

Generation and Validation of Event Processing Rules in EnStream

Alexandra Moraru, Klemen Kenda, Blaž Fortuna, Luka Bradeško,
Maja Škrjanc, Dunja Mladenić, Carolina Fortuna

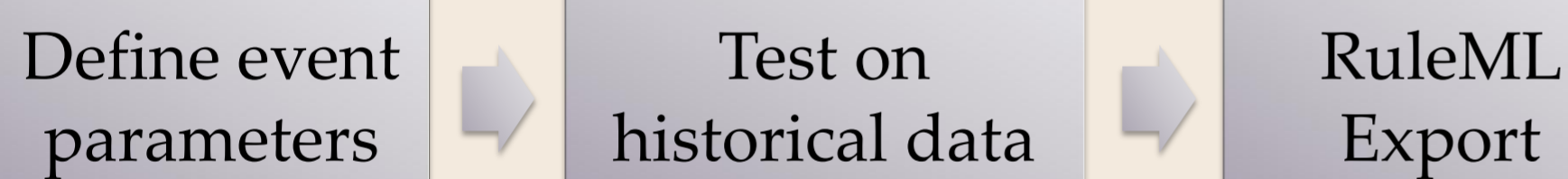
What are we demonstrating?

- A system which support domain experts in exploring large datasets
- **Our goal:** provide to the experts an environment where they can create and validate the rules

EnStream Infrastructure

- integrated and scalable software modules
- allows development of various applications
- used in the ENVISION project for stream data mining
 - in environmental scenarios: **landslides, oil spills and river floods**

Rule Creation Process



➤ Illustrative example

- *if the amount of rainfall exceeds 250 mm per day in 3 consecutive days then a landslide can occur*

```

<And> <Atom> <op> <Rel iri="openCyc:sensorObservation"/> </op>
<Var> sensor </Var>
<Ind iri="openCyc:Raindrop"/> </Atom>
<Atom> n <op> <Rel iri="openCyc:doneBy"/> </op>
<Var> sensor </Var>
<Var> measurement </Var> </Atom>
<Atom> <op> <Rel iri="openCyc:measurementResult"/> </op>
<Var> measurement </Var>
<Var> val1 </Var> </Atom>
<Atom> <op> <Rel iri="openCyc:duration"/> </op>
<Var> measurement </Var>
<Ind type="xs:time">24:00:00</Ind> </Atom>
<Atom> <op> <Rel iri="openCyc:greaterThanOrEqualTo"/> </op>
<Var> val1 </Var>
<Ind type="xs:float">250</Ind> </Atom> </And>
    
```

Event definition

precipitation [24h] >= 250

current day

OR - AND

precipitation [24h] (2 days ago) >= 250 AND precipitation [24h] (1 day ago) >= 250

