XSDGuide – Automated Generation of Web Interfaces from XML Schemas

A Case Study for Suspicious Activity Reporting

Fabrizio Gotti – University of Montreal
Kevin Heffner – Pegasus Research & Technologies, Montreal
Guy Lapalme – University of Montreal

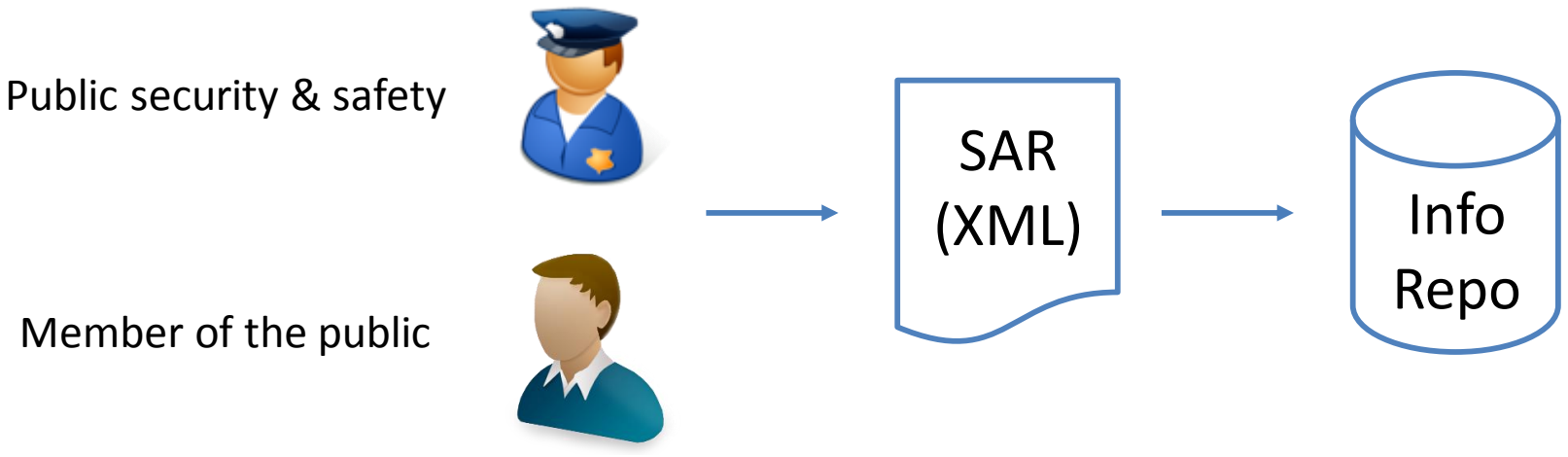
Balisage 2015

Pegasus RESEARCH & TECHNOLOGIES

Université de Montréal

Context of the work

- In the public security and safety communities, information exchange is a priority
- XSDGuide is a software prototype attempting to facilitate **Suspicious Activity Reporting (SAR)**



Date	2007-09-16	Reference ID	12345678
Status	Open	Dissemination Criteria	CHK
Point of Contact	Lou, Eddie	Phone Number	111 555 1212 x 14
Email	Lou@spd.gov	Organization	Springfield Police - Records Bureau

Target

Target Category	Structure	Critical?	true
Infrastructure Sector	Nuclear Reactors, Materials, and Waste Energy Utilities Power Plant		
Structure Category	Utility structure/ other commercial		
Description	Springfield Nuclear Power Plant is a fictional nuclear power plant in the television animated cartoon series The Simpsons. The plant, owned by Charles Montgomery Burns, is located at 100 Industrial Way. The plant has a monopoly on the city of Springfield's energy supply, and the carelessness of Mr. Burns and the plant's employees often endangers the residents and natural environment of Springfield.		

