

## POWDER: THE W3C RECOMMENDATION, USE CASES, FUTURE DIRECTIONS

Stasinos Konstantopoulos

National Centre for Scientific Research “Demokritos”

Athens, Greece

# Introduction

2/9

- Protocol for Web Description Resources (POWDER)
  - Serving metadata about Web resources
  - Without retrieving the resources
- POWDER is akin to:
  - PICS, the precursor to RDF
  - `rdf:aboutEachPrefix`
  - `robots.txt`
  - URI template

# Description Resource

3/9

- A Description Resource
  - Is the basic information unit of POWDER documents
  - Assigns properties to a set of resources
- Scope:
  - A specification of a set of resources
- Description:
  - Properties and property values

```
<dr>
  <iriset>
    Scope
  </iriset>

  <descriptorset>
    Description
  </descriptorset>
</dr>
```

# Description

4/9

- A Descriptor Set
  - Describes resources
  - Both object and datatype properties
- typeof
  - assigns the `rdf:type` property
- displaytext and icon
  - Human-understandable description and depiction

```
<descriptorset>

  <ex:colour>red</ex:colour>
  <ex:shape rdf:resource=
    "http://ex.com/voc#square"/>

  <typeof src=
    "http://ex.com/voc#object"/>
  <displaytext>
    Red and square
  </displaytext>
  <displayicon src=
    "http://ex.com/redsqa.png" />

</descriptorset>
```

# IRI Set

5/9

- An IRI Set
  - A specification of a set of resources
- Access to all IRI parts:
  - includeschemes
  - includehosts
  - includeiripattern
  - includeports
  - includeexactpaths
  - includepathcontains
  - includepathstartswith
  - includepathendswith
  - includequerycontains
- And their complements:
  - excludeschemes ...

```
<iriset>
  <includehosts>
    example.com
  </includehosts>
  <includepathstartswith>
    /red/square/
  </includepathstartswith>
</iriset>
<iriset>
  <includeiripatterns>
    *.example.com
  </includeiripatterns>
  <includepathstartswith>
    /red/square/
  </includepathstartswith>
</iriset>
```

# More complex IRI sets

6/9

- **Multiple values:**
  - It is sufficient that any one is satisfied
- **Sub-elements of iriset:**
  - All must be satisfied
- **DR with multiple IRI sets:**
  - It is sufficient that any one is satisfied

```
<iriset>
  <includehosts>
    example.com example.net
  </includehosts>
  <includepathstartswith>
    /red/square/
  </includepathstartswith>
</iriset>
<iriset>
  <includehosts>
    example.org
  </includehosts>
  <includepathstartswith>
    /square/red/
  </includepathstartswith>
</iriset>
```

# Grouping by regular expression

7/9

- in/excluderegex
  - Element content is a single regular expression
  - Following regular expression syntax of XQuery 1.0 and XPath 2.0 “*Functions and Operators*”
- All IRI groups can be expressed as in/excluderegex
  - This point will be revisited

```
<iriset>
  <includeregex>
    \.example\.org/red/square/
  </includeregex>
</iriset>

<iriset>
  <includeiripatterns>
    *.example.com
  </includeiripatterns>
  <includepathstartswith>
    /red/square/
  </includepathstartswith>
</iriset>
```

## References

P. Archer, A. Perego, and K. Smith, (eds), *POWDER: Grouping of Resources*, W3C Recommendation, 1 Sep 2009. URL <http://www.w3.org/TR/powder-grouping>

# POWDER document structure

8/9

- We have seen what a single DR is
- A POWDER document:
  - Has an unordered set of DRs
    - All must be applied
  - Has an ordered list of DRs
    - First match wins
  - Has attribution metadata covering all the DRs in the document

```
<powder>
  <attribution>
    Provenance
    Validity period
  </attribution>
  <dr> description resource </dr>
  ...
  <dr> description resource </dr>
</ol>
  <dr> description resource </dr>
  ...
  <dr> description resource </dr>
</ol>
</powder>
```

## References

P. Archer, K. Smith, and A. Peregó (eds), *POWDER: Description Resources*, W3C Recommendation, 1 Sep 2009. URL <http://www.w3.org/TR/powder-dr>



# POWDER-BASE

9/15

- IRI grouping expressions:
  - All can be reduced to regular expression matching
- POWDER-BASE:
  - a POWDER-BASE document is a POWDER document that only uses `<includeregexp/>` and `<excluderegexp/>`
- The XSLT transformation from POWDER to POWDER-BASE formally defines POWDER
  - <http://www.w3.org/2007/powder/powder2powderBase.xsl>