Validate

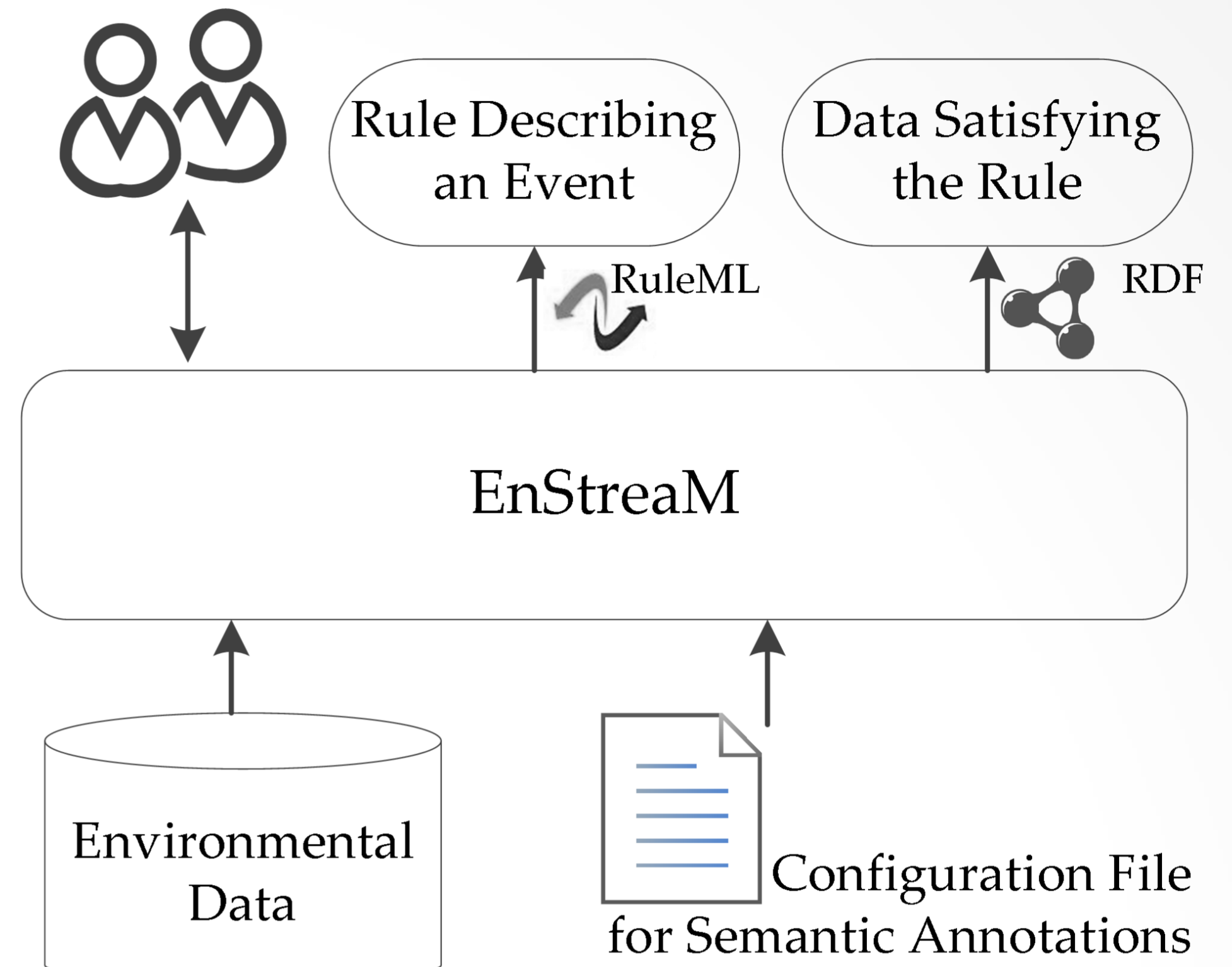
Event name: _____

Export event rule | Export event data

Events

13. 05. 2010: Landslide [select]

28. 10. 2010: Landslide [select]



➤ EnStream input

- environmental datasets: sensor measurements and past events
 - e.g.: volume of rainfall for a given geographical location and landslides
- configuration files for semantic annotations

➤ EnStream output

- **rules** created and validated on real datasets through the GUI
 - exported in RuleML format
- **datasets** to which the rules apply
 - exported in RDF format

➤ Semantic Annotations

- provide a unified view over different sources of data
- **OpenCyc ontology** – very large and contains concepts for many specific domains
 - any other ontology could be used
 - **alternative:** Semantic Sensor Network ontology
- **RDF export** of dataset corresponding to a rule
- **RuleML export** is depended on the vocabulary used for the relation names
 - Specialized domain ontologies can simplify the RuleML representation

➤ Future work

- Extend EnStream for **real-time monitoring** of streaming data in order to **detect** the **events** described in the rules generated
- **Integrate the rules** discovered into **knowledge bases** used by specific reasoning engines.
 - semi-automatic extension of knowledge bases
 - support for advanced reasoning for problems such as complex events processing, anomaly detection or automatic monitoring

References

<http://www.opencyc.org>

<http://www.envision-project.eu/>

Kenda, K., Fortuna, C., Fortuna, B., Grobelnik, M. Videk: A Mash-up for Environmental Intelligence. AI Mashup Challenge, ESWC (2011)

Škrjanc, M., Mladenić, D. Stream mining on environmental data. In Proceedings of Information Society conference IS-2010, volume A, pp. 184-187, Ljubljana, Slovenia, (2010)