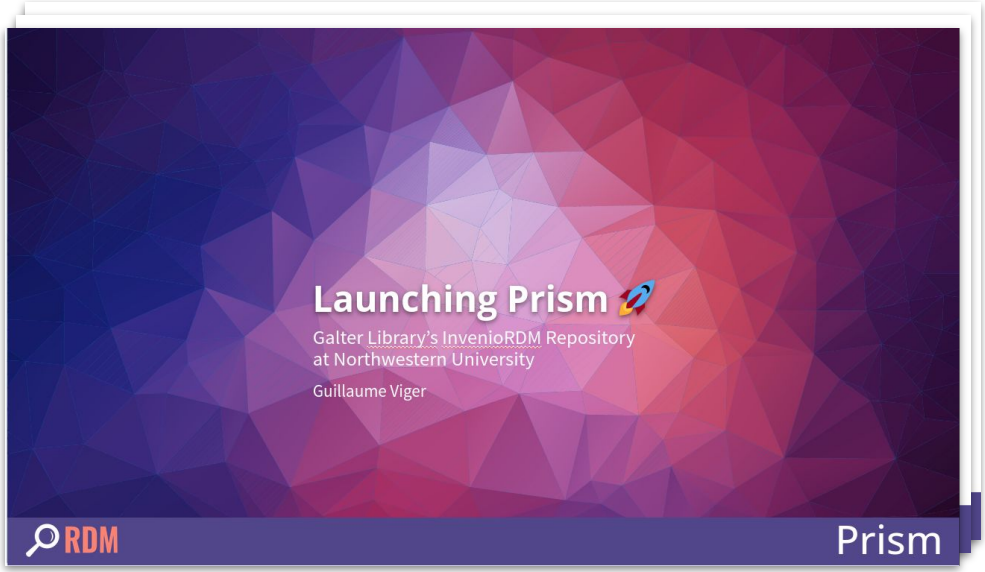
Signposting InvenioRDM 

Guillaume Viger
guillaume.viger@northwestern.edu

RDM

Past Presentation



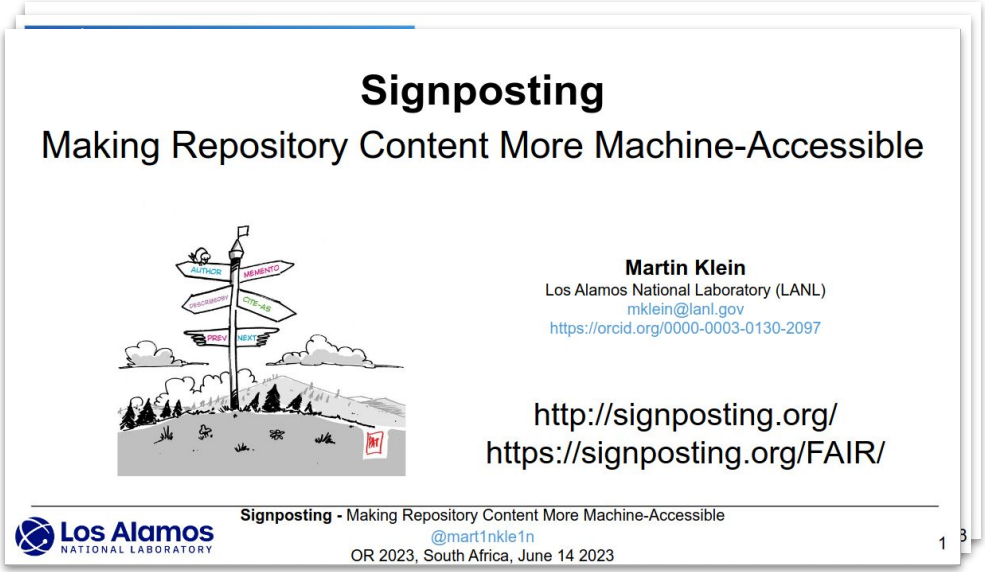
OR2023 Presentation

https://doi.org/10.5281/zenodo.8023551

Signposting

Signposting at OR2023

https://doi.org/10.5281/zenodo.8091605



OR2023 Presentation

