An Architecture for Automatic Generation of Computer Interpretable Guidelines

Daniel R. Schlegel, PhD – Department of Computer Science – daniel.schlegel@oswego.edu

OSWEGO STATE UNIVERSITY OF NEW YORK

Introduction

We propose a system architecture based on the application of bottom-up natural language understanding techniques which would be capable of handling sophisticated guideline recommendations, grounding them in the most specific terms available so that they may easily integrate with clinical decision support systems, and producing a computer interpretable guideline.
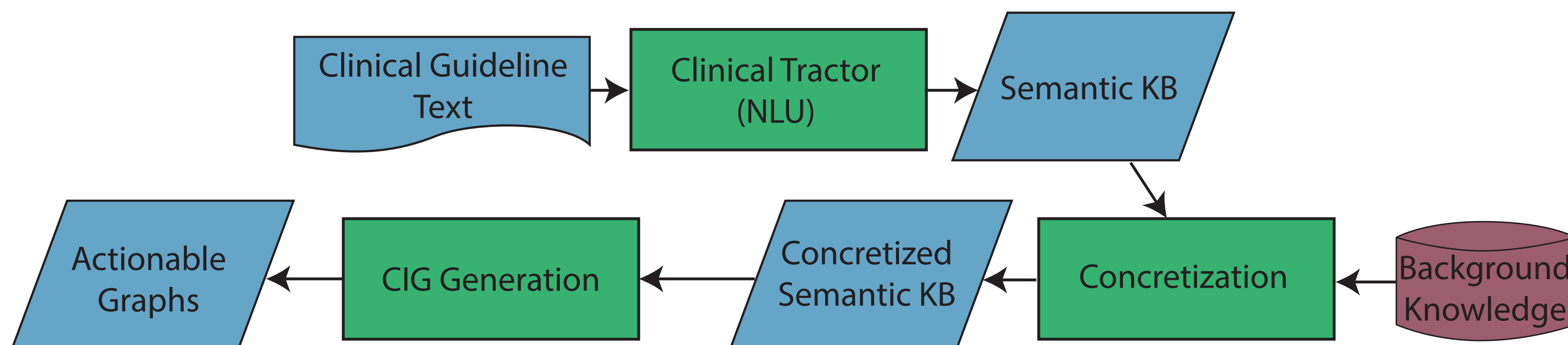

Figure 1. Overall system architecture for the proposed guideline understanding and formalization system.

NLU with Clinical Tractor

Repurposing of the Tractor[1] natural language understanding system for the clinical domain

