**Title**: OpenSearch and JSON-LD for enhanced Earth observation data and service discovery

**Presenter/Organization**: Ingo Simonis, OGC

**Type**: long preferred, but short possible as well

**Abstract**:

Discovery of Earth observation (EO) data is an art that requires efforts from several sides to achieve satisfying performance for customers looking for specific data. Data needs to be annotated on the provider side, interfaces need to support simple yet powerful filtering and access mechanisms, and users need to speak the language of the data in order to discover what they are looking for. Recent developments now try to simplify in particular the situation for the data explorer, the user looking for data. These developments use landing pages that provide an entry point into the whole dataset, links from this landing page to the various resources offered by the server, and standardized models for the essential metadata items.

Discovery in the Earth observation domain has featured a different approach in the past. The focus was on maximizing the information provided to the data explorer to allow fine-granular searches in massive datasets using powerful filter languages. This focus led to information models that make sense in their dedicated namespaces and for selected communities only. Several specifications have been developed in this context by ISO and OGC, such as e.g. ISO 19115 Geographic Information: Metadata with its XML serializations and extensions defined in ISO 19139 and ISO 19139-2 to describe geospatial data, and catalog interface specifications with their filtering languages to register and explore metadata entries.

Though proven to be powerful, these approaches cause several challenges for standard search engines. As search engines need to accommodate a broad range of domains, domain specific solutions are usually not or only rudimentarily supported. For that reason, the geospatial community is now exploring approaches that are based on Web APIs with landing pages that provide all required information to explore all resources offered at a specific interface, feature OpenSearch formats and JSON encodings, and try to use domain-independent metadata models extensively. The goal is to integrate as much domain-independent technology and information models and to enrich only where necessary. The goal is to establish a similar user experience for Earth observation data discovery as for website discovery.

This presentation describes the results of OGC's latest discovery experiment: "Earth Observation Process and Application Discovery". The experiment explores the potential of GeoJSON(-LD) metadata encodings combined with OpenSearch formats to enhance the discovery of Earth observation data, processes, and processing services. The initiative features a holistic approach, focusing not only on data discovery, but extends to questions such as "which application can produce the data I need?", “which data is supported by which application?”, “which applications can I chain together, given their data input/output capabilities?” By defining a Web-API landing page and corresponding link structure, the initiative explores the level of search-engine friendliness without losing too much discovery capabilities.

The initiative has started exploring the potential of the metadata encodings defined in OGC document 17-084,the GeoJSON(-LD) metadata encoding for EO collections. OGC 17-084 re-uses GeoJSON and OWS Context properties but is designed towards ease of use. It minimizes redesign of properties by re-using properties from several existing namespaces, including Data Catalog Vocabulary (dcat), Friends of a Friend (foaf), Location Core Vocabulary (locn), PROV Data Model (prov), Resource Description Format (rdf), Simple Knowledge Organization System (skos), vCard Ontology (vcard), and several others. This metadata side is complemented with OpenSearch response models as specified in OGC document 17-047, the OpenSearch-EO GeoJSON(-LD) Response Encoding Standard. The goal is to use technologies, encodings, and information models that are supported by current search engines to the maximal extent in order to provide a sophisticated user experience to the data and service customer.