

SNOMED CT Research Webinar  
May 20, 2020

# SNOMED CT and research

*Why should the SNOMED community  
and the research community care?*

Olivier Bodenreider

National Library of Medicine, Bethesda, Maryland, USA



U.S. National Library of Medicine



# Disclaimer

The views and opinions expressed do not necessarily state or reflect those of the U.S. Government, and they may not be used for advertising or product endorsement purposes.



# Outline

- Before SNOMED CT
- “The early days of SNOMED CT” – Aspects of SNOMED CT research in PubMed 2001-2006
- KR-MED 2008 – *Representing and sharing knowledge using SNOMED*
- My own research journey with SNOMED CT
- Final thoughts

# Before SNOMED CT



# SNOMED CT was born in a research lab...

- Representation of clinical data  
(using logical definitions; KRSS)
- Concurrency control  
(distributed editing  
and reconciliation)
- Configuration management  
(versioning)

DISTRIBUTED DEVELOPMENT  
OF A LOGIC-BASED  
CONTROLLED MEDICAL TERMINOLOGY

A DISSERTATION  
SUBMITTED TO THE PROGRAM IN MEDICAL INFORMATION SCIENCES  
AND THE COMMITTEE ON GRADUATE STUDIES  
OF STANFORD UNIVERSITY  
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

Keith Eugene Campbell

June 1997



# Beyond lexical features

Journal of the American Medical Informatics Association Volume 5 Number 2 Mar / Apr 1998

203

*Research Paper* ■

## Evaluation of a “Lexically Assign, Logically Refine” Strategy for Semi-automated Integration of Overlapping Terminologies

---

ROBERT H. DOLIN, MD, STANLEY M. HUFF, MD, ROBERTO A. ROCHA, MD, PHD,  
KENT A. SPACKMAN, MD, PHD, KEITH E. CAMPBELL, MD, PHD

**Abstract** **Objective:** To evaluate a “lexically assign, logically refine” (LALR) strategy for merging overlapping healthcare terminologies. This strategy combines description logic classification with lexical techniques that propose initial term definitions. The lexically suggested initial definitions are manually refined by domain experts to yield description logic definitions for each term in the overlapping terminologies of interest. Logic-based techniques are then used to merge defined terms.

# Other early research efforts

- Language independent concept representation systems
- Description logics (GRAIL)
- Separation between
  - concept model
  - linguistic mechanisms
- Terminology server



Computer Methods and Programs in Biomedicine 45 (1994) 75–78

computer methods  
and programs  
in biomedicine

## The Galen Project<sup>☆</sup>

A.L. Rector\*, W.A. Nowlan and the GALEN Consortium

*Medical Informatics Group, Department of Computer Science, University of Manchester, Manchester M13 9PL, UK*