- Previously applied to the counter-insurgency domain – short intelligence messages
- Converted plain-text messages to a knowledgebase with 92% semantic relations
- Rule-based syntax to semantics transformation after text processing and importing background knowledge
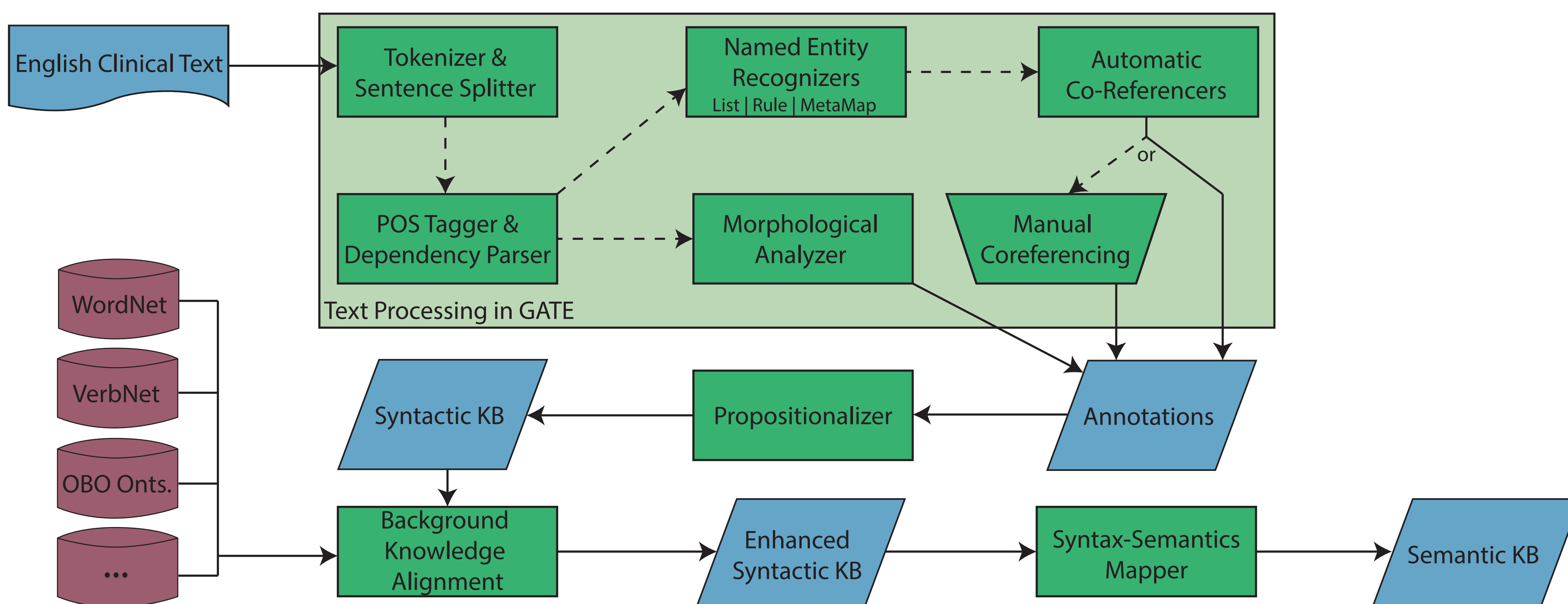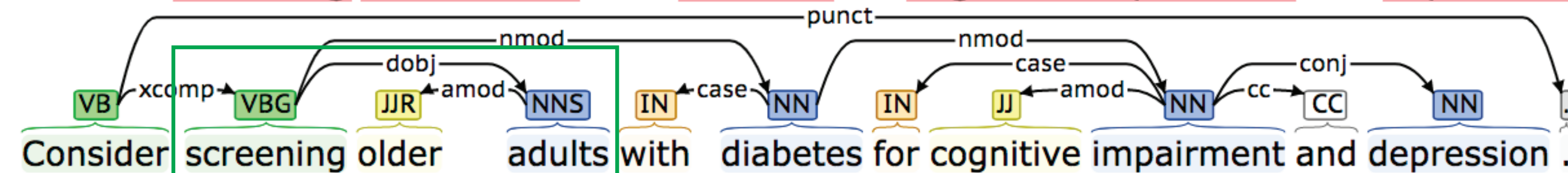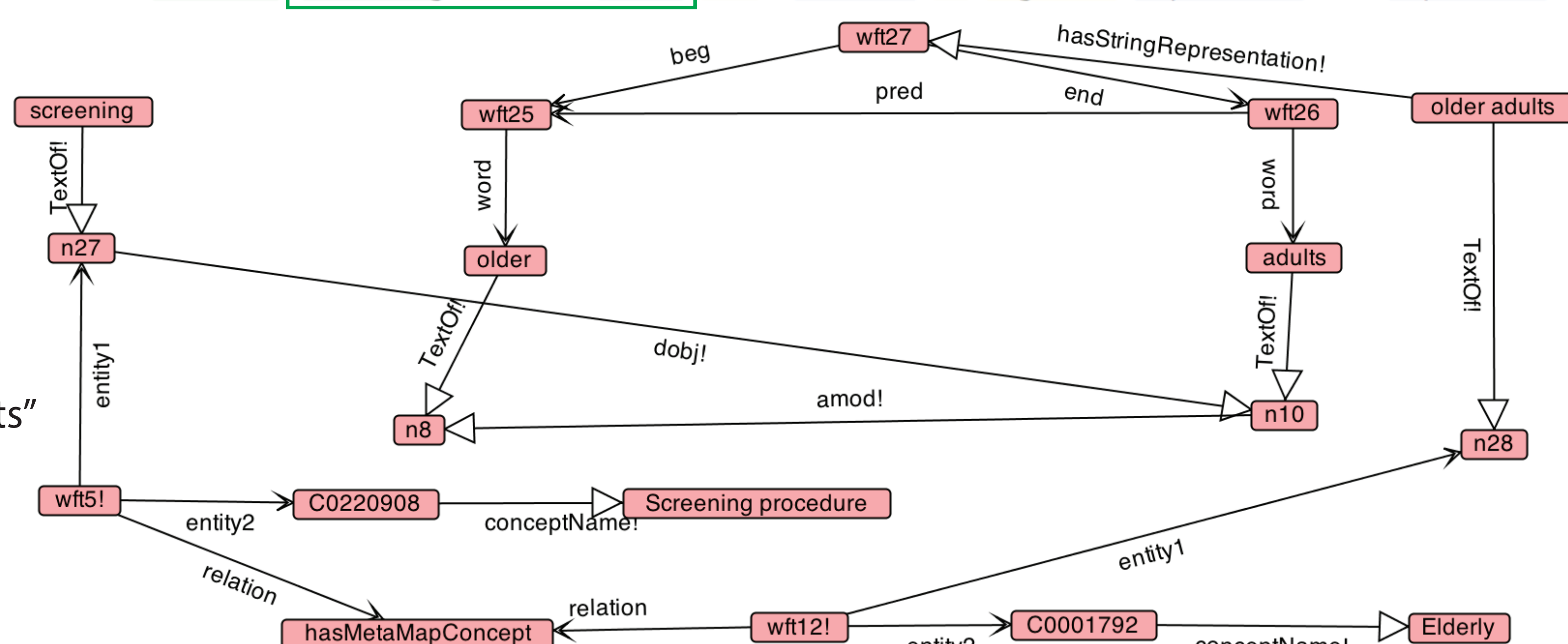

