Explanations on the Web: A Provenance-based Approach

Luc Moreau
King's College London

Outline

- Food, art and AI
- PROV 101
- Provenance Analytics
- Provenance based Explanations
- Conclusion

Acknowledgement

- Dong Huynh, Ayah Helal, David Kohan Marzagao, Senka Krivic, Gerard Canal, Archie Drake, Quratul-Ain Mahesar, Menisha Patel, Andrew Coles, Paul Luff, Perry Keller, Rita Borgo, Simon Parsons, (*King's College London*)
- Sophie Stalla-Bourdillon, Niko Tsakalakis (*University of Southampton*)

PLEAD Provenance-driven & Legally-grounded Explanations for Automated Decisions

EPSRC - EP/S027238/1

Provenance Analytics for Command and Control

ONR-G - N62909-18-1-2079

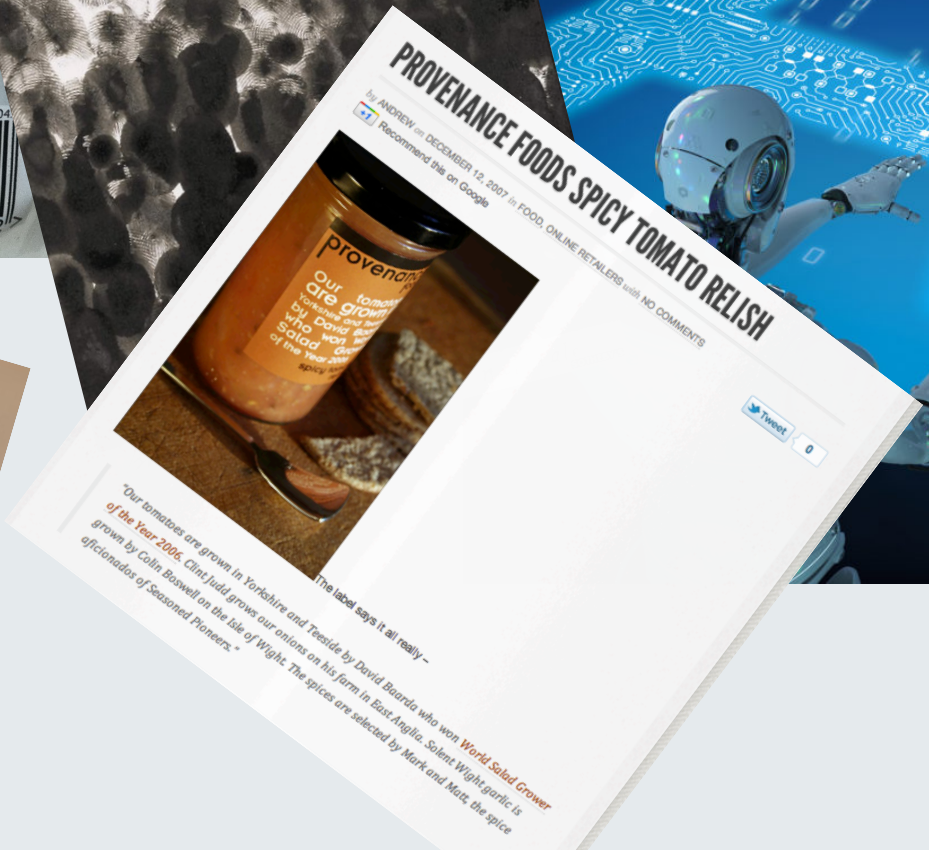
THuMP | trust in human-machine partnerships

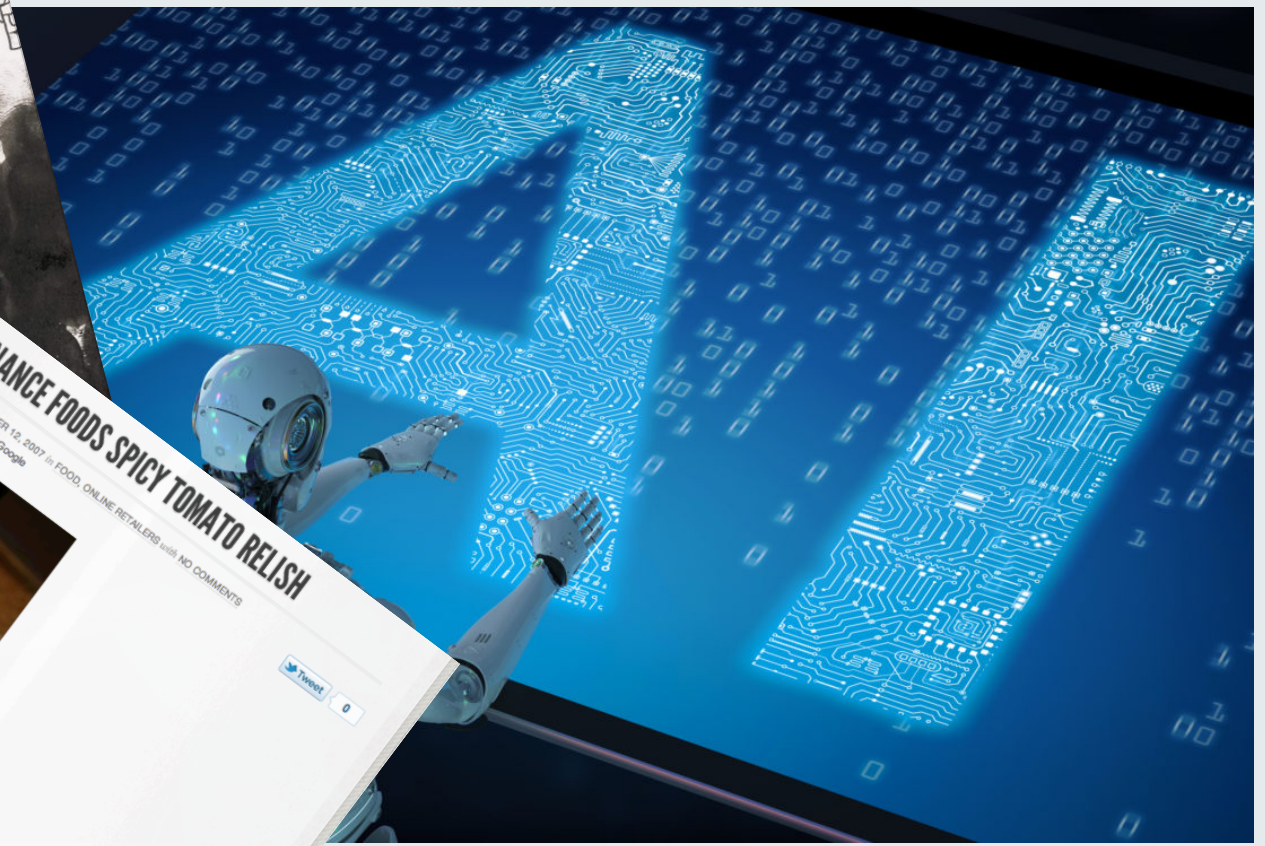
EPSRC - EP/R033722/1

UKRI Trustworthy Autonomous Systems Hub

EPSRC - EP/V00784X/1

What do they have in common?





Provenance of Food

Food provenance means:

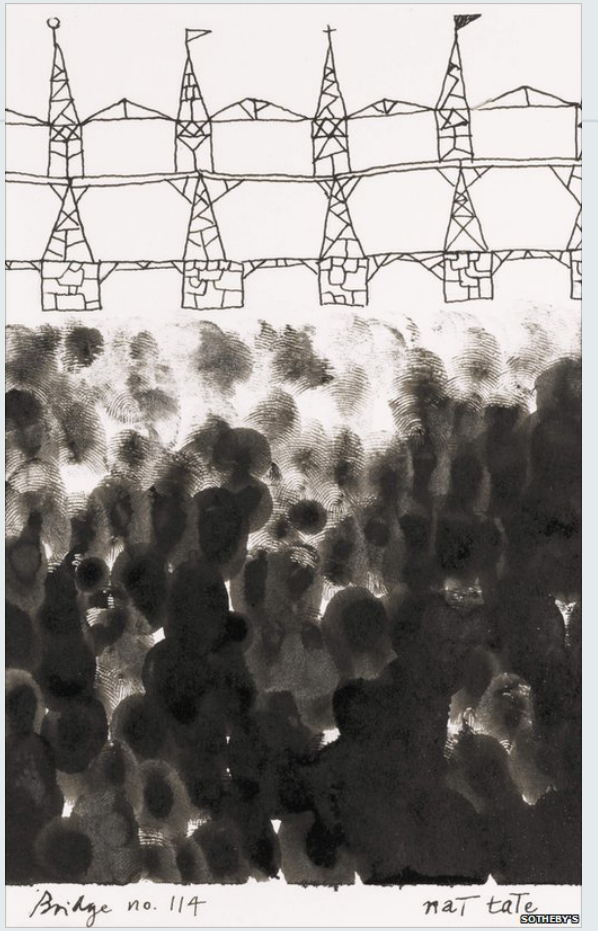
- knowing where food was grown, caught or raised
- knowing how food was produced
- knowing how food was transported

(source BBC bitesize)

Provenance of Art

Provenance of art:

- is the documentation that authenticates a particular art piece
- provides details like the work's creator, history, and appraisal value.



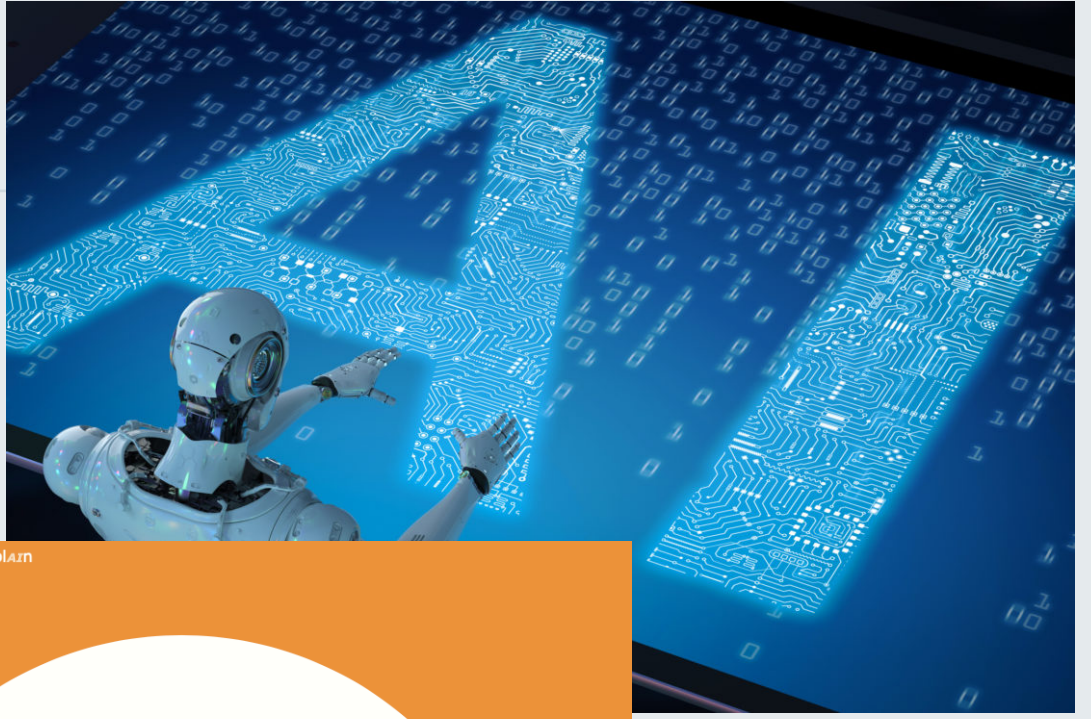
Nat Tate, a fictional dead artist, whose drawings
are being sold at real auction houses
https://www.theguardian.com/books/2011/oct/14/nat-tate-artist-hoax-william-boyd

Provenance of Data — A timeline

- We are poor at tracking the provenance of our data! But it is essential!
  - 2002 UK e-Science programme, myGrid project
  - But how?
- "Good curation demands good provenance. Provenance is no longer merely the nicety of artists, academics, and wine makers. It is an ethic we expect." (Jeff Jarvis, 2010)
  http://buzzmachine.com/2010/06/27/the-importance-of-provenance/
- Provenance is a record that describes the people, institutions, entities, and activities, involved in producing, influencing, or delivering a piece of data or a thing in the world
  - Provenance working group, World Wide Web Consortium, 2013
  - W3C PROV is a standard for provenance on the Web

W3C PROV

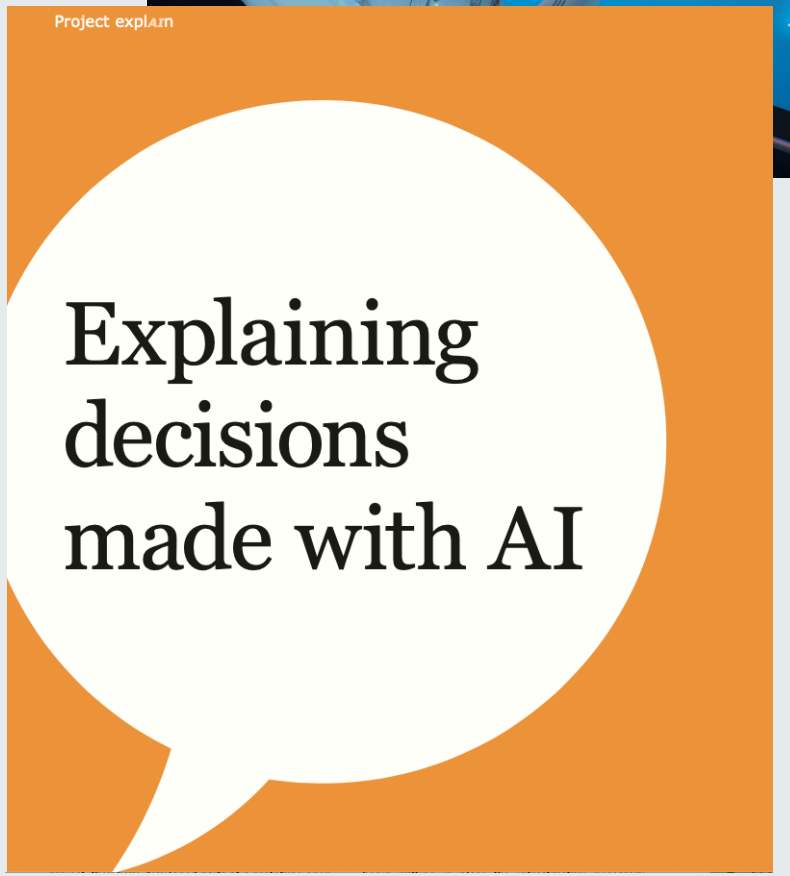
Provenance in AI



"When PROV is adopted as a way of uniformly encoding the provenance of a decision within or across organisations

…. You can then extract the relevant information to construct the desired explanation

