SnoLyze – a SNOMED CT Expression Constraint Language Execution Engine in R

Ronald Cornet, **Sander Laverman**

Academic Medical Center – University of Amsterdam, The Netherlands

Daniel Karlsson

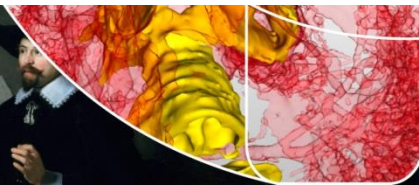Linköping University & SocialStyrelsen, Sweden

Background

- Question at April 2016 IHTSDO Business meeting: "How to analyse SNOMED-encoded data?"

London, we have a problem

ANALYTICS

SNOMED CT

The global
language of
healthcare

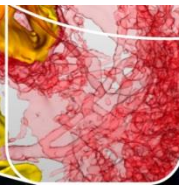# London, we have a problem
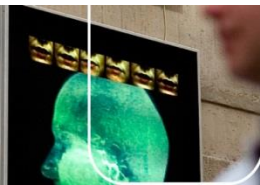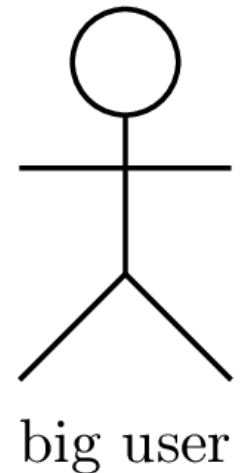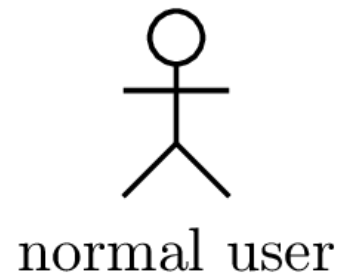

Theory

Practice

Aim

- Enable analysis of data that is encoded using SNOMED CT

- Integration with common analytics tooling

- Proof-of-concept with potential

Use Case

- Analysts / statisticians should be able to perform at least basic functions
  - All descendants of concept X
  - All concepts with attribute Y

Approach

- Develop a module for R

- What is R?
  - R is a **free** software environment for **statistical computing** and graphics
  - On UNIX platforms, Windows and MacOS
  - GNU (GPL) version 2

These 'R' the looks…

Implementation – language

- Expression Constraint Language v1.2
  - Official SNOMED International specification
  - enables queries over SNOMED CT content to be expressed
  - could be embedded within record-based query languages (such as SQL) to represent the terminological aspects of these queries

[Expression Constraint Language - Specification and Guide]

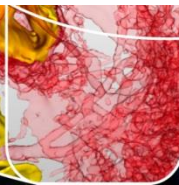# Expression Constraint Language Examples