Suspicious Activity

Start	2007-09-16T19:24:00.0Z	End	2007-09-16T20:21:00.0Z
Status	Follow-Up has been initiated	Description	Espionage flights over Springfield Nuclear Power Plant.
Coordinates	44°3'11", 122°59'28"	Address	100 N Industrial Way. SW Sector 7 Springfield, Oregon 97477

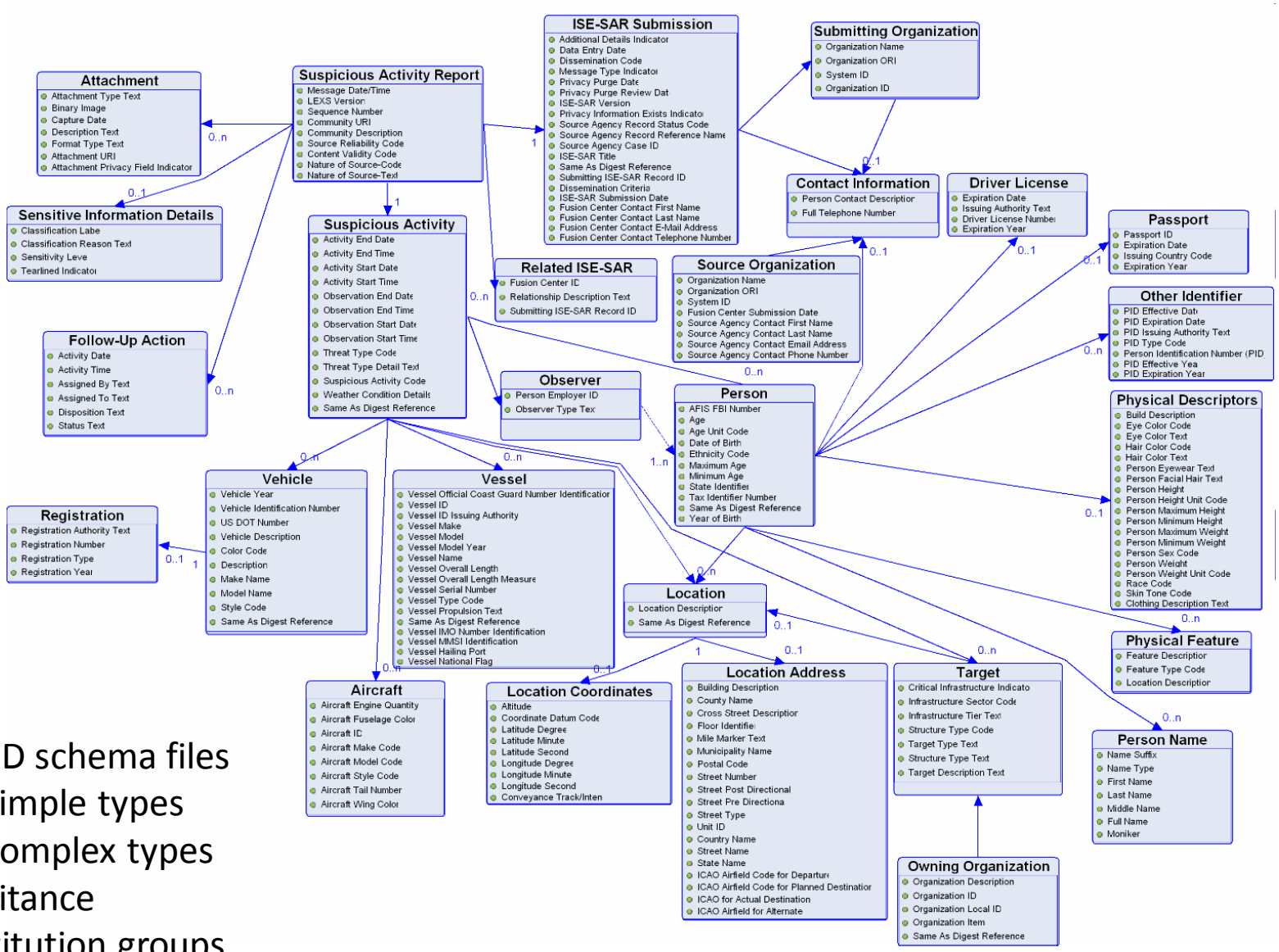
SAR and NIEM

- SAR one of set of messages supported by National Information Exchange Model (NIEM)
  - Created by the US Departments of Justice and Homeland Security
  - Interoperability standard approved by the US & Canada
  - XSD schemas (data model)

- NIEM-SAR
  - A SAR XSD importing/including NIEM-SAR XSDs
  - Operational, promising, but faster response times required

