SNOMED AI Symposium
Building Foundations

Atlanta, USA
October 2023
Don Sweete
Chief Executive Officer
SNOMED International
Today’s Agenda

Keynote presentation - Paula Braun

VCR - Dr. Halamka

Panel Discussion: SNOMED CT and the Role of Terminologies
- Will Hardman
- Anthony Shek
- Igor Couto
- Christian Reich

BREAK

VCR - 3Data (nuvie)

Presentation - Charlie Harp

Entity Linking Challenge - Rory Davidson
Paula Braun
Entrepreneur in Residence
CDC/ONC
Standards + AI = :-)  
It’s Not an Either / Or Situation
Why We Do What We Do

Help Deliver Care That Is

- Proactive
- Collaborative
- Highly Personalized

Learning Health Systems

Systematically gather and create evidence.
Apply the most promising evidence to improve care.
Standards Help Make Data Retrievable, Understandable, and Actionable
United Federal Action Supports Standards
Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1) Proposed Rule

ONC's HTI-1 proposed rule seeks to implement provisions of the 21st Century Cures Act and make updates to the ONC Health IT Certification Program (Certification Program) with new and updated standards, implementation specifications, and certification criteria. Implementation of the proposed rule's provisions will advance interoperability, improve transparency, and support the access, exchange, and use of electronic health information.

Key provisions of the proposed rule include:

- Implementing the Electronic Health Record Reporting Program as new Condition and Maintenance of Certification requirements (Insights Condition) for developers of certified health information technology (health IT) under the Certification Program.
- Modifying and expanding exceptions in the information blocking regulations to support information sharing and certainty for regulated actors.
- Revising several Certification Program certification criteria, including existing criteria for clinical decision support (CDS), patient demographics and observations, electronic case reporting, and application programming interfaces for patient and population services.
- Raising the baseline version of the United States Core Data for Interoperability (USCDI) from Version 1 to Version 3.
- Updating standards adopted under the Certification Program to advance interoperability, support enhanced health IT functionality, and reduce burden and costs.
Predictive Decision Support Interventions

**Objective:** Enable improved information transparency on the trustworthiness of predictive DSIs to support their responsible and widespread use in health care.

**Improve Transparency**
Regarding how a predictive DSI is designed, developed, trained, evaluated, and should be used

**Enhance Trustworthiness**
Through transparency on how certified health IT developers manage potential risks and govern predictive DSIs that their certified Health IT Modules enable or interface with

**Support Consistency**
In the availability of predictive DSI information to users, so that users may determine the DSI’s quality and whether its recommendations are fair, appropriate, valid, effective, and safe (FAVES)

**Advance Health Equity by Design**
By addressing bias and health disparities, potentially propagated by predictive DSIs, to expand the use of these technologies in safer, more appropriate, and more equitable ways
New Frontiers for AI

Learn Complex Patterns Faster Across Larger and More Diverse Data
New Models Unleash New Possibilities

Artificial Intelligence
AI involves techniques that equip computers to emulate human behavior, enabling them to learn, make decisions, recognize patterns, and solve complex problems in a manner akin to human intelligence.

Machine Learning
ML is a subset of AI, uses advanced algorithms to detect patterns in large data sets, allowing machines to learn and adapt. ML algorithms use supervised or unsupervised learning methods.

Deep Learning
DL is a subset of ML which uses neural networks for in-depth data processing and analytical tasks. DL leverages multiple layers of artificial neural networks to extract high-level features from raw input data, simulating the way human brains perceive and understand the world.

Generative AI
Generative AI is a subset of DL models that generates content like text, images, or code based on provided input. Trained on vast data sets, these models detect patterns and create outputs without explicit instruction, using a mix of supervised and unsupervised learning.

(PDF) The Application of AutoML Techniques in Diabetes Diagnosis: Current Approaches, Performance, and Future Directions (researchgate.net)

One Model, One Task
Models Learn, Adapt, & Generate New Content
TLDR: It’s All About Transformers & Attention

What data do you need to focus on?

What data distract or are less important?