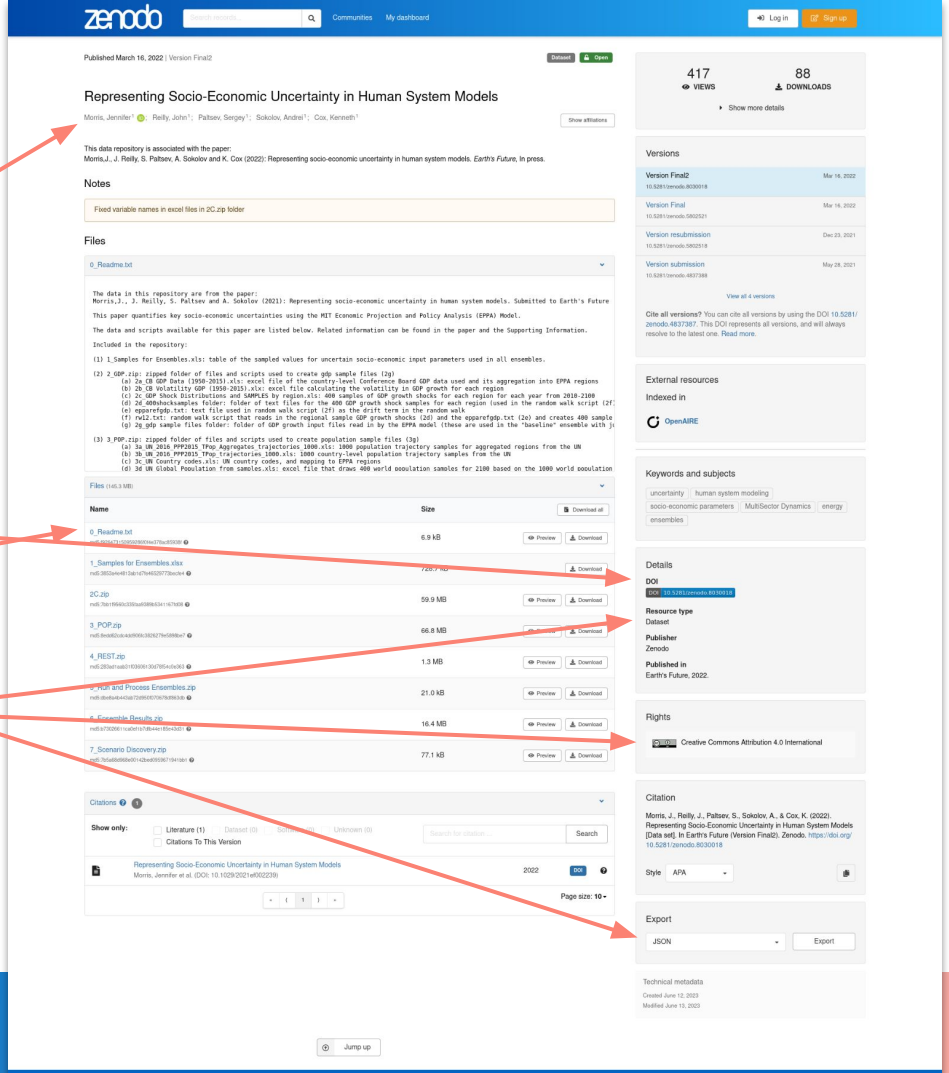
What is Signposting?

Signposting is an interoperable means to structure the discovery of scholarly web resources for machine agents

What problem does it seek to solve?

Humans can identify

- Authors
- DOI
- Formats
- Files
- License
- Type



RDM

What problem does it seek to solve?

Robots ...