### Abstract

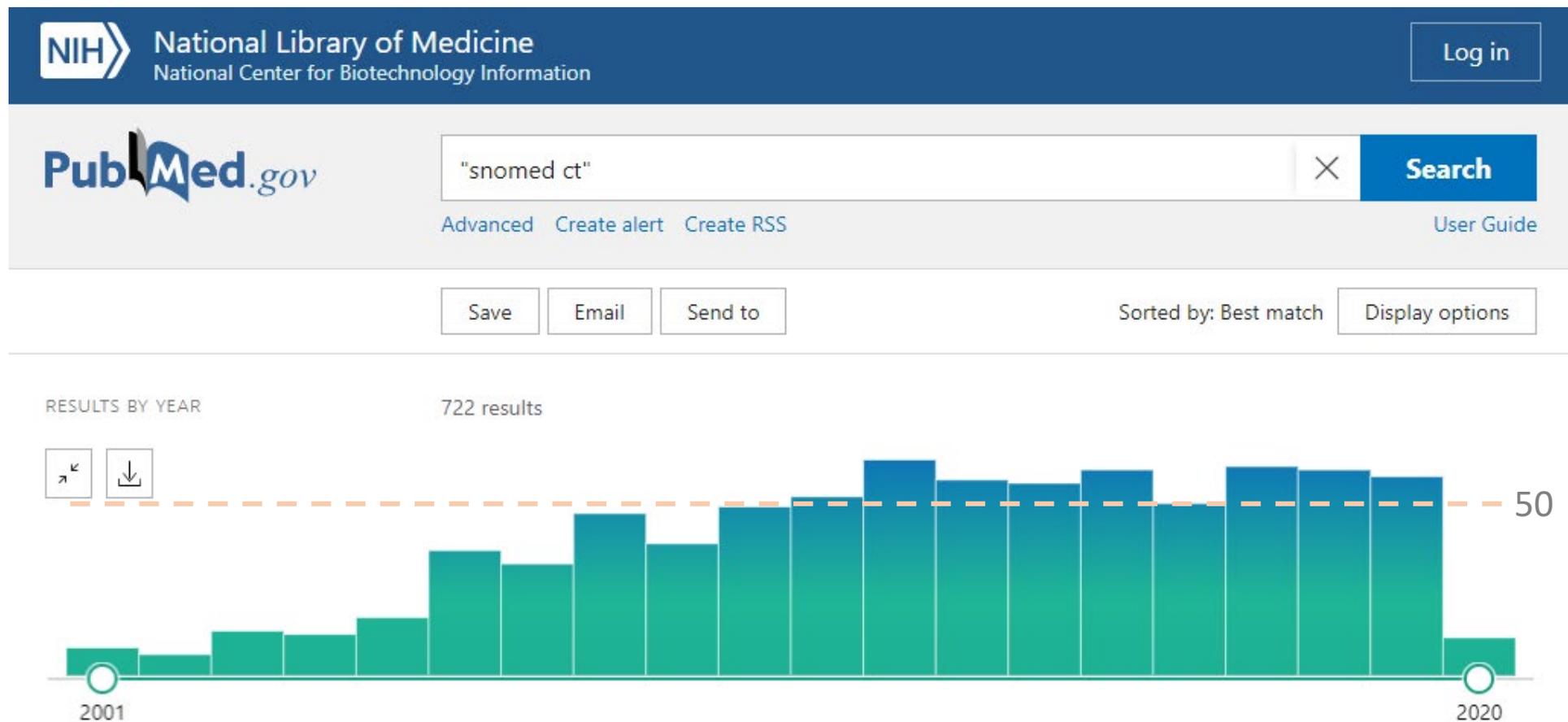
The GALEN project is developing language independent concept representation systems as the foundations for the next generation of multilingual coding systems. It aims to support the flexibility required to cope with the diversity amongst medical applications, while ensuring the coherence necessary for integration and re-use of terminologies. GALEN is developing a fully compositional and generative formal system for modelling concepts: the GALEN Representation and Integration Language (GRAIL) Kernel. Its goal is to overcome many of the problems with traditional coding and classification systems, in particular the combinatorial explosion of terms in enumerative systems and the generation of nonsensical terms in partially compositional systems. It will also provide a clean separation between the concept model and linguistic mechanisms which interpret that model (i.e., the words in a specific language, syntax, alternative phrasings, etc.) in order to allow the development of multilingual systems. GRAIL aims to be formally sound and produce models that are verifiable and contain no contradictions or ambiguities, with realistic human effort. A Coding Reference (CORE) Model of medical terminology covering is being developed which aims to represent the core concepts in for example pathology, anatomy and therapeutics, that have widespread applicability in medical applications. It should also provide the basis for specialist extensions according to the formal principles of GRAIL. The main results of GALEN will be delivered as a Terminology Server (TeS) which encapsulates and coordinates the functionality of the concept module, multilingual module, and code conversion module, and also provides a uniform applications programming interface and network services for use by external applications.

# “The early days of SNOMED CT”

*Aspects of SNOMED CT research in PubMed 2001-2006*



# “SNOMED CT” in PubMed



# The early days – Building it

Mapping between SNOMED RT and Clinical terms version 3: a key component of the **SNOMED CT** development process.

Wang AY, Barrett JW, Bentley T, Markwell D, Price C, Spackman KA, Stearns MQ.

Proc AMIA Symp. 2001:741-5.

PMID: 11825284

The SNOMED clinical terms development process: refinement and analysis of content.

Wang AY, Sable JH, Spackman KA.

Proc AMIA Symp. 2002:845-9.

PMID: 12463944

The SNOMED RT Procedure Model.

