WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

WMO WIS 2.0 Discovery Metadata exchange, harvesting and search pilot project

Tom Kralidis
Meteorological Service of Canada

WMO OMM
World Meteorological Organization
Organisation météorologique mondiale

Project objectives

- This project aims to experiment implementing WMO discovery metadata as DCAT using the OGC API - Records draft standard. This project will also experiment actionable linkages with demonstration project 1 (AMQP/MQTT), search/access of collections of variables of NWP data, as well as enabling search capability against WIS 2.0 topics

WMO OMM

Project team

Name	Why / Role	Country
Task Team on WIS Metadata (TT-WISMD)		
Mr Tom KRALIDIS	Chair	Canada
Mr Guillaume AUBERT	Member (RA VI)	EUMETSAT
Mr Xinqiang HAN	Member (RA II)	China
Ms Hanane KAMIL	Member (RA I)	Morocco
Mr Jan OSUSKY	Member (RA VI)	HMEI
Ms Julia SIELAND	Member (RA VI)	Germany
Mr Ioannis MALLAS	Member (RA VI)	ECMWF
Mr Steve OLSON	Member (RA IV)	USA
WMO Secretariat		
Ms Anna MILAN		
Mr Enrico FUCILE		

Name	Role	Country
Jeremy Tandy	Chair ET-W2AT	UK
Chris Holmes	OGC (external consultation)	USA
Douglas Fils	Consortium for Ocean Leadership (external consultation)	USA
OGC MetOcean Domain Working Group	Meteorological data discovery APIs (discussion, standards alignment)	

WMO OMM

Project plan

- Collaboration
  - SC-IMT
  - ET-Metadata/TT-WISMD
  - OGC MetOcean Domain Working Group
  - Ocean Info Hub
  - Open Geospatial Consortium / STAC

Milestone	Delivery Date
metadata design (types, crosswalk)	2021-03-31
demonstration (harvesting, search)	2021-09-30
final report	2021-10-31

WMO OMM

STANDARDS AND WEB SERVICES LANDSCAPE / ECOSYSTEM EVOLUTION

WMO OMM

First Generation Web Services and Standards

- XML-RPC/CORBA
- SOAP/WSDL/UDDI
- Service Oriented Architecture (SOA)
- Strong concept of RDBMS backend
- OGC first generation services (WMS, WFS, WCS, WPS, CSW, etc.)
- XML Payloads

WMO OMM

First Generation Web Services Realities

- XML Payloads
  - Heavy for web/mobile applications
- Overloading of native HTTP functionality / tunneling
  - GET /api?request=GetRecordById&id=5
  - GET /api?action=search&query=sea+ice
- Not using native HTTP status codes
  - Returning a 200 for an error/exception

WMO OMM

First Generation Standards Realities

- Building for the 100% use case
  - Given the 80/20 rule, the last 20% is usually difficult
- Data: Deep, complex content models
- APIs: overloaded methods, not “of the web”
- See OGC GML

WMO OMM

First Generation Services and Standards Realities

- Lack of mass market integration
  - Complex machinery/architecture for services to be indexed/crawlable
  - Challenging for web developers
  - Challenging for mass market search

WMO OMM

CURRENT STATE WEB SERVICES AND STANDARDS

WMO OMM

REST/JSON/OpenAPI

- REST
  - HTTP verbs
    - GET /collections/foo/items/5
    - GET /collections/foo/items?q=sea+ice
  - HTTP status codes (200, 201, 400, 404, etc.)
  - URIs to identify resources
  - Content negotiation (media types)
  - Stateless
- OpenAPI
  - How to describe a REST API
  - Endpoints, methods, request parameters, responses
- JSON (JavaScript Object Notation)
  - Small payload (no closing tags/keys)
  - Very popular for web/mobile application data handling
  - JSON Schema
  - Same data structure / constructs in popular programming languages (Python, JavaScript, etc.)
  - Rich tooling / ecosystem
  - JSON is a foundational building block
    - GeoJSON (RFC 7946)
    - OGC API
    - STAC

WMO OMM

Standards Evolution

- 2016: GeoJSON
- 2017: W3C Data on the Web Best Practices
  - https://www.w3.org/TR/dwbp
- 2017: W3C Spatial Data on the Web Best Practices
  - https://www.w3.org/TR/sdw-bp
- 2017: OGC API Whitepaper
- 2018: STAC
- 2018: OGC API development (https://ogcapi.org)
  - https://ogcapi.ogc.org/apiroadmap.html

- Being webby/of the web
- Specifications developed on the Web
  - GitHub (issues/discussion/collaboration)
  - https://github.com/opengeospatial
  - https://github.com/w3c
- Standards developed in AsciiDoc and released as HTML (first) and PDF

WMO OMM

METADATA RECORDS

WMO OMM

The Record Model

- OGC API – Records, Metadata Record Schema
  - https://github.com/opengeospatial/ogcapi-records/blob/master/core/openapi/schemas/recordGeoJSON.yaml
- Example (MSC NWP Global Model, 15km)
  - https://github.com/OGCMetOceanDWG/ogcapi-records-metocean-bp/blob/master/core/examples/msc.gdps.json