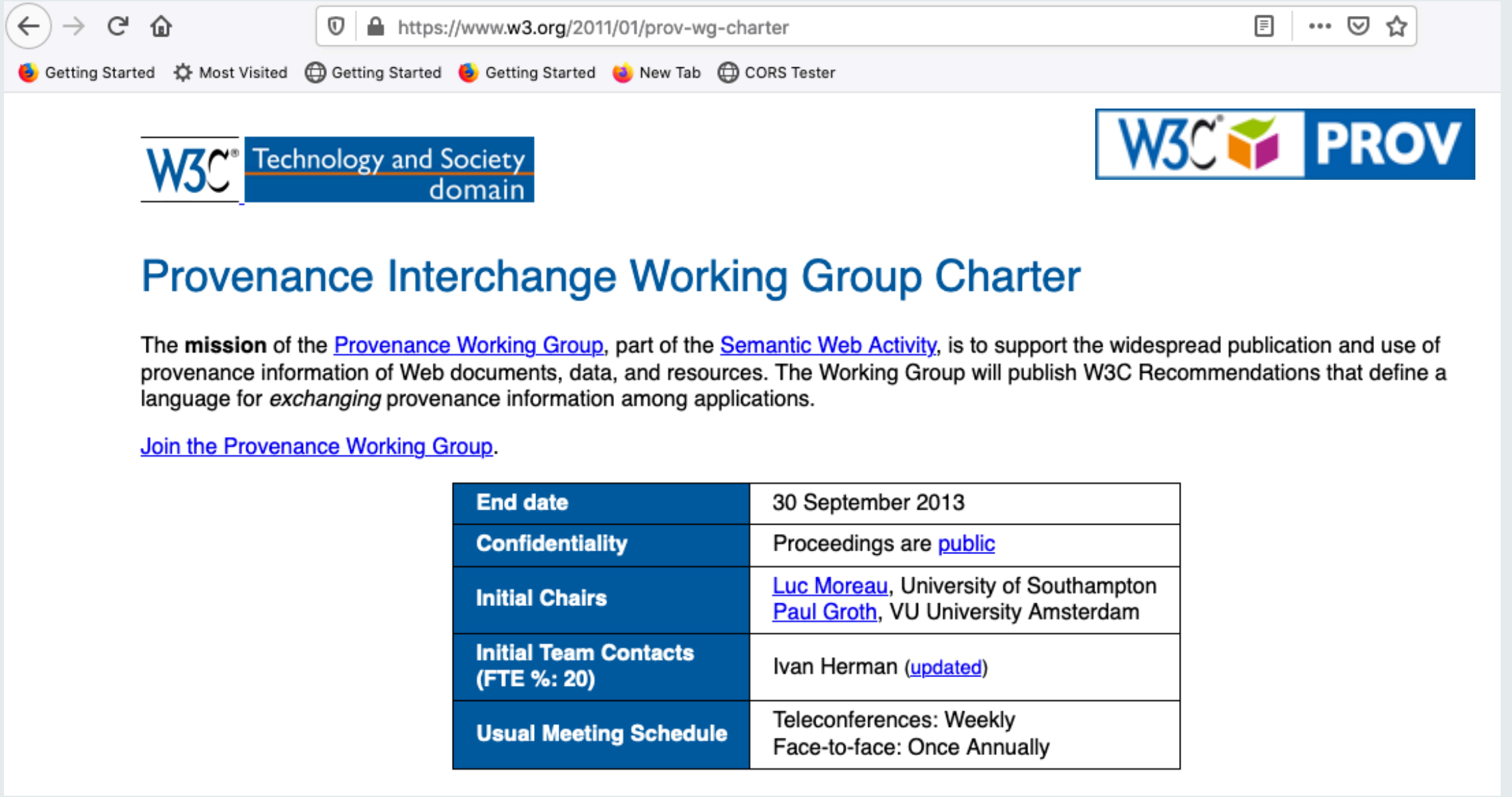
the approach will help automate the process of extracting explanations about the pipeline around an AI model."



The UK Information Commissioner's office
https://ico.org.uk/for-organisations/guide-to-data-protection/key-data-protection-themes/explaining-decisions-made-with-artificial-intelligence/

PROV 101

Provenance Working Group



https://www.w3.org/2011/01/prov-wg-charter

Getting Started | Most Visited | Getting Started | Getting Started | New Tab | CORS Tester

W3C Technology and Society domain

W3C PROV

Provenance Interchange Working Group Charter

The **mission** of the Provenance Working Group, part of the Semantic Web Activity, is to support the widespread publication and use of provenance information of Web documents, data, and resources. The Working Group will publish W3C Recommendations that define a language for *exchanging* provenance information among applications.

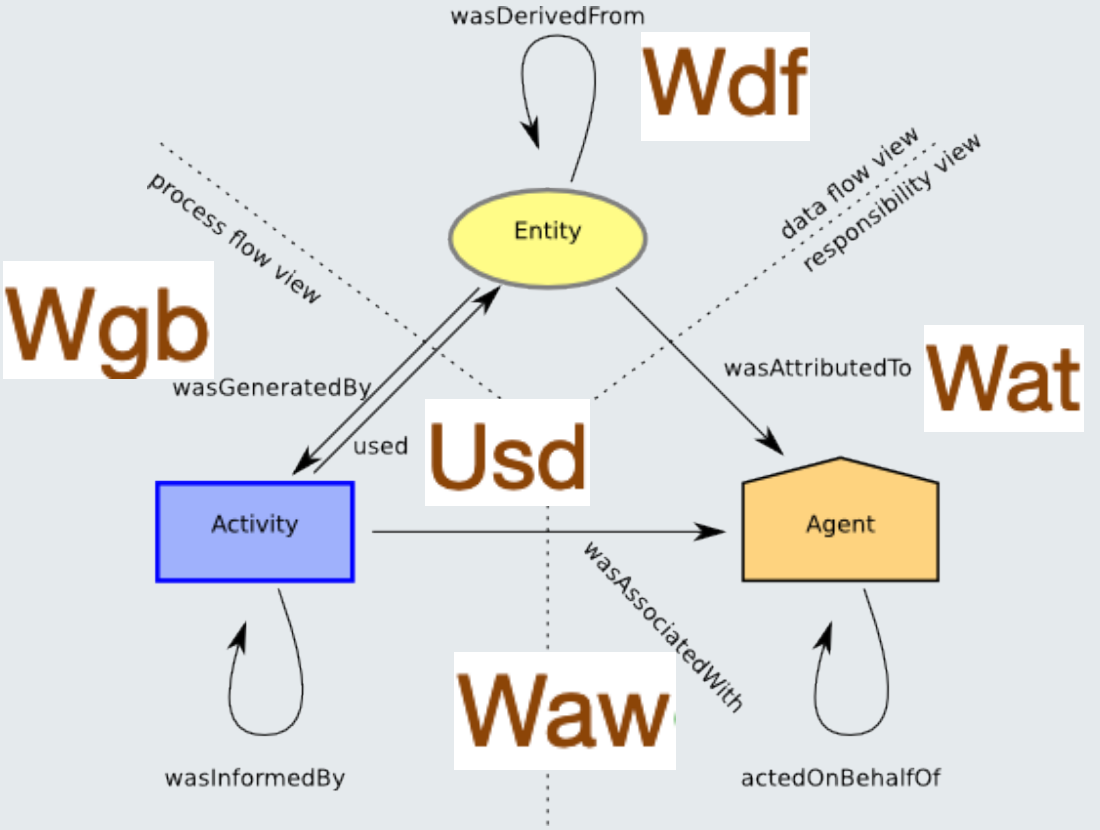
Join the Provenance Working Group.

End date	30 September 2013
Confidentiality	Proceedings are public
Initial Chairs	Luc Moreau, University of Southampton Paul Groth, VU University Amsterdam
Initial Team Contacts (FTE %: 20)	Ivan Herman (updated)
Usual Meeting Schedule	Teleconferences: Weekly Face-to-face: Once Annually

Provenance 101: Three Core Concepts and Associations

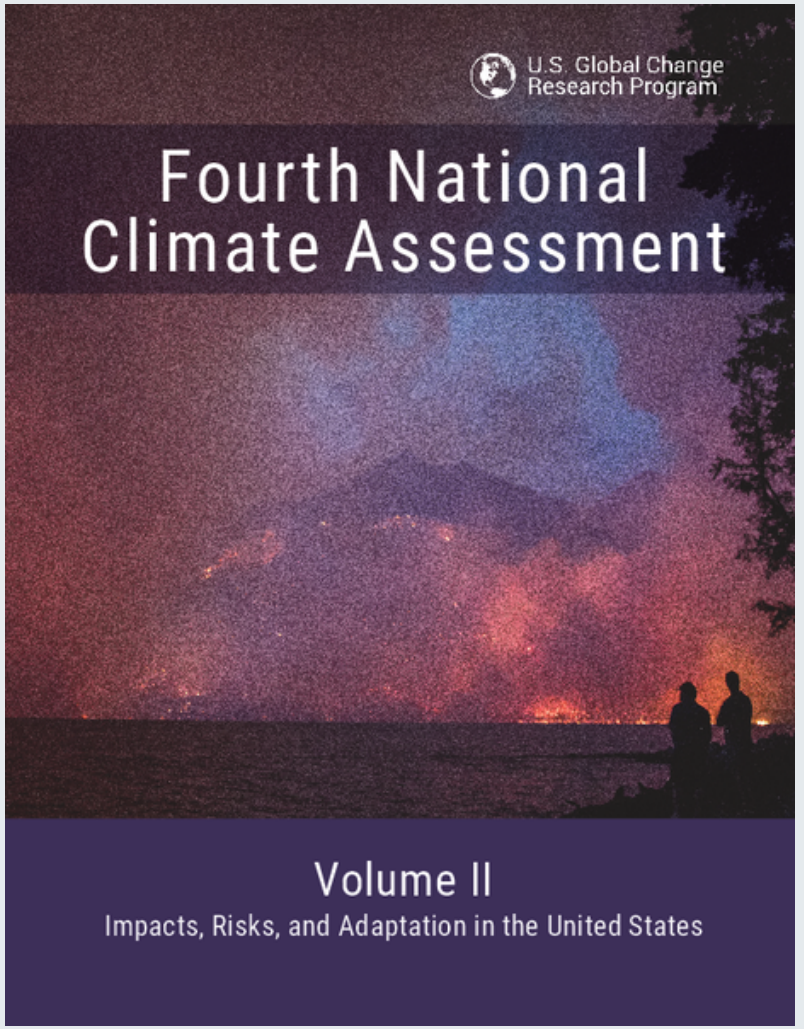
The PROV Data Model http://www.w3.org/TR/prov-dm/

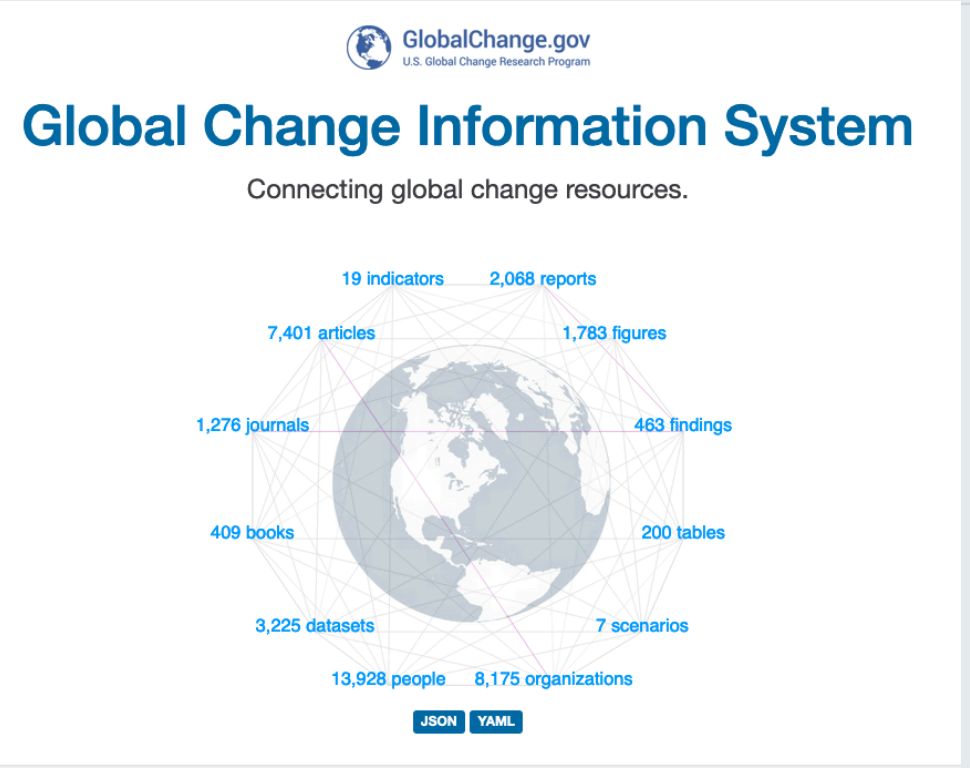
Application	PROV
Piece of information, decision, vote, document, position reading, plan, data set, model,	entity
Actions such as planning, monitoring, voting, writing, reporting, commenting, approving, training, classifying	activity
Person, service, system, organization, collective	agent



W3C PROV

National Climate Assessment Reports (2018,2014)





GCIS utilizes PROV to express provenance of resources

https://data.globalchange.gov/image/87622e9d-eca7-450a-9970-c9d79035b442

USGCRP GCIS

Search

image : 87622e9d-eca7-450a-9970-c9d79035b442

Late 21st Century, Higher Scenario (RCP8.5)

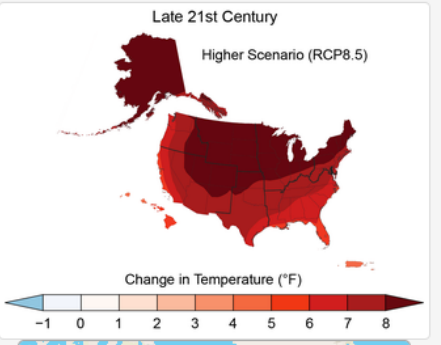


Figure may be copyright protected and permission may be required. Contact original figure source for information

The time range for this image is January 01, 1986 (00:00 AM) to December 31, 2099 (00:00 AM).

This image was created on June 07, 2018.

This image was submitted on December 03, 2018.

The spatial range for this image is 16.92° to 71.38° latitude, and 173.00° to -65.09° longitude.

Provenance

This image was derived from dataset World Climate Research Program's (WCRP's) Coupled Model Intercomparison Project phase 5 (CMIP5) multi-model ensemble using the activity nca4-projected-global-temperature-panel-4-wi97g-process

This image is part of this figure :

```
<https://data.globalchange.gov/image/87622e9d-eca7-450a-9970-c9d79035b442>
   prov:wasGeneratedBy <https://data.globalchange.gov/activity/nca4-projected-global-temperature-panel-4-wi97g-process>;
   prov:wasDerivedFrom <https://data.globalchange.gov/dataset/nca3-cmip5-r1>.
```

https://data.globalchange.gov/image/87622e9d-eca7-450a-9970-c9d79035b442.thtml

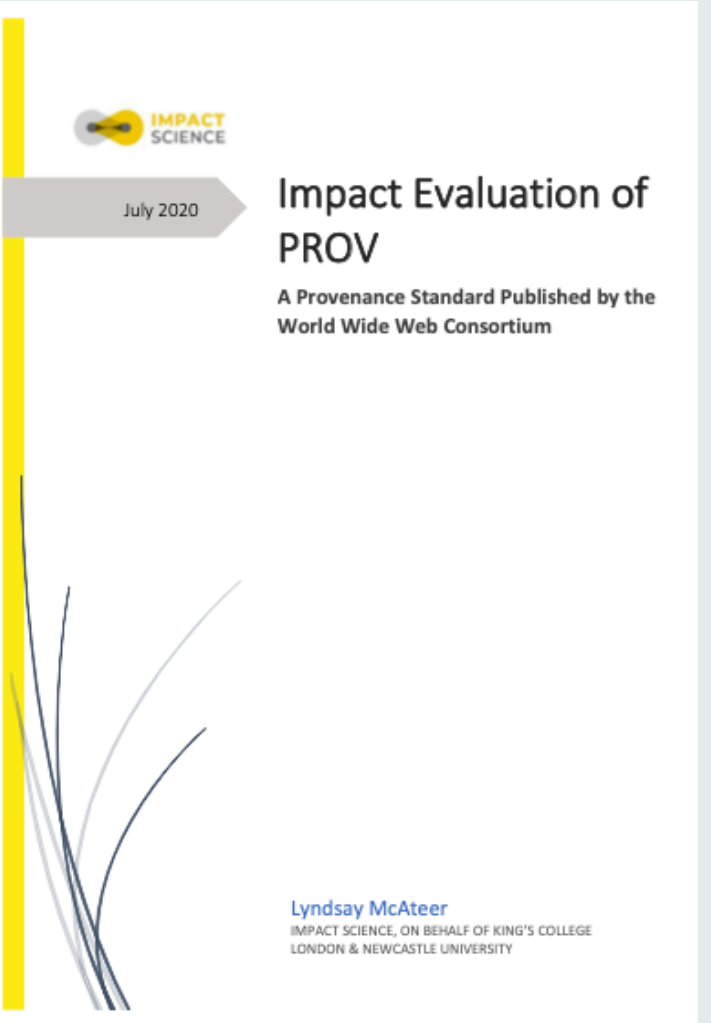
Alternatives : JSON | YAML | Turtle | N-Triples | JSON Triples | RDF+XML | RDF+JSON | Graphviz | SVG

GlobalChange.gov
U.S. Global Change Research Program

Impact of PROV

- **“PROV improves access to information through the use of linked knowledge and the knowledge graph”**
- **“PROV works in the background to provide a clear, ethical and transparent data source”**
- **“PROV ensures that information released to the public domain is accurate”**

https://www.impact.science/case-study/evaluating-impact-of-the-prov-data-model-how-impact-science-gathered-evidence-for-ref-2021-for-kings-college-london-and-newcastle-university/



Provenance Analytics

Why analytics? Common sense motivation

Origin, Timing, Provenance and Quality: Is Data the New Wine?