RDM

What it looks like concretely

1

```
> GET https://zenodo.org/records/8030018
> ...

< 200 OK
< link: <https://zenodo.org/api/records/8030018> ; rel="linkset"
; type="application/linkset+json"
< ...
```

What it looks like concretely

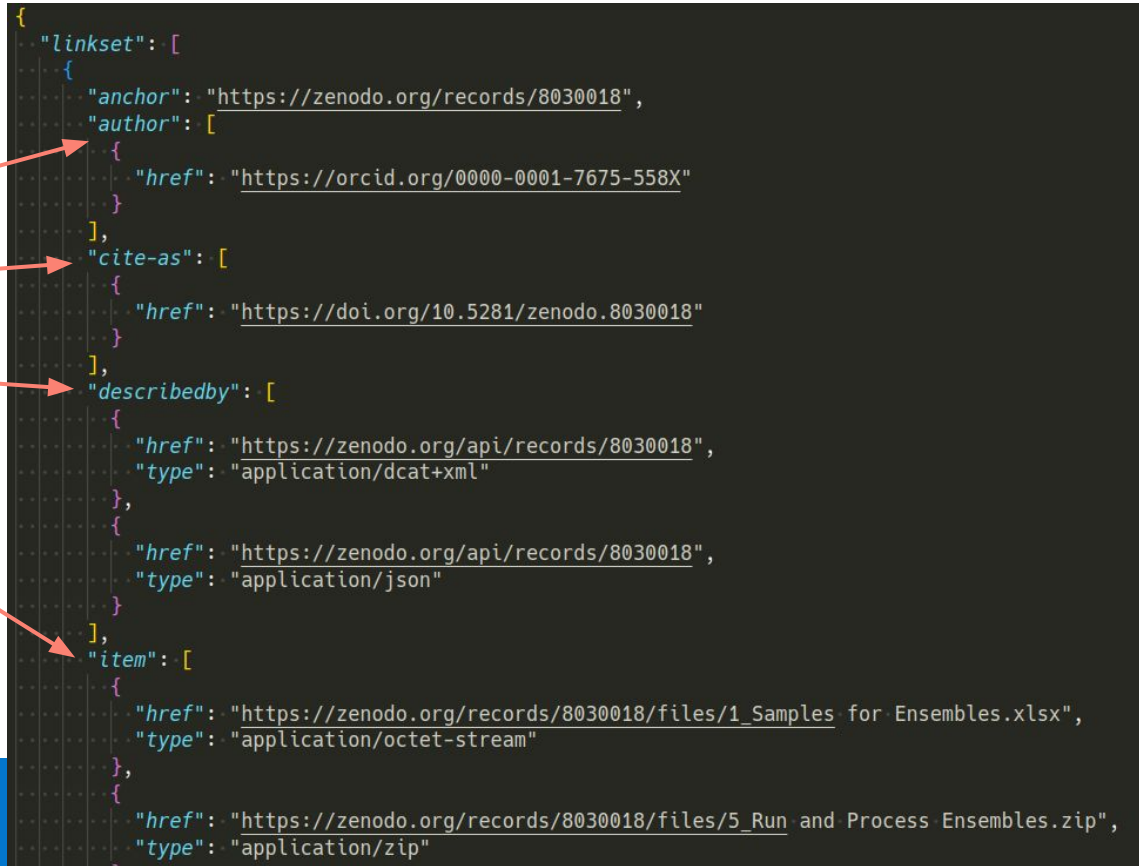
2

```
> GET https://zenodo.org/api/records/8030018
> Accept: application/linkset+json
> ...
< 200 OK
< ...
< {"linkset":[{"anchor":"https://zenodo.org/records/8030018", … }
```