- XSDGuide attempts to reduce this workload

ISE-FS-200-version-1.5 Suspicious Activity Reporting



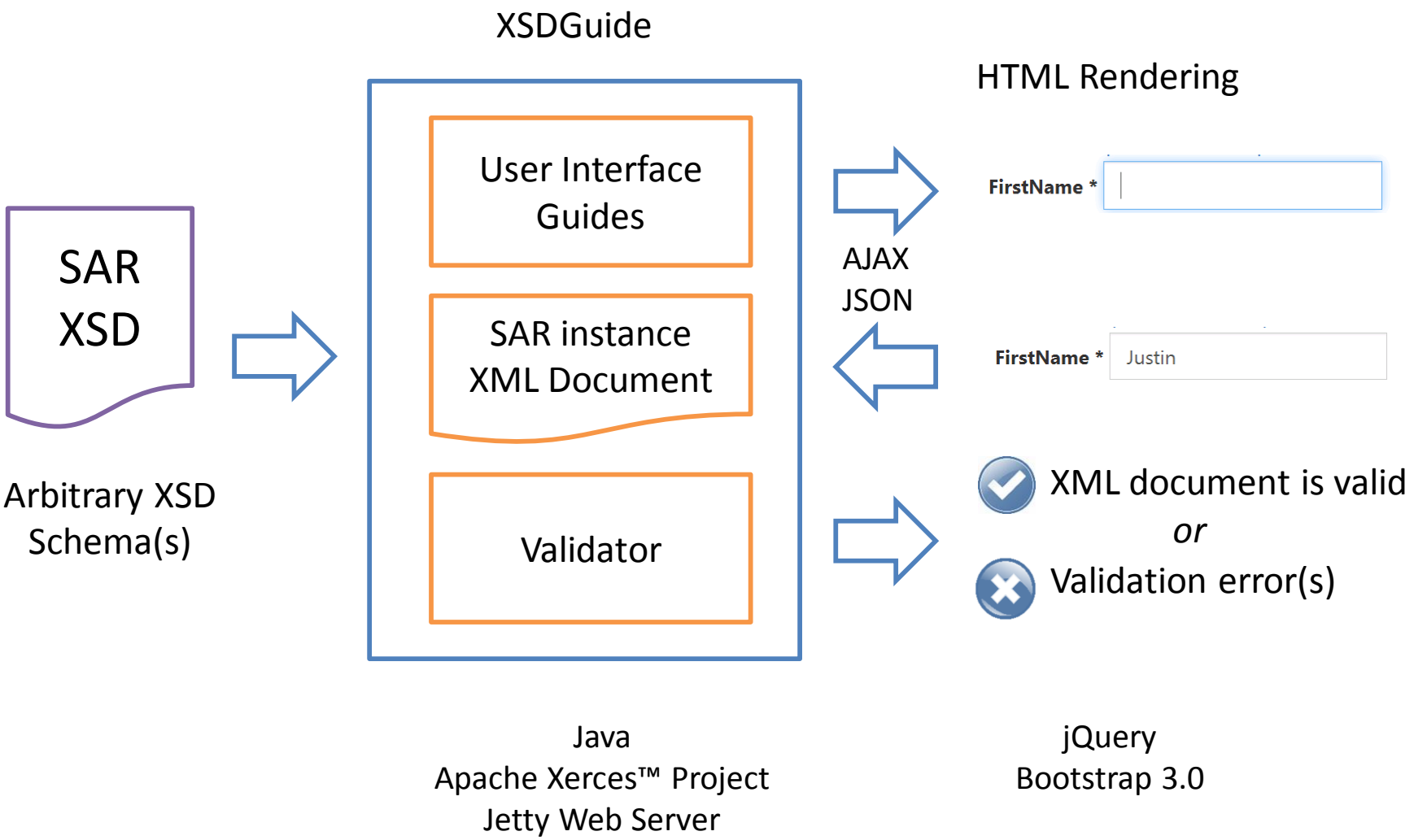
74 XSD schema files
196 simple types
658 complex types
Inheritance
Substitution groups

XSDGuide: a compromise

- Very difficult to reconcile conformance to such a standard and timely creation of SARs
- Powerful tools already exist (e.g. <oXygen/> XML Editor)

- XSDGuide: an effort to explore potential solutions

XSDGuide: general architecture



SAR-RALI.xsd used as the running example

```
<xs:element name="SuspiciousActivityReport">
    <xs:annotation>
      <xs:documentation>A structure that describes a SAR Report
</xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="sarrali:Metadata" />
        <xs:element ref="sarrali:Data"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>

  <xs:element name="Metadata">
    <xs:annotation>
      <xs:documentation>A structure that describes Metadata about a
related SAR</xs:documentation>
    </xs:annotation>
```

New report | Schema manager | Switch to simple view | Validate XML | Save to XML | Help

SuspiciousActivityReport

A structure that describes a SAR Report.

