ShEx and FHIR/RDF

Eric Prud'hommeaux – W3C/MIT
(presenter)
Harold Solbrig – Mayo Clinic
Guoqian Jiang – Mayo Clinic
Clinical objectives

• Improve research, efficacy and personalization.
• Integrate clinical, medical and scientific data.
• Unify clinical and terminological expressions.

1\textsuperscript{st} step: Model FHIR in RDF.
What’s ShEx?

• Shape Expressions;
• Structural schema language for RDF.
• Analogous to XML Schema.
• Closed or Open validation.
• Accounts for all mentioned predicates.
What’s it look like?

• 3 formats:
  • Human (compact) syntax: ShExC
  • JSON (abstract) syntax: ShExJ
  • RDF (graph) syntax: ShExR
• See test suite
What's it do?

```
<MedicationRequest> {
  :dosageInstruction @<DosageInstruction>
}

<DosageInstruction> {
  :timing @<Timing>
}

<Timing> {
  ( :frequency xsd:integer
   | :when ["AC" "ACM" "ACD" "ACV"] )
}

<MedicationRequest/123>
  :dosageInstruction <DosageInstruction/456>.

<DosageInstruction/456>
  :timing <Timing/789> .

<Timing/789>
  :when "ACD" .
```
It complains
(before others do)

<MedicationRequest> { 
    :dosageInstruction @<DosageInstruction>
}

<DosageInstruction> { 
    :timing @<Timing>
}

<Timing> { 
    ( :frequency xsd:integer 
    | :when ["AC" "ACM" "ACD" "ACV"] )
}

<MedicationRequest/123> 
    :dosageInstruction <DosageInstruction/456>.

<DosageInstruction/456> 
    :timing <Timing/789> .

<Timing/789> 
    :when "AD" .
What’s FHIR?

- Resource-oriented,
- fast-moving,
  clinical information model.
### MedicationOrder

<table>
<thead>
<tr>
<th>Name</th>
<th>Card.</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>DosageInstruction</td>
<td>0..*</td>
<td>BackboneElement</td>
</tr>
<tr>
<td>Timing</td>
<td>0..1</td>
<td>Timing</td>
</tr>
</tbody>
</table>

### Timing

<table>
<thead>
<tr>
<th>Name</th>
<th>Card.</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event</td>
<td>0..*</td>
<td>dateTime</td>
</tr>
<tr>
<td>Repeat</td>
<td>0..1</td>
<td>Element</td>
</tr>
<tr>
<td>FrequencyMax</td>
<td>0..1</td>
<td>integer</td>
</tr>
<tr>
<td>When</td>
<td>0..1</td>
<td>code</td>
</tr>
</tbody>
</table>
Publication process

• Resource-oriented
  • HL7 governance for core Resources

• Fast-moving
  • Tooling works with authoritative sources
  • No one-time translations!

DSTU2
Alternate definitions: Schema/Schematron, Resource Profile (XML, JSON), Questionnaire

STU3
Alternate definitions: Master Definition (XML, JSON), XML Schema/Schematron (for ) + JSON Schema, ShEx (for Turtle)
- **Structure**
  - Ordered attributes – (XML Schema)
  - Datatypes – (XML Schema)
  - Co-occurrence – (Schematron)
- **Intra-resource coherence** – (reference library)

