

# Scout

# Design and architecture of a novel preservation watch system

Luís Faria Ifaria@keep.pt



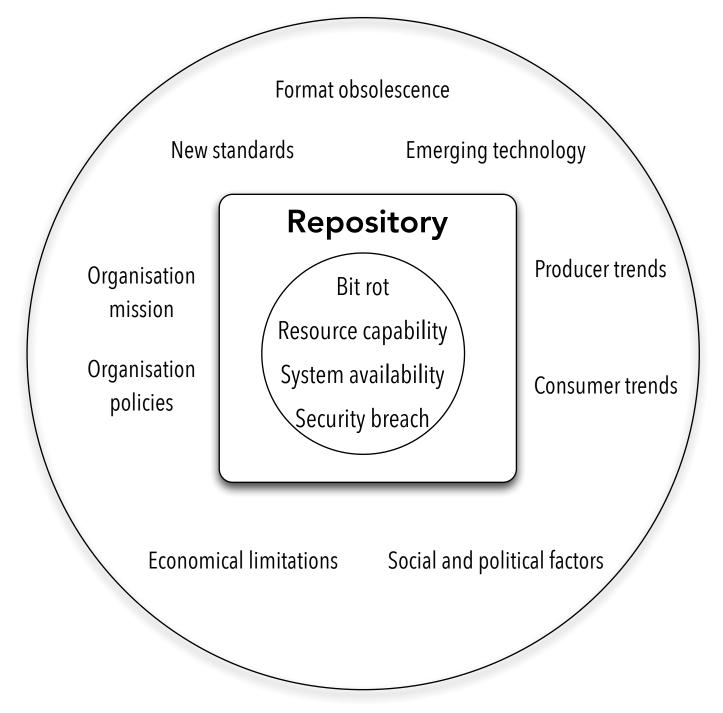
#### Outline



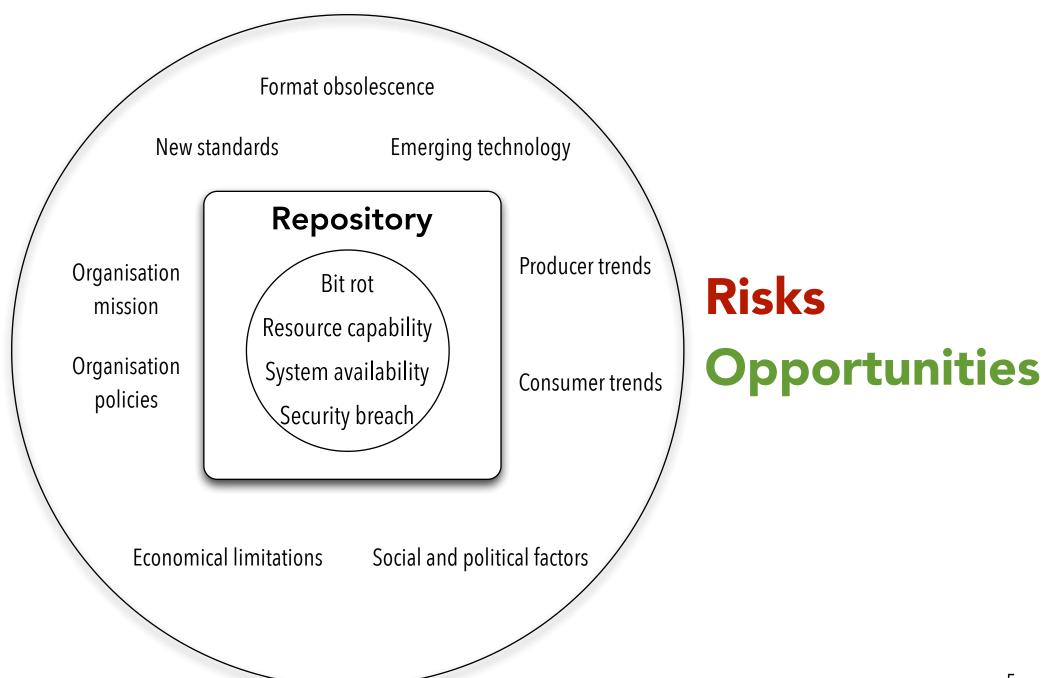
- Preservation monitoring
  - Why and what is needed
  - State of the art
- A novel approach
  - Methodology
  - Architecture
  - How can you participate?