Waitrose focuses on quality and provenance in marketing push

The v … praisal

Lidl to laun… quality' ads

Wight Marque - Isle of Wight Food Provenance Scheme

Can we relate provenance to properties (quality, reliability, etc) of data and underpinning processes?

Distinguishing entities by their Provenance Type (i.e. their history)

Compare:

(1) Harry Potter and the Philosopher's Stone,

(2) Harry Potter à l'école des sorciers

Similar because they are both entities: (**same** Level 0)

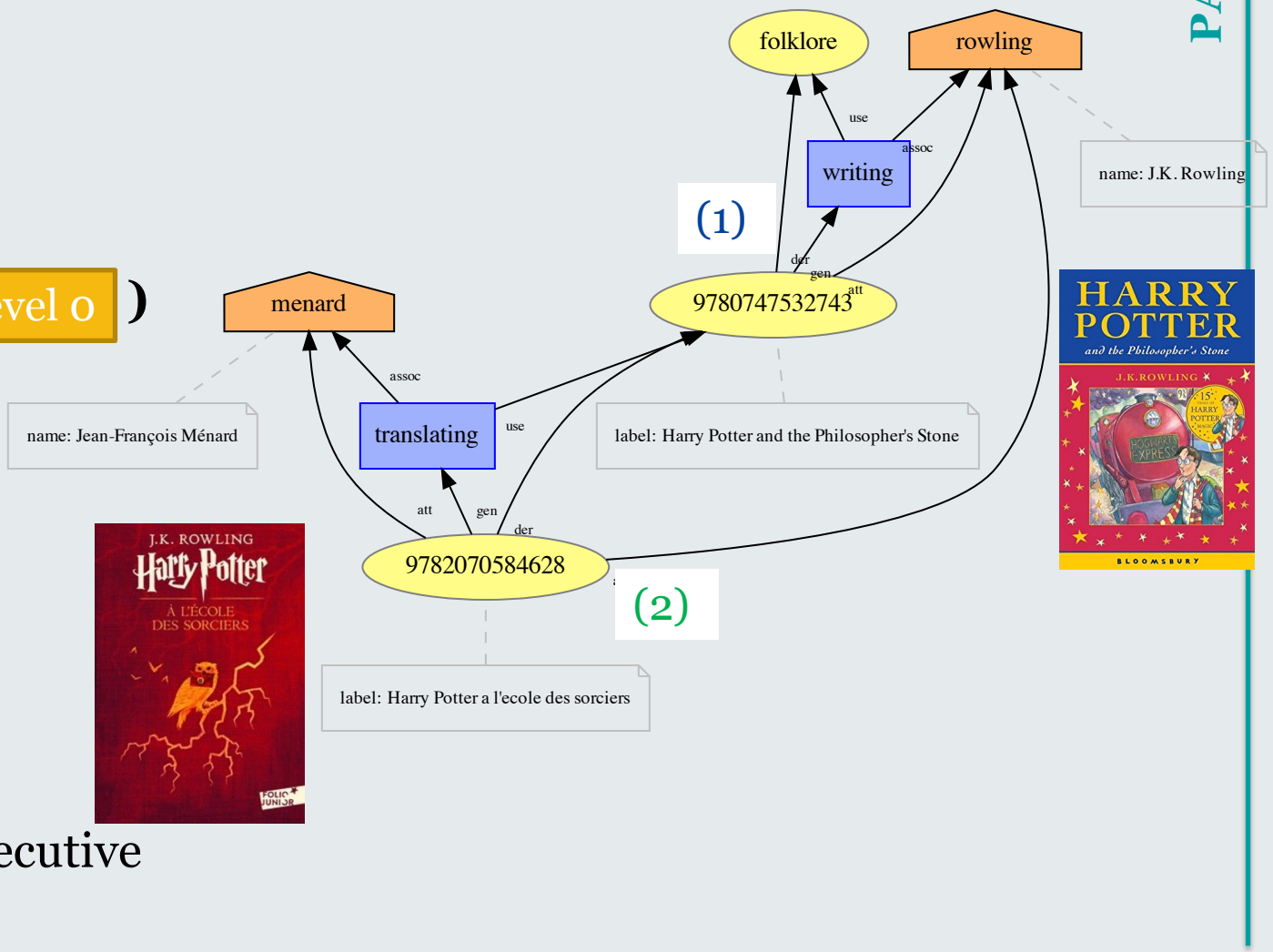
Even more similar because they were both:

- Generated by an activity,
- Attributed to an agent, and
- Derived from other entity.

(**same** Level 1)

Going one step further, (1) is a result of two consecutive derivations, whereas (2) isn't.

(different Level 2)



Computing Provenance Types

Provenance types are inferred **recursively**:

- A node v 's Level n type by given by combining all edges starting at v with the Level n-1 type of the nodes v connects to.

The computational complexity is thus **linear** on the number of edges of the graph.

Level 0 **and** Level 1

```
"0":[Ent] (count: 1)
"1":[Ag] (count: 2)
"2":[Ent,Wat([Ag]),Wdf([Ent],[]),Wgb([Act])] (count: 2)
"3":[Act,Usd([Ent]),Waw([Ag],[])] (count: 2)
```

Provenance Type "2" can be paraphrased as:

The type of entities attributed to an agent, derived from an entity, and generated by an activity

Ent,Wat([Ag]), Wdf([Ent],[]), Wgb([Act])]

Provenance Types (level 0, 1, 2, … , k)

Level 0

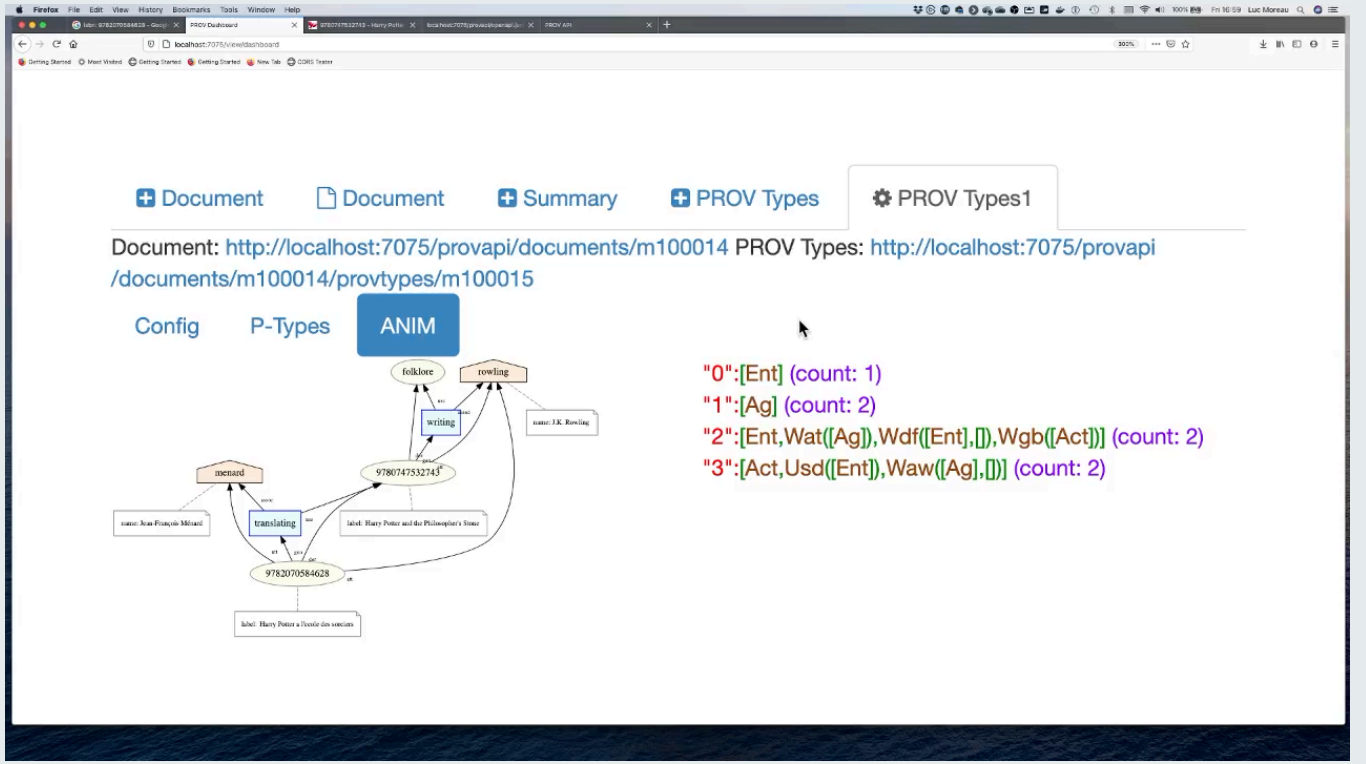
"0":[Act] (count: 2)
"1":[Ent] (count: 3)
"2":[Ag] (count: 2)

Level 1

"0":[Ent] (count: 1)
"1":[Ag] (count: 2)
"2":[Ent,Wat([Ag]),Wdf([Ent],[]),Wgb([Act])] (count: 2)
"3":[Act,Usd([Ent]),Waw([Ag],[])] (count: 2)

Level 2

"0":[Ent] (count: 1)
"1":[Ag] (count: 2)
"2":[Act,Usd([Ent]),Waw([Ag],[])] (count: 1)
"3":[Ent,Wat([Ag]),Wdf([Ent],[]),Wgb([Act]),Wgb([Usd([Ent]),Waw([Ag],[])])] (count: 1)
"4":[Act,Usd([Ent]),Usd([Wat([Ag]),Wdf([Ent],[]),Wgb([Act])]),Waw([Ag],[])] (count: 1)
"5":[Ent,Wat([Ag]),Wdf([Ent],[]),Wdf([Wat([Ag]),Wdf([Ent],[]),Wgb([Act])],[]),Wgb([Act]),Wgb([Usd([Ent]),Waw([Ag],[])])] (count: 1)



The deeper the type, the more discriminating it is

Provenance Kernel

David Kohan Marzagão, Trung Dong Huynh, Ayah Helal, Luc Moreau, Provenance Graph Kernel
Arxiv 2020, https://arxiv.org/abs/2010.10343

- For a graph, we calculate each node's provenance types up to a given level k.

- **We count how often each type appears in the graph**

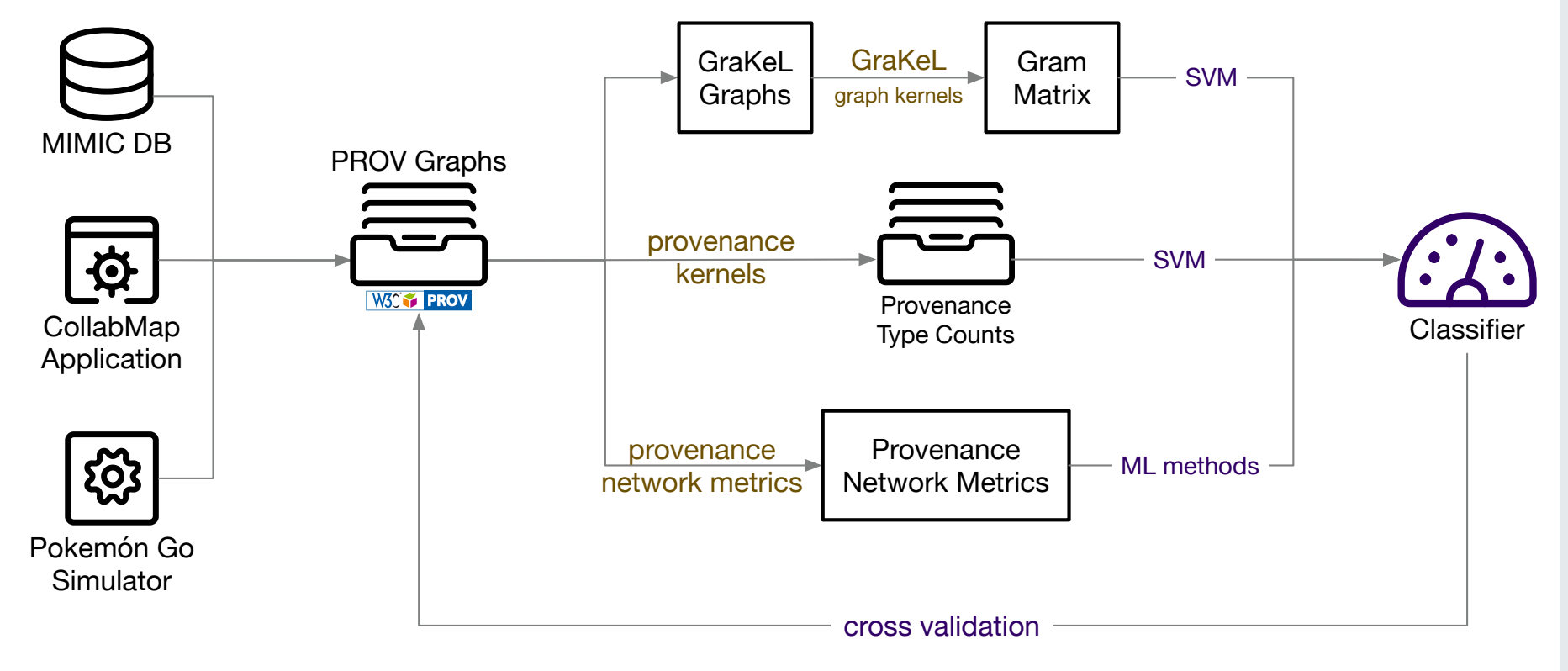
- Two graphs are **similar** if they have similar combinations of provenance types.
  - A graph's **feature vector** counts the number of occurrence of each of its provenance types.
  - The dot product between two feature vectors is a measure of how **similar** the graphs are (**provenance kernel**): the higher the dot product, the more similar the graphs

"0":[Ent] (count: 1)
"1":[Ag] (count: 2)
"2":[Ent,Wat([Ag]),Wdf([Ent],[]),Wgb([Act])] (count: 2)
"3":[Act,Usd([Ent]),Waw([Ag],[])] (count: 2)