Metadata

A structure that describes Metadata about a related SAR.

Add new RelatedSarList (optional)

UniqueId *

Title *

SubmissionSystem

Element that uniquely identifies an organization and a system where date originated, was submitted from, or is being sent.

OrganizationName *

Identifier *

Author

Contact information for the system owner. Includes a person and organization to contact and their phone number and email address.

AuthorPerson

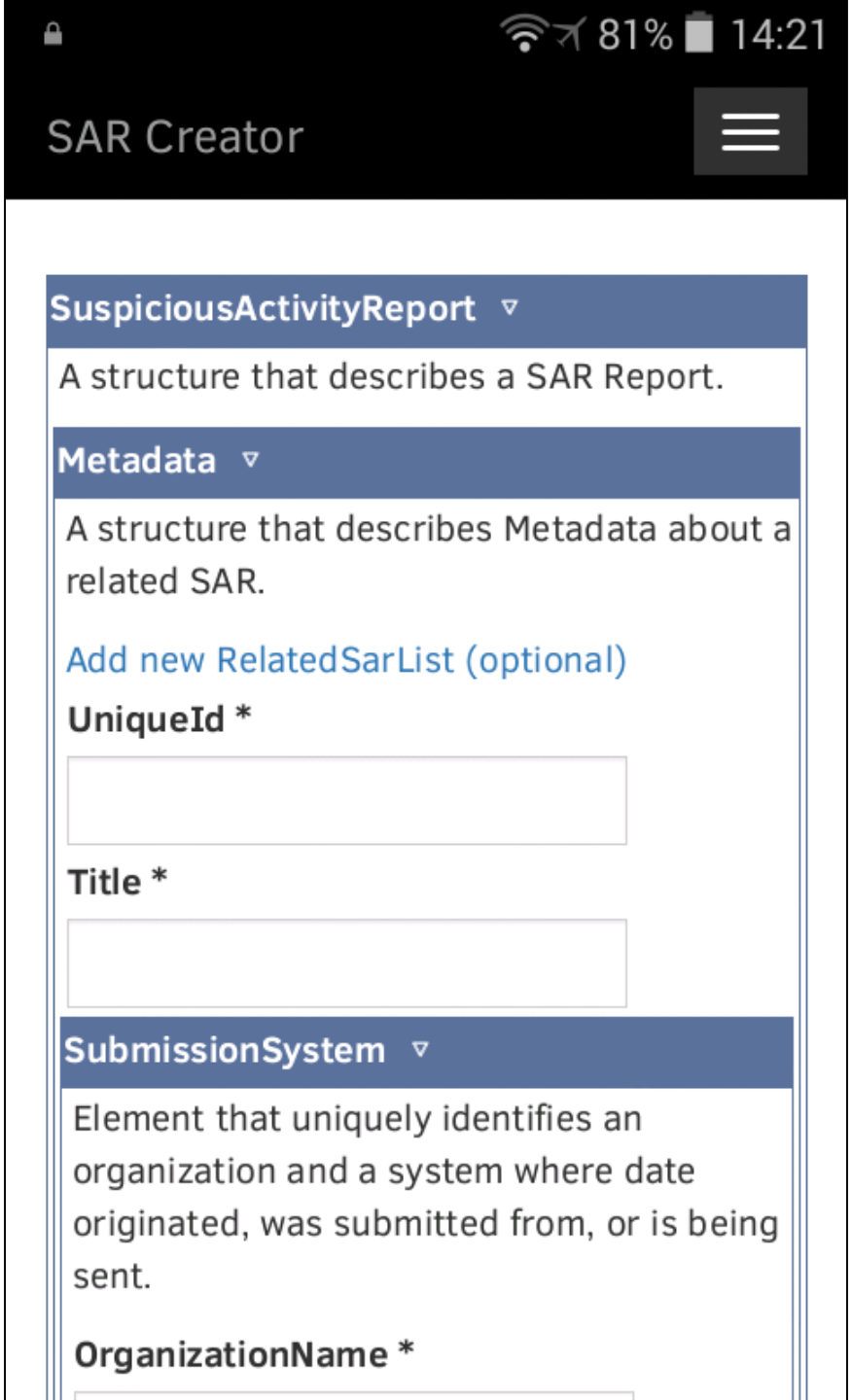
The person who authors the report.