# Preservation monitoring

## Why do we need monitoring?



## Why do we need monitoring?



#### State of the Art



- Digital Format Registries
- AONS
- Technology watch reports

#### State of the Art



- Digital Format Registries
  - Lack of coverage
  - Statically-defined generic risks
  - Lack of structure in risks
  - Focus on format obsolescence
- AONS
  - Total dependency on format registries
- Technology watch reports
  - Machine unreadable

#### **Risk Assessment**

Yes but manual and adhoc

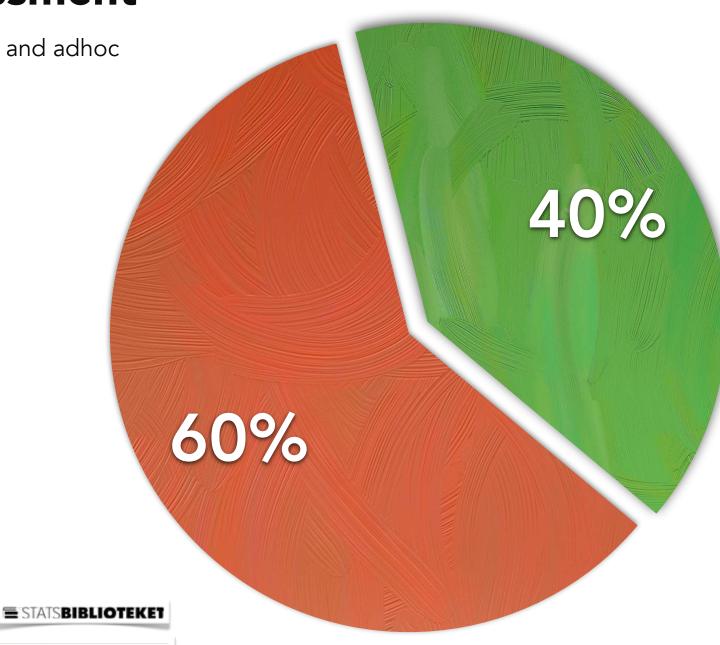
None

Survey on:





Koninklijke Bibliotheek National Library of the Netherlands

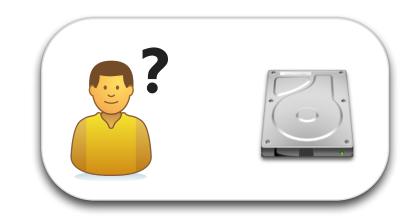


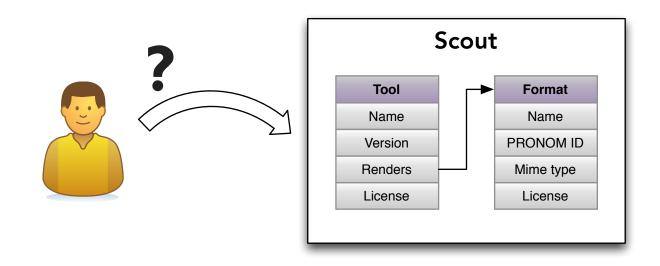
#### What is needed?

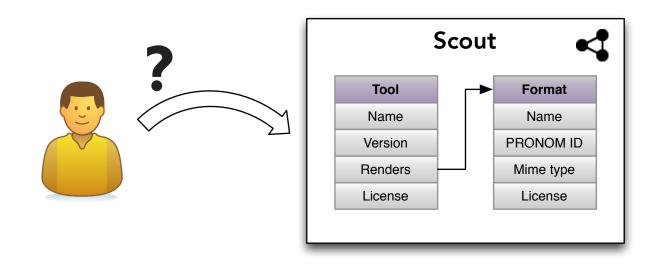


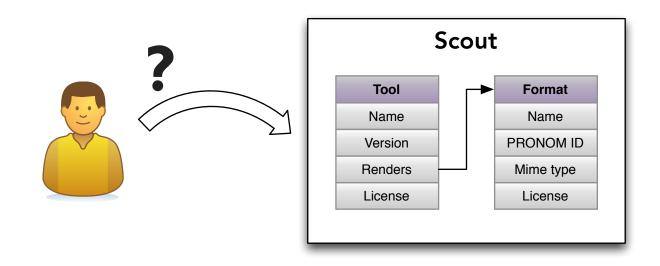
- We need data!
  - From anywhere and everywhere
  - Sharing
- Usability & Scalability
  - Structured data
  - Controlled vocabulary