# Formal Semantics: description

10/15

- DRs:
  - Can be reduced to the subsumption of the IRI set under the descriptor set
- Descriptor sets:
  - Can be reduced to the set of resources that have the properties in the <descriptorset/>
- IRI sets:
  - Cannot be expressed using standard Semantic Web technologies

```
iriset_1 rdfs:subClassOf descriptorset_1 .  
  
descriptorset_1 rdfs:subClassOf [  
  owl:intersectionOf ( [  
    [ rdf:type owl:Restriction ;  
      owl:onProperty ex:colour ;  
      owl:hasValue "red" ]  
    [ rdf:type owl:Restriction ;  
      owl:onProperty ex:shape ;  
      owl:hasValue ex:square ]  
  )  
].
```

# Formal Semantics: grouping by IRI

11/15

- **Grouping elements:**
  - Reduced to classes of resources that have an IRI that matches a reg. exp.
  - The `wdrs:matchesregex` property exists in the POWDER-S extension of RDF semantics

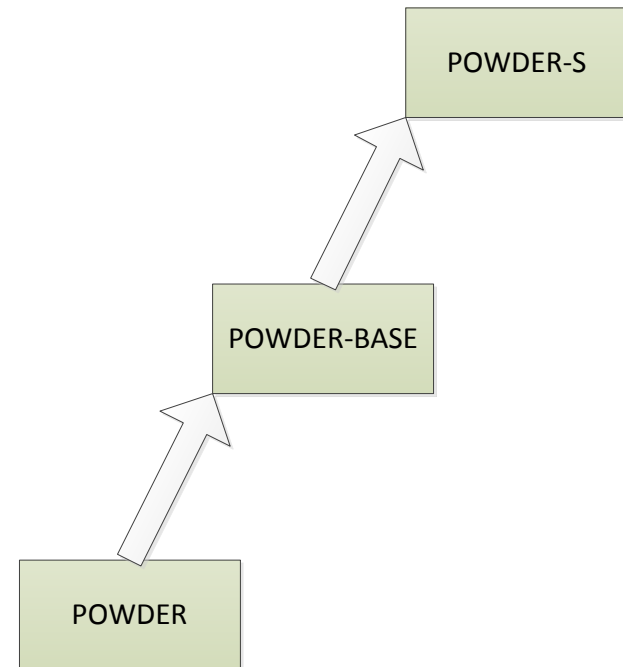
```
iriset_1 owl:equivalentClass [  
  rdf:type owl:Class ;  
  owl:intersectionOf ([  
    rdf:type owl:Restriction ;  
    owl:hasValue "example\\.com"^^xsd:string ;  
    owl:onProperty wdrs:matchesregex ])  
].
```

- **Both description and grouping elements:**
- The XSLT transformation from POWDER-BASE to POWDER-S formally defines POWDER-BASE
  - <http://www.w3.org/2007/powder/powderBase2powderS.xsl>

# From POWDER to POWDER-S

12/15

- A hierarchy of increasingly specific specifications:
  - POWDER extends POWDER-BASE with grouping elements specific to the resolvable Web
  - POWDER-BASE is defined using the POWDER-S extension of RDF



## References

S. Konstantopoulos and P. Archer (eds), *POWDER: Formal Semantics*, W3C Recommendation, 1 Sep 2009. URL <http://www.w3.org/TR/powder-formal>

# Processing POWDER documents

13/9

- A POWDER processor:
  - Receives as input an IRI
  - Accesses one or more POWDER documents
    - No need to transform, may directly interpret any flavour it knows about
  - Provides as output property names and values
    - RDF triples, assuming the input IRI as subject
- Implementations:
  - May directly process POWDER in XML
  - May implement the POWDER-S extension

# Publishing POWDER documents

14/9

- Anyone can publish a POWDER doc describing anything
- Web resources claim being described by POWDER doc(s):
  - RDF(a) wdrs:describedby
  - @rel describedby

<http://www.w3.org/2007/powder/powder-errata#describedby>

## References

POWDER: *Description Resources*, Section 4.1. See also:

