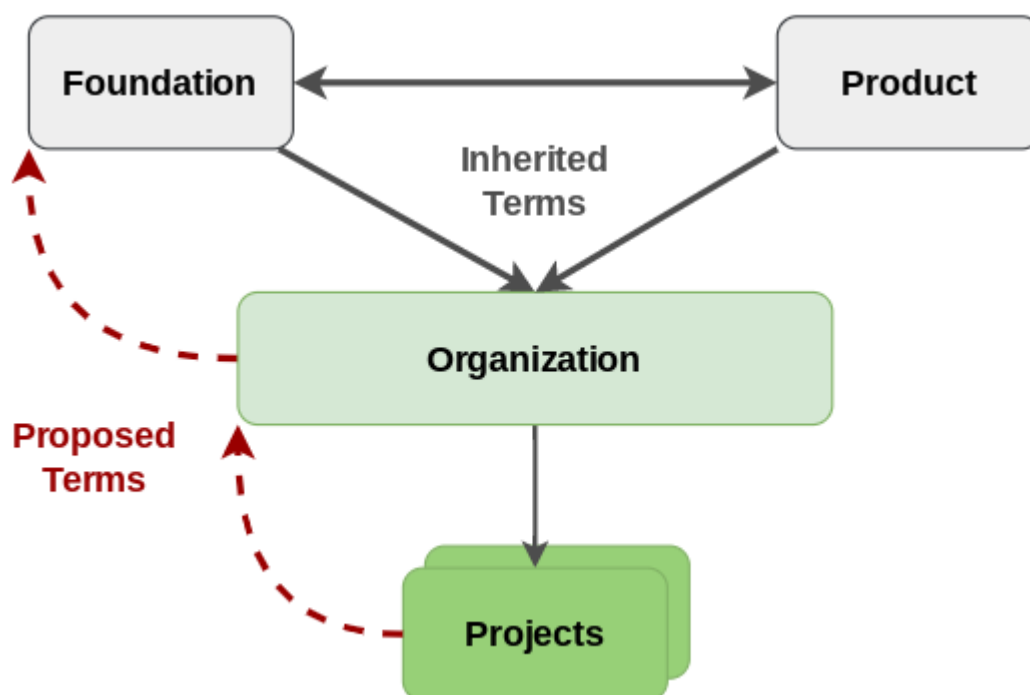


Sourced from:

<https://github.com/thegooddocsproject/website-hugo/tree/glossaries/content/en/docs/glossaries>

Sharable glossaries

The *sharable glossaries project* is developing standards, tools and processes to help establish interoperable glossaries between organizations and projects.



Glossary inheriting terms.

About sharable glossaries

Glossaries are easy to set up for simple examples but very hard to scale - especially when you try to scale across use cases, across domains and across different organizations.

Unfortunately, in 2022, the world is yet to:

1. Agree on a common glossary format.
2. See the format widely adopted.

We aim to change this by breaking down technical barriers, raising awareness of solutions, and supporting widespread adoption of *sharable glossaries*.

Why are glossaries hard?

As projects and their teams grow, documentation becomes increasingly important to enable clear communication. At the same time, it becomes harder to maintain unambiguous terminology due to competing source material and team priorities.

While there are mature standards and supporting tools for publishing and maintaining controlled vocabularies, there is currently a high technical barrier to learning these - something we hope to help simplify.

More details in the [modeling primer](#).

Use cases

Initial high value use cases to target include:

- Auto-generate your glossary by passing your documents against an authoritative source.
- Enable popup descriptions for your project's acronyms.
- Personalize spell checking tooling.

Refer to [glossary use cases](#) for more use cases and more details.

Maturity levels

A glossary should start small, with minimal metadata fields. As the glossary matures, more metadata fields can be added to address advanced use cases.

Level	Value added
0 No glossary	
1 Non-sharable	Help a reader understand locally defined terms.
2 Machine readable	Enables tooling, like domain specific hover-over popups.
3 Traced to source	Know which glossaries terms come from.
4 Governed	Manage glossary maintenance.
5 Extra metadata	Support power use cases.