Machine Learning for Provenance

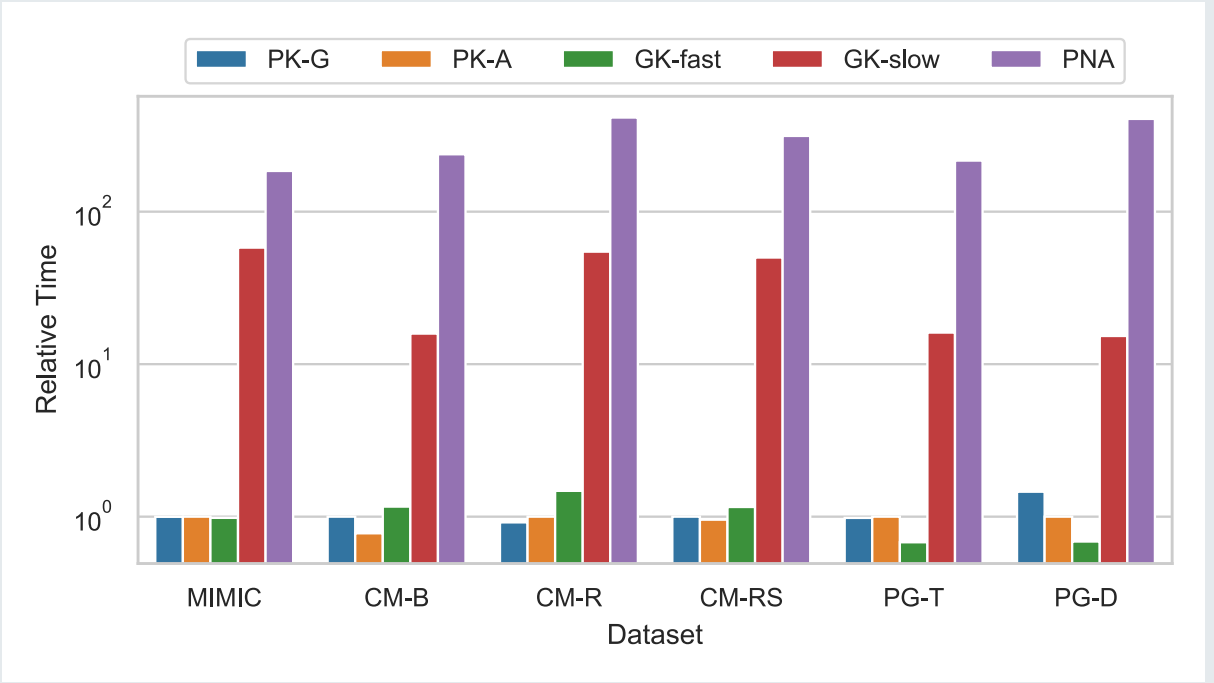
- Why ML for Provenance?
  - Classification task for provenance … to predict data/application characteristics

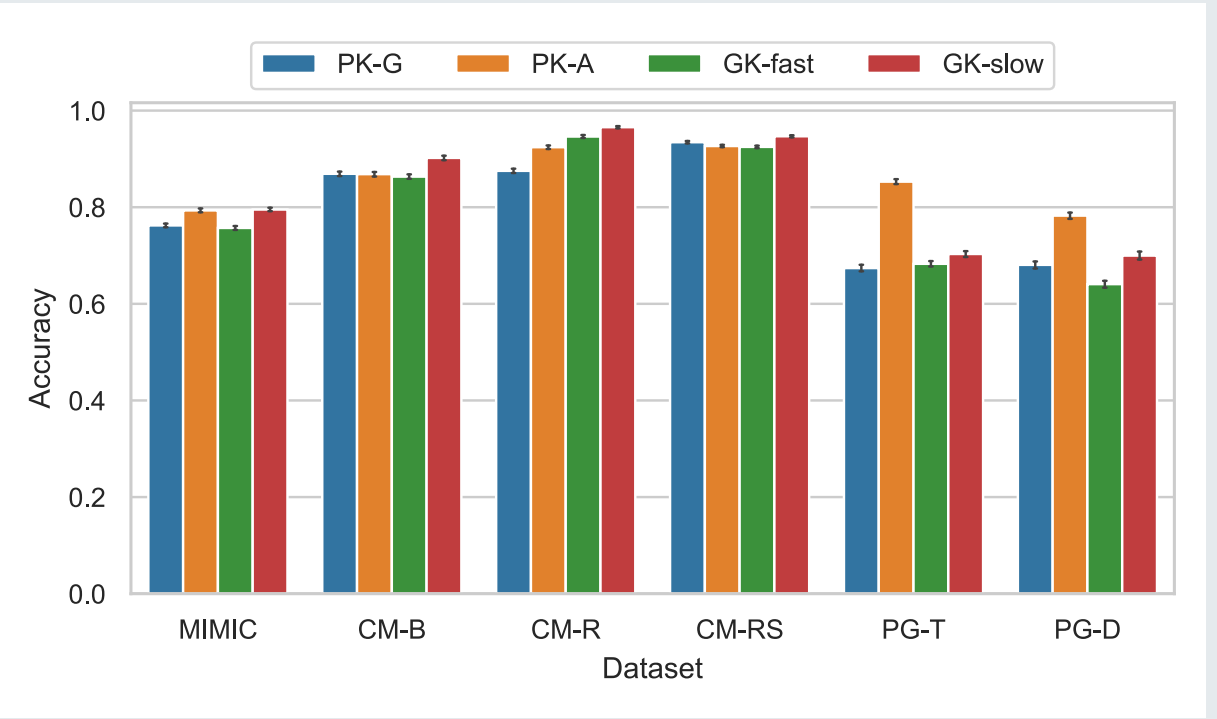
- Method that doesn't need to be tailored to each different application.



Machine Learning for Provenance

- Comparison of Provenance Kernels (PK), with Graph Kernels (GK) and Provenance Network Analytics (PNA)
- Variety of datasets: human processes, huma-machine processes, machine-only processes (simulations)



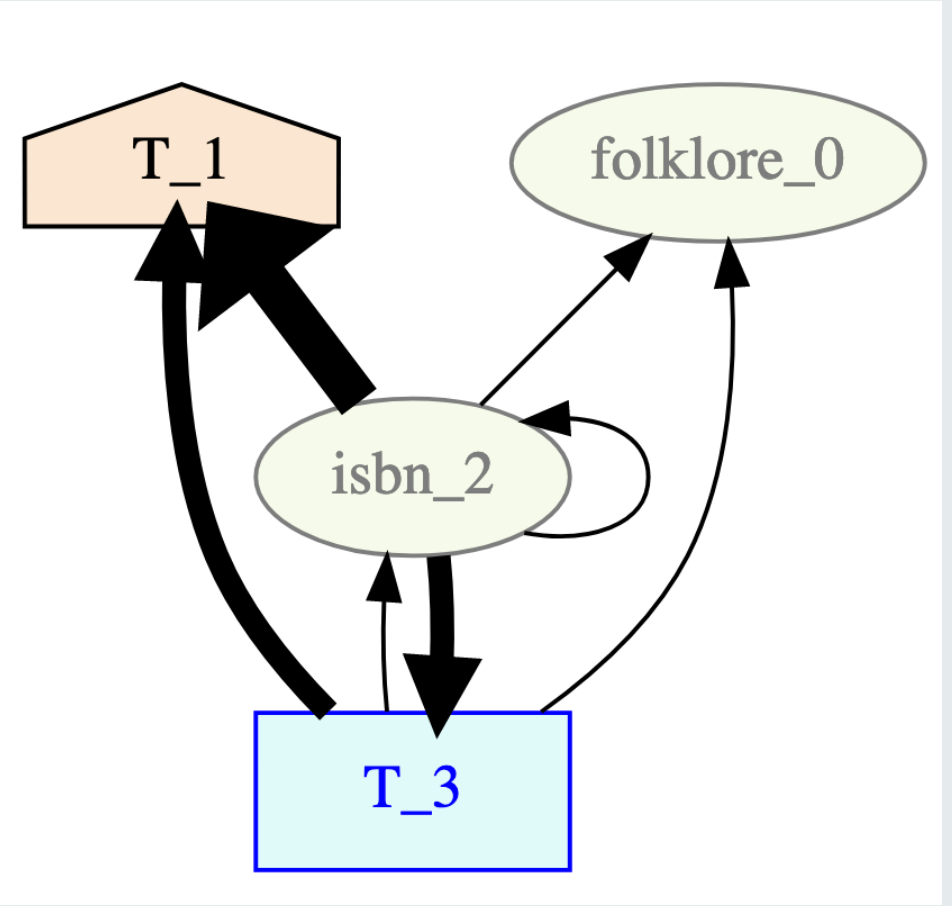


PK is fast and outperforms fast GK methods, is providing similar performance to slower methods (GK-slow and PNA)

Provenance Graph Summarisation

Luc Moreau. Aggregation by provenance types: A technique for summarising provenance graphs. In *Graphs as Models 2015*, Electronic Proceedings in Theoretical Computer Science 2015

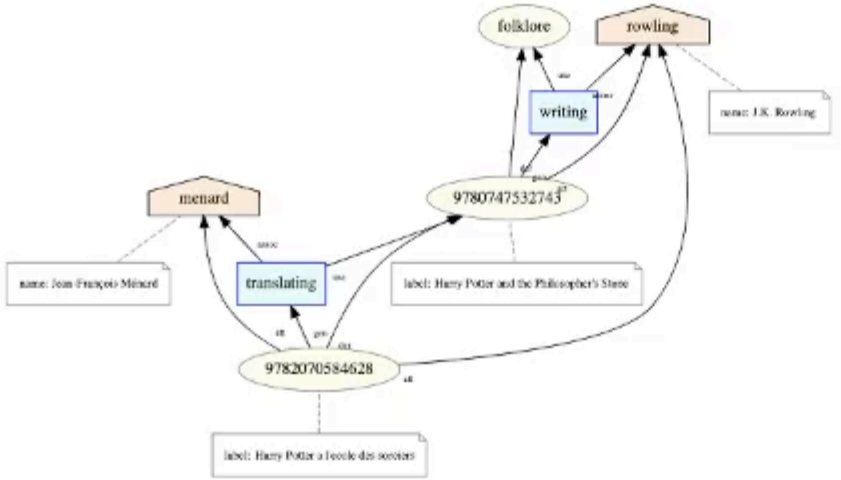
- Summarisation = Provenance Graph Transformation
- Each node in Summary graph corresponds to a provenance type in the original graph
- Map each graph node (n1) to a summary node (t1) with its provenance type
  - + count of the number of graph nodes for each summary node
- Map each graph edge (n1,n2) to a summary edge (t1, t2)
  - + count of the number of graph edges for each summary edge
- Visual rendering of the counts by the thickness of edges/nodes
  - Facilitates outliers and common pattern detection

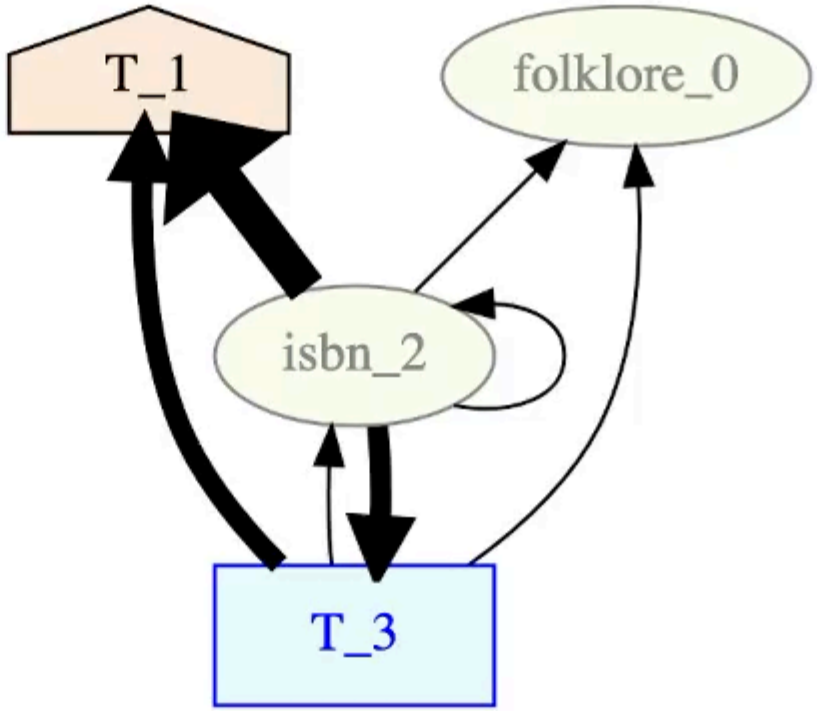


Document | Document | Summary | PROV Types | Summary1

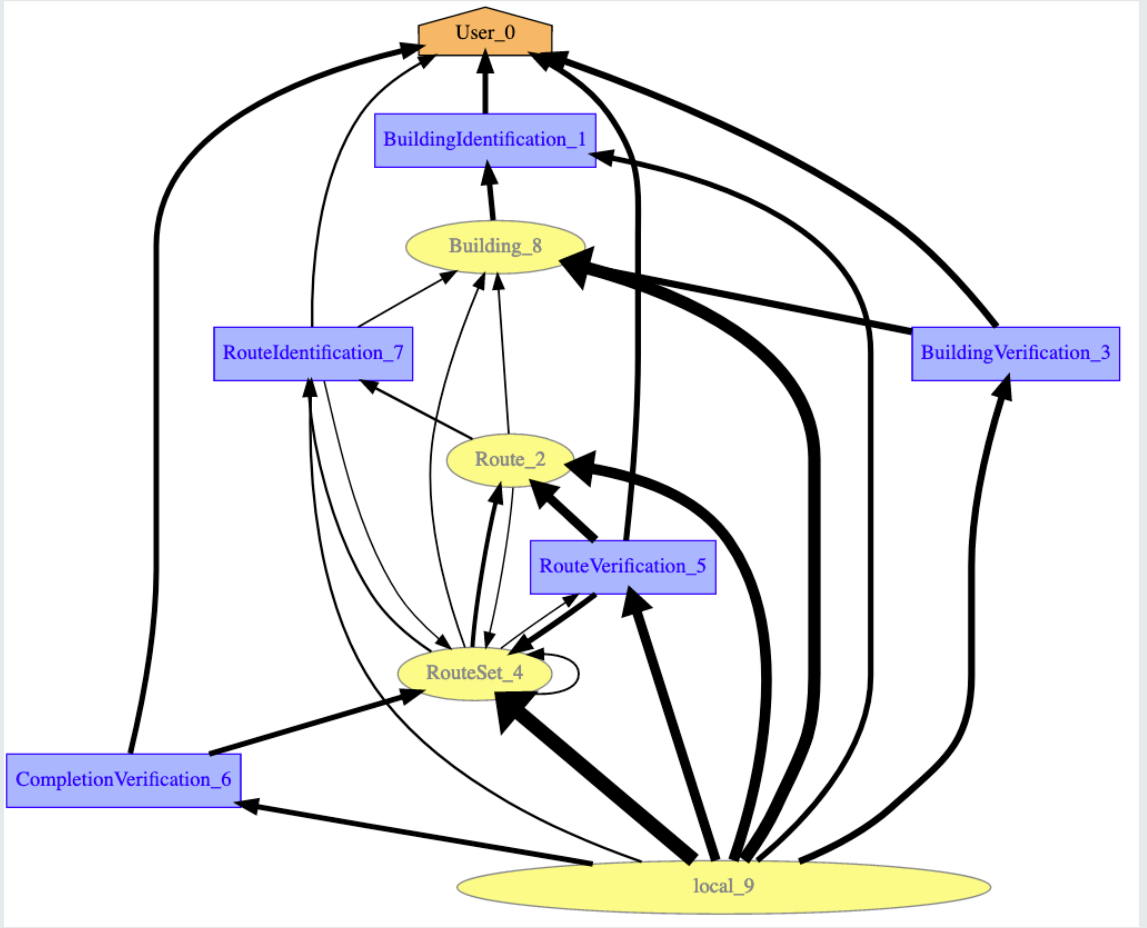
Document: http://localhost:7075/provapi/documents/m100013 Summary: http://localhost:7075/provapi/documents/m100016

