HL7 AI Standards – Laying the Foundation

EHR WG Project – AI Data Lifecycle
Gary Dickinson FHL7
Mark Janczewski MD MPH

HL7 37th Annual Plenary and Work Group Meeting
11 September 2023
Phoenix, Arizona USA
The HL7 EHR WG has a well-established learning project focused on **Reducing Clinician Burden**

Key Objectives are to:
- Understand the substance, extent and impact of clinician burden
- Recognize root causes
- Identify success stories
- Support novel and disruptively innovative advances that will allow healthcare IT to reduce burden and improve care quality by better supporting clinical workflow
Considering AI...
Where We Started

- Starting in January 2022, we began to receive presentations/reports on AI – with emphasis not so much on artificial, but rather on assistive or augmented, intelligence
- We began to seriously consider the potential of AI to reduce burden – recognizing that AI has applications/benefits beyond this objective
- We then established an AI Focus Team to evaluate the optimum role for HL7 in AI standards development
- With expert input from many sources, we identified 21 AI Topics/Areas of Focus
- With further review, we prioritized the list and identified AI Data Lifecycle as our initial priority and project
<table>
<thead>
<tr>
<th>Item</th>
<th>Topic/Area of Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Contribute SDO expertise at standards development or providing a framework for such. Use SDO consensus process to review and approve AI artifacts/standards.</td>
</tr>
<tr>
<td>2</td>
<td>Establish “Contexts” within which AI Algorithms are developed and ultimately applied in health and healthcare.</td>
</tr>
<tr>
<td></td>
<td>a) ISO 13119 – Clinical knowledge resources - Metadata</td>
</tr>
<tr>
<td></td>
<td>b) Patient/Provider/Action Contexts</td>
</tr>
<tr>
<td></td>
<td>c) Problem Context</td>
</tr>
<tr>
<td>3</td>
<td>Create a label for every algorithm — analogous to a nutrition label, or a drug label — describing the data used to develop an algorithm, its usefulness and limitations, its measured performance, and its suitability for a given population.</td>
</tr>
<tr>
<td>4</td>
<td>Test and monitor the performance of algorithm-guided care within the settings in which it is deployed in an ongoing way.</td>
</tr>
<tr>
<td>5</td>
<td>Create best practices for establishing the usefulness, reliability, and fairness of AI algorithms that bring together different organizations to develop and test AI on data sets drawn from diverse and representative groups of patients.</td>
</tr>
<tr>
<td>6</td>
<td>Create a standard way for government, academia, and industry to monitor the behavior of AI algorithms over time.</td>
</tr>
<tr>
<td>7</td>
<td>Understand clinical context and goals of each algorithm and know what attributes — quality, safety, outcomes, cost, speed, and the like — are being optimized.</td>
</tr>
<tr>
<td>8</td>
<td>Learn how local variations in lifestyle, physiology, socioeconomic factors, and access to health care affect both the construction and fielding of AI systems and the risk of bias.</td>
</tr>
<tr>
<td>9</td>
<td>Assess the risk that AI might be used, intentionally or not, to maintain the status quo and reinforce, rather than eliminate, discriminatory policies.</td>
</tr>
<tr>
<td>10</td>
<td>Develop approaches for appropriate clinical use of AI in combination with human expertise, experience, and judgment, and discourage overreliance on, or unreflective trust of, algorithmic recommendations.</td>
</tr>
<tr>
<td>11</td>
<td>AI Data Lifecycle: AI algorithms depend on data quality and context. Consider how to bolster accountability, data quality, context and provenance at each step in the data lifecycle (capture, share, use). Develop Provenance Resource Profile (FHIR for AI).</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>12</td>
<td>Develop standards for auditing AI</td>
</tr>
<tr>
<td>13</td>
<td>Develop standards for CDS software to identify AI tools potentially applicable to a patient problem, combine data about the applicable tools (including but not limited to metadata in ISO 13119) with data about the patient/problem, and generate options (i.e. alternative tools) with details about the pros and cons of each option specific to the patient (see paragraph I.A.4 of the “Problem Context” memo in Item 2c above).</td>
</tr>
<tr>
<td>14A</td>
<td>Publication guidance for medical literature including precise markup optimized for extraction of knowledge constructs to support machine learning</td>
</tr>
<tr>
<td>14B</td>
<td>Translation of knowledge fragments into structured patterns, assuring validation and fidelity, and enabling input mapping to support the AI user interface and Natural Language Processing</td>
</tr>
<tr>
<td>15</td>
<td>Everyday patient care record, rigorously formatted – extracted for diagnosis, treatment, decisions, use for clinical trials</td>
</tr>
<tr>
<td>16</td>
<td>Data standards for data brokers</td>
</tr>
<tr>
<td>17</td>
<td>Create / enhance standards for capturing results from AI, which are linked back to the AI source</td>
</tr>
<tr>
<td>18</td>
<td>Create / enhance standards to translate AI results into existing medical domains while ensuring that they are correctly flagged as coming from AI and linked back to the source AI</td>
</tr>
<tr>
<td>19</td>
<td>Use HL7 FHIR resources/data structures for AI data strata</td>
</tr>
<tr>
<td>20</td>
<td>Develop a minimum viable standard for AI ethics specific to healthcare. This might incorporate the idea of making data FAIR and establishing a standard for developing AI ready datasets. (#10 seems to touch on Ethics and therefore it may be possible to address both)</td>
</tr>
<tr>
<td>21</td>
<td>Develop standardized approach towards the development of AI ready data sets, including research design, data generation, data processing</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Underway</td>
</tr>
<tr>
<td>Deferred to Other Organizations</td>
</tr>
<tr>
<td>Potential Concurrent Track or Subsequent Phase of Current Project</td>
</tr>
<tr>
<td>For Future Consideration</td>
</tr>
</tbody>
</table>
What Do We Aim to Achieve?