Dolin RH, Spackman K, Abilla A, Correia C, Goldberg B, Konicek D, Lukoff J, Lundberg CB.

Proc AMIA Symp. 2001:139-43.

PMID: 11825170 **Free PMC article.**

# The early days – Assessing its value [1/2]

Improved coding of the primary reason for visit to the emergency department using SNOMED.

McClay JC, Campbell J.

Proc AMIA Symp. 2002:499-503.

PMID: 12463874 [Free PMC article.](#)

An evaluation of the usefulness of two terminology models for integrating nursing diagnosis concepts into SNOMED Clinical Terms.

Bakken S, Warren JJ, Lundberg C, Casey A, Correia C, Konicek D, Zingo C.

Int J Med Inform. 2002 Dec 18;68(1-3):71-7. doi: 10.1016/s1386-5056(02)00066-7.

PMID: 12467792

Evaluation of SNOMED coverage of Veterans Health Administration terms.

Penz JF, Brown SH, Carter JS, Elkin PL, Nguyen VN, Sims SA, Lincoln MJ.

Stud Health Technol Inform. 2004;107(Pt 1):540-4.

PMID: 15360871

Evaluation of the content coverage of **SNOMED CT**: ability of SNOMED clinical terms to represent clinical problem lists.

Elkin PL, Brown SH, Husser CS, Bauer BA, Wahner-Roedler D, Rosenbloom ST, Speroff T.

Mayo Clin Proc. 2006 Jun;81(6):741-8. doi: 10.4065/81.6.741.

PMID: 16770974



# The early days – Assessing its value [1/2]

Use of **SNOMED CT** to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research.

Richesson RL, Andrews JE, Krischer JP.

J Am Med Inform Assoc. 2006 Sep-Oct;13(5):536-46. doi: 10.1197/jamia.M2093. Epub 2006 Jun 23.

PMID: 16799121

Development and evaluation of methods for structured recording of heart murmur findings using **SNOMED-CT** post-coordination.

Green JM, Wilcke JR, Abbott J, Rees LP.

J Am Med Inform Assoc. 2006 May-Jun;13(3):321-33. doi: 10.1197/jamia.M1973. Epub 2006 Feb 24.

PMID: 16501179

Representing natural-language case report form terminology using Health Level 7 Common Document Architecture, LOINC, and **SNOMED-CT**: lessons learned.

Hunscher D, Boyd A, Green LA, Clauw DJ.

AMIA Annu Symp Proc. 2006;2006:961.

PMID: 17238580 [Free PMC article.](#)

# The early days – Assessing its quality

## Ontology-based error detection in **SNOMED-CT**.

Ceusters W, Smith B, Kumar A, Dhaen C.

Stud Health Technol Inform. 2004;107(Pt 1):482-6.

PMID: 15360859

## Reliability of **SNOMED-CT** coding by three physicians using two terminology browsers.

Chiang MF, Hwang JC, Yu AC, Casper DS, Cimino JJ, Starren JB.

AMIA Annu Symp Proc. 2006;2006:131-5.

PMID: 17238317 [Free PMC article.](#)

## The semantics of procedures and diseases in **SNOMED CT**.

Schulz S, Hanser S, Hahn U, Rogers J.

Methods Inf Med. 2006;45(4):354-8.

PMID: 16964349

## Comparing the representation of anatomy in the FMA and **SNOMED CT**.

Bodenreider O, Zhang S.

AMIA Annu Symp Proc. 2006;2006:46-50.

PMID: 17238300 [Free PMC article.](#)

# The early days – Mapping/integration

Terminological mapping for high throughput comparative biology of phenotypes.

Lussier YA, Li J.

Pac Symp Biocomput. 2004;202-13. doi: 10.1142/9789812704856\_0020.

PMID: 14992504 [Free PMC article.](#)

Integrating **SNOMED CT** into the UMLS: an exploration of different views of synonymy and quality of editing.

Fung KW, Hole WT, Nelson SJ, Srinivasan S, Powell T, Roth L.

J Am Med Inform Assoc. 2005 Jul-Aug;12(4):486-94. doi: 10.1197/jamia.M1767. Epub 2005 Mar 31.

PMID: 15802483 [Free PMC article.](#)

Standardized nursing language in the systematized nomenclature of medicine clinical terms: A cross-mapping validation method.