Config | PROV-N | P-Types | SVG | ANIM





Summarisation Application: Outlier detection

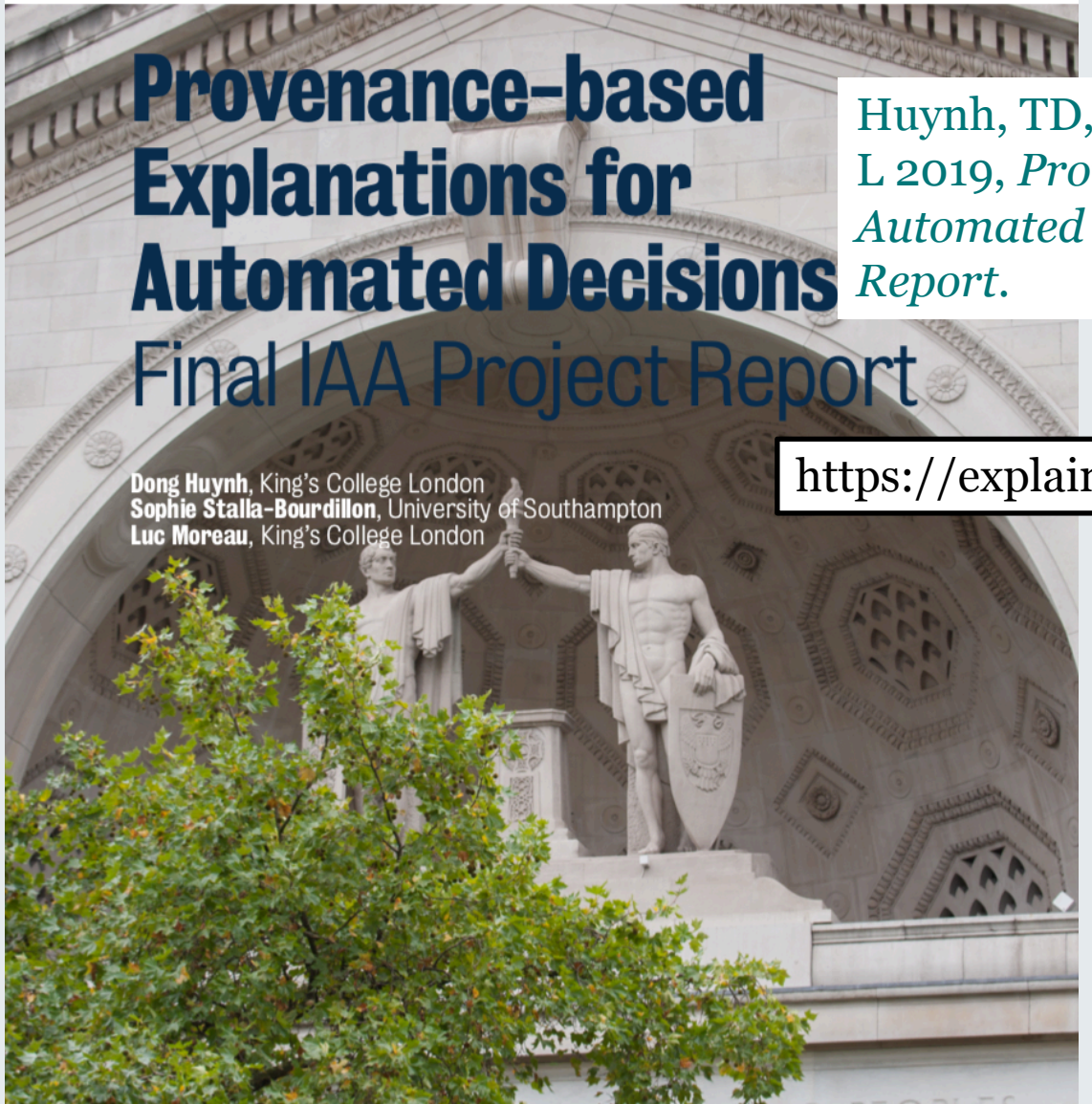


"0":[Ag,Prim([collabmap:User])] (count: 85)
"1":[Act,Prim([collabmap:BuildingIdentification])] (count: 1082)
"2":[Ent,Prim([collabmap:Route])] (count: 370)
"3":[Act,Prim([collabmap:BuildingVerification])] (count: 1393)
"4":[Ent,Prim([collabmap:RouteSet])] (count: 483)
"5":[Act,Prim([collabmap:RouteVerification])] (count: 1135)
"6":[Act,Prim([collabmap:CompletionVerification])] (count: 1077)
"7":[Act,Prim([collabmap:RouteIdentification])] (count: 371)
"8":[Ent,Prim([collabmap:Building])] (count: 1082)
"9":[Ent,Prim([collabmap:Vote])] (count: 5952)

The original has over 10000 nodes, and cannot be visualised in a useful way. The summary however shows key aspects.

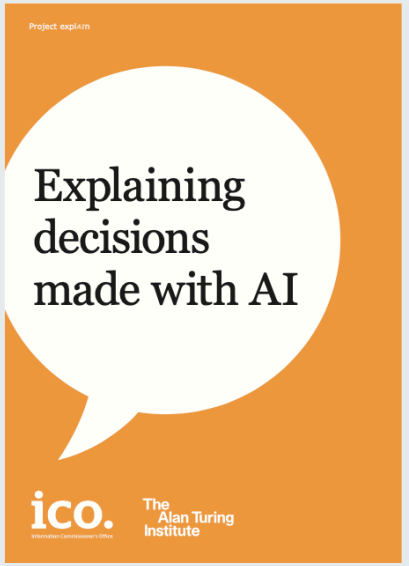
Ramchurn, Sarvapali D., Huynh, Trung Dong, Venanzi, Matteo and Shi, Bing (2013) **Collabmap**: crowdsourcing maps for emergency planning. *The 5th Annual ACM Web Science Conference, France. 02 - 04 May 2013*. pp. 326-335 .

Explanations



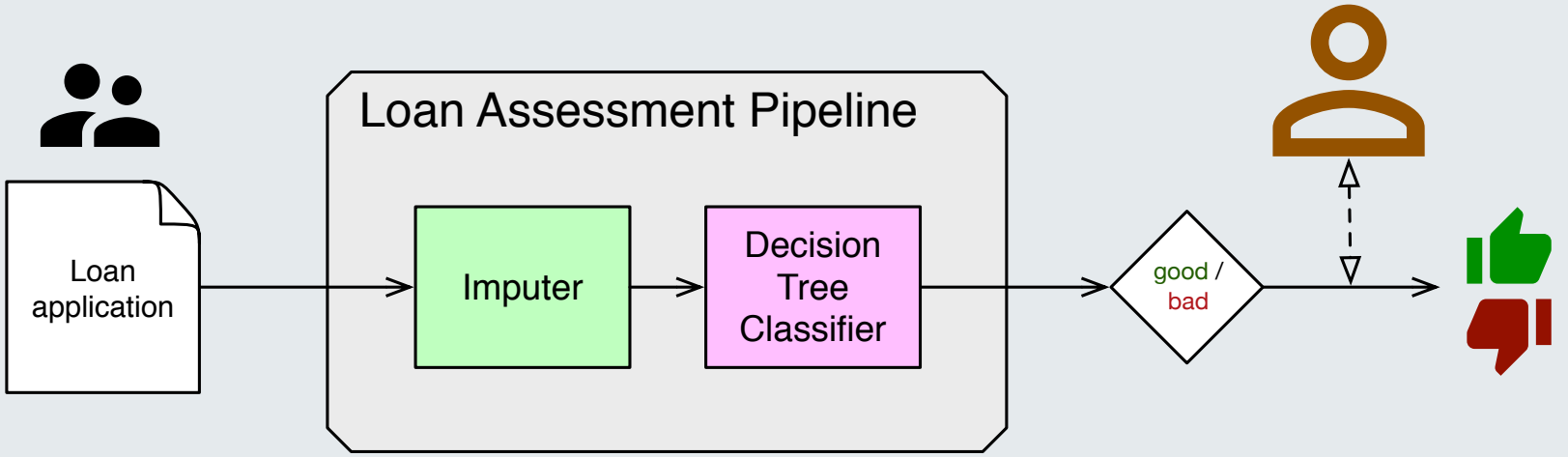
Huynh, TD, Stalla-Bourdillon, S & Moreau, L 2019, *Provenance-based Explanations for Automated Decisions: Final IAA Project Report*.

https://explain.openprovenance.org/report/



Loan Assessment Pipeline

https://explain.openprovenance.org/

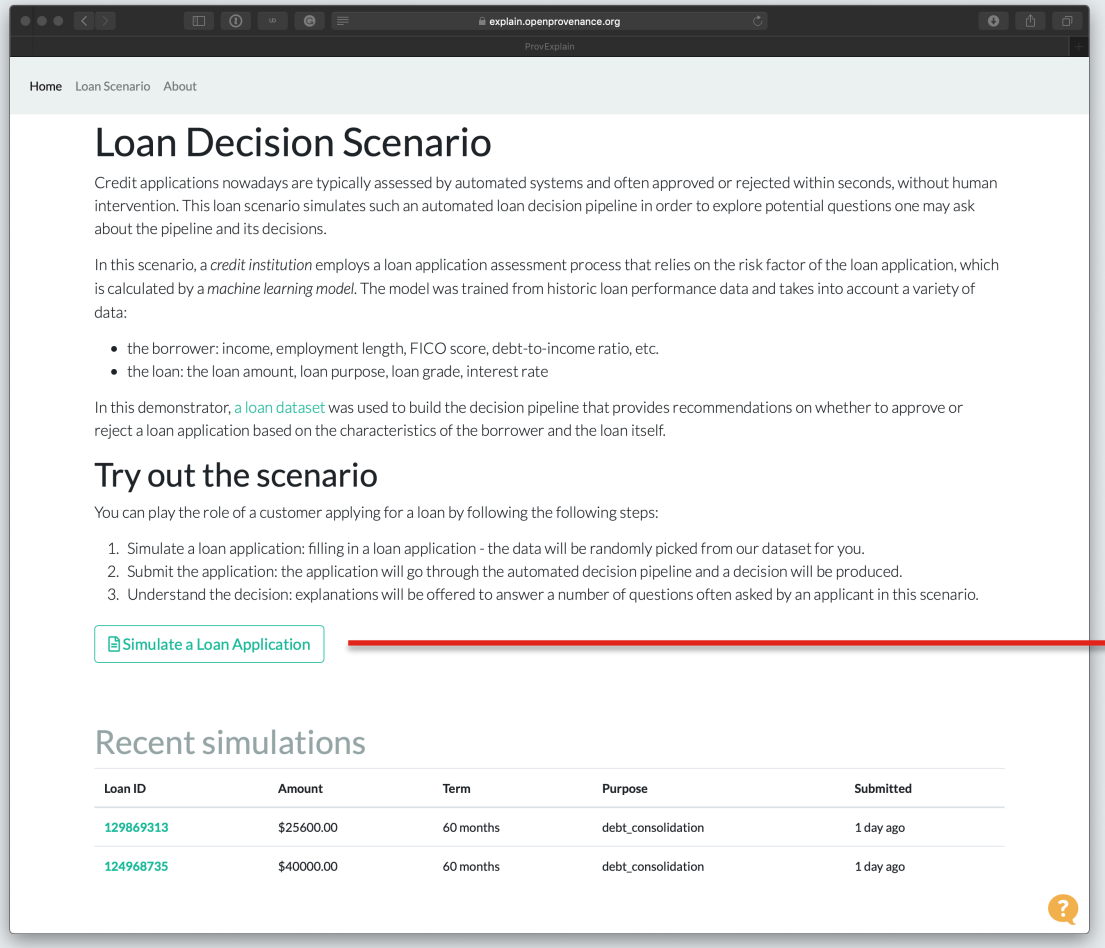


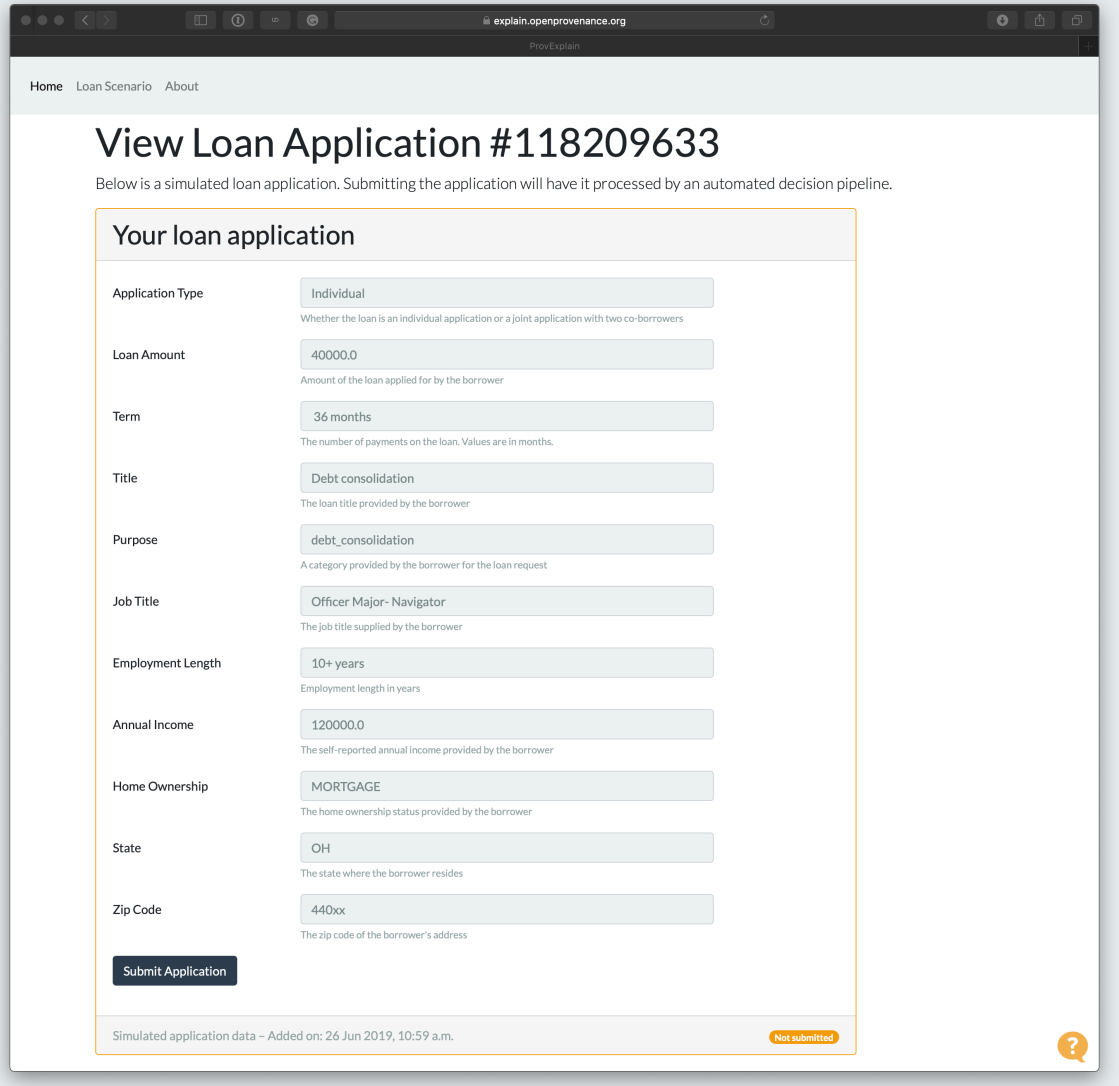
Provenance-enabled Decision Pipeline

- User submits a loan application
- Decision making pipeline makes a recommendation
- Classifier trained over a dataset of loan applications and associated decisions

The provenance of each step is recorded and stored after each run

Simulate a loan application





Provenance-based Explanations for Data Subjects

We recorded the provenance of the above decision, from which explanations about the decision can be generated. If you have queries about the above decision, some explanations can be found below by clicking on the corresponding questions below.

- Has the loan decision been reached solely via automated means?

  Whether a decision made solely by automated means without any meaningful human involvement.
- What types of data were used to assess my loan application?

  A loan application assessment may consider several types of data about the applicant, such as credit scores, or other publicly available information.
- Which data was excluded from the decision process?

