Part III
PROV Implementations and Extensions

Paolo Missier, Newcastle University
Khalid Belhajjame, University of Manchester
James Cheney, University of Edinburgh
Outline

- Implementation Report
- PROV Collection and Dictionary
- PROV and Scientific Workflows
- PROV and Executable Papers
- PROV and Smart Cities
• As of February the 25th 2013, 64 PROV implementations were reported to the W3C Provenance Working Group.
• These implementations take the form of:
  – Applications: 30
  – Framework/APIs: 10
  – Services: 9
  – Vocabularies (Usage and Extension): 21
  – Constraints Validators: 3
PROV USAGE: Coverage of PROV-DM Concepts and Relations in implementations of type Application, Framework / API, or Service.
PROV USAGE: Coverage of PROV-DM Concepts and Relations in implementations of type Application, Framework / API, or Service.
PROV USAGE: Coverage of PROV-DM Concepts and Relations in implementations of type Application, Framework / API, or Service.
PROV USAGE: Coverage of PROV-DM Concepts and Relations in implementations of type Application, Framework / API, or Service.
PROV USAGE: Coverage of PROV-DM Concepts and Relations in implementations of type Application, Framework / API, or Service.
PROV USAGE: Coverage of PROV-DM Concepts and Relations in implementations of type Application, Framework / API, or Service.
PROV terms **used** by vocabularies.

PROV terms **extended** by Vocabularies
PROV Implementations Report

PROV terms **used** by vocabularies.

![Bar chart showing usage of PROV terms](chart1.png)

PROV terms **extended** by Vocabularies

![Bar chart showing extended usage of PROV terms](chart2.png)
PROV Implementations and Extensions

Outline

- Implementation Report
- PROV Collections and Dictionaries
- PROV and Scientific Workflows
- PROV and Executable Papers
- PROV and Smart Cities
• PROV defines a collection as an entity that has some members, which themselves are entities,

• PROV dictionary is an extension that specializes PROV collections and provides the relationships necessary for expressing the lineage of dictionaries.
• PROV defines a collection as an entity that has some members, which themselves are entities,

• PROV dictionary is an extension that specializes PROV collections and provides the relationships necessary for expressing the lineage of dictionaries.
• Derivation by insertion

\[
derivationByInsertionFromExpression ::= \text{derivedByInsertionFrom} \left( \text{optionalIdentifier } ";" \right) ? \text{dIdentifier} , \text{dIdentifier} "\,\{" \text{keyValuePairs} "}\,"," \text{optionalAttributeValuePairs } "\)\]

Example

\[
\text{derivedByInsertionFrom(id; d1, d, \{"k1", e1\}, \{"k2", e2\}, [])}
\]

• Derivation by removal

\[
derivationByInsertionFromExpression ::= \text{derivedByInsertionFrom} \left( \text{optionalIdentifier } ";" \right) ? \text{dIdentifier} \,\"\,\text{dIdentifier} \,\"\{\text{keySet}\}\,\"\,\text{optionalAttributeValuePairs } \"\)
\]

Example

\[
\text{derivedByRemovalFrom(id; d3, d, \{"k1", "k2"\}, [])}
\]
Outline

• Implementation Report
• PROV Collections and Dictionaries
• **PROV and Scientific Workflows**
• PROV and Executable Papers
• PROV and Smart Cities
Scientific workflows are for specifying and automating the execution of scientific experiments.

Using a workflow, an experiment is modeled as a network of analysis operations, connected together by data links describing the flow of data between operation outputs and inputs.

Workflow systems can be instrumented in a straightforward manner to capture provenance information about the experiment execution.
- E.g., Taverna, Kepler, SciCumulus, Vistrails and Wings

Workflow provenance traces have numerous applications.
- E.g., Workflow debugging, data lineage, inspecting result reproducibility.
Currently, each workflow system uses its own proprietary model for encoding provenance traces.

With PROV, there is an opportunity to achieve a certain degree of interoperability.

Wf4Ever (EU FP7 Project) [1] and D-PROV (DataONE) [2] provide PROV extensions to capture workflow specifications and their provenance.

In what follows, we will present Wf4Ever PROV extension.