More details at [sharable glossary maturity levels](#).

Roadmap to sharable glossaries

Some may think it optimistic to nudge the world toward adopting sharable glossaries. We believe it is achievable, within a few years, by: breaking the problem into solvable components, and collaborating with domain experts for each part of the problem.

These components include:

- Establish a development community.
- Establish a *shareable glossary* standard format.
- Publish prominent glossaries as a web service using this standard.
- Scale, by making it super-easy for projects to auto-generate a glossary file for their websites from authoritative glossaries.
- Incrementally add tooling to realize glossary use cases.
- Introduce governance best practices.
- Provide user, and implementation guides.

Refer to our [detailed roadmap](#) and [current status](#) for more details.

Further reading

- [Rules for making a vocabulary Findable Accessible Interoperable and Reusable \(FAIR\)](#).
- [SKOS Primer](#).

Glossary maturity levels

This page proposes maturity levels for *sharable glossary* use cases, along with expected fields for each level's schema profile.

Last updated: January 2022

Version: 0.1

Status: Unreviewed. (Errors expected in the schema formats.)

Maturity levels

A glossary should start small, minimizing barriers to implementation. As the glossary matures, more metadata fields can be added to address advanced use cases.

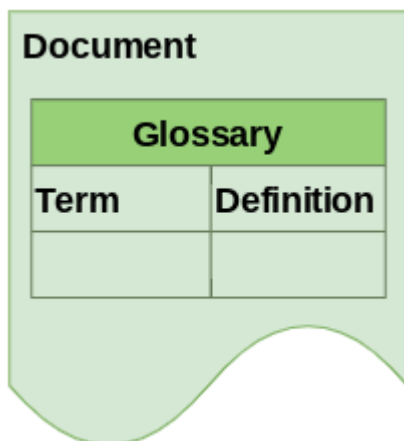
Level	Value added
-------	-------------

0 No glossary	
1 Non-sharable	Help a reader understand locally defined terms.
2 Machine readable	Enables tooling, like domain specific hover-over popups.
3 Traced to source	Know which glossaries terms come from.
4 Governed	Manage glossary maintenance
5 Extra metadata	Support power use cases

Level 1: Non-sharable

What

A non-sharable glossary may be included as a table within a document or website.



Non-sharable glossary.

Value added

A glossary is relatively straightforward to establish for a document or small website. It help readers understand terms used.

How

Copy domain-specific terms from your documentation into a glossary table. Include columns for Term and Definition. Ideally, source definitions from a more authoritative glossary.

Example

Term	Definition
mercury	metal element which is liquid standard temperatures

Level 2: Machine readable and reusable

What

- Store terms in a standard, machine readable format. Typically the glossary is stored as a static file, accessible via a URL, such as: <https://mywebsite.org/glossary.json>.



Glossary stored in a machine readable .json file in website.

Value added

Adding machine readability enables multiple domain specific use cases, such as:

- Hover-over popups for terms within a document.
- Downstream glossaries referencing terms within this glossary.

How

- Level 2 of the *shared glossary profile* involves copying the terms and definitions from your Level 1 table into a glossary file. (Fields from higher maturity levels may additionally be included as needed.)

Fields for level 2

- skos:Collection # Glossary's top level, covering a collection of terms.
- skos:Concept # A term to describe
- skos:prefLabel@en # The preferred term name
- skos:prefLabel@fr # Optional: Other languages can be supported
- skos:altLabel # Optional: Store acronyms and synonyms
- skos:definition # Textual definition

Fields from higher maturity levels may optionally be included.