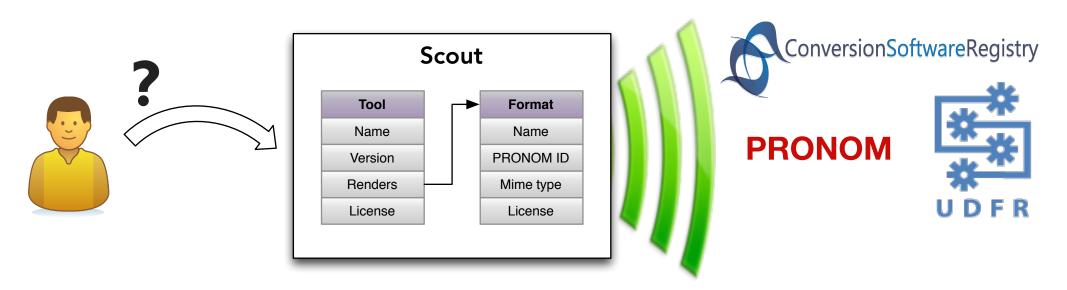
# **Scout** A novel approach

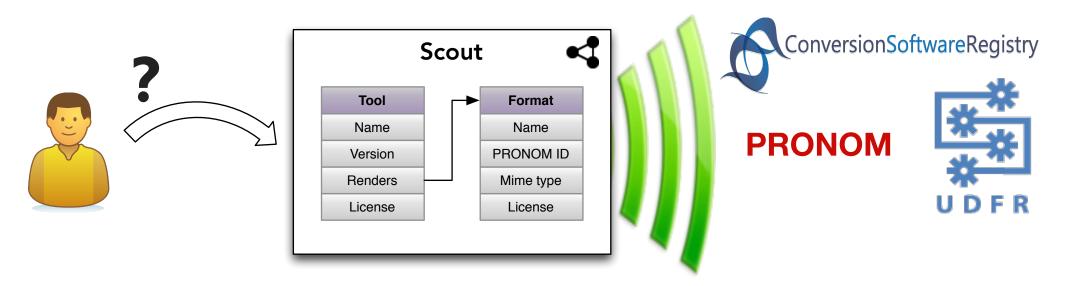


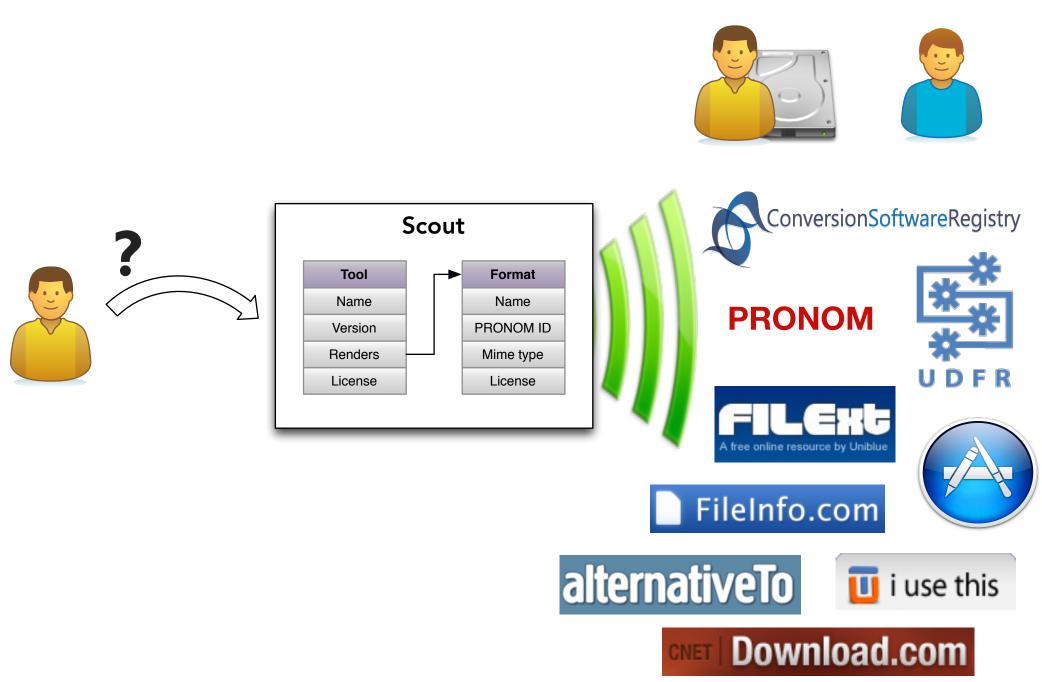


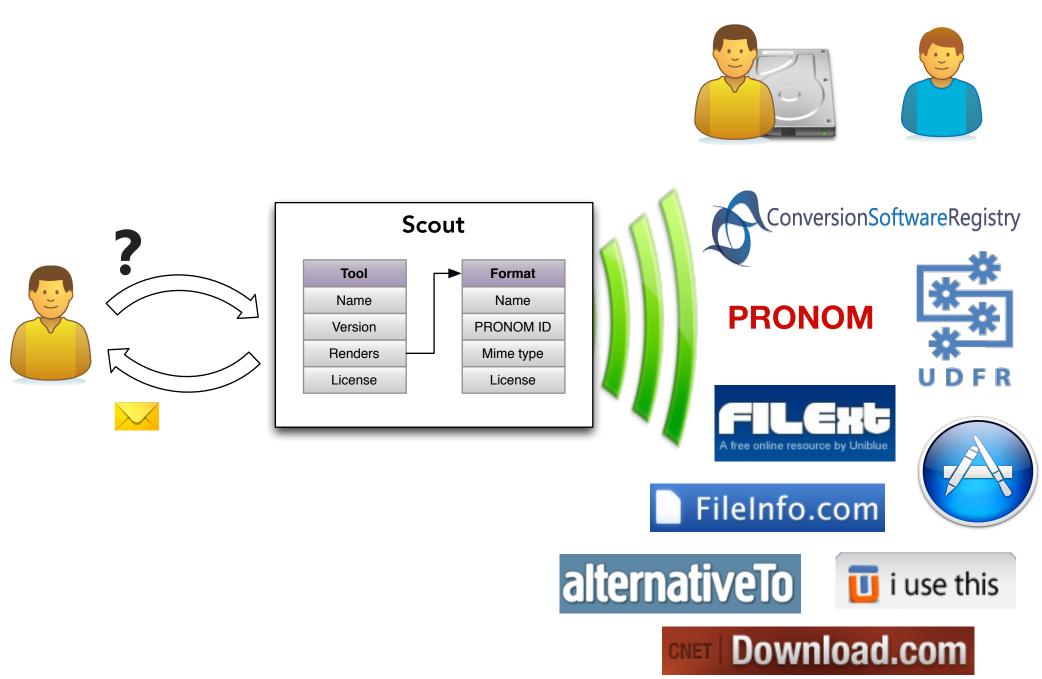












#### Methodology



- Survey practitioners on information to be monitored
- Create structured data model
- Find sources of information (try to automize)
- Define notification triggers
- Frequent monitoring of sources of information
- Frequent assessment of triggers and notification