WMO OMM

Metadata Granularity Examples

- datasetcollection: NWP model (OARec record metadata)
  - dataset: NWP model output: air temperature (OARec record metadata)
  - product options
    - API endpoint to interrogate the data
    - x/y/z/t (granule) (STAC Item with link to actual data asset)
- datasetcollection: surface weather observations (OARec record metadata)
  - dataset: air temperature (OARec record metadata)
  - product options
    - API endpoint to interrogate the data
    - x/y/z/t (granule) (STAC Item, with link to actual data asset, link to WIGOS metadata)
- datasetsetcollection: METAR
  - dataset: air temperature (OARec record metadata)
  - product options
    - API endpoint to interrogate the data
    - product: single message (granule) (STAC Item with link to actual data asset)

WMO OMM

WIS 2.0 DISCOVERY, HARVEST, AND SEARCH

WMO OMM

OGC API – Records

- OGC API effort/successor to CSW specification
- Extends OGC API – Features
- HTML, JSON
- Core record model (Dublin Core/DCAT/CSW 3)
  - GeoJSON dialect
- Can be implemented as standalone catalogue or via API
- STAC: focus on EO, lower level metadata/granules
- Public RFC: Q1 2022

WMO OMM

WMO Search Updates: WIS 2.0

- 1950s: data exchange via WMO Global Telecommunications System (GTS)
- 1970s: Manual on GTS
- 2007: WMO Information System (WIS)
  - metadata and catalogue atop GTS
- Today:
  - Earth System Monitoring and Prediction
  - Big data
  - Cloud
- WIS 2.0
  - Simple data exchange
  - Open Standards
  - APIs
  - PubSub
  - Cloud