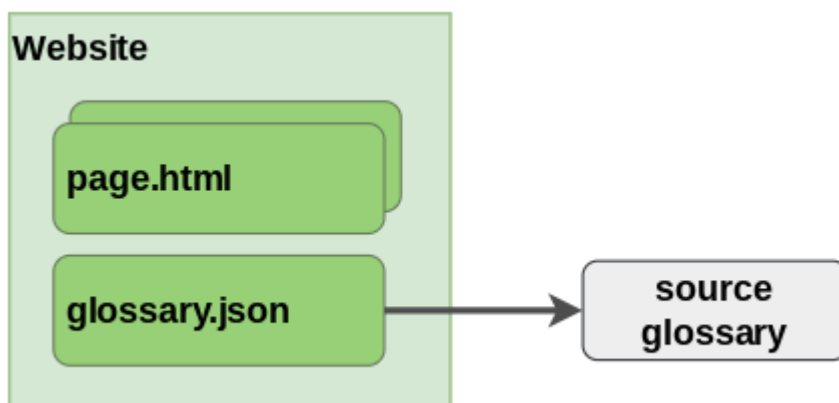
Example

```
{
  ex:chemistry rdf:type skos:Collection : {
    ex:mercury rdf:type skos:Concept : {
      skos:prefLabel "mercury"@en;
      skos:prefLabel "mercure"@fr; # Optionally support multiple
languages
      skos:altLabel "he"@en; # altLabel is optional
      skos:definition "metal element which is liquid standard
temperatures"@en;
    }
  }
}
```

Level 3 Reference source glossaries

What

- Source terms from upstream glossaries and retain hyperlink to the source.
- Assign a license which allows repurposing of the glossary by downstream users, such as [Creative Commons By Attribution \(CC-BY\)](#).



Glossary terms referencing source.

Value added

Adding a reference to the source glossary enables:

- Increased interoperability.
- Suitable attribution provided to address license conditions.
- Term definitions can be refreshed from updated source glossaries.
- A glossary's credibility will benefit from referencing a more authoritative source.

Extra fields for level 3

- `dcterms:license`
- `skos:inScheme`
- `dcterms:identifier`

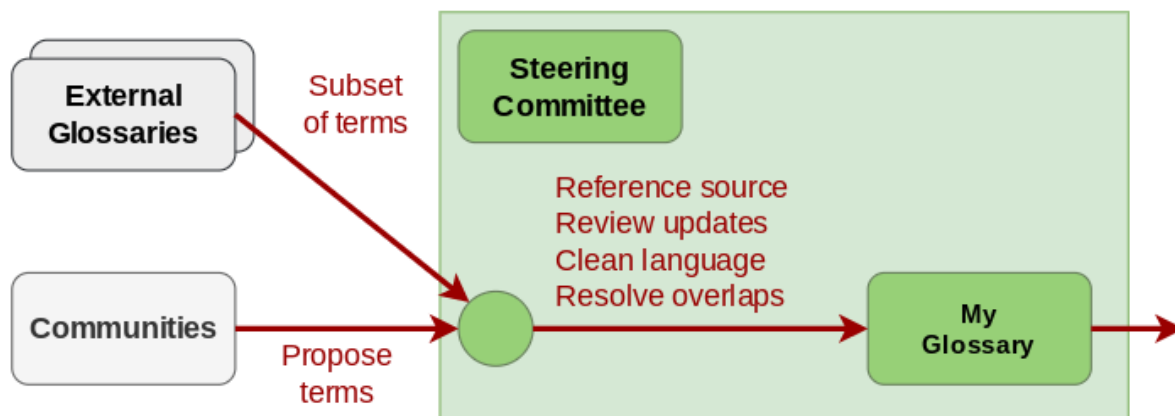
Example

```
{
  ex:chemistry rdf:type skos:Collection : {
    dcterms:license "https://creativecommons.org/licenses/by/4.0/";
    ex:mercury rdf:type skos:Concept : {
      skos:prefLabel "mercury"@en;
      skos:prefLabel "helio"@es; # Optionally support multiple
languages
      skos:altLabel "he"@en;
      skos:definition "metal element which is liquid standard
temperatures"@en;
      dcterms:identifier "mercury"; #unique identifier
      skos:inScheme "https://authoritative1.org/glossary.json"
      # If term doesn't have a source, then skos:inScheme isn't
mentioned
    }
  }
}
```

Level 4 Governed and versioned

What

- Apply governance processes to manage glossary updates.
- Identify a glossary custodian
- Apply a unique identifier for each term.
- Apply versioning to terms (`skos:Concepts`).
- Periodically release a baselined version of the glossary (`skos:Collection`).
- Publish glossary license information.