Problem solved

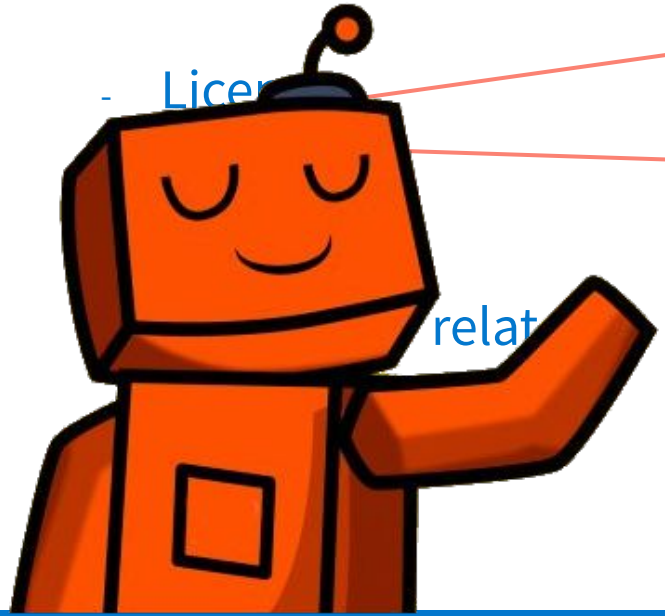
Robots can identify

- Authors
- DOI
- Formats
- Files



RDM

- Licer

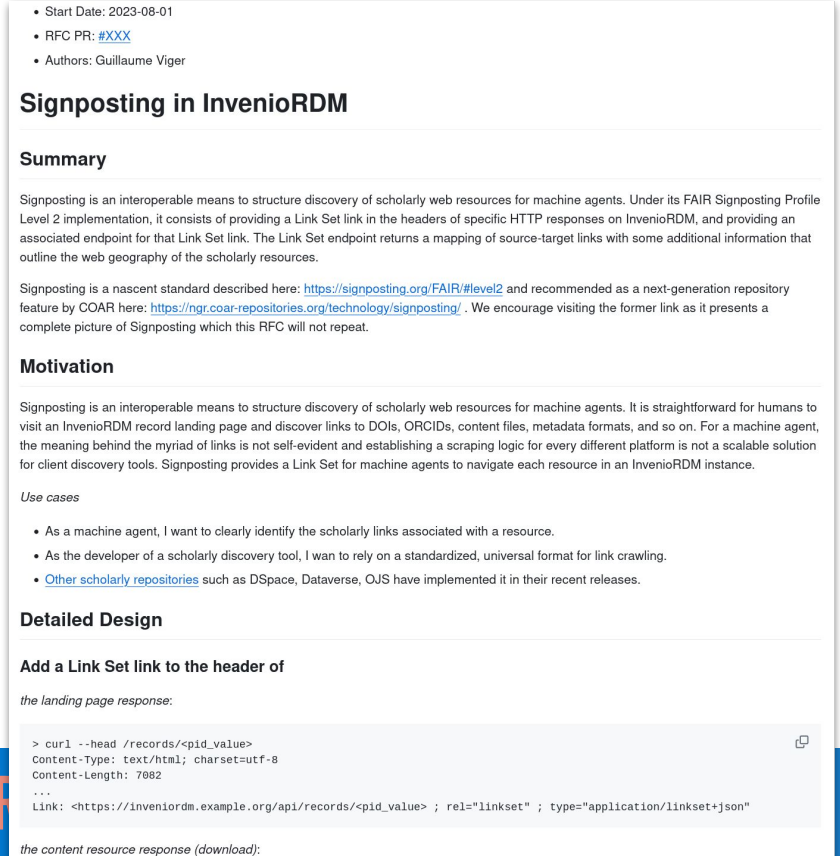


relat

RDM


```
    "item": [
      {
        "href": "https://zenodo.org/records/8030018/files/1_Samples for Ensembles.xlsx",
        "type": "application/octet-stream"
      },
      {
        "href": "https://zenodo.org/records/8030018/files/5_Run and Process Ensembles.zip",
        "type": "application/zip"
      },
      {
        "href": "https://zenodo.org/records/8030018/files/6_Ensemble Results.zip",
        "type": "application/zip"
      }
    ],
    "license": [
      {
        "href": "https://creativecommons.org/licenses/by/4.0/legalcode"
      }
    ],
    "type": [
      {
        "href": "https://schema.org/Dataset"
      },
      {
        "href": "https://schema.org/AboutPage"
      }
    ]
  },
  // From other direction
  {
    "anchor": "https://zenodo.org/records/8030018/files/1_Samples for Ensembles.xlsx",
    "collection": [
      {
        "href": "https://zenodo.org/records/8030018",
        "type": "text/html"
      }
    ]
  }
 ]
}
```

Adoption



- Start Date: 2023-08-01
- RFC PR: #XXX
- Authors: Guillaume Viger

Signposting in InvenioRDM

Summary

Signposting is an interoperable means to structure discovery of scholarly web resources for machine agents. Under its FAIR Signposting Profile Level 2 implementation, it consists of providing a Link Set link in the headers of specific HTTP responses on InvenioRDM, and providing an associated endpoint for that Link Set link. The Link Set endpoint returns a mapping of source-target links with some additional information that outline the web geography of the scholarly resources.

Signposting is a nascent standard described here: https://signposting.org/FAIR/#level2 and recommended as a next-generation repository feature by COAR here: https://ngr.coar-repositories.org/technology/signposting/ . We encourage visiting the former link as it presents a complete picture of Signposting which this RFC will not repeat.

Motivation