[1] www.wf4ever-project.org
Wf4Ever PROV Extension

Legend

- PROV
- Workflow
Wf4Ever PROV Extension

Legend
- PROV
- Workflow
- Provenance

Diagram:
- Workflow
- DataLink
- Input
- Output
- Parameter
- Artifact
- ProcessRun
- Process
- Plan
- hasSource
- hasSink
- hasInput
- hasOutput
- hasDataLink
- hasSubProcess
- used
- usedInput
- describedByParameter
- describedByProcess
- wasGeneratedBy
- wasOutputFrom
- wasEnactedBy
- wasOutputFrom
- WorkflowEngine
- SoftwareAgent
PROV Implementations and Extensions

Outline

• Implementation Report
• PROV Collections and Dictionaries
• PROV and Scientific Workflows
• PROV and Executable Papers
• PROV and Smart Cities
• Scholarly articles have played (and will continue to play) a primordial role in dissemination and exchange of research investigations.
• However, they present limitations.

• The executable paper initiative aims to address this issue.
• Existing executable papers proposals include:
  – Deep (e-Stat project) [3]
  – Research Objects (Wf4Ever EU Project) [4]
  – Sole (Argonne National Laboratory) [5]

DEEP: A Provenance Aware Executable Document

- Aims at enhancing the dissemination of scientific results
- Generates scientific results dynamically and interactively, and records the provenance of these results.
DEEP: A Provenance Aware Executable Document

- Aims at enhancing the dissemination of scientific results.
- Generates scientific results dynamically and interactively, and records the provenance of these results.

Diagram:

- Resource Binding
- Resource
- Activity
- Usage
- Input
- Output

Estat reader
Upload
Save
Expert
Debug
Resources

Modelling binary responses

Chapter 1: 1-level model

Input, Model and Equations

Graph beta 1
Graph beta 2

Chapter 2: Multi-level models

Introduction, input and model

Results summary

Results: Summary

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ESS</th>
<th>sd</th>
<th>mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>beta0</td>
<td>2470</td>
<td>0.053841</td>
<td>-0.413393</td>
</tr>
<tr>
<td>beta1</td>
<td>2470</td>
<td>0.094173</td>
<td>0.012129</td>
</tr>
<tr>
<td>deviance</td>
<td>2048</td>
<td>1.972096</td>
<td>5.446457</td>
</tr>
</tbody>
</table>
PROV Implementations and Extensions

Outline

• Implementation Report
• PROV Collections and Dictionaries
• PROV and Scientific Workflows
• PROV and Executable Papers
• PROV and Smart Cities
• Smart cities rely on citizens participation and contribution to address a variety of problems, e.g., mobility, governance, economy, environment.

PROV and Smart Cities

- Smart cities rely on citizens participation and contribution to address a variety of problems, e.g., mobility, governance, economy, environment.
- **UrbanMatch** [6] is a location-based Game with a Purpose in the form of a mobile application.

• Smart cities rely on citizens participation and contribution to address a variety of problems, e.g., mobility, governance, economy, environment.

• **UrbanMatch** [6] is a location-based Game with a Purpose in the form of a mobile application.

Outline

- Implementation Report
- PROV Collections and Dictionaries
- PROV and Scientific Workflows
- PROV and Executable Papers
- PROV and Smart Cities
PROV Extensions

- PROV Dictionary
  - Core PROV
  - Collection
- Scientific Workflows
  - Core PROV
  - Plan
  - Bundle
- Executable Paper
  - Core PROV
  - Bundle
- Smart Cities
  - Core PROV
Conclusions

• **Previous Tutorials:**
  – IPAW 2012 - 4th International Provenance and Annotation Workshop (Santa Barbara)
  – ISWC 2012 – 11th International Semantic Web Conference (Boston)

• **Upcoming Tutorial**
  – ESWC 2012 - 9th Extended Semantic Web Conference (Montpellier)

• **Prov Status**
  – As of March the 8th 2013, **PROV-DM, PROV-N, PROV-O** and **Prov-Constraints** are **Proposed Recommendations**

• So join the club in promoting and adopting PROV! After all, this is really a community effort…
  – Call for implementation can be found here: [http://www.w3.org/2011/prov/wiki/Call_For_Implementations](http://www.w3.org/2011/prov/wiki/Call_For_Implementations)
Thank you for your attention!