Possible glossary governance process.

Value added

- Managed quality control processes.
- Planned release cycles.
- Reduced accidental inconsistencies.
- Maintainers can recommend updates to source glossaries.

Extra fields for level 4

- `ex:collectionVersion`
- `ex:sourceGlossaryVersion`
- `ex:sourceConceptVersion`

Example

```

{
  ex:chemistry rdf:type skos:Collection : {
    dcterms:license "https://creativecommons.org/licenses/by/4.0/";
    ex:collectionVersion rdf:type owl:versionInfo "1.2.1";
    ex:priorVersion rdf:type owl:priorVersion "1.2.0";
    ex:mercury rdf:type skos:Concept : {
      skos:prefLabel "mercury"@en;
      skos:altLabel "he"@en;
      skos:definition "inert gas"@en;
      dcterms:identifier "mercury"; #unique identifier within Collection
      skos:inScheme "https://authoritative1.org/glossary.json"
      ex:sourceGlossaryVersion rdf:type owl:versionInfo "2.0";
      ex:sourceConceptVersion rdf:type owl:versionInfo "3";
    }
  }
}
  
```

Level 5 Extra metadata

More extensive use cases may involve adding extra metadata fields to the `ConceptCollection` and `Concepts`.

As at January 2022, this section requires further definition.

Possible extra fields for level 5 skos:note # Comments about the term.
skos:scopeNote # Clarifications on usage. dcterms:creator dcterms:created
dcterms:modified dcterms:source dcterms:replaces rdfs:seeAlso skos:broader
skos:narrower skos:related skos:broadMatch skos:closeMatch skos:exactMatch
skos:narrowMatch skos:relatedMatch skos:ConceptScheme

Further reading

- [Rules for making a vocabulary Findable Accessible Interoperable and Reusable \(FAIR\)](#)

Sharable glossaries roadmap

This roadmap describes steps planned by open source collaborators to realize the potential for large-scale sharing of glossary terms between disparate organizations and projects.

It is designed to help collaborators understand how they can join and benefit from the initiative.

How can we bootstrap sharable glossaries?

How can we push through existing hurdles and bootstrap a *sharable glossary* framework, and achieve widespread adoption? It will involve a multi-phased approach, addressing key challenges which have hindered prior attempts:

- Establish a core *sharable glossary* format, with help from domain modelers from ISO and OGC standards communities.
- Publish prominent glossaries as a web service, starting initially with committed stakeholders from within the geospatial domain.
- Scale, by making it super-easy for projects to auto-generate a glossary file for their websites from authoritative glossaries.
- Incrementally add tooling to realize high value glossary use cases.
- Help projects mature by introducing glossary management processes.

Implementation phases

Phase 1 Define problem and establish a community

Challenge

Define the problem and attract collaborators.

How

- Within this website, we describe the business, architecture, collaboration and technical challenges that require addressing.
- We have broken problems into modular components, are engaging domain experts to focus on their area, and coordinating contributions into a holistic solution.

Phase 2 Establish a defacto glossary standard

Challenge

Defining a glossary format is illusively difficult.

- It needs to be simple and easily understood by the masses for common use cases.
- It needs to expand to support edge use cases, including referencing source terms.
- It should be interoperable with existing Knowledge Organization Systems (KOS).

How

- We are working with domain modeling experts, from ISO and OGC standards organizations to define a _sharable glossary _schema format.
- This will extend to establishing a defacto web service API.
- We need to reconcile different sentence structure guidance for writing glossary definitions.