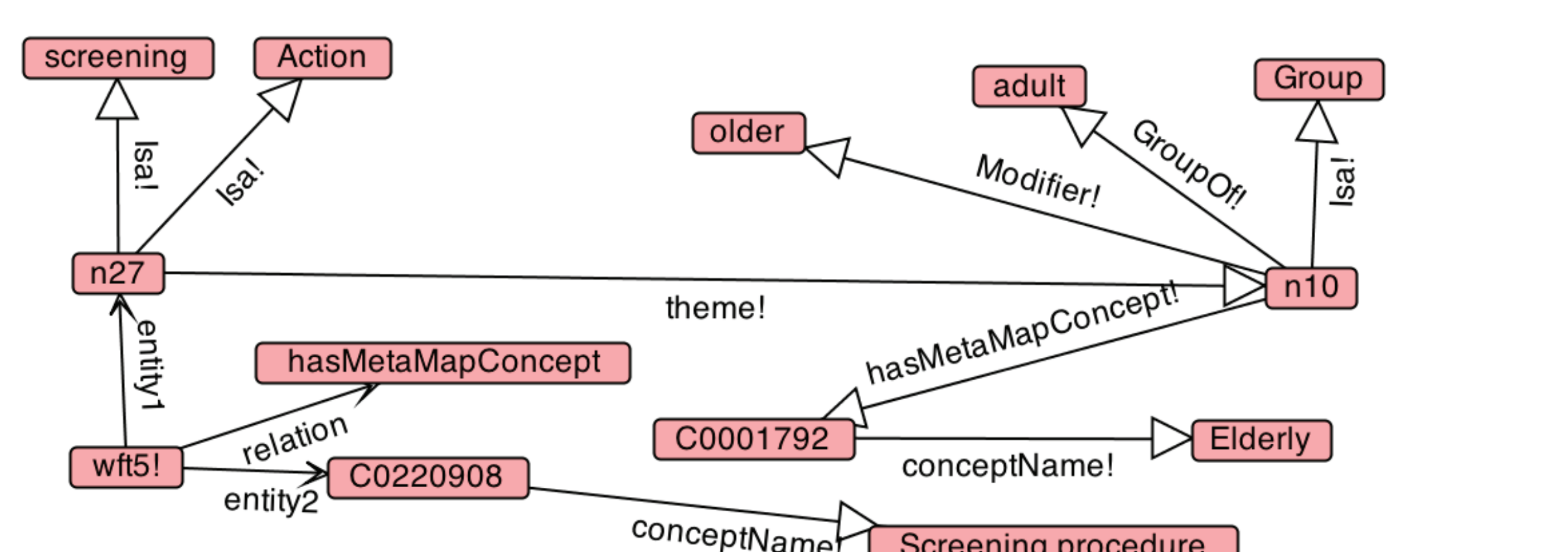**Figure 2.** Clinical Tractor system architecture.

