|  |  |  |
| --- | --- | --- |
| **Title** | Embedded metadata for landing pages |  |
| **Version** | 1 |
| **Author(s)** | Dieter Van Uytvanck |
| **Date** | 2019-11-05 |
| **Status** | Draft |
| **Distribution** | SCCTC, NCF |
| **ID** | CE-2019-1544 |

# Introduction

In September 2018, Google announced its [Dataset Search](https://toolbox.google.com/datasetsearch). In the subsequent NCF meeting, it was discussed whether or not to support this. The SCCTC was asked for technical advice on how a technical connection to this search engine, by embedding metadata in landing pages, would relate to the CLARIN infrastructure. This document tries to answer that question.

# Suggested way forward

* The **choice** whether to embed metadata into landing pages should be with the **individual CLARIN centres**. If a centre wants to enable this, it can easily do so by following the instructions at <https://developers.google.com/search/docs/data-types/dataset>
* **Metadata catalogues**, such as the VLO, **should not embed metadata into landing pages**, since doing so might be conflicting with the centre’s choice.

# Relation to existing metadata infrastructure

The possibility to embed metadata in landing pages does not influence the requirement for CLARIN B- and C-centres to provide an OAI-PMH endpoint and CMDI (for B-centres) or at least Dublin Core (for C-centres) metadata.

# Background information

* [RDA Research Metadata Schemas WG](https://www.rd-alliance.org/groups/research-metadata-schemas-wg) – an RDA working group on the topic of metadata schemas for embedding in web page. Both technically informative and also a place where heated discussions take place, with important players from the research metadata community.
* Vocabularies supported for metadata embedding:
  + [Schema.org](https://en.wikipedia.org/wiki/Schema.org) (initiated by search engines, simpler, less extensible)
  + [DCAT](https://en.wikipedia.org/wiki/Data_Catalog_Vocabulary) (academically inspired, more complex and extensible)
* [Draft mapping between schema.org, DCAT and WikiData](https://www.wikidata.org/wiki/Wikidata:WikiProject_Datasets/Data_Structure/DCAT_-_Wikidata_-_Schema.org_mapping)