Further information

- [Rules for making a vocabulary Findable Accessible Interoperable Reusable \(FAIR\)](#).
- Schema standards: [SKOS](#), [Dublin Core Terms](#), [DCAT](#), [JSON-LD](#) and [SHACL](#).
- [SKOS Primer](#).

Phase 3 Publish prominent glossaries

Challenge

To bootstrap a sharable glossary network, we want to stand up large, authoritative glossaries behind a web service using the *common glossary* format. This will enable downstream projects to build derived, project-specific glossaries.

How

The authors of the open source [Paneron](#) glossary have committed to publishing glossaries via the *common glossary* format, once developed.

This software is used by large glossary publishers, such as:

- The 180,000+ entries in the [IEC Electropedia](#), the world's authoritative electrotechnical terminology-set.
- The glossary of spatial terms from the Open Geospatial Consortium and ISO TC211.

We will reach out to other glossary publishers and encourage their participation. If you own a glossary you'd like to publish, then please connect with us.

Further information

- [Proposal for a glossary publishing pilot at larger organizations.](#)

Phase 4 Scale: a glossary for every website

Goal

Provide a compelling reason for all technical websites to create a glossary for their website, and make it very easy to do so.

Challenge

Building and maintaining a non-trivial glossary is difficult. It typically requires rallying and continued engagement from many stakeholders. However, if we have a sharable glossary framework, and access to an authoritative glossary, we can search our website for glossary terms.

Example use of a script to build a glossary:

```
build-glossary
```

```
--source https://my-standards.org/glossary \
```

```
--source https://my-domain.org/glossary \
```

```
--website https://my-website.com \
```

```
--out my-glossary.json
```

For the simple use case, the created glossary need only be stored as a file in the root directory of the website and viewed as a table, or as in-line popup help.



Glossary stored in a machine readable .json file in website.

How

- We aim to attract a developer to create a *build-glossary* tool, and introduce a means to publish the glossary within a website. Initially the tool could be a simple command-line script. Over time, it will likely gain more features and grow into a web service.
- These will be accompanied by how-to guides, written by technical writers from The Good Docs Project.
- Initially we will pilot this approach with ~ 50 open source geospatial projects we have a relationship with within The Open Source Geospatial Foundation. Later we will evangelize to technical writers and website builders through forums such as Write the Docs and The Good Docs Project.

Phase 5 Tooling to address high value use cases

Once a sharable glossary framework is established, and a critical mass of websites starts to build, there will be a compelling business case to integrate tools with glossaries, such as:

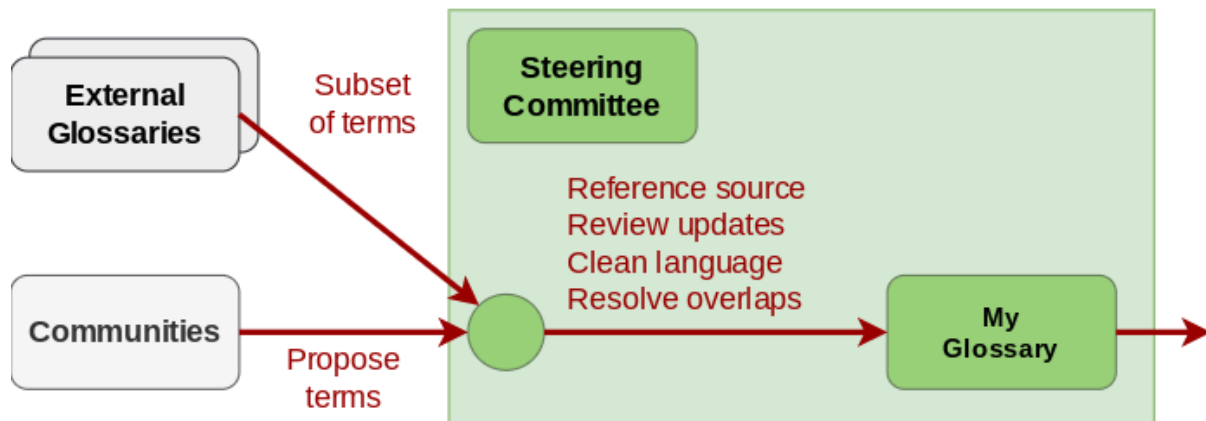
- Web-hosting platforms including hover-over popups for terms.
- Spell and grammar check tools to source glossary terms.
- More accurate translation applied to translation tools.
- Smarter search algorithms.