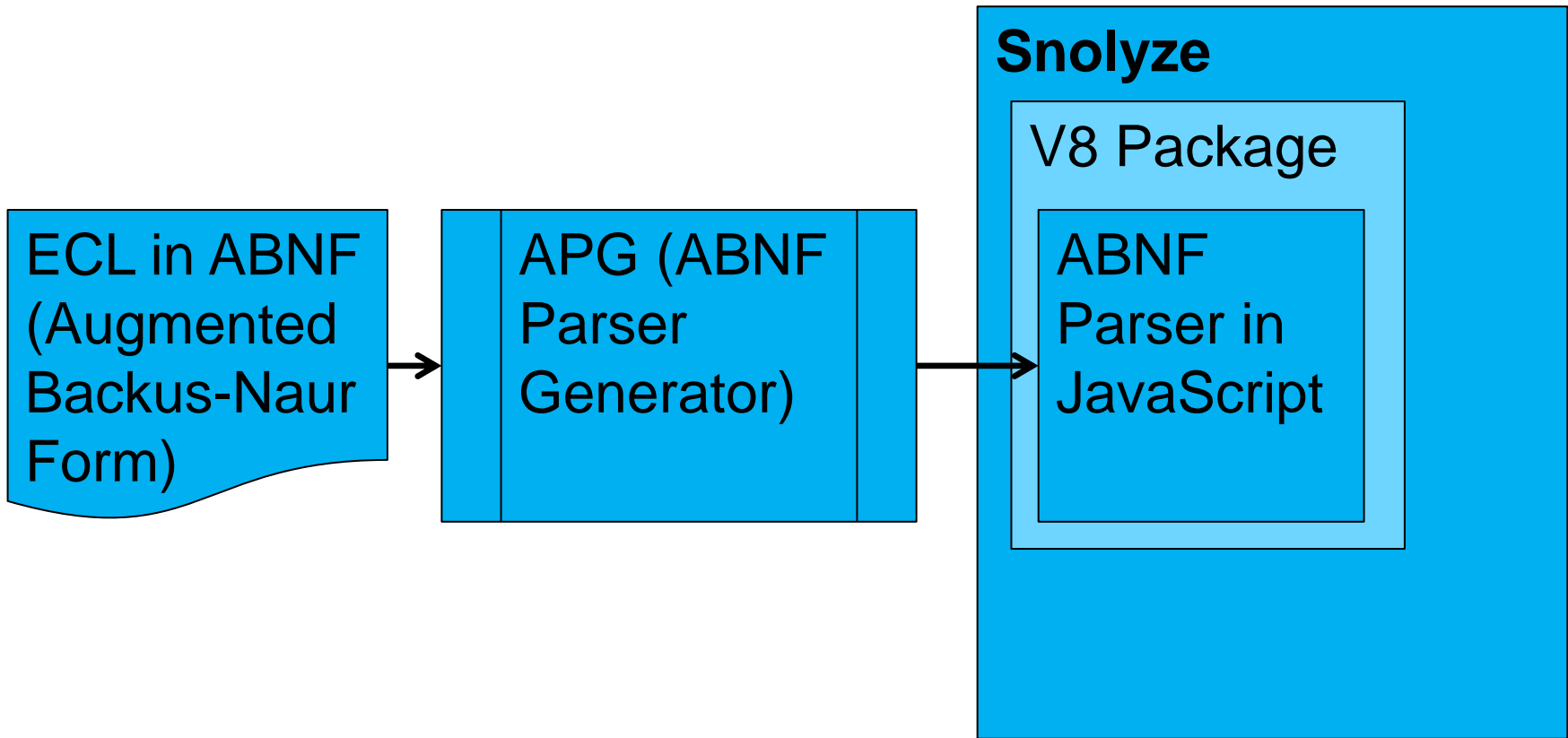<   descendantOf

<<   descendantOrSelfOf

<!   childOf

- descendantOf 19829001 |Disorder of lung| : 116676008 |Associated morphology| = descendantOrSelfOf 79654002 |Edema|

- << 19829001 |Disorder of lung| MINUS << 301867009 |Edema of trunk|

Implementation – technical

Snolyze

V8 Package

ECL in ABNF (Augmented Backus-Naur Form) → APG (ABNF Parser Generator) → ABNF Parser in JavaScript

Implementation – functions

- launch(relationship_snapshot, [trans_closure])
  - Initialize SnoLyze

- getTransitiveClosure()
  - Calculate TC for the snapshot

- execute(query)
  - Run an ECL query, returns a vector of SNOMED IDs

Demo

Conclusion

- SnoLyze performs well for research practice
  - Generally within 0.05 s, up to 0.2 s
- Easy integration for performing queries on data encoded using SNOMED CT

The next challenge…

"Data! Data! Data!"
he cried impatiently.
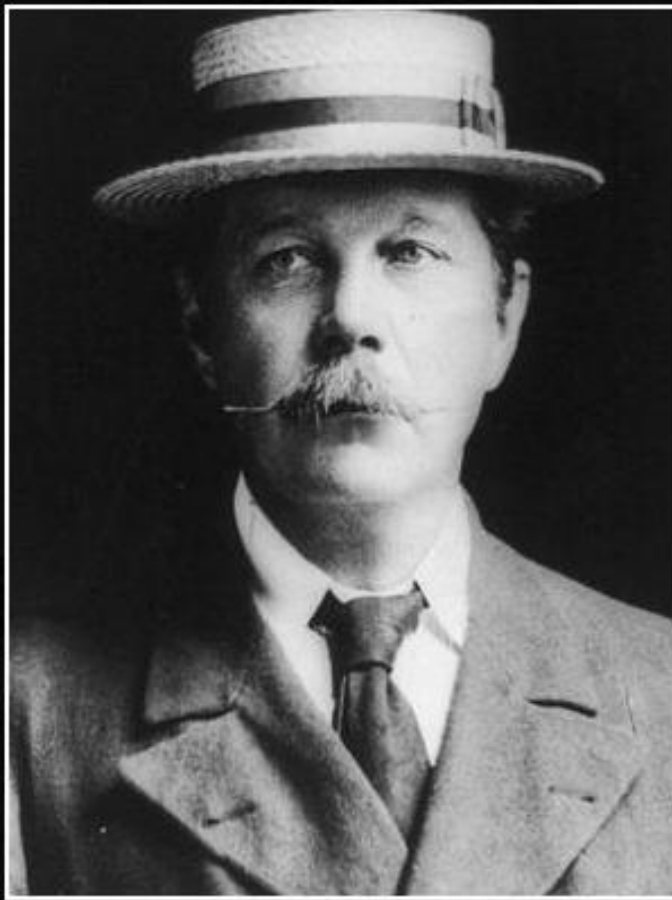
"I can't make bricks
without clay!"

Sherlock Holmes, The Adventure in the Copper Beaches

Acknowledgments

- Janice Watson, raising the question
- Sander Laverman, for closing the theory-practice bridge

Now it's your turn!

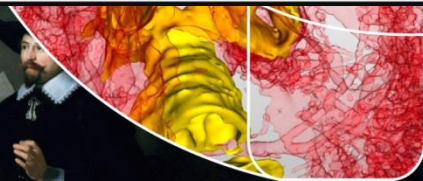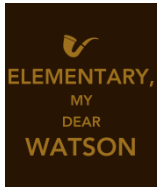What one man can invent, another can discover.

— Arthur Conan Doyle —

https://github.com/slaverman/SnoLyze

Image Credits

- https://userscontent2.emaze.com/images/e6ba5ed2-f324-4d07-bfe2-e34f70f287f4/f39390d5e1a53d73bffbff4db4d5c07c.png

- http://vizts.com/london-bridge/

- http://www.360logica.com/blog/peek-proof-concept-poc/

- https://perso.ensta-paristech.fr/~kielbasi/tikzuml/var/files/html/web-tikz-uml-userguidech3.html

- https://es.slideshare.net/manolofrias/introducing-a/2

- http://www.azquotes.com/picture-quotes/quote-what-one-man-can-invent-another-can-discover-arthur-conan-doyle-37-42-35.jpg