<table>
<thead>
<tr>
<th>Name</th>
<th>Flags</th>
<th>Card.</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>MedicationRequest</td>
<td></td>
<td></td>
<td>DomainResource</td>
</tr>
<tr>
<td>status</td>
<td></td>
<td></td>
<td>code</td>
</tr>
<tr>
<td>intent</td>
<td></td>
<td></td>
<td>proposal</td>
</tr>
<tr>
<td>medication[x]</td>
<td></td>
<td></td>
<td>code</td>
</tr>
<tr>
<td>medicationCodeableConcept</td>
<td></td>
<td></td>
<td>CodeableConcept</td>
</tr>
<tr>
<td>medicationReference</td>
<td></td>
<td></td>
<td>Reference(Medication)</td>
</tr>
<tr>
<td>subject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>requester</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>agent</td>
<td></td>
<td></td>
<td>Reference(Organization)</td>
</tr>
<tr>
<td>onBehalfOf</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reasonCode</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>reasonReference</td>
<td></td>
<td>0..*</td>
<td>Reference(Condition</td>
</tr>
<tr>
<td>dispenseRequest</td>
<td></td>
<td>0..1</td>
<td>BackboneElement</td>
</tr>
<tr>
<td>numberOfRepeatsAllowed</td>
<td></td>
<td>0..1</td>
<td>positiveInt</td>
</tr>
<tr>
<td>quantity</td>
<td></td>
<td>0..1</td>
<td>SimpleQuantity</td>
</tr>
<tr>
<td>performer</td>
<td></td>
<td>0..1</td>
<td>Reference(Organization)</td>
</tr>
</tbody>
</table>
<MedicationRequest> {
  :status @<code> AND
  {fhir:value ["active" "on-hold" "cancelled" "completed"]}?;
  :intent @<code> AND
  {fhir:value ["proposal" "plan" "order" "instance-order"]};
( :medicationCodeableConcept @<CodeableConcept> |
  :medicationReference @<MedicationReference> |
  :subject @<Patient> |
  :request @<Request> |
  :request @<Plan>
  :reasonCode @<Code> |
  :reasonReference @<Reference>*;
  :dispenseRequest {
    :dispenseRequest.numberOfRepeatsAllowed @<positiveInt>?;
    :dispenseRequest.quantity @<SimpleQuantity>?;
    :dispenseRequest.performer @<Reference>? } 

• Structure
  - Ordered attributes – (ShEx)
  - Datatypes – (ShEx)
  - Co-occurrence – (ShEx)
• Intra-resource coherence – (ShEx)
Why RDF?

• Clinical reality is a graph, not a tree.
• Linked data is engineered for integration
  • e.g. chemical or environmental
• Terminological OWL ontologies:
  - single query language to access the information model and inferences on the terminological model.

The question is not whether RDF is the right graph technology but whether preemptive normalization pays for itself.
Information model + Terminology

• Information model:
  • who
  • when
  • code for what
  • (code for why)

• Terminology model:
  • what
  • why
  • parents and siblings
  • formal semantics
Fruitful marriage

- Structured, subsumptive queries...
- Procedure:
  - 84203001 | Method |=32413006 | surgical transplantation
- Medication:
  - 69431002 | Immunosuppressant (product)

Still needed:
  Date comparison for (inferred) causality
Next steps

- Compositional terms in terminology.
- Extend ShEx to support FHIR Profiles.
- Linked data profile.
  - Applies to XML and JSON as well
Profiles

• Interoperability through constraints.
• Further refine core Resources:
  - cardinality
  - value sets

• Add inheritance to ShEx
  - already prototyped but not standardized
## Linked Data Profile

### Table 1: MedicationRequest

<table>
<thead>
<tr>
<th>Name</th>
<th>Card.</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>MedicationRequest</td>
<td></td>
<td>DomainResource</td>
</tr>
<tr>
<td>dosageInstruction</td>
<td>0..*</td>
<td>Dosage</td>
</tr>
</tbody>
</table>

### Table 2: Dosage

<table>
<thead>
<tr>
<th>Name</th>
<th>Card.</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dosage</td>
<td></td>
<td>Element</td>
</tr>
</tbody>
</table>

### Table 3: Timing

<table>
<thead>
<tr>
<th>Name</th>
<th>Card.</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timing</td>
<td></td>
<td>Element</td>
</tr>
</tbody>
</table>

#### Timing Attributes

- event: 0..* dateTime
- repeat: 0..1 Element
- duration: 0..1 decimal
- durationMax: 0..1 decimal
- durationUnit: 0..1 code
- timeOfDay: 0..* time
- when: 0..* code
ShEx ROI

- Enabled standardization of FHIR/RDF
- 400+ FHIR resource instances fixed after validation
- ShEx engine swapped in before STU3
  - No changes to resources – fully interoperable
  - V2.0 release Monday 27 March 2017
Questions?

(see also: JBI paper: “Modeling and validating HL7 FHIR profiles using semantic web Shape Expressions (ShEx)”)