Meta-Transformer: A Unified Framework for Multimodal Learning (kxgong.github.io)
# Models, Models Everywhere - Example Use Cases

## Generative AI Models in Healthcare

<table>
<thead>
<tr>
<th>Language Models</th>
<th>Medical Models</th>
<th>Audio Models</th>
<th>Image Models</th>
<th>Chemical Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Notes / Papers)</td>
<td>(Codes / Concepts)</td>
<td>(Waveforms)</td>
<td>(Path / Radiology)</td>
<td>(RNA, DNA, Protein)</td>
</tr>
<tr>
<td>Chatbots</td>
<td>Call Center Automation</td>
<td>Chatbots</td>
<td>Call Center Automation</td>
<td>Chatbots</td>
</tr>
<tr>
<td>Formulary Design</td>
<td>Fraud Detection and Prevention</td>
<td>Formulary Design</td>
<td>Fraud Detection and Prevention</td>
<td>Formulary Design</td>
</tr>
<tr>
<td>Provider Referral and Engagement Management</td>
<td>Drug Discovery and Development</td>
<td>Provider Referral and Engagement Management</td>
<td>Drug Discovery and Development</td>
<td>Provider Referral and Engagement Management</td>
</tr>
<tr>
<td>Pharmacovigilance and Adverse Events</td>
<td>Digital Twins</td>
<td>Pharmacovigilance and Adverse Events</td>
<td>Digital Twins</td>
<td>Pharmacovigilance and Adverse Events</td>
</tr>
<tr>
<td>Medical Imaging and Diagnostics</td>
<td>Automation of Provider Notes</td>
<td>Medical Imaging and Diagnostics</td>
<td>Automation of Provider Notes</td>
<td>Medical Imaging and Diagnostics</td>
</tr>
<tr>
<td>Patient Matching</td>
<td>Episodes of Care</td>
<td>Patient Matching</td>
<td>Episodes of Care</td>
<td>Patient Matching</td>
</tr>
<tr>
<td>Education and Content Creation</td>
<td>Language Translation</td>
<td>Education and Content Creation</td>
<td>Language Translation</td>
<td>Education and Content Creation</td>
</tr>
</tbody>
</table>

# Generative AI in Healthcare: The Definitive Guide

[GenHealth.ai Blog](https://www.genhealth.ai)
Help People Understand What Drives their Health, with Guidance from their Physicians and Care Team
Navigating Choices About One’s Health Can be Hard and Confusing; Available Options Have Limitations

"OpenAI's models are not fine-tuned to provide medical information," a company spokesperson said. "You should never use our models to provide diagnostic or treatment services for serious medical conditions."

ChatGPT may be more accurate than other online medical advice: Shots - Health News: NPR
Imagine If . . .

What if every patient had access to a digital ally – like a mix of Waze and WebMD – solely dedicated to navigating their healthcare journey and improving their health outcomes?
Tailored, Trustworthy Information in the Pocket of Every Patient

Deliver patients a better understanding of what their doctors recommend and what their insurance covers, to help improve their care regime and increase their quality of life.
Leverage Existing Platforms and Standardized APIs that Help People Access and Organize their Data

Fetch Data via APIs + Make it Understandable using AI

- EHR
- Claims
- Sensors
- Imaging
- Molecular
- Patient Contributed
Step 1: Ensure Secure Access to Sensitive Data

Keep data secure from unauthorized access and tampering while also granting patients (and others they trust) access to their own data (and AI acting on their behalf).
Step 2: Train AI Across Larger and More Diverse Patient Profiles

Train generative AI models across these “honeycombs” to learn patterns and make discoveries across more representative, complex patient profiles.
Step 3: Assess AI’s Potential to Increase Patients’ Understanding and Improve Their Health Outcomes

Assess how safely, accurately, and reliably AI trained across the “honeycombs” can help communicate vital information about disease progression, treatment options, odds of success, and risks across diverse patient subpopulations.
Put Patients at the Center

- A more complete picture of what happening in your body – both during and in between doctors visits.
- Clear explanations and action steps that are tailored to your needs and are accessible to you.
- A care plan that makes sense and is based on your input. It spells out what’s happening, what’ll likely work, and what likely won’t based on your unique biology, personal situation, and real-world results from people like you.
- Reminders and feedback to help you see your progress, stay on track, and pivot when things change.
- Immediate and insightful responses to your questions, from credible sources.
What Role(s) Will Standards Play in AI?