<http://www.w3.org/2007/powder/powder-errata#describedby>

# Extendibility

15/15

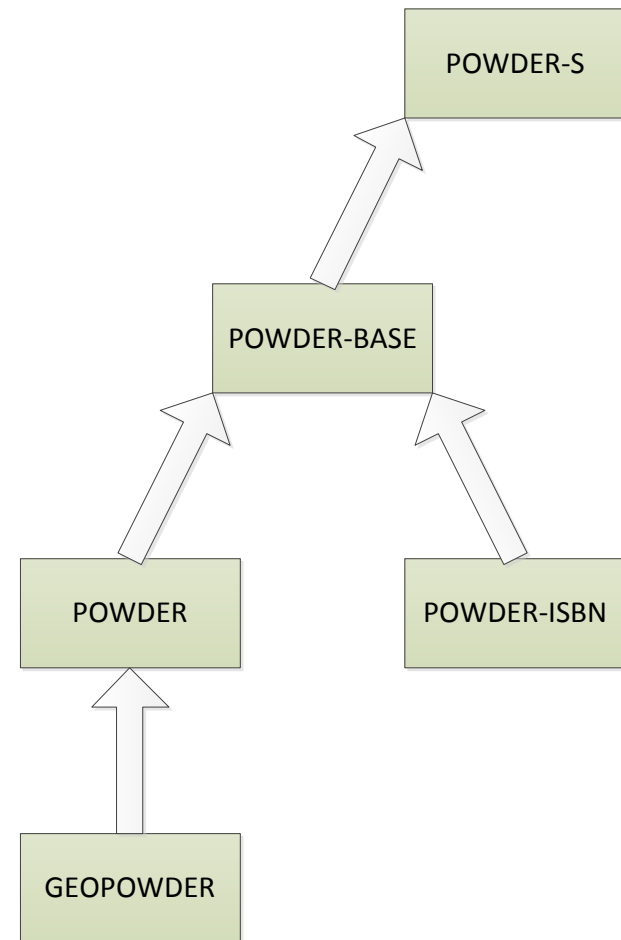
- Site-specific extensions
  - ▣ Avoid clutter in POWDER documents
- E.g., UK Ordnance Survey:
  - ▣ <http://data.ordnancesurvey.co.uk/id/postcodeunit/AL11AE>
- Define `<postcodestartswith/>`
  - ▣ Associate `<geopowder/>` to XSLT that transforms into `<powder>`
- Easier to express mapping to, e.g., geonames

```
<geopowder>
  <dr>
    <iriset>
      <includepostcodestartswith>
        AL
      </includepostcodestartswith>
    </iriset>
    <descriptorset>
      <gn:locatedIn rdf:resource="http://
sws.geonames.org/2647043/" />
      <displaytext>
        Hertfordshire
      <displaytext>
    </descriptorset>
  </dr>
</geopowder>
```

# Extendibility

16/15

- POWDER extensions
  - Not restricted to the resolvable Web
  - Other structured identifiers can also be expressed as URIs
    - Typically in the urn: scheme
- E.g., urn:isbn:3-642-12841-7
  - Group: 3, *German group*
  - Publisher: 642, *Springer Verlag*
  - Edition: 12841, *Proceedings of SETN 2010, print edition*
  - Control digit: 7





# Implementations and deployments

17/15

- Reference implementations
  - Both POWDER XML and POWDER-S processors
  - Available at <http://www.w3.org/2007/powder/#tools>
- QUATRO Plus project (<http://www.quatro-project.org>)
  - Content labelling use case
  - POWDER document server
  - Browser plug-ins
    - Browser retrieves and visually depicts annotations

## References

P. Archer, E. Ferrari, V. Karkaletsis, *et al.* "QUATRO Plus: Quality you can trust?", *Workshop on Trust and Privacy on the Social and Semantic Web (SPOT 2009)*.

# POWDER as inference

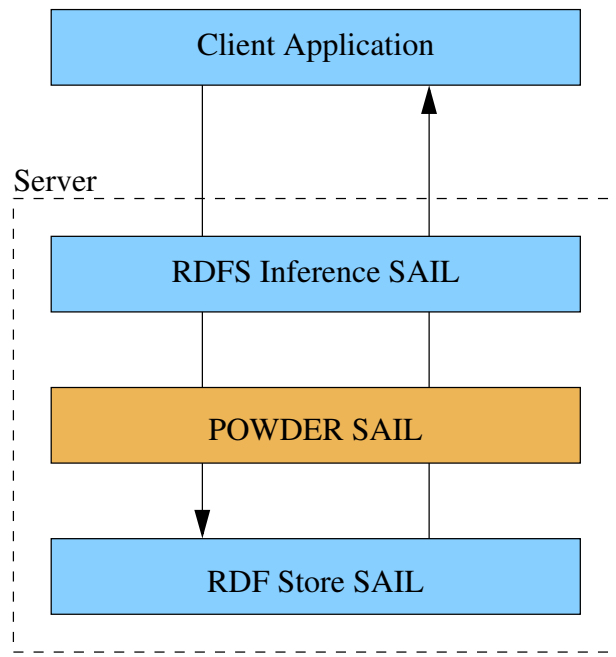
18/15

- SYNC3, <http://www.sync3.eu>
  - Large-scale on-line news items indexing
  - Extensive gazeteer, served as REST
    - <http://www.sync3.eu/.../person/73>
    - <http://www.sync3.eu/.../organization/42>
    - ...
  - POWDER processor used to infer types
    - Internal compression mechanism, rather than metadata publishing
    - Effectively, an inference engine
    - POWDER results input for RDFS inference
      - No cyclicity: POWDER **only** takes into account URI patterns, not properties