#### **Example questions**



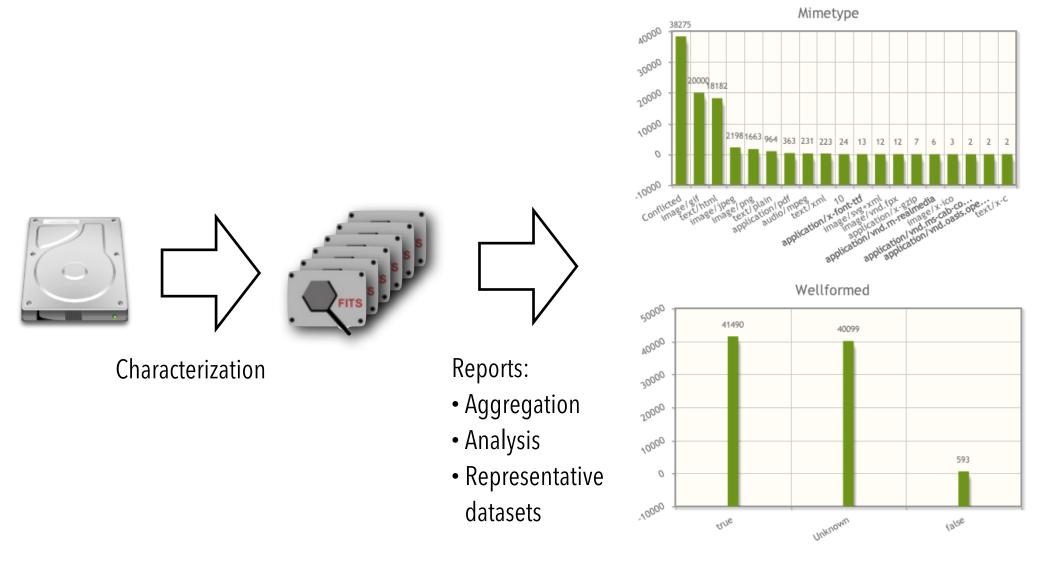
- Are there any tools that can render the format X?
- Is my repository the only one that has format Y?
- Are my preservation plans still valid?
- Are my repository policies being enforced?

#### **Information Sources**



- Format registries & software catalogues
- Digital repositories & web archives
- Organizational objectives
- Experiments
- Simulation
- Human knowledge

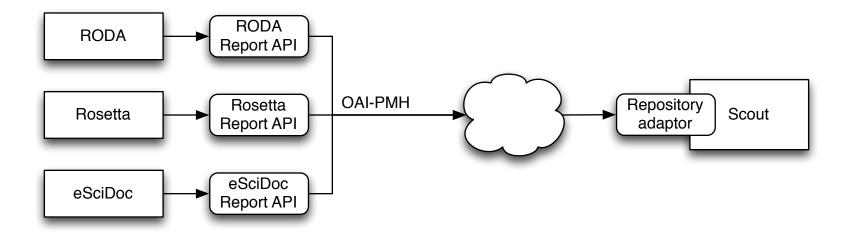
#### C3PO Content profile tool



https://github.com/peshkira/c3po

Petar Petrov petrov@ifs.tuwien.ac.at

#### **Repository events adaptor**



- OAI-PMH with PREMIS
- Normalize events
- Fine-grain events
- History

- Events example
  - Ingest started/ended
  - Representation downloaded
  - Plan executed

#### **Define triggers**

• Notify me when there are tools that can render the format X.

ELECT ?s WHERE { ?s rdf:type watch:Entity .	
?s watch:type ?type .	
?type watch:name "tools"^^xsd:String .	
?value watch:entity ?s.	
?value watch:property ?property	
<pre>?property watch:name "renders"^^xsd:String .</pre>	
?value watch:value "format X"^^xsd:String	

### Receive notifications

#### Email

00

To: Luis Faria Scout notification

Scout <scout@scape-project.eu>

One of your requests has been triggered. SCOUT Preservation Watch System There is a notification for you! What is SCOUT? SCOUT is a preservation watch There are tools that can render format X. system that monitors the whole world to warn you of preservation risks or Question opportunities. The assessment of a question has triggered this Why did I received this email? notification: A trigger was created to warn SPARQL: ?s watch:entity watch-Entity:entity1. ?s you when something important watch:property watch-Property:tests-property1. FILTER(? happened. Your email was defined as a notification that s < 200) this important event has Target: PROPERTY VALUE occured. Type: Test entities

Scout notification - Inbox

Plan The assessment of a plan has triggered this notification: Plan ID: plan123

This software is copyrighted by the SCAPE Project Consortium. The SCAPE project is co-funded by the European Union under FP7 ICT-2009.4.1 (Grant Agreement number 270137). This is an open-source project available at <u>GitHub</u>