- Explainable AI
- Black Box gives way to Clarity
  - Trustworthy, Accountable, Fit for Purpose
  - Transparent and Fully Traceable – from Machine Learning to Algorithm Derivation to Live Operation
- AI Conclusions/Recommendations are Substantiable
  - Evidence-based
  - Measurable as to Data Accuracy and Quality
- Ethically Sound
- Initially Focused on Training and Developing Machine Learning (AI/ML) Models
  - NOT Generative AI leveraging Large Language Models
AI Project 11 - AI Data Lifecycle Framework

- Incorporates AI Use Case Examples
- Utilizes Trigger Events and Audit Trails (capturing key Provenance and Metadata elements)
- Patterned after FHIR R5 Record Lifecycle Event Implementation Guide
- Uses US General Accounting Office (GAO) Framework for an AI Project Lifecycle, with Distinct Phases:
  - Design Phase
  - Development Phase – including Machine Learning/Training and Algorithm Creation
  - Deployment Phase – including Testing
  - Production Phase – Live Operation with Continuous Monitoring
Project 11 – AI Data Lifecycle

GAO AI Accountability Framework – Phases 1-4

The Phases in the AI Life Cycle

- **Design**: Involves articulating the system’s concept and objectives, underlying assumptions, context and requirements, and potentially building a prototype.
- **Continuous monitoring**: Involves operating the AI system and continuously assessing its recommendations and impacts (both intended and unintended) in light of objectives and ethical considerations. This phase identifies problems and adjusts by reverting to other phases or, if necessary, retiring the AI system from production.
- **Development**: Involves planning and design, including establishing technical requirements, data collection and processing, model building and interpretation, and system verification and validation.
- **Deployment**: Involves piloting, checking compatibility with legacy systems, ensuring regulatory compliance, managing organizational change, and evaluating user experience.

US General Accountability Office (GAO) – “Artificial Intelligence An Accountability Framework for Federal Agencies and Other Entities” (2021) – Guidance for the AI Project Life Cycle, including AI for healthcare, which involves several critical phases, each requiring specific standards to ensure consistent high-quality implementation.
Project 11 – AI Data Lifecycle

Timeline

- Project Underway early 2023
- Project Proposal (Title and Description)
  - 22 Mar 2023: Approved
- Project Scope Statement (PSS)
  - 23 May 2023: Approved by HL7 EHR WG
  - 30 June 2023: Passed Consensus Review
  - 17 July 2023: Approved by Technical Steering Committee (TSC)
- 1st Deliverable: Laying the Foundation – White Paper
  - Draft in development
  - Ballot Cycle: 2024JAN

- Over 50 Individuals have Volunteered
Project 11 – AI Data Lifecycle – Initial Deliverable

White Paper

- So far, about 50 pages long
- Table of Contents (High-Level)
  - Introduction (Forward, Caveats, Scope)
  - Use Cases and Case Studies
  - Background (Definitions, Challenges Associated with the Use of AI in Healthcare, Benefits of AI Standardization
  - Discussion
  - Recommendations / Guidance
- Currently in Review and Draft Development
Many Thanks to:

- AkéLex
- Availity/Diameter Health
- Beth Israel Deaconess/Lahey Health
- Book Zurman
- Booz Allen Hamilton
- Centers for Medicare and Medicaid Services
- Cognotekt (Germany)
- Computer Network Architects
- EHR Standards Consulting
- Google Health
- Harvard University/Massachusetts Institute of Technology
- ISO TC215/Task Force 5 on AI
- JP Systems
- Mitre
- Mayo Platform
- Oak Ridge National Laboratories
- Philips
- Prosumer Health
- Stanford University
- University of Nebraska
- US Department of Defense
- US Food and Drug Administration
- US National Institutes of Health
- US General Accountability Office
- US National Institute of Standards and Technology
- US Office of National Coordinator
- US Veterans Administration
- Wolters-Kluwer
- Yale University
- and Many Other Individuals

→ Key Instigators
AI Focus Team and P11 Project – AI Data Lifecycle

Upcoming Sessions

- **This WGM:** Tuesday Q4 – hosted by the EHR WG
- **Recurring Meetings**
  - AI Focus Team: 4th Tuesday each month, 1600-1700 US ET
  - AI Project 11: Every second Monday, 1100-1200 US ET
  - Reducing Clinician Burden: 2nd and 4th Mondays each month, 1200-1300 US ET

- **Your Participation is Welcome!**
  - See POCs on next slide if you wish to be included on email distribution lists
  - To receive meeting announcements, links and materials
Points of Contact

- **AI Focus Team**
  - Gary Dickinson FHL7, EHR Standards Consulting
    gary.dickinson@ehr-standards.com
  - Steve Datena MD, AkéLex
    steve.datena@akelex.com

- **Project 11 – AI Data Lifecycle**
  - Rhonna Clark, US Department of Veterans Affairs – Project Lead
    rhonna.clark@va.gov
  - Ioana Singureanu, US Department of Veterans Affairs – Project Advisor
    ioana.singureanu@va.gov
  - Mark Janczewski MD MPH, Medical Networks LLC – Project Facilitator
    mark.janczewski@gmail.com