  Some information you provided may not be used, either because it is not legal to do so orthe organisation deemed it is not relevant to the decision of approving your loan.
- Where did you get those data about me?

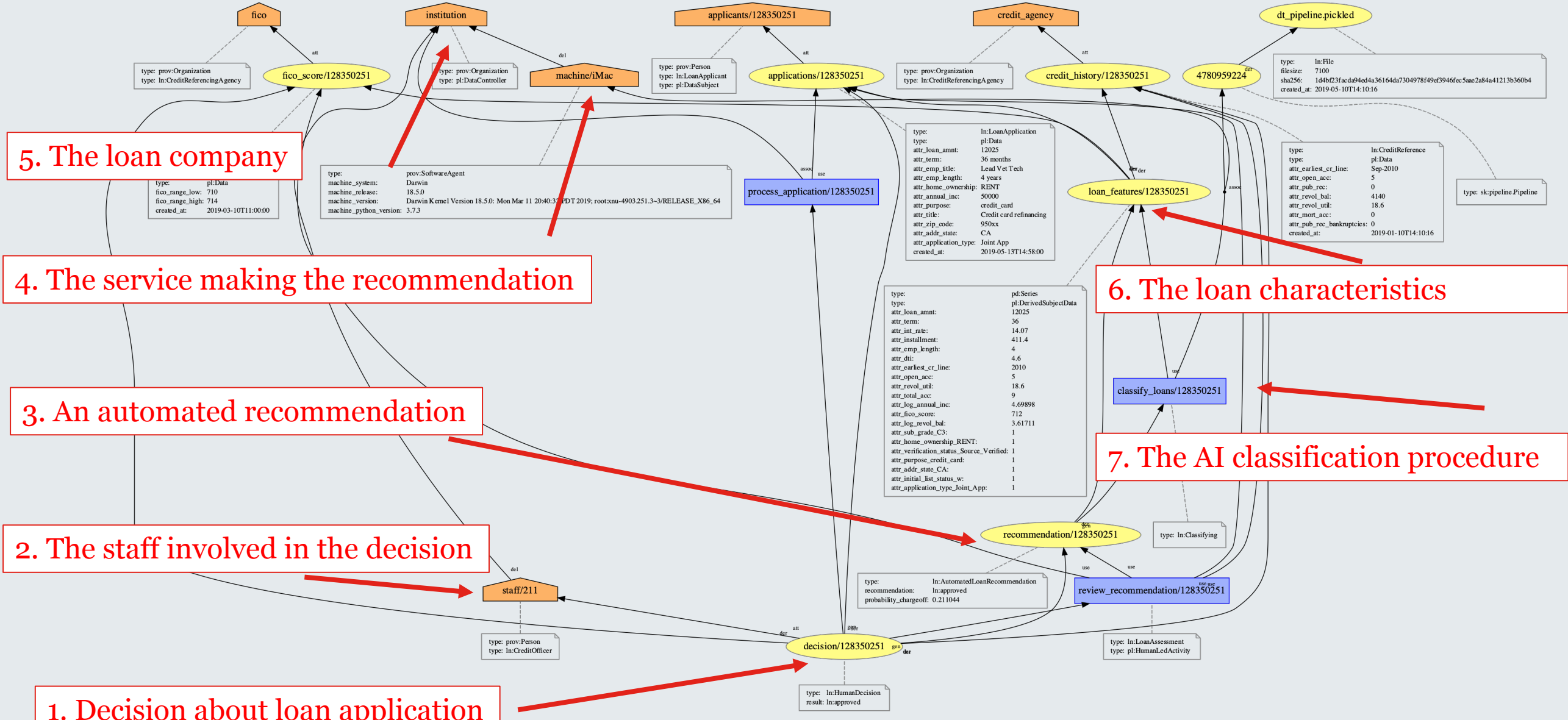
  Data considered by a credit institution may come from a variety of sources.
- How timely relevant is the data used for assessing my loan?

  Data used in loan decision making may be collected a long time ago and no longer relevant.
- Are the data used for assessing my loan application correct?

  Data correctness may not be guaranteed: the applicant may have made a typo in their application or the data provided by a third-party may be inaccurate.
- Is there bias introduced in the decision by my home ownership status?

  An automated decision may be sensitive to a particular demographic such as whether the loan applicant owns a home or not, for instance.

Loan Decision



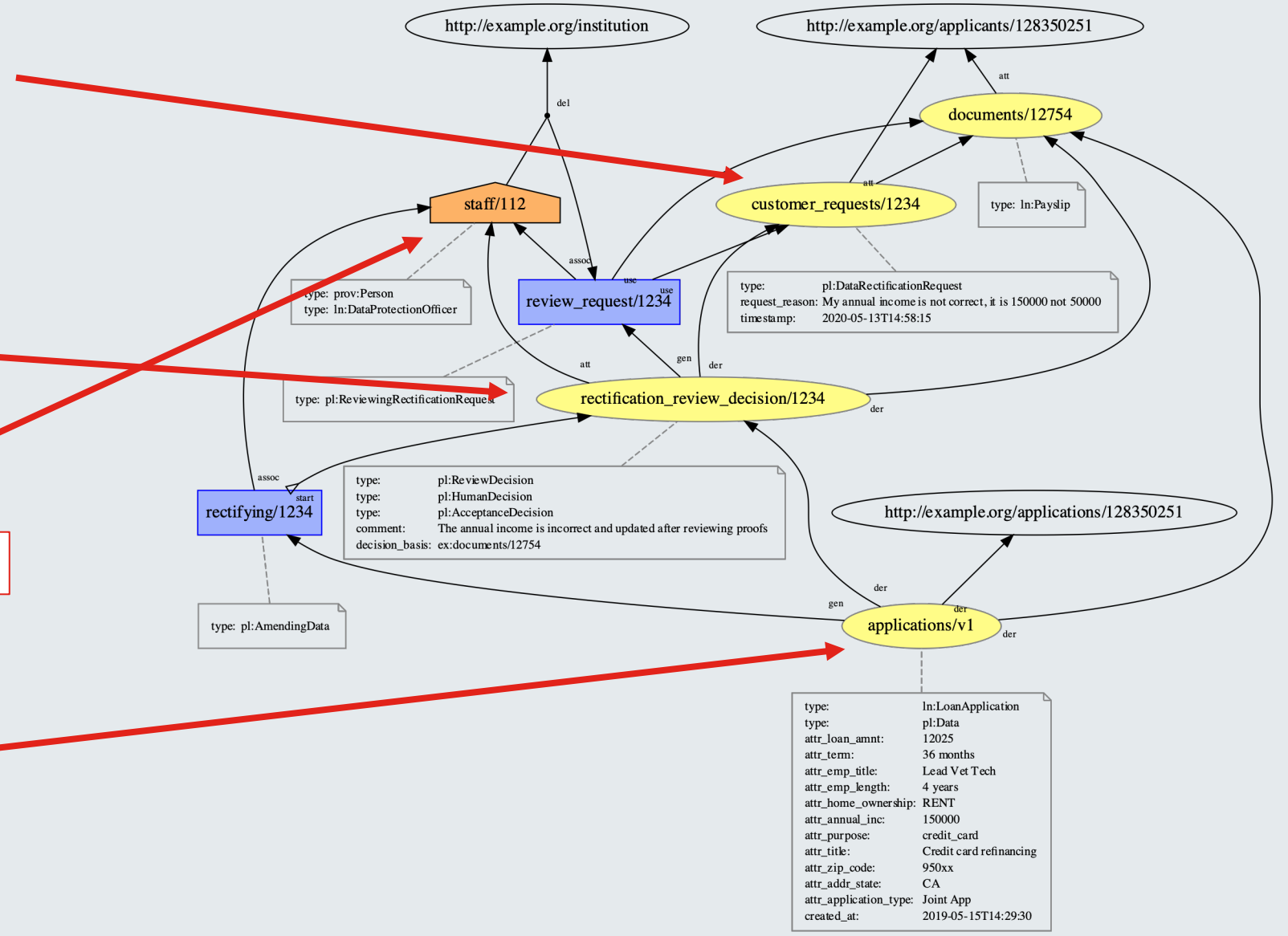
Rectification Request

1. Request to rectify data

2. Decision to rectify data

3. The staff involved in the decision

4. The revised data

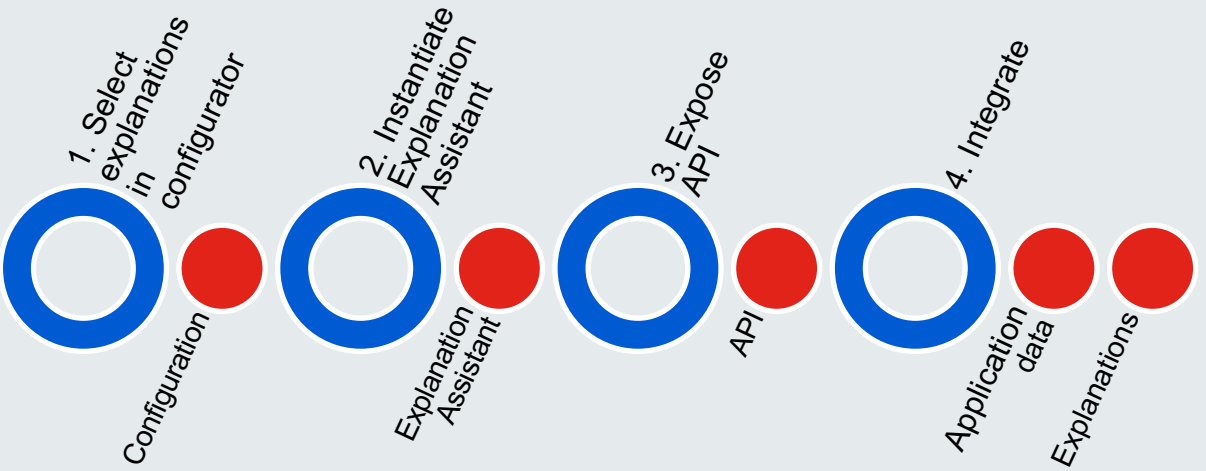


Towards a methodology for legally-grounded explanations

work with Niko Tsakalakis and Sophie Stalla-Bourdillon

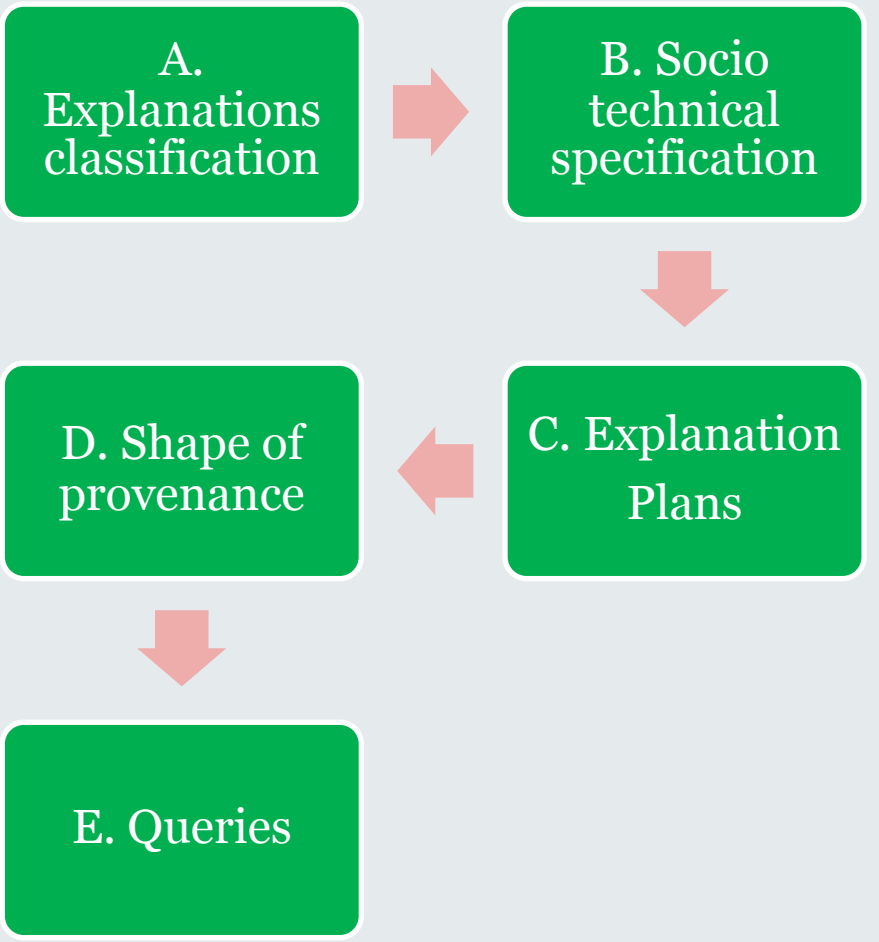
Methodology overview

From the point of view of an organisation

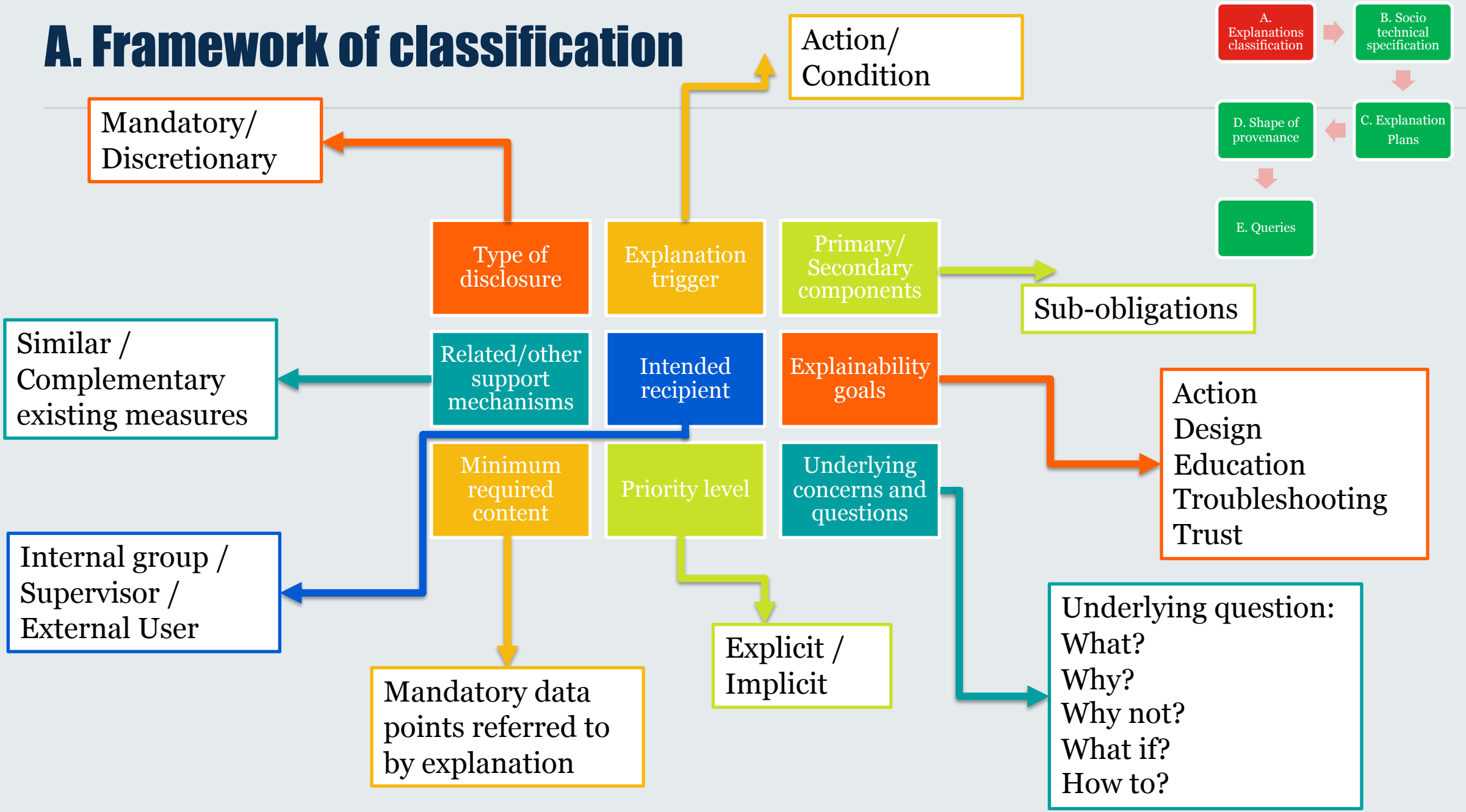


Explanation Assistant tool for organisations to provision their applications with provenance-based explanations capabilities.