Figure 3. MetaMap matches and the dependency parse for "*Consider screening older adults with diabetes for cognitive impairment and depression*" are shown at the top, with a subset of the CSNePS syntactic KB after propositionalization, and semantic KB after the mapping rules have been applied shown for the phrase "*screening older adults.*" .

Concretization

Maps terms to their definitions / elaborations found

- Within guideline;
- Within background knowledge.

Goal: ease interop with existing EHR systems.

Definitions / elaborations appear in:

1. Appositional phrases, e.g., "*overweight or obese (BMI ≥ 25kg/m2 or ≥ 23kg/m2)*" [3];
2. Explicit definitions, e.g., "*modest weight loss, defined as sustained reduction of 5% of initial body weight*" [3];
3. Discourse elaboration e.g., "*... identify risk factors for ulcers...*" and later "*The risk of ulcers or amputations is increased in people who have the following risk factors...*" [3];
4. Tables / figures, e.g., "*A1C goals are presented in Table 12.1.*" [3]

CIG Generation

Target format: Actionable Graphs

- Used in multi-CIG mitigation framework[4]
- Nodes: context (root), actions, decisions
- Arcs: transitions

Requires understanding actions / conditions in text.

- Preconditions often require background knowledge.

"Every tobacco user should be advised at every visit to quit." (ACC)

Precondition: *The patient* is a tobacco user

- Unstated, but needs to be determined.

Clinical Tractor approach:

- Maintain a list of simple roles, including *user*.
- Notice that *tobacco* modifies *user*.
- Therefore *tobacco user* is a role.
- Roles are filled by persons.
- We can generate a rule that every person who has the role of *tobacco user* should be advised to quit at every visit.
- We know the person the recommendation will apply to is the patient.

Multiple potential actions indicated in the text will be translated into decision nodes.

Written order in the guideline will be used as temporal order, unless defined otherwise.

References

1. Stuart C. Shapiro and Daniel R. Schlegel. Natural language understanding for soft information fusion. In *Proceedings of Fusion 2013*. IFIP, July 2013. 9 pages.
2. Daniel R. Schlegel and Stuart C. Shapiro. Inference graphs: Combining natural deduction and subsumption inference in a concurrent reasoner. In *Proceedings of AAAI-15*, 2015.
3. American Diabetes Association. Standards of Medical Care in Diabetes. *Diabetes Care*, 40(suppl 1), 2017.
4. Szymon Wilk, Martin Michalowski, et al. Comprehensive mitigation framework for concurrent application of multiple clinical practice guidelines. *JBI*, 66:52–71, 2017.