Add new TelephoneNumber (optional)

FirstName *

LastName *

OrganizationName *



Android smart phone

A demo

- A short demonstration video 

- More on our website
  http://rali.iro.umontreal.ca/rali/en/xsdguide

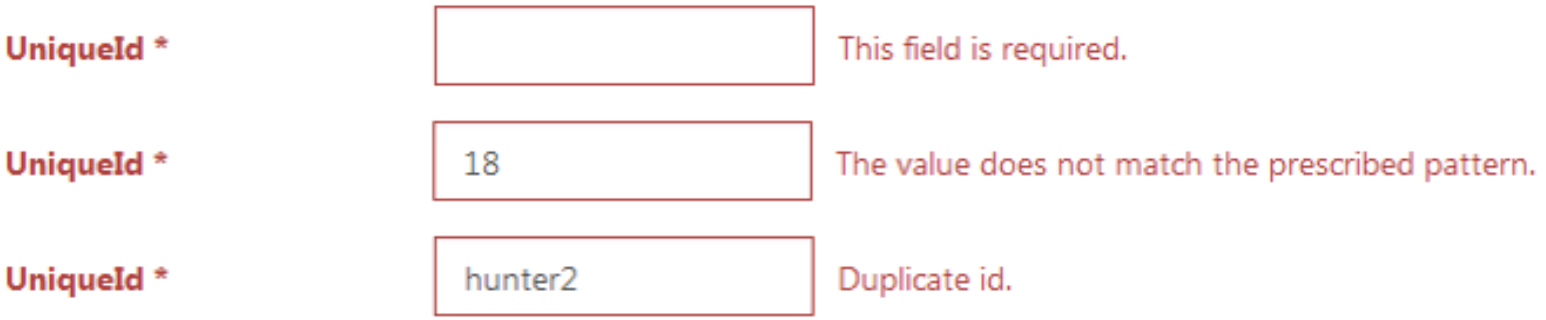
HTML Rendering: Form libraries

- At the core of the UI: good forms
  - Intuitive
  - Good input sanitation
- Numerous library possibilities, among which
  - XForms (well thought-out, poorly supported)
  - HTML5 controls (limited features)
  - Others (not fully adapted to our specific needs)
- In-house implementation

XML validation: A two-step process

1. Client-side (in the browser)

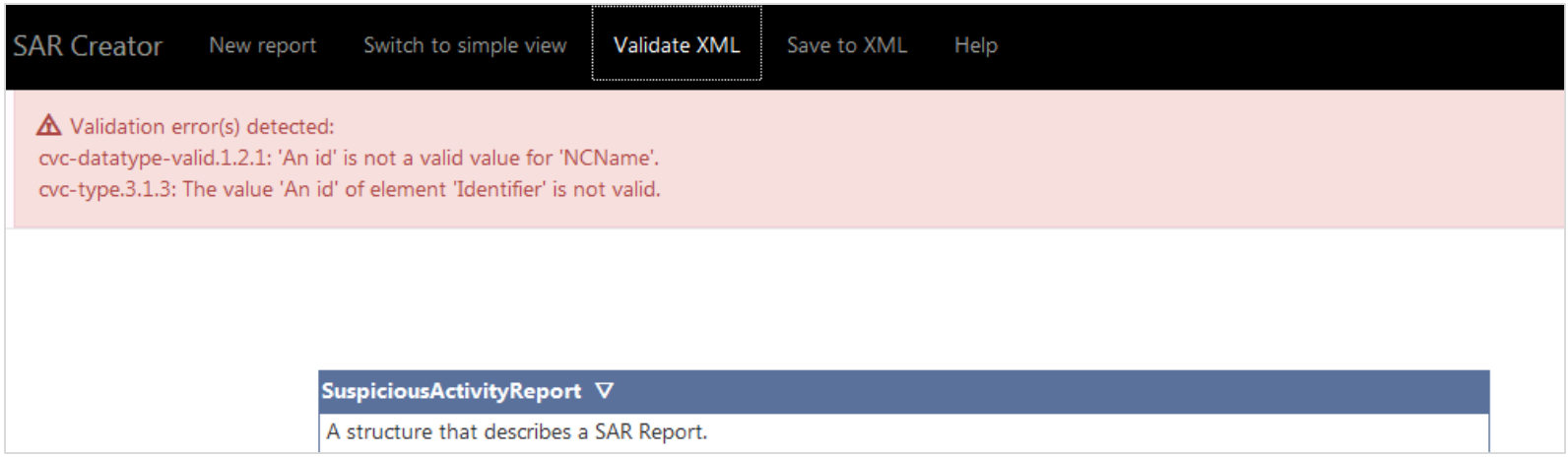
- Interested in the data entered, not the structure
- Elements checked
  - Required fields
  - Most *facets*
  - xs:id and xs:idref
- Detailed feedback to the user