From the research point of view

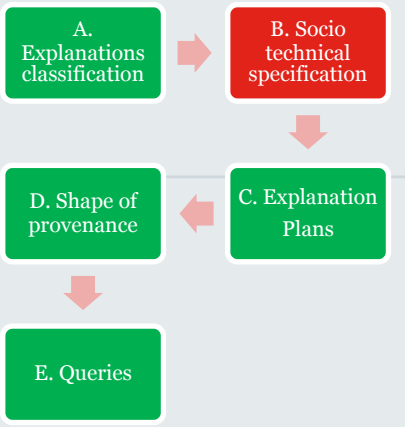


A. Framework of classification



(Sophie Stalla-Bourdillon, Niko Tsakalakis, PLEAD Project)

B. Socio-technical specification

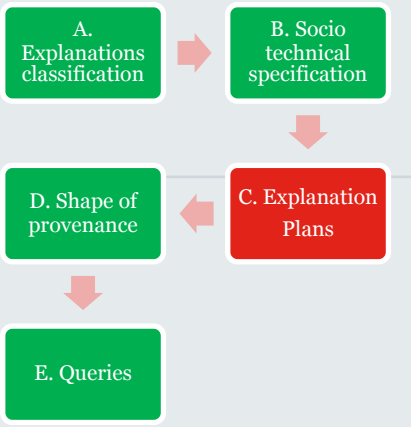


Requirements are matched with explainability goals and broken down based on

- triggers
- audience
- minimum content
- priority…

Identifier	Description	Type of disclosure	Time of disclosure	Explanation trigger	Intended recipient	Explainability goal category	Minimum required content	Priority level	Underlying concerns and questions
DP.GDPR.2	Data subjects are adequately informed about the processing	Plain (with the exception of the sources of data)	Before process	Processing of personal data	Data subjects	Understandability: Accountability	See below 2.1 - 2.6	Mandatory	
DP.GDPR.2.1	Data subjects are informed about the sources of data, the	Plain (with the exception of the sources of data)	Before process	Processing of personal data	Data subjects	Understandability: Accountability	See below 2.1.1 - 2.1.5	Mandatory	
DP.GDPR.2.1.1	Data subjects are informed about the sources of data	Aggregated values	Before process	Processing of personal data	Data subjects	Understandability: Accountability	List of data sources (e.g. public registers)	Mandatory	Where will you obtain my personal data from? How will you collect personal data about me?
DP.GDPR.2.1.2	Data subjects are informed about the recipients of data	Plain (in practice usually aggregated, e.g. 'advertising partners')	Before process	Processing of personal data	Data subjects	Understandability: Accountability	List of recipients of data (aggregated form acceptable(?))	Mandatory	Who can view, edit or receive my personal data?
DP.GDPR.2.1.3	Data subjects are informed about the purposes of processing and the legal basis	Plain (has to be explicit)	Before process	Processing of personal data	Data subjects	Understandability: Accountability	List of purposes	Mandatory	Why do you need to process my personal data? Can you achieve the same purpose without processing my personal data?
DP.GDPR.2.1.4	Data subjects are informed about the categories of data	Aggregated values	Before process	Processing of personal data	Data subjects	Understandability: Accountability	Frequency of usage for data categories and attributes	Mandatory	What type of data are used on average by this processing?
DP.GDPR.2.1.5	Data subjects are informed about transfers to third countries	Aggregated (usually general clause)	Before process	Processing of personal data	Data subjects	Understandability: Accountability	- General clause of possibility to transfer to third countries - Applicable safeguards	Mandatory	Are any data transferred to areas where the GDPR does not apply? Are transferred data protected?

C. Explanation Plans

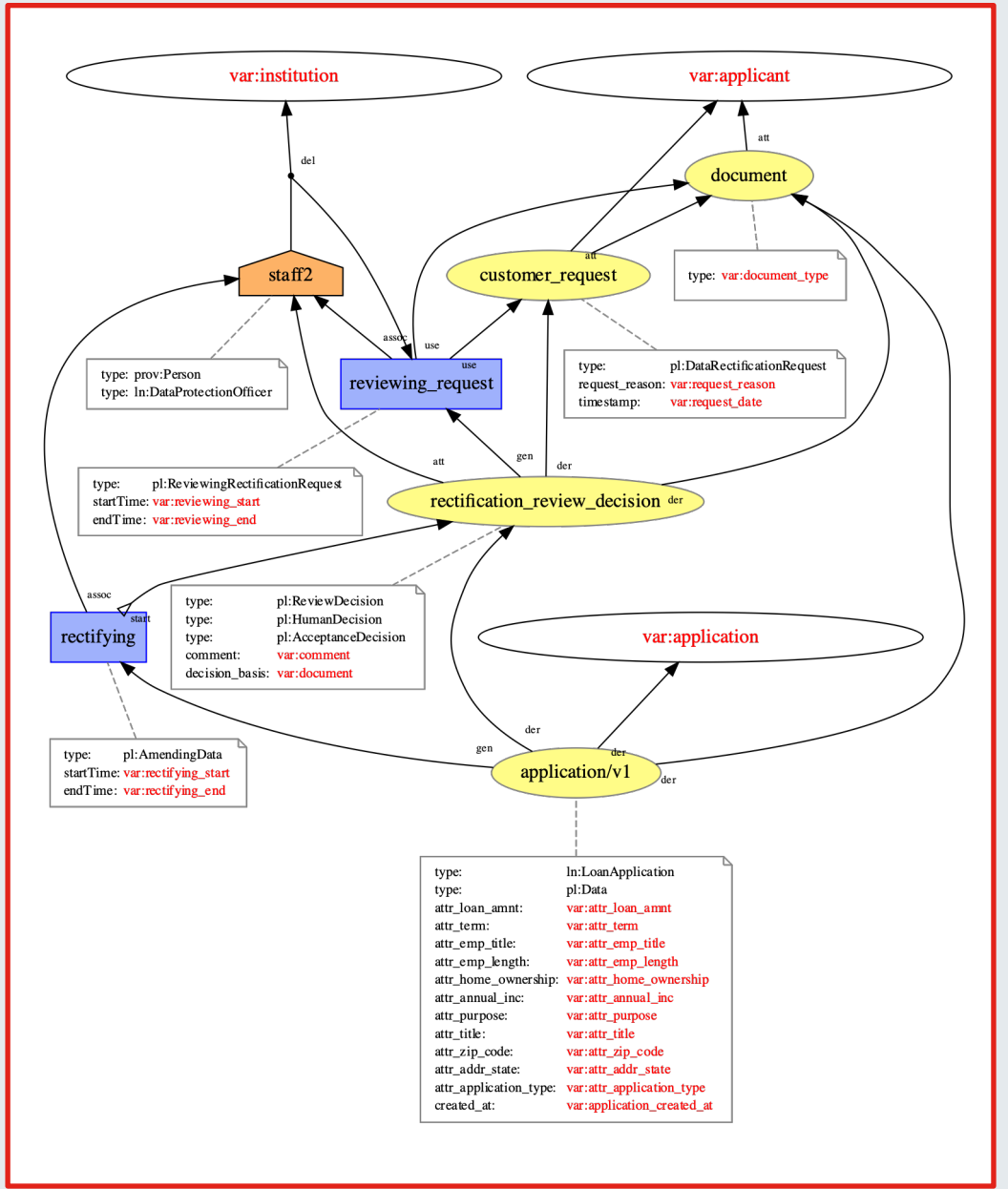


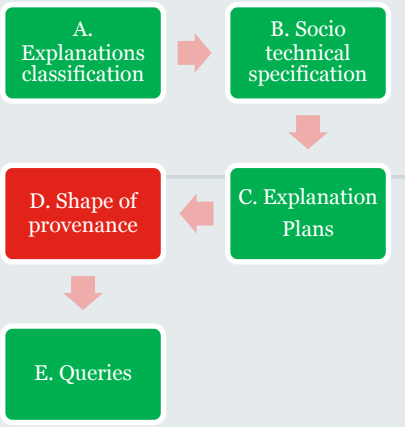
The data subject rectification request *<reference to request>* was reviewed by an agent <reference to staff> who decided to accept it based on the data subject provided reason <subject reason>.

- Includes reference to PROV elements variables
- Grammatical structure
- Surface generation left to NLG library

```
PhraseElement: category=CLAUSE, features={expletive_subject=false verb_phrase={realisation=null, category=
  |-PhraseElement: category=NOUN_PHRASE, features={specifier=WordElement[a:DETERMINER] head_markup_elemen
  |  |-WordElement: base=a, category=DETERMINER, {realisation=null, category=DETERMINER, features={defaul
  |  |-StringElement: content="agent (staff/211)", features={number=SINGULAR, gender=NEUTER, acronym=fals
  |  \-PhraseElement: category=CLAUSE, features={postmodifiers=[{realisation=null, category=PREPOSITIONAL_
  |     \-PhraseElement: category=VERB_PHRASE, features={negated=false perfect=false passive=false head=W
  |        |-WordElement: base=decide, category=VERB, {realisation=null, category=VERB, features={default_
  |        |-PhraseElement: category=NOUN_PHRASE, features={head=WordElement[it:PRONOUN] specifier=WordEl
  |        |  |-WordElement: base=your, category=DETERMINER, {realisation=null, category=DETERMINER, feat
  |        |  \-WordElement: base=it, category=PRONOUN, {realisation=null, category=PRONOUN, features={nu
  |        \-PhraseElement: category=PREPOSITIONAL_PHRASE, features={head=WordElement[based on:PREPOSITIO
  |           |-WordElement: base=based on, category=PREPOSITION, {realisation=null, category=PREPOSITION
  |           \-PhraseElement: category=NOUN_PHRASE, features={head=provided reason specifier={realisatio
  |              |-PhraseElement: category=NOUN_PHRASE, features={head=WordElement[subject:NOUN] specifie
  |              |  |-WordElement: base=the, category=DETERMINER, {realisation=null, category=DETERMINER,
  |              |  |-PhraseElement: category=ADJECTIVE_PHRASE, features={head=data elided=false }
  |              |  |  \-StringElement: content="data", features={discourse_function=PRE_MODIFIER, elided
  |              |  \-WordElement: base=subject, category=NOUN, {realisation=null, category=NOUN, feature
  |              |-StringElement: content="provided reason", features={number=SINGULAR, gender=NEUTER, ac
  |              |-PhraseElement: category=NOUN_PHRASE, features={head=[ number=SINGULAR gender=NEUTER ac
  |              |  \-StringElement: content="[", features={number=SINGULAR, gender=NEUTER, acronym=false
```

D. Shape of Provenance

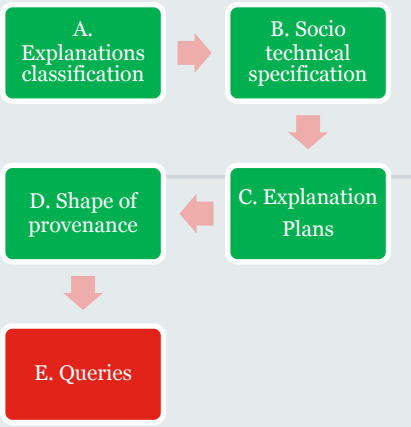




Term	Super Class	same-as	Description
DataSubject	prov:Agent	https://w3id.org/GDPRtEXT#DataSubjec	The data subject is the person with data at the current organisation
DataController			The organisation that controlls personal data
PrivacyNotice			The rules that the organisation apply for the handling any data
Rentation			The amount of time those rules are being used in the organisation
Purpose			The information and the rules included in this privacy notice
Publication			The activity of publishing the privacy notice on the organisation website
Data		https://w3id.org/GDPRtEXT#Data	The data holded on data subject by the current organisation
data_source			The source of the data holded about the data subject, either from another organisations or from the data subject themselves
Attribute			A specific feature/column/section in the data that is holded about the data subject
Rectification			The process of update/fix data attributes based on the data subject request
DecisionRefusal			The process of refusing data subject request, with provided data
RequestErasure			Data subject requesting a data erasure with provided reasons.
Erasure			The process of removing data attributes based on the data subject request
ErasureDecision			The data controller decision on a data subject request of erasure
RequestRestrict			Data subject requesting a ristriction on this data with provided reasons
Restrict			The process of restricting the data acesss based on the data subject request
RestrictDecision			The data controller decision on a data subject request of restriction
ObjectionDecision			The data controller decision on a data subject request of objection of process his data
RequestObjection			Data subject requesting an objection to processing his data
Objection			The process of stopping the use of data subject's data
Staff			

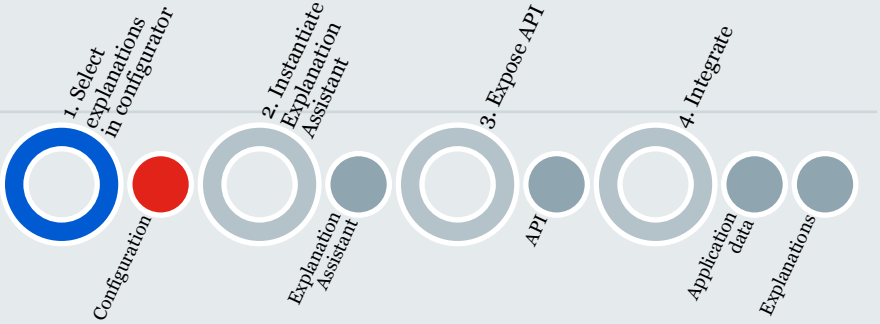
A lightweight vocabulary used to annotate the provenance and exploited by our queries

E. Provenance Queries




```
prefix ln <https://plead-project.org/ns/loan#>
prefix pl <https://plead-project.org/ns/plead#>
select * from waw a prov:WasAssociatedWith
from officer a prov:Agent
 join waw.agent = officer.id
from reviewing a prov:Activity
 join waw.activity=reviewing.id
from der1 a prov:WasDerivedFrom
 join reviewing.id=der1.activity
from decision a prov:Entity
 join der1.generatedEntity=decision.id
from request a prov:Entity
 join der1.usedEntity=request.id
where officer[prov:type]     >= 'prov:Person'
  and decision[prov:type]    >= 'pl:ReviewDecision'
  and decision[prov:type]    >= 'pl:AcceptanceDecision'
  and request[prov:type]     >= 'pl:DataRectificationRequest'
```

1. Configuring the Explanation Assistant



PLEAD Socio-technical Specifications

Search and select supported explanations to configure the Explanation Assistant

Regulations:

GDPR | Equality Act 2010 | Human Rights Act 1998 | Consumer Credit Act (CCA) 1974 | Data Protection Act (DPA) 2018 | Investigatory Powers Act (IPA) 2016

Select the regulations to be supported

Priority Level:

Mandatory | Discretionary

Search

DP.GDPR.2

Data subjects are adequately informed about the processing **Mandatory**

The controller shall take appropriate measures to provide any information referred to in Articles 13 and 14 and any communication under Articles 15 to...