- 4 (=

12

July 19, 2012 5:59 PM



#### HTTP Push API

Interfaces			⊕						
Web page			Scout A preservation watch system						
		1	Scout is a you to be	a digital preservation watch system that allows notified whenever some significant event in a can bring you harm or profit.	How does it work? It gathers information from all around the world and curates it using link data technology. Then you can define questions that cross-link information. The system frequently monitors the your questions and notifies you when something important happens.				
O O Swagger API Explorer ×			R <sub>M</sub>						
← → C [] scout.keep.pt/apidoc/#!/sourceadaptor/listSourceAdapto	or_GET		☆ =						
http://scout.keep.pt/api	api_key	Explore							
source	Show/Hide List Operations	Expand Operations	Raw						
property	Show/Hide List Operations	Expand Operations	Raw						
sourceadaptorevent	Show/Hide List Operations	Expand Operations	Raw						
request	Show/Hide List Operations	Expand Operations	Raw						
sourceadaptor	Show/Hide List Operations	Expand Operations	Raw						
Put /sourceadaptor.json/update	Updat	e an existing Source Ad	aptor						
GET /sourceadaptor.json/{instance}	Get a source adaptor based on	the instance unique ide	ntifier						
GET /sourceadaptor.json/list		List all source ada	ptors	REST API					

#### Parameters

Parameter active Try it out! Description Source adaptor active state Value



#### How to be a part of Scout

- Join the surveys
  - Send your email to me <<u>lfaria@keep.pt</u>>
- Integrate your content
  - Send your content profile with C3PO
  - Send repository events with Report API (soon)
- Contribute with information (soon)
  - Use Scout form for manual input of knowledge

## Scout

0 ( popenplanet	s/scout ×					
→ C	Inc. [US] https://git	hub.com/openplanets/scou	ıt			📀 🙂 🏠
github 🔹	Search or Type a Comm	nand () S Explore	Gist Blog Help		👰 luis100	¢ × ₽
openplanets	/ scout		ູ່ໃງ Pull Request	🏵 Unwatch 🝷 ★	Unstar 8	ŷ Fork 1
Code	Network	Pull Requests 0	Issues 53	Wiki Graj	phs	Admin
SCOUT - A preservati	ion watch system — R oject.eu	lead more			_	
🔹 Clone in Mac	Ф ZIP НТТР	SSH Git Read-Only http	os://github.com/openp	lanets/scout.git	Read+Wr	ite access
🎾 branch: master	Files Com	nmits Branches (5)			Tags 2	Downloads
scout /					G	136 commits
Update README.md						
Juis100 authored a	month ago			1	🚹 latest comm	it 480ab2db8d
adaptors	7 months ago	adding minor change befo	re cleanup (peshkira)			
code-style	8 months ago	fixes 51 and getting rid of	build tools config and fix	ing potential bugs [peshk	ira]	
in core	7 months ago	Fixing one unhandled exc	eptionin CollectionProfile	mMonitor [kduretec]		
		language of an ability of a				
im model	7 months ago	improved multithreaded co	orrectness in DAO single	tons, improved code s	[luis100]	
im model	7 months ago 8 months ago	configuring site (not finish	-	Mons, improved code s	[luis100]	
-			ned) [luis100]		(luis100)	
src	8 months ago	configuring site (not finish	ned) [luis100] e to the CentralMonitor vi	ia the setter [kduretec]	[luis100]	
in src	8 months ago 7 months ago	configuring site (not finish adding NotificationService	ned) [luis100] e to the CentralMonitor vi file adaptor stub [peshkira	ia the setter [kduretec]	[luis100]	

Luís Faria <a href="mailto:style="color: blue;"><a href="mailto:style="color: blue;">lluís Faria</a> <a href="mailto:style: style: blue;"></a> <a href="mailto:style: style: style: blue;">style: style: styl

#### SCAlable Preservation Environments http://www.scape-project.eu

SCAPE

#### Watch us on GitHub!

https://github.com/openplanets/scout



#### http://www.keep.pt