WMO

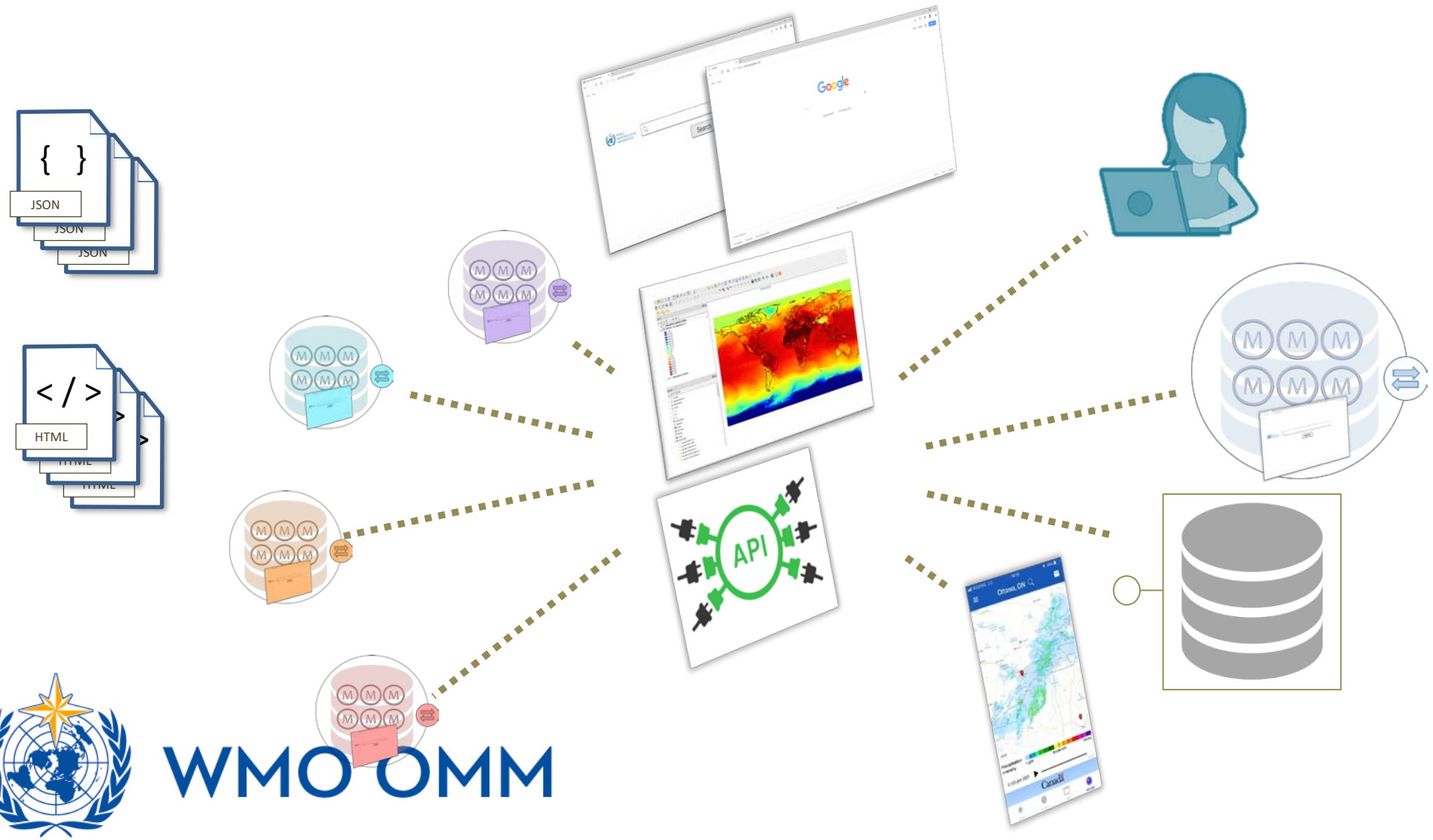
WMO OMM

User stories

- NWP centre operator (find/bind/analyze/integrate/publish)
- Forecaster (find/analyze)
- Start-up (find)
- Developer (find/bind)
- Casual user (find)

WMO OMM

WIS 2.0 Architecture



WMO

WMO OMM

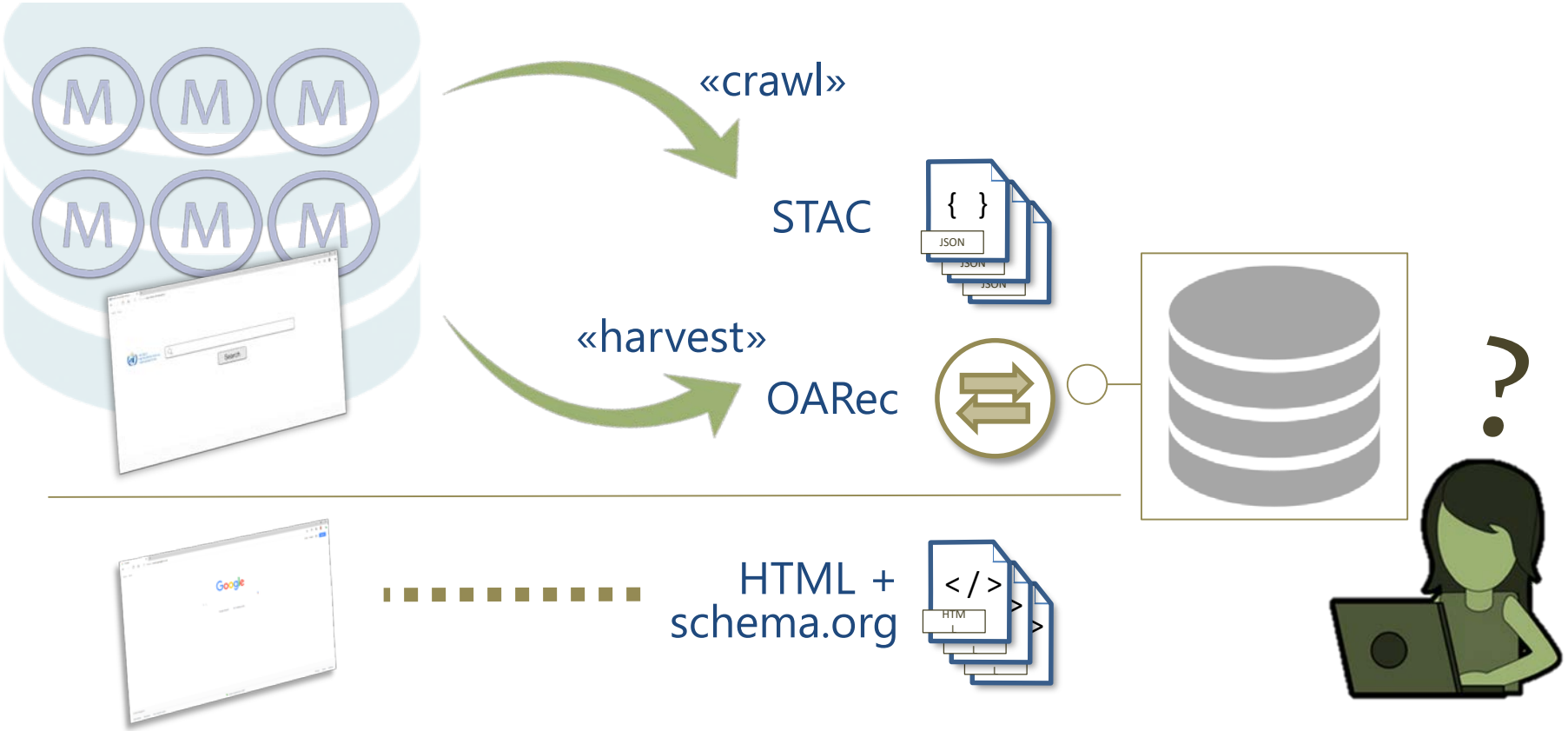
Data publication mechanisms

Simple files: GRIB, BUFR, NetCDF, (COG, Zarr?)

Web API: OGC-API, OpenAPI, (AsyncAPI)