Phase 6 Glossary management processes and tooling

As projects mature, glossary owners will adopt tools and processes to support the maintenance, versioning, and cross-domain management of terms.

This will typically involve moving glossary terms from a glossary file into a specialized hosting tool such as the [Paneron](#) open source glossary management tool.

This will be supported by documented best practice processes.



Possible glossary governance process.

Sharable glossaries status

We kicked off the sharable glossary project in mid-2020, and by January 2022 we have:

- Early pilot implementations of core technology, with more work needed.
- An experienced team working on multiple aspects of the project.

Goal	Status at January 2022
Define problem	Within this website, we have defined the problem and suggested solutions.
Establish a development community.	Assembled a core team, meeting fortnightly. Team has decades of collective experience from: <ul style="list-style-type: none"> * Data modeling in ISO and Open Geospatial Standards organizations. * Documentation communities. * Google tech writers. * Glossary and open source software projects. * Requirements management community.

Establish a <i>shareable glossary</i> standard schema.	Early version developed. Still to be published, tested and refined. Target pilot by 1st half 2022.
Establish a sentence structure for glossary definitions.	Piloted using ISO's definition conventions . Expect we will need to extend this guidance to address conventions used in other glossaries. Target 2022.
Build glossary software to publish the <i>shareable glossary format</i> .	Open source implementation developed, but needs ruggedizing for mass market, and needs to integrate the shareable glossary format. Target pilot by 2022.
Pilot implementations	Have established a pilot cross-project glossary within the geospatial domain, with contributors from: * The Open Geospatial Consortium (OGC), responsible for spatial standards. * ISO TC211, responsible for ISO's geospatial glossary. * The Open Geospatial Foundation, and multiple open source geospatial projects. Target refinements in 2022/23.
Publish prominent glossaries as a web service.	Have established relationships with some prominent glossary holders. Waiting on software to publish the <i>shareable glossary standard</i> . Target 2022.
Scale, by making it super-easy for projects to auto-generate a glossary file for their websites from authoritative glossaries.	Not started. Target 2022/23.
Incrementally add tooling to realize high value use cases.	Not started, but there are some quick-win tools which will be easy to implement for a software developer. Target 2022-24.

Introduce governance best practices.	Understood, but barely started. Target 2022/23.
Provide user, and implementation guides.	Barely started. Depends on implementation to stabilize. Target 2022/23.

Prior status reports:

- [2020-12](#)
- [2020-08](#)

Sharable glossary use cases

This page lists the various use cases enabled by glossaries.

Discover term meaning

- As a general document reader, I want to find definitions for the terms and acronyms in the document I am reading.
 - Traditionally, glossaries have been presented as a table or list within the associated document or website.
 - Some systems support readers hovering over terms to get a popup with more information.
- As a technical writer, I want to find the preferred spelling, capitalization and word choice for a term.
- As an advanced document reader, I want to know of alternate definitions for a term.
- As an advanced document reader, I want to understand the inheritance path back to upstream source definitions, where I may find more information.

Translation

- As a document translator, I want glossary terms to be translated into my target languages, so I can consistently translate a source term to the same target term.

Information mining

- As a search engine or software algorithm building knowledge graphs, I want to use glossaries to help extract meaning from textual information sources.

- As a researcher, I want to be able to find related information even if it uses different terms for the same concepts.

