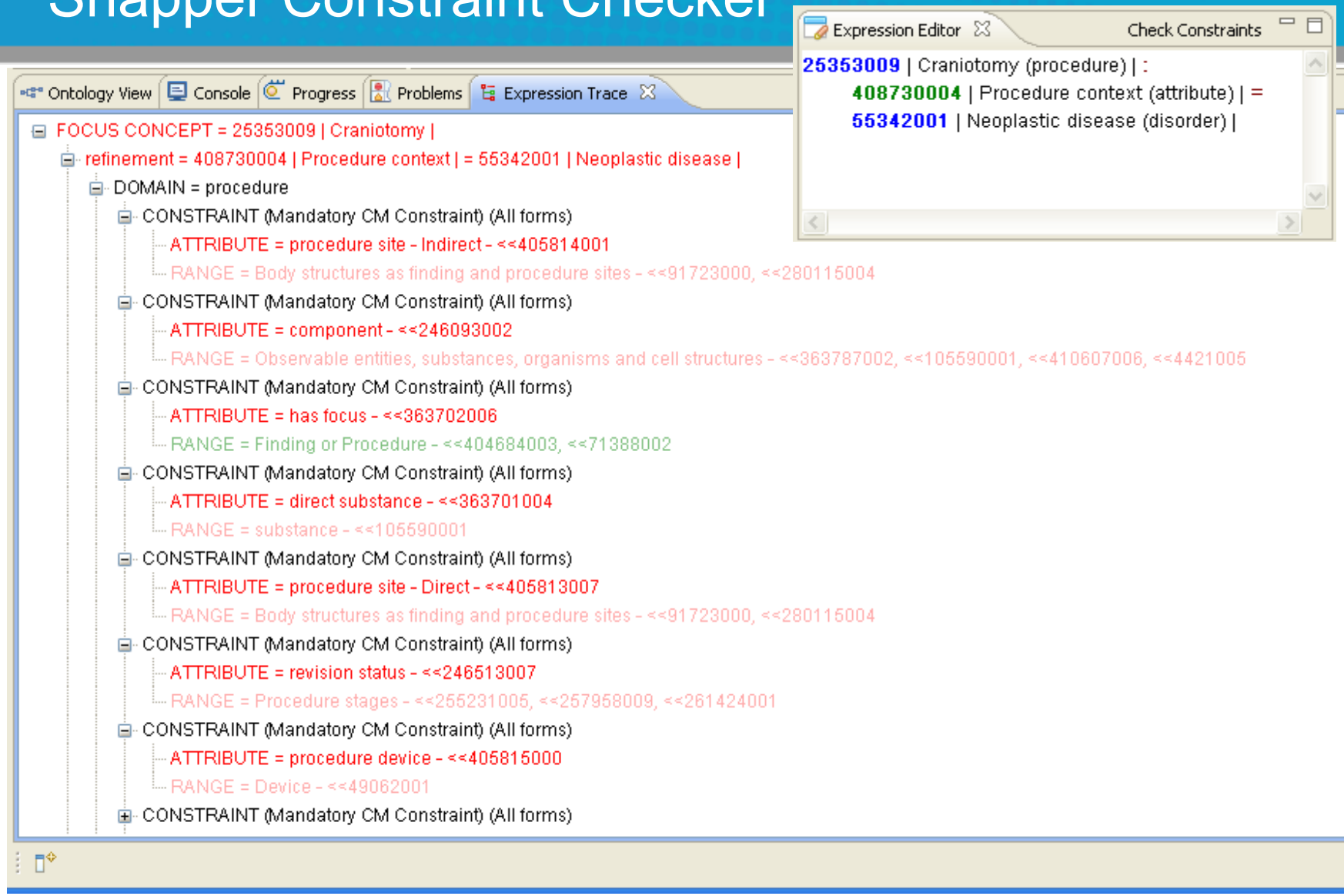
Interview with Data Analyst

- Issues the Data Analyst raised
  - Amount of information in the style guides was initially so overwhelming that they were only skimmed at first.
  - The quality of the mappings improved with experience and as their knowledge of the style guides improved.
  - They noted how different their mappings would be without the assistance of the ICU clinician to explain intended meanings and emphasis of codes
  - The Data Analyst's tendency was to do nothing rather than try something and get it wrong – potentially for a lot of mappings

Interview with Data Analyst (cont.)

- Issues we noted from the discussion
  - Some knowledge of SNOMED CT is needed
    - Inferred refinements from parent concepts
    - Terms cannot always be mapped to SNOMED CT
      - E.g. negation and disjunction
    - The distinction between primitive concepts and fully specified concepts and how full modelling is not always required
    - Existing concepts can be used as a pattern for constructing new concepts
  - The broader the terminology, the greater the Style Guide knowledge the mapper must have
    - ANZICS diagnosis codes are broad and are a difficult starting point for a novice

Snapper Constraint Checker



Interview with Data Analyst (cont.)

- Mappings checked using Snapper's constraint checker
  - ALL expressions that were checked were wrong!
  - Helped the Data Analyst understand why the expressions were wrong
  - Made the Data Analyst fully appreciate the need to understand and follow the style guides.
  - Gave the Data Analyst the confidence to attempt the mappings again
    - Immediate feedback when trying something a little more complex

What the SNOMED CT Concept Model can't do

- Examine mappings to single concepts (no refinements) for semantic correctness
  - Peer review of mappings can help
- Check expressions against editorial guidelines
  - No machine readable version

Conclusions

- In general ..
  - Mappings should be performed by trained, experienced people
    - Documents recently added to the IHTSDO website support this
  - A domain expert needs to be involved

- For tools ..
  - Should maintain the hierarchy that can exist in terminologies
    - Maintaining the context of child and sibling terms makes mapping easier
  - Provide feedback when expressions violate the expression grammar
    - Highlight part of the expression that is causing the violation
  - Provide feedback when expressions violate the concept model
  - Support peer-review of mappings

Wish List for the Future

- All of the Style Guides to be machine readable
- An even better constraint checker
  - Easy to understand
  - Easy to locate source of problem and see alternatives
- Intelligent editor that prevents you from making syntax and semantic mistakes in the first place!

The Australian e-Health Research Centre
Deanne Vickers
Software Engineer

Phone: 07 3253 3633
Email: deanne.vickers@csiro.au
Web: http://aehrc.com

www.csiro.au

Thank you

THE AUSTRALIAN E-HEALTH RESEARCH CENTRE

CSIRO