XML validation: A two-step process

2. Server-side (full validation)

- Only when first validation is ok
- Elements checked
  - Full validation with Apache Xerces™ engine
- Difficult to pinpoint the problem to the user



Not a text-based XML instance: no column/row

Other features

- Saving the SAR report (with its `xsi:schemaLocation`)
- Schema uploads (including IEPDs)

Limitations: XSD rules

- We focused our efforts on rules found in NIEM-SAR IEPDs
- XSD rules are not all implemented

- Easy to implement
  - Distinctions in text-like types (`string, ncname, nmtoken`…)
  - Some facets (`totaldigits, whitespace`)
  - `xsd:any`

- Fixable with effort
  - `xs:sequence`, `xs:choice` and `xs:all` cannot have cardinality more than 1
  - Regular expressions from XSD patterns ≠ JavaScript regexps
    e.g. range subtraction `[...-[...]]` does not exist in JavaScript

Limitations: XSD rules

- Harder problems, unseen in NIEM-SAR
  - Identity constraints
    - `unique` element
    - `key` and `keyref` elements
    - Etc.
  - `mixed` content

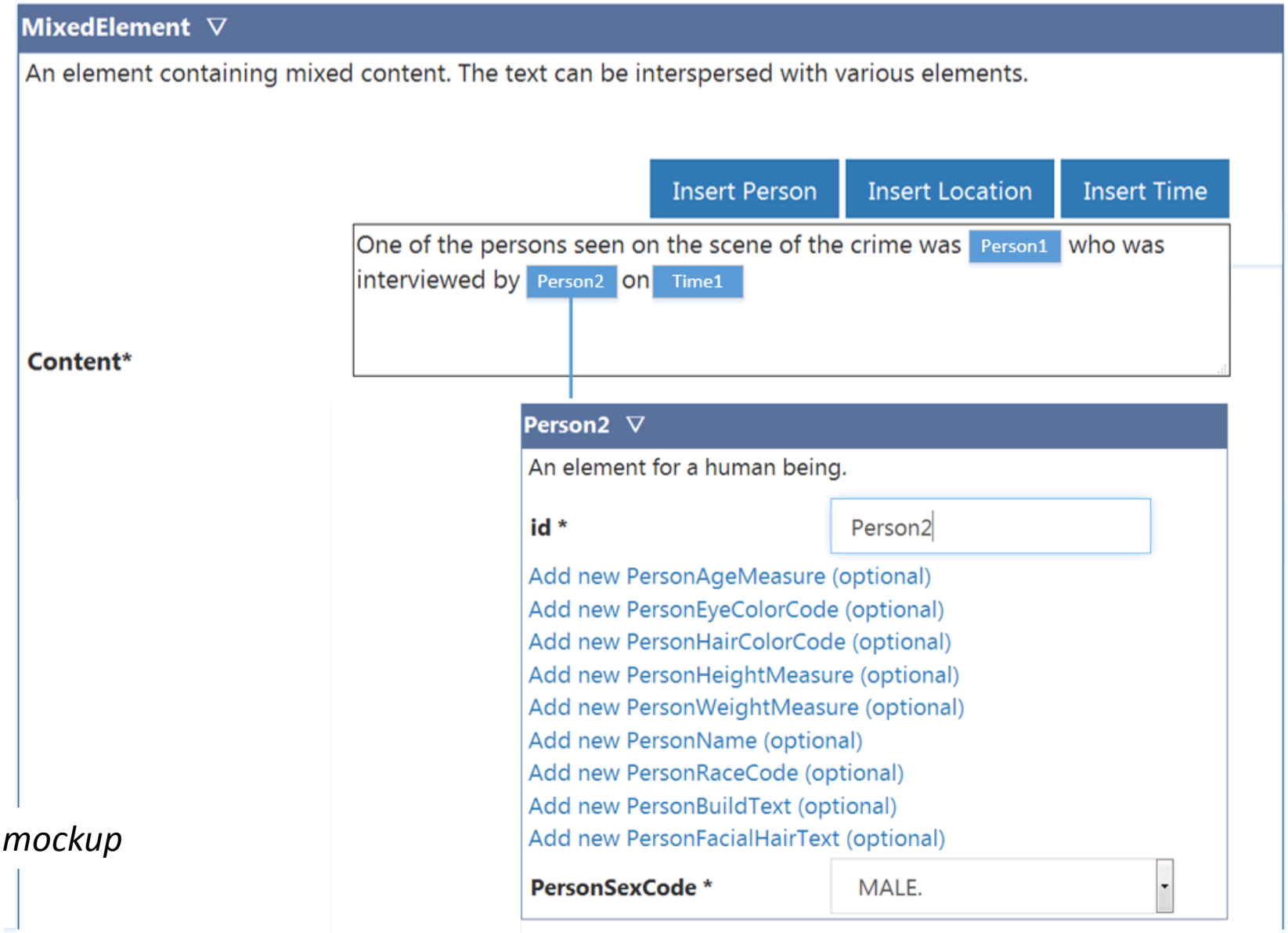