Selected

DP.GDPR.2.2

Data subjects are informed about the retention periods **Mandatory**

he period for which the personal data will be stored, or if that is not possible, the criteria used to determine that period

Selected

DP.GDPR.5.2

Data subject requests the erasure of data **Mandatory**

Selected

DP.GDPR.2.1

Data subjects are informed about the sources of data, the recipients of data, the

DP.GDPR.2.3

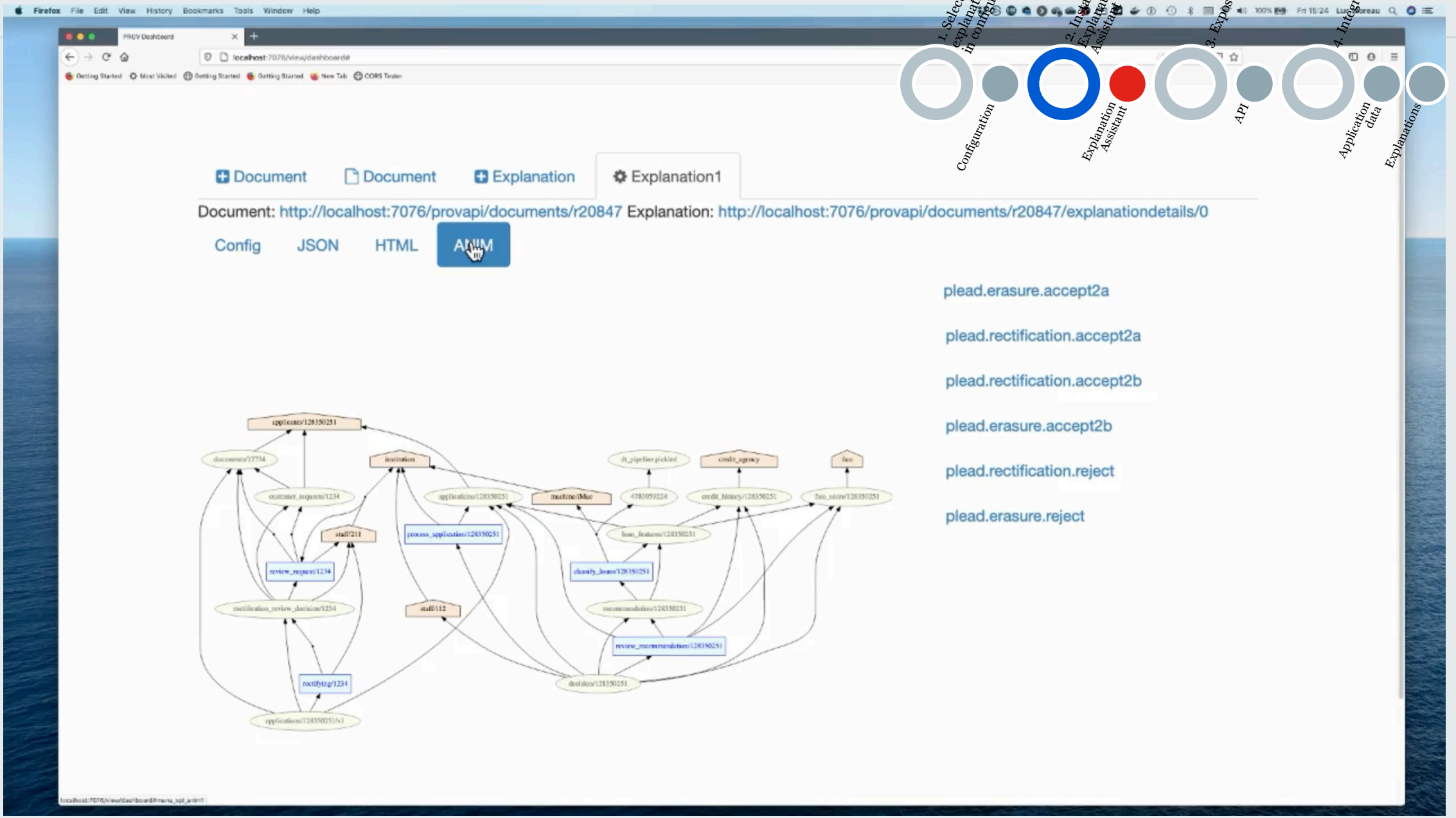
Data subjects are informed about the legitimate interests pursued **Mandatory**

DP.GDPR.5.2.1

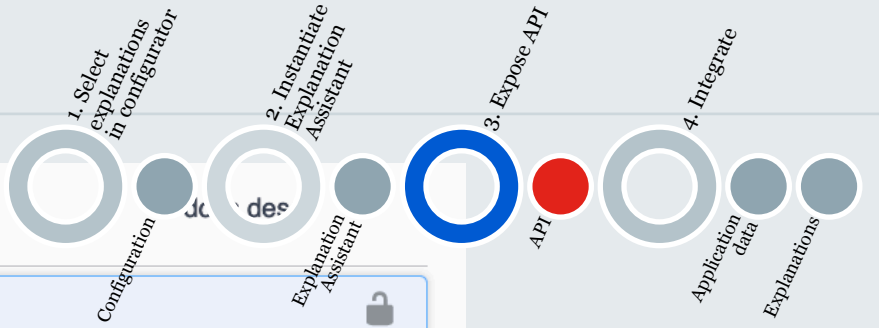
Recipients of data are informed about the erasure **Discretionary**

The controller shall communicate any rectification or erasure of personal data or restriction of processing

2. Instantiate Explanation Assistant



3. Expose API



narrative provenance api (narrative)

GET	`/provapi/documents/{docId}/random_narrative`	Random Narrative of Document
GET	`/provapi/documents/{docId}/explanation/{expid}`	Explanation from Document
GET	`/provapi/documents/{docId}/explanationdetails/{expid}`	Explanation from Document
GET	`/provapi/documents/{docId}/animation.svg`	SVG for animation from Document
GET	`/provapi/documents/{docId}/explanationdetails/{expid}/config`	Configuration of Explanation for Document
GET	`/provapi/documents/{docId}/explanation/{expid}/config`	Configuration of Explanation for Document
POST	`/provapi/documents/{docId}/explanation`	Create explanation from existing Document
POST	`/provapi/documents/{docId}/explanationdetails`	Create explanation, and returns all details for existing Document
GET	`/provapi/documents/{docId}/linear_narrative`	Linear Narrative of Document

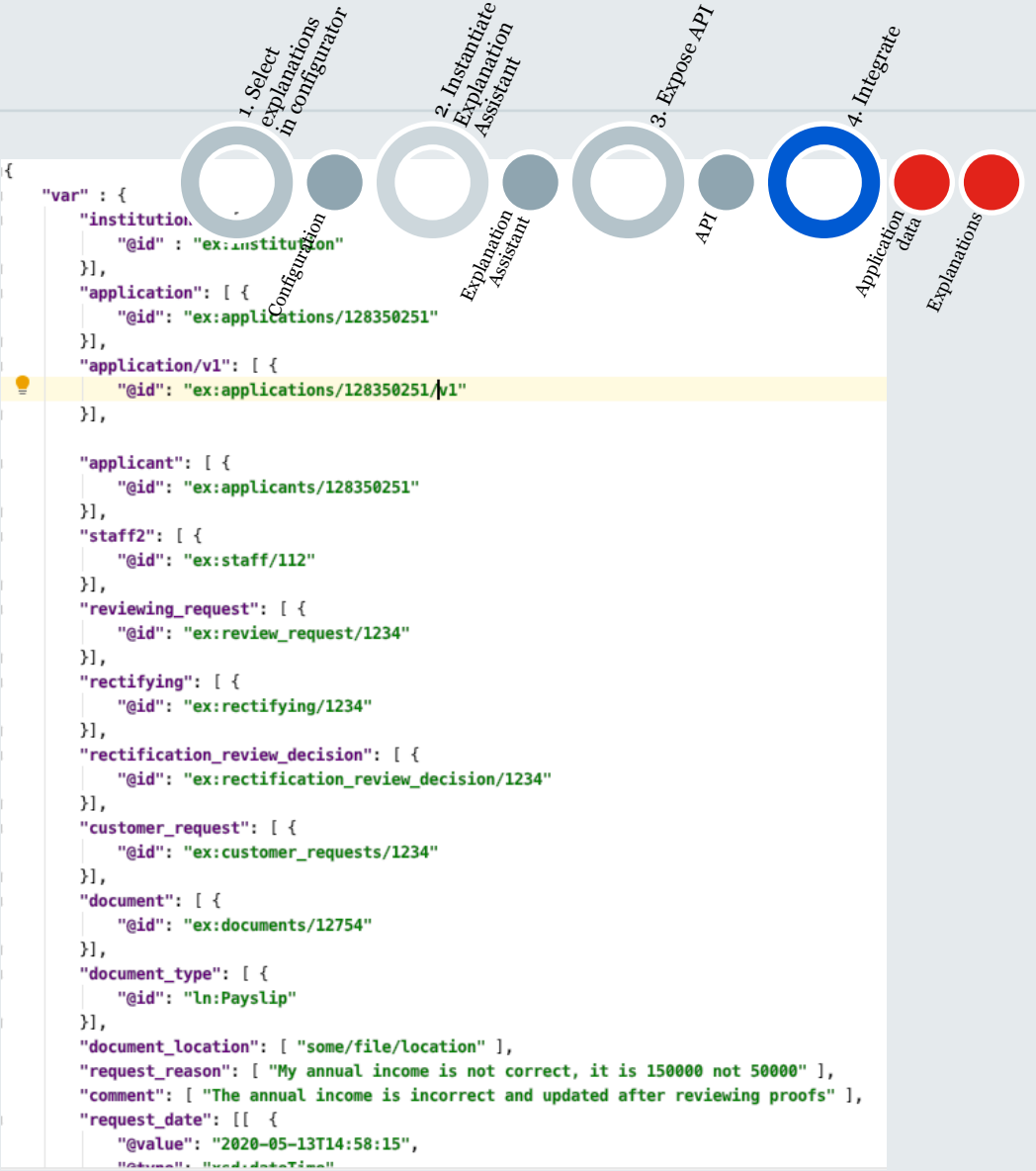
4. Integrate

```
{
  "plead.rectification.accept2a":
"Your rectification request
(customer_requests/1234) was
reviewed by an agent (staff/211) who
decided to accept it based on your
provided reason [ My annual income
is not correct, it is 150000 not
50000 ]." ,

  "plead.rectification.accept2b":
"Your loan application
(applications/128350251/v1) was
corrected on 2019-05-15T14:29:30."

}
```

Explanations as JSON (without markup)



Logged valued to instantiate provenance templates

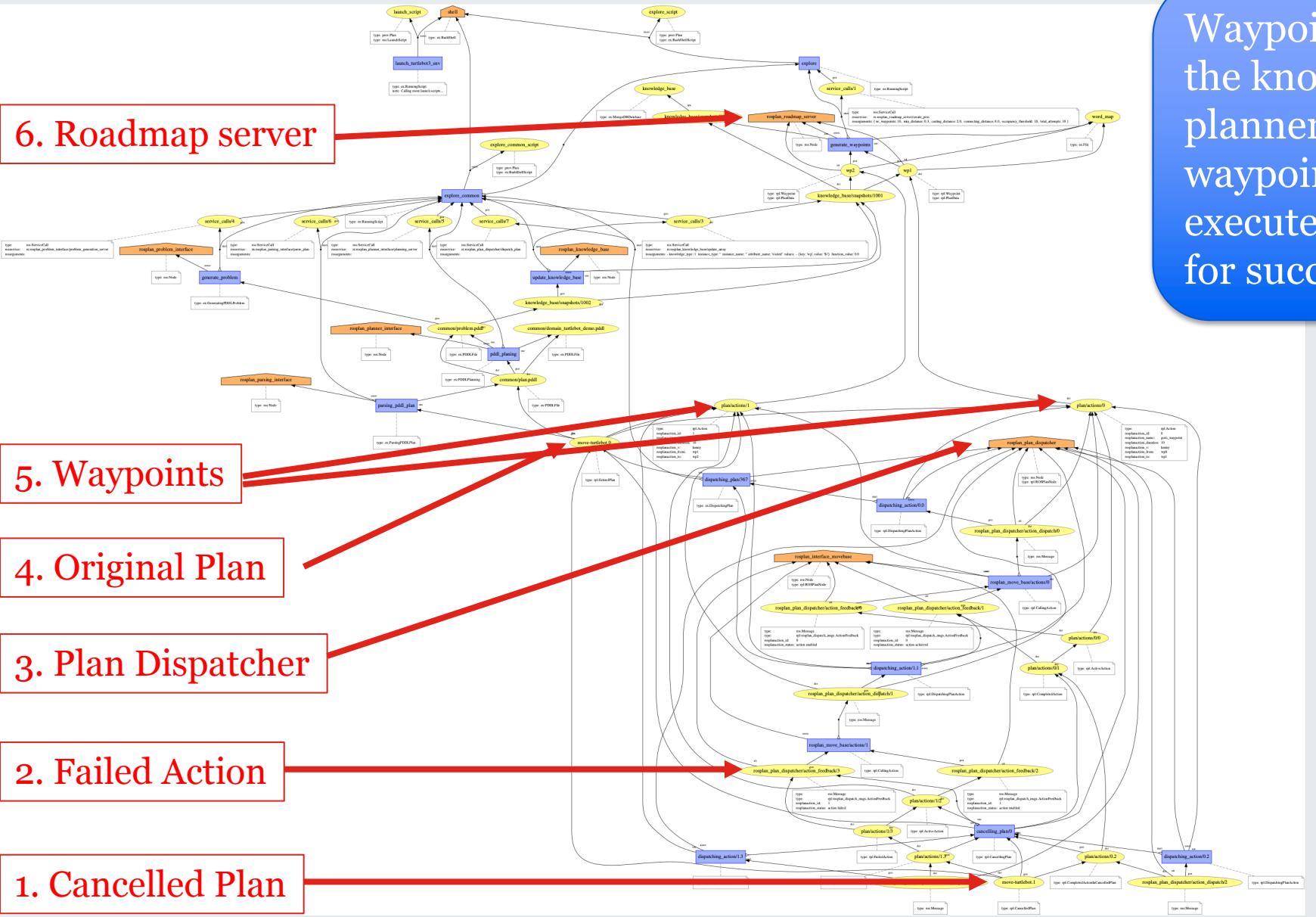
Robotic Use Case

Work with Senka Krivic, Gerard Canal, Dong Huynh, Thump Project

And robots in the mix ...

In an environment where the knowledge base changes frequently, how can we explain the actions of the robot?

Provenance Enabling ROS Plan



Waypoints are generated and stored in the knowledge base of robot. The robot planner creates a plan to visit all the waypoints. Plans actions are issued, executed by the robot and monitored for successful completion.

First Explanations

Plan description

```
"thump.plan.actions1a": "Plan (move-
turtlebot.0) has action
(plan/actions/0) and action
(plan/actions/1)."
```

Statement of failure

```
"thump.plan.failure1a": "The plan
(move-turtlebot.0) which was
dispatched during activity
(dispatching_plan/367) by the
dispatcher (rosplan_plan_dispatcher)
was canceled."
```

Root cause analysis, choice of explanations with increasing details

```
"thump.plan.failure1b": "Plan (move-
turtlebot.1) was failed by action
(plan/actions/1.3)."
```

```
"thump.plan.failure1d": "Plan (move-
turtlebot.1) was failed by action
(plan/actions/1.3) derived from
(plan/actions/1) with waypoint (wp2)
created by ROS Node
(rosplan_roadmap_server)."
```

Conclusions

Conclusion

- PROV
  - National climate assessment, loan assessment, robotic scenario
- PROV
  - Allows for traceability of artifacts, activities, and responsibility
  - Allows for all versions to be managed and searched
  - Allows for user navigation
- PROV
  - Allows for essence of provenance to be found
  - Allows for explanations to be constructed
- User benefits of PROV
  - Increased trust in data and processes
  - Provenance is actionable
    - Navigability
    - Explainability
  - Explanations are meeting a clear purpose (legal or other)

Provenance has become a fundamental data governance tool!

Contact details/for more information
Prof Luc Moreau
King's College London
Bush House, 30 Aldwych, London, WC2B 4BG
+44 (0)20 7848 6808
luc.moreau@kcl.ac.uk
https://nms.kcl.ac.uk/luc.moreau/