Signposting is an interoperable means to structure discovery of scholarly web resources for machine agents. It is straightforward for humans to visit an InvenioRDM record landing page and discover links to DOIs, ORCIDs, content files, metadata formats, and so on. For a machine agent, the meaning behind the myriad of links is not self-evident and establishing a scraping logic for every different platform is not a scalable solution for client discovery tools. Signposting provides a Link Set for machine agents to navigate each resource in an InvenioRDM instance.

Use cases

- As a machine agent, I want to clearly identify the scholarly links associated with a resource.
- As the developer of a scholarly discovery tool, I wan to rely on a standardized, universal format for link crawling.
- Other scholarly repositories such as DSpace, Dataverse, OJS have implemented it in their recent releases.

Detailed Design

Add a Link Set link to the header of

the landing page response:

```
> curl --head /records/<pid_value>
Content-Type: text/html; charset=utf-8
Content-Length: 7082
...
Link: <https://inveniordm.example.org/api/records/<pid_value> ; rel="linkset" ; type="application/linkset+json"
```

the content resource response (download):

- COAR next-generation feature
- DSpace, Dataverse and others have implemented it already
- Straightforward specification
- Independent feature
- Interest from partners institution and me

- Request For Comment
- 2 repos affected

Discoveries

- Specification-wise:
  - Anchor to target for formats needed types
  - Redundancy
  - Potential absence of header link on third-party served files is ok
- Implementation-wise:
  - Uncovered fundamental test misunderstanding leading to significant performance gains
  - Investigate better link generation

There you have it...

- Covered signposting
  - What it is
  - What its motivations are
  - How it's implemented in InvenioRDM
  - Adoption process
  - Discoveries along the way
- Possibilities from there
  - Signposting: Versions and other relationships
  - InvenioRDM: openness to other standards

References

- Klein, M., Walk, P., Shearer, K., & Rodrigues, E. (2023, June 14). Signposting: Making Repository Content More Machine-Accessible. Open Repositories 2023 (OR2023), Stellenbosch, South Africa. Zenodo. https://doi.org/10.5281/zenodo.8091605
- Original robot image: https://www.linkedin.com/posts/player-xp_ai-gamingindustry-gamingcommunity-activity-7151539769691316224--8cm
- Example Zenodo page: https://zenodo.org/records/8030018