- Difficulty: conceptual & from a UI perspective

UI challenge: the `mixed` element

- Conceptually simple:
  A `mixed` element can contain text interspersed with nested elements

```
<MixedElement>
   One of the persons on the scene of the crime was
   <Person>…</Person> who was interviewed by
   <Person>…</Person> on <Time>…</Time>
</MixedElement>
```

UI challenge: the `mixed` element



UI mockup

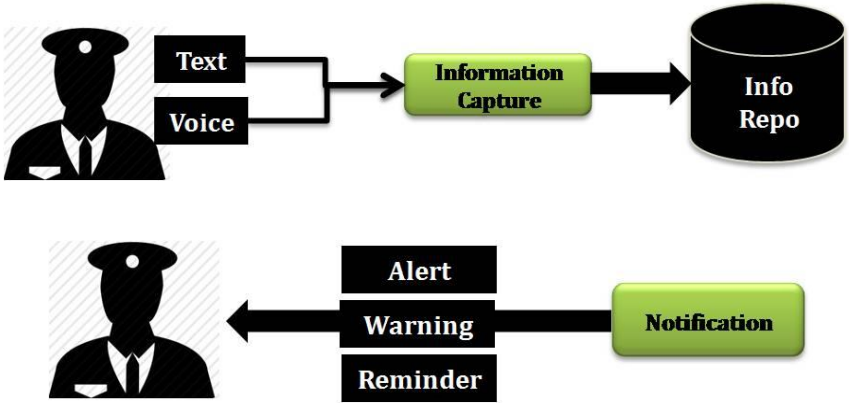
Additional features

- Necessary improvements
  - SAR loading
  - More precise validation feedback

Conclusions

- Our goal was to facilitate the creation of NIEM-SARs
- The prototype XSDGuide
  - Shows promise
  - Has clear limitations (easy & hard to fix)

- A formal evaluation is necessary in a realistic setting



Conclusions

- An unintended consequence: An exploration tool for NIEM-SAR
  - Allows to see the standard “in action”
  - Some debatable choices
  - Exhaustive to the point of being very difficult to use

- A perspective: Collection of SAR data
  - Creation of new UI guides (autocomplete…)
  - XSDGuide would be a “bootstrapping” tool

Merci