Lu DF, Eichmann D, Konicek D, Park HT, Ucharattana P, Delaney C.

Comput Inform Nurs. 2006 Sep-Oct;24(5):288-96. doi: 10.1097/00024665-200609000-00011.

PMID: 16980782



# The early days – Use as a knowledge source

## Classifying diseases with respect to anatomy: a study in **SNOMED CT**.

Burgun A, Bodenreider O, Mouglin F.

AMIA Annu Symp Proc. 2005;2005:91-5.

PMID: 16779008 [Free PMC article.](#)

## Inter-patient distance metrics using **SNOMED CT** defining relationships.

Melton GB, Parsons S, Morrison FP, Rothschild AS, Markatou M, Hripcsak G.

J Biomed Inform. 2006 Dec;39(6):697-705. doi: 10.1016/j.jbi.2006.01.004. Epub 2006 Feb 24.

PMID: 16554186 [Free article.](#)

## Contribution to terminology internationalization by word alignment in parallel corpora.

Deléger L, Merkel M, Zweigenbaum P.

AMIA Annu Symp Proc. 2006;2006:185-9.

PMID: 17238328 [Free PMC article.](#)

# The early days – Summary

- A few SNOMED CT papers at the very beginning (“building it”)
- Adoption by the research community
  - Applied research – Coverage, utility
    - Mostly clinical communities
    - Pushing the envelope – e.g., clinical research
  - “Basic” research
    - Terminology research – New methods for quality assurance, mapping
    - Uses beyond terminology – Semantic similarity, knowledge source
- Sustained research: ~50 articles each year in PubMed

# KR-MED 2008

*Representing and sharing knowledge using SNOMED*





- home
- sponsoring opportunities
- call for papers
- instructions for authors
- program
- important dates
- committees
- location
- registration
- proceedings
- previous conferences

## KR-MED 2008

### Representing and sharing knowledge using SNOMED

May 31 - June 2, 2008, Phoenix, Arizona, USA

Conference organized by the International Health Terminology Standards Development Organisation (IHTSDO®) and the Working Group on Formal (Bio-)Medical Knowledge Representation of the American Medical Informatics Association (AMIA).

Collocated with the 2008 AMIA Spring Congress.

The vision of a universal clinical terminology, covering a broad range of health-related domains and meeting the needs of all health professionals has stimulated numerous health informatics research activities in the last two decades.

SNOMED CT® (Systematized Nomenclature of Medicine-Clinical Terms) is emerging as a comprehensive, multilingual clinical healthcare terminology, since 2007 under a new international ownership. KR-MED 2008 will follow up the successful first SNOMED conference organized in Copenhagen in October 2006, and will address health policy makers, clinicians, nurses, system developers, informatics researchers, computer scientists, terminologists and translators. KR-MED 2008 is also the third event of the AMIA working group KR-MED. It will therefore challenge the current state and the planned roadmap of SNOMED CT development from a perspective of knowledge representation and formal ontologies.

A number of prominent invited speakers will provide an overview of current efforts and developments in the context of clinical terminologies and state-of-the art ontology and terminology development.

#### Sponsors:



<https://www.kr-med.org/2008/index.html>



## May 31

|         |  |
|---------|--|
| 1:30 pm | <b>Tutorials</b><br><br>1. Terminological, Ontological and Knowledge Representation Aspects<br>Stefan Schulz, Kent Spackman  <br><br>2. Translating SNOMED CT<br>Ulrich Andersen, Asta Høy   |
| 3:00 pm | Coffee Break   |
| 3:30 pm | <b>Tutorials (continued)</b>   |
| 5:00 pm | Opening  |
| 5:30 pm | <b>Invited Speech</b><br>Building medical ontologies using Description Logics: what does it buy us?<br>Franz Baader, University of Dresden, Germany  |
| 6:30 pm | Reception  |

8:30 am **Invited Speech**  
 Managing Clinical Terminology in a Post-Classification Era  
 James R. Campbell, University of Nebraska 