Governance

- As a glossary owner, I want a governance framework to help resolve terminology management conflicts between terminology sources and stakeholders.
- As a glossary owner, I want access to version history to allow predictable baseline updates and understand historical updates.
- As a downstream glossary owner, I want source glossary terms licensed such that I can republish the terms, (such as Creative Commons By Attribution).
- As a glossary owner, I want tooling to build and maintain my glossary by matching terms used in my documentation with source glossaries.
- As a term maintainer, I want tooling which allows me to suggest new terms or update existing terms.

Note: A glossary might be targeting terms from a document, a website, a project, an organization, a domain, or a foundation.

Note: Sources for glossaries can be both:

- *More authoritative (such as from upstream standards), and*
- *Less authoritative (such as from downstream community users).*

Tooling

- As a software developer, I want terms and relationships between glossaries in a machine-readable form so that I can integrate glossary functionality into software.
- As a data modeler, I want to align the terms I use with others in my domain so that we can seamlessly integrate our data models.

Modeling Primer

This page includes background information about data modeling, to help future implementers.

Version: 0.1

Status: Incomplete and unreviewed. (This page still needs to be fleshed out.)

Vocabulary formats

Glossaries are a specific instance of Knowledge Organization Systems (KOS). These vocabulary formats provide models for:

- Thesauri
- Classification schemes
- Taxonomies
- Folksonomies, and
- Other controlled vocabularies.

There are mature standards and supporting tools for publishing and maintaining controlled vocabularies.

- [Simple Knowledge Organization System \(SKOS\)](#)
- [Resource Description Framework \(RDF\)](#)
- [Dublin Core Terms](#)
- [Data Catalog Vocabulary \(DCAT\)](#)
- [JSON for Linking Data \(JSON-LD\)](#)
- [Shapes Constraint Language \(SHACL\)](#)
- [Web Ontology Language \(OWL\)](#)
- [schema.org](#)

Our *sharable glossary* schema aggregates fields from these standards.

More details at:

- [Sharable glossary maturity levels.](#)
- [Rules for making a vocabulary Findable Accessible Interoperable and Reusable \(FAIR\).](#)
- [SKOS Primer.](#)

Target Audience

Potential developers / sponsors / collaborators

- Wants to assess whether to engage:
 - Wants to know vision
 - Wants to understand feasibility
 - Wants to know timeframe
 - Wants to know whether to invest
- Wants to know roadmap and how to engage
- Wants to know about pilot implementation

Glossary owner

- Wants to know how they should publish their glossary
- Wants governance framework for updating the glossary

- Wants to customize governance rules
- Wants to know metadata to collect
- Wants to know maturity levels

Tech writer wanting tools to establish a glossary

- Wants to know current state of glossaries
- Wants to know what tools to use
- Wants to know best practice for writing existing glossary

General writer

- Want to know best practice for structuring

Sources

- [Kick off manifesto](#)
- [Halfway status](#)
- [Sharable glossaries roadmap](#)
- Ron's [Proposal for a glossary publishing pilot at larger organizations](#)
- [osglex_doc_n5_Procedures_for_drafting_terminology_entries_2020-09-07](#)
- [How to build a cross-domain glossary](#)
- [Glossary use cases](#)
- [Glossary Pilot Doc Deliverables](#)
- [Glossary maturity levels](#)
- [Glossary Governance Template](#)
- [Defining the Glossary Project \(Ankita's version\)](#)
- [Research Paper on Glossary Schemas](#)
-

Website layout

- Index.html
 - Vision
 - Sharable glossaries
 - Key use cases
 - Status
 - TOC
- Theory
 - Background
 - Easy to understand, hard to solve
 - Use cases
 - Maturity levels
 - Governance
- Roadmap (for implementation)
- Design
 - Metadata
 - Status
 - Description

- Who's working on this
-
- Data Schema
 - Schema challenges
 - Implementation status
- Flat file format
- API
- Web Server
- Governance
- Guides
 - Status of guides (lack of)

CONTRIBUTING

Contact

Website content