"Let your hopes, not your hurts, shape your future." ~ Robert H. Schuller
Questions and Discussion
John Halamka on AI’s Future

Panel Discussion: SNOMED CT and the Role of Terminologies

Will Hardman
Data Scientist
Veratai

Anthony Shek
Clinical Data Scientist
Guy’s and St Thomas’ NHS Foundation Trust & King’s College

Igor Couto
CEO
Sofya

Christian Reich
CEO
Odysseus
SNOMED AI Symposium

Break
We will return shortly
A3Data - nuvie

https://a3data.com.br/
www.nuvie.ai
Charlie Harp
Chief Executive Officer
Clinical Architecture
Rory Davidson
Chief Information Officer
SNOMED International
Artificial Intelligence Technology Landscape

- Neuromorphic Computing
- Autonomous System
- Cognitive Cyber Security
- Machine Learning
- Robotic Personal Assistants
- Deep Learning
- Autonomous Surgical Robotics
- Neural Networks
- Next Gen Cloud Robotics
- Pattern Recognition
- Thought Controlled Gaming
- Natural Language Processing
- Real Time Universal Translation
- Chatbots
- Virtual Companions
- Real Time Emotion Analysis
Hype Cycle for Artificial Intelligence, 2023

- Smart Robots
- Responsible AI
- Neuromorphic Computing
- Prompt Engineering
- Artificial General Intelligence
- Decision Intelligence
- AI TRSM
- Operational AI Systems
- Composite AI
- Data-Centric AI
- AI Engineering
- AI Simulation
- Causal AI
- Neuro-Symbolic AI
- Multiagent Systems
- First-Principles AI
- Automatic Systems
- Generative AI
- Synthetic Data
- Foundation Models
- ModelOps
- EdgeAI
- Knowledge Graphs
- AI Maker and Teaching Kits
- Autonomous Vehicles
- Computer Vision
- Cloud AI Services
- Data Labeling and Annotation
- Intelligent Applications

Source: Gartner
© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. 2079794
Members, Implementers, Community

AI

Organization

Product
Explore how AI can be used to support implementation:

- Translation
- Generated Extension Content

Show how AI can enhance the “End User Experience”:

- LLM improved contextual search

Drive improvements in machine-assisted clinical coding:

- Entity Linking

Promote the incorporation of SNOMED CT in AI models:

- SNOMED CT providing “guardrails” for an LLM.

Use AI to support modelling and Quality Assurance:

- QA models & templates
Help uncover how entity linking models can improve healthcare delivery

The SNOMED CT Entity Linking Challenge will provide an opportunity to advance the development of efficient and reliable tools for **automating the coding of patient data**, facilitating **interoperability**, **decision support**, and improving healthcare delivery.
PhysioNet

- Data publishing platform at MIT
- Established in 1999
- Funded by NIH as an outreach arm of a research project
- Rebuilt in 2019
  - give datasets equal standing to the articles that they underpin

- https://physionet.org/

MIMIC-IV

- Highly-detailed critical care database
- >40k patients admitted to Beth Israel Medical Center
  - ECGs, waveforms
  - Vital signs, medications, labs tests
  - Free text notes
  - Chest X-rays
  - Echocardiograms
- Extensively used across education, research, and industry
The Dataset

- Using MIMIC-IV in cooperation with PhysioNet
- Annotation completed with:
  - 6 annotators from the SNOMED community
  - 70,599 annotations across 272 documents
  - Estimated the overall inter-annotator agreement (IAA) – as measured by the Jaccard Similarity score – of 0.83 over the entire annotation dataset.
  - Annotation results document produced with lessons learnt
  - SNOMED CT Annotation Guideline also produced

- This was an ambitious and challenging annotation project, more complex than either of the nearest comparators from the literature.
- To our knowledge, this is the largest openly available clinical annotation dataset, with SNOMED CT used as the target terminology
The algorithms that perform best on additional real-world data will be publicly recognized on the challenge leaderboard, and $25,000 USD in cash prizes will be awarded to the top teams!
Timeline

• Registration opens on DrivenData.org: Today, October 25, 2023
  • snomed.org/entity-linking-challenge

• Register for access competition data through PhysioNet.
  • Real-world patient data is highly sensitive and difficult to share safely. The data for this competition has identifying factors removed, but still details the care of vulnerable and still-living people.
  • To access the competition data, each participant will need to register with PhysioNet under MIT’s agreement and complete a short online training course. Then you will have access to the world’s largest publicly available repository of patient data for this challenge and more!
Timeline

- **Live challenge period:** Starts early December, 2023 until mid February, 2024
  - Data scientists, researchers, and engineers from around the world will use their skills to build accurate predictive models using cutting-edge approaches in data science.
  - Participants will build ML models and submit containerized code for executing inference on the cloud. The results of the inference will be compared against real-world data.

- **Post-challenge prize awarding** at the SNOMED International April Business meetings in London, April 13-17, 2024

- At the end of the challenge, all prize-winning solutions will be shared under **an open source license for anyone to use and learn from.**
- The annotated dataset will also be **donated to PhysioNet to be part of the MIMIC-IV dataset.**
And... Go

snomed.drivendata.org
Thank You for Attending