9:30 am Coffee Break

10:00 am **Scientific Session I (two parallel sessions)**

Session I A - Formalization and Classification

1. Debugging SNOMED CT Using Axiom Pinpointing in the Description Logic EL+  
 Boontawee Suntisrivaraporn and Franz Baader 
2. Exploiting Fast Classification of SNOMED CT for Query and Integration of Health Data  
 Michael Lawley 
3. Why do it the hard way? The Case for an Expressive Description Logic for SNOMED  
 Alan Rector, Sebastian Brandt and Jay Kola 

Session I B - Applications

1. Leveraging SNOMED CT with a General Purpose Terminology Server  
 Robert Weida, Jack Bowie, Robert McClure and David Sperzel 
2. LinkBase® and SNOMED: some distinct features and impact on NLP  
 Maria van Gorp, Marnix Holvoet and Mariana Casella dos Santos 
3. SNOMED CT: Browsing the Browsers  
 Jeremy Rogers and Olivier Bodenreider 

3:30 pm **Scientific Session II (two parallel sessions)**

Session II A - Mapping and comparing

1. Comparing SNOMED CT and the NCI Thesaurus through Semantic Web Technologies  
 Olivier Bodenreider 
2. Exploratory Reverse Mapping of ICD-10-CA to SNOMED CT  
 Dennis Lee and Francis Lau 

Session II B - SNOMED CT and Information Models

1. Interoperability of Data Models and Terminology Models : Issues with using the SNOMED CT terminology  
 Rahil Qamar Siddiqui, Jay Kola and Alan Rector 
2. Essential SNOMED: Simplifying SNOMED-CT and supporting Integration with Health Information Models  
 Peter MacIsaac, Donald Walker, Rachel Richesson, Heather Grain, Peter Elkin and Jon Patrick 

June 2<sup>nd</sup>

8:00 am **Scientific Session III (two parallel sessions)**

Session III A - Representation

1. Comparing the Effects of Two Semantic Terminology Models on Classification of Clinical Notes: A Study of Heart Murmur Findings  
Guoqian Jiang and Christopher G Chute 
2. Creation and Usage of a "Micro theory" for Long Bone Fractures: An Experience Report  
Howard Goldberg, Vipul Kashyap and Kent Spackman
3. Representing clinical information using SNOMED Clinical Terms with different structural information models  
David Markwell, Laura Sato and Edward Cheetham 

Session III B - Mapping and Subsetting

1. Strategies for Updating Terminology Mappings and Subsets  
John Mapoles, Corey Smith, Jane Cook and Brian Levy 
2. Using SNOMED CT as a Mediation Terminology: Mapping Issues, Lessons Learned, and Next Steps Toward Achieving Semantic Interoperability  
Sarah Maulden, Patty Greim, Omar Bouhaddou, Pradnya Warnekar, Laura Megas, Fola Parrish and Michael Lincoln 
3. Using SNOMED CT For Translational Genomics Data Integration  
Joel Dudley, David P. Chen and Atul Butte 

10:00 am **Scientific Session IV (two parallel sessions)**

Session IV A

1. A Methodology for Encoding Problem Lists with SNOMED CT in General Practice  
Francis Lau, Raymond Simkus and Dennis Lee 
2. Post-Coordination in the Mapping of Interface Terms of a Clinical Wound Documentation System to SNOMED CT  
Martin Boeker, Stefan Schulz and Thilo Schuler 
3. ~~Using SNOMED CT Concepts for PAIRS~~  
Madan Rao

Session IV B - State of Affairs in SNOMED CT Member Countries (panel)

11:30 am **Invited Speech**  
SNOMED CT Adoption: Lessons and Challenges  
Howard S. Goldberg, Partners Healthcare 