## References

- S. Konstantopoulos, P. Archer, *et al.*, The POWDER protocol as infrastructure to serving and compressing semantic data. *Intl Journal of Metadata, Semantics and Ontologies* **7**(1):1-15. 2012.
- S. Konstantopoulos, Three ways to sprinkle POWDER. *ISWC 2011, Bonn*, 2011.

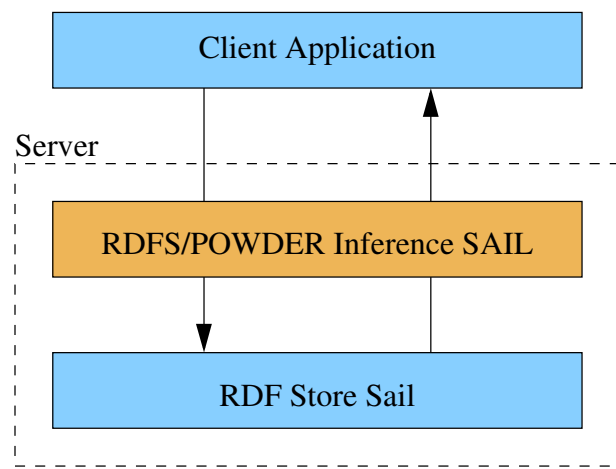
# The POWDER Sesame store



P-store architecture:  
FILTER clauses

Tuples retr.	Time (in seconds)		
	RDFS only	P-store	RDFS/P
706	22	19	14
2233	109	86	106
3418	117	96	114

13% compression  
18% speedup



RDFS/P architecture:  
Forward chaining inference

# Implementing POWDER-S

20/15

- Semagrow, <http://www.semagrow.eu>
  - Direct implementation of the P-store
    - Regular-expression index, not post-filtering as in SYNC3
    - Retrieves `wdrs:matchesregex` instances without checking all URLs
  - Resource discovery
    - Exploits regularities in resource naming conventions
    - Responds to URLs with potential data sources
  - SPARQL endpoint federation
    - Efficiently distribute SPARQL queries
    - Does not require control over the federated end points

## References

S. Konstantopoulos, A. Koukourikos, P. Karampiperis, 'Improving the Real-time Performance of Heterogeneous Extremely Large Datasets'. Accepted to *Special Session on The Web of Data, Panhellenic Conference on Informatics (PCI 2013)*, 19-21 Sep 2013, Thessaloniki, Greece.

# Bottom line

21/15

- URIs are, in principle, meaningless labels, but:
  - Humans tend to organise data in tree structures
    - Sub-directories are often also subsumed in a semantic hierarchy
    - URI design guidelines tend to support POWDER's assumptions  
<http://philarcher.org/diary/2013/uripersistence>
  - Large-scale repositories employ systematic naming conventions
- This can be exploited
  - To mass-annotate Web resources
    - Access metadata without having to access the resources
    - Including accessing metadata provided by third parties
  - To capture the intension behind an information structure
    - Resource discovery

# Thank you for your attention!

22/15

Stasinou Konstantopoulos

([konstant@iit.demokritos.gr](mailto:konstant@iit.demokritos.gr))

Institute of Informatics and  
Telecommunications

National Centre for Scientific  
Research “Demokritos”

Phil Archer

([phila@w3.org](mailto:phila@w3.org))

W3C Team Member

## Recommendations:

*POWDER Description Resources*

<http://www.w3.org/TR/powder-dr>

*POWDER Grouping Resources*

<http://www.w3.org/TR/powder-grouping>

*POWDER Formal Semantics*

<http://www.w3.org/TR/powder-formal>

## Notes:

*POWDER Primer*

<http://www.w3.org/TR/powder-primer>

*POWDER Test Suite*

<http://www.w3.org/TR/powder-test>

## Relevant projects:

<http://www.quatro-project.org>

<http://www.sync3.eu>

<http://www.semagrow.eu>