12:30 pm Closing Remarks

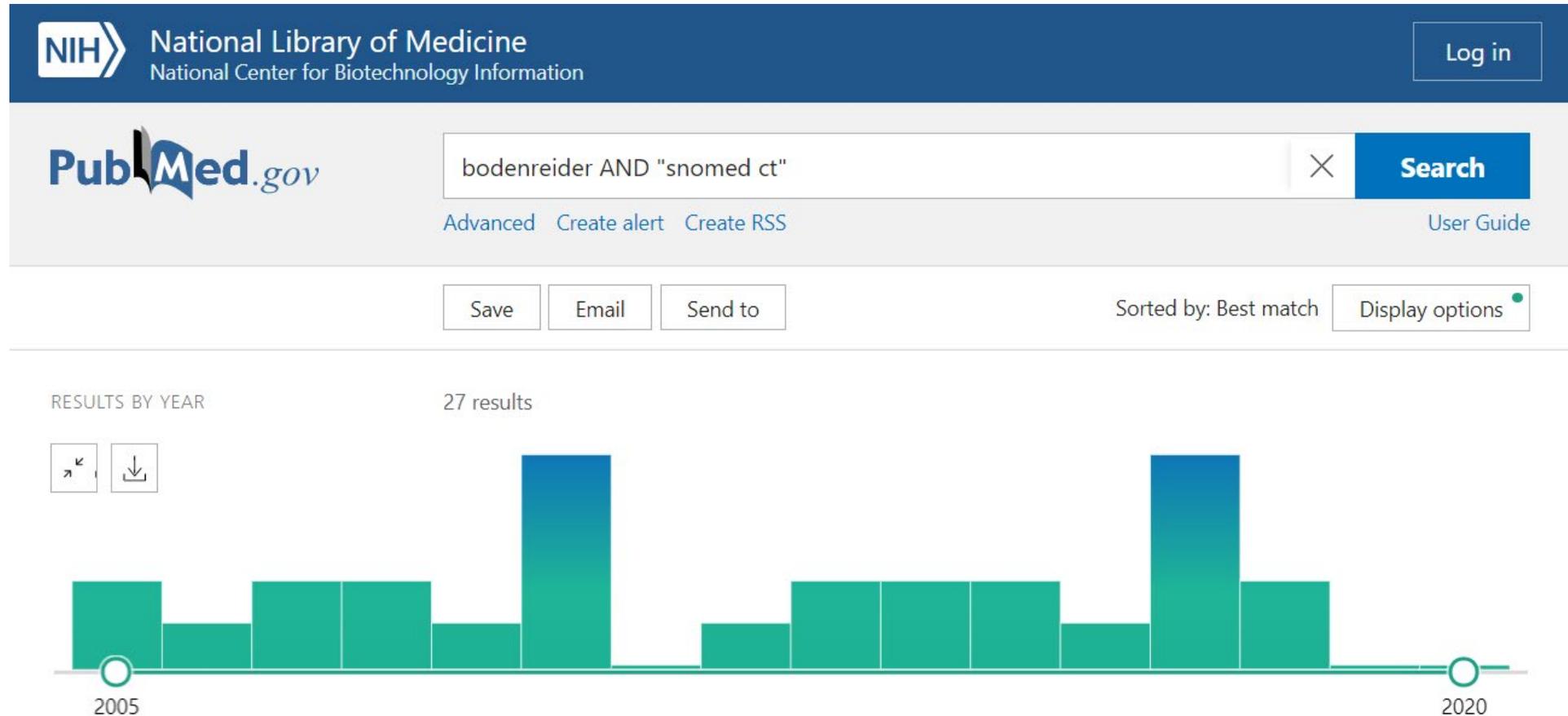
# KR-MED 2008 – Summary

- The medical informatics research community was very engaged and driving the agenda (Sponsored by AMIA)
- Even attracted the computer science community
- The only KR-MED entirely dedicated to SNOMED CT
  - Previous KR-MED had some SNOMED CT papers
  - Later on, ICBO has been dominated by OBO, with less clinical ontology and occasional SNOMED CT papers
    - Until ICBO 2019, where SNOMED International was a sponsor

# My own research journey with SNOMED CT



# “SNOMED CT” in PubMed – Bodenreider edition



# Review/education

Biomedical ontologies in action: role in knowledge management, data integration and decision support.

Bodenreider O.

Yearb Med Inform. 2008;67-79.

PMID: 18660879 [Free PMC article.](#)

Recent Developments in Clinical Terminologies - **SNOMED CT**, LOINC, and RxNorm.

Bodenreider O, Cornet R, Vreeman DJ.

Yearb Med Inform. 2018 Aug;27(1):129-139. doi: 10.1055/s-0038-1667077. Epub 2018 Aug 29.

PMID: 30157516 [Free PMC article.](#)

# Quality assurance

Investigating subsumption in **SNOMED CT**: an exploration into large description logic-based biomedical terminologies.

Bodenreider O, Smith B, Kumar A, Burgun A.

Artif Intell Med. 2007 Mar;39(3):183-95. doi: 10.1016/j.artmed.2006.12.003. Epub 2007 Jan 22.

PMID: 17241777 [Free PMC article.](#)

Using SPARQL to Test for Lattices: application to quality assurance in biomedical ontologies.

Zhang GQ, Bodenreider O.

Semant Web ISWC. 2010;6497:273-288. doi: 10.1007/978-3-642-17749-1\_18.

PMID: 25699294

Auditing **SNOMED CT** hierarchical relations based on lexical features of concepts in non-lattice subgraphs.

Cui L, Bodenreider O, Shi J, Zhang GQ.

J Biomed Inform. 2018 Feb;78:177-184. doi: 10.1016/j.jbi.2017.12.010. Epub 2017 Dec 20.

PMID: 29274386 [Free PMC article.](#)



# “before it was cool”/trial balloons [1/2]

## Comparing the representation of anatomy in the FMA and **SNOMED CT**.

Bodenreider O, Zhang S.

AMIA Annu Symp Proc. 2006;2006:46-50.

PMID: 17238300 [Free PMC article.](#)

## Issues in mapping LOINC laboratory tests to **SNOMED CT**.

Bodenreider O.

AMIA Annu Symp Proc. 2008 Nov 6;2008:51-5.

PMID: 18999311 [Free PMC article.](#)

## Using **SNOMED CT** in combination with MedDRA for reporting signal detection and adverse drug reactions reporting.

Bodenreider O.

AMIA Annu Symp Proc. 2009 Nov 14;2009:45-9.

PMID: 20351820 [Free PMC article.](#)

# “before it was cool”/trial balloons [2/2]

Coverage of rare disease names in standard terminologies and implications for patients, providers, and research.

Fung KW, Richesson R, Bodenreider O.

AMIA Annu Symp Proc. 2014 Nov 14;2014:564-72. eCollection 2014.

PMID: 25954361 [Free PMC article.](#)

Interoperability between phenotypes in research and healthcare terminologies--  
Investigating partial mappings between HPO and **SNOMED CT**.

Dhombres F, Bodenreider O.

J Biomed Semantics. 2016 Feb 9;7:3. doi: 10.1186/s13326-016-0047-3. eCollection 2016.

PMID: 26865946 [Free PMC article.](#)

# My own journey – Summary

- Interesting collaborations
- Lots of fun investigations
  - Cool methods
  - Applied at scale
- Privileged to investigate issues before they became mainstream
- Minor impact on SNOMED CT overall
  - Not integrated in the SNOMED CT development process (e.g., QA)

# Final thoughts

*Why should the SNOMED community and the research community care?*



# Different communities

- Two active, but separate communities with different goals and cultures
  - SNOMED CT community of practice
    - Results-driven, pragmatic
    - On-time, at-scale
    - Long-term
    - SNOMED CT Expo
  - Research community (biomedical informatics)
    - Methods-driven
    - Toy examples, limited scale
    - Short-term (e.g., grant)
    - Scientific conferences

# How the research community perceives SNOMED CT

- Pros
  - Rich and voluminous material
  - Interesting problems for driving research
- Cons
  - Constraints imposed by the license
    - Cannot distribute materials freely
  - Perceived lack of interest from the SNOMED CT community of practice
  - Lack of a process for engaging the research community\*
  - More interested in working with other communities of practice or SDOs

\*until recently

# How the SNOMED CT community perceives research\*

\*until recently

- Pros
  - Potential source of innovation and experimentation
- Cons
  - Not always involved clinically / disconnected from practice
  - Often not engaged in the long term
  - Distraction; not essential to business
  - Mild hostility: Some researchers have been openly critical of SNOMED CT

# Suggestions for moving forward? (researcher's perspective)

- SNOMED CT Research Webinar series – Thanks, Suzy!
- SNOMED International convening/sponsoring research workshops
- Research license
- Organize research efforts (e.g., for clinical analytics)
  - E.g., testing SNOMED CT at scale in clinical data warehouses
- Provide support for evaluation
  - Access to SNOMED CT experts
  - In-kind contribution to research efforts



[olivier.bodenreider@nih.gov](mailto:olivier.bodenreider@nih.gov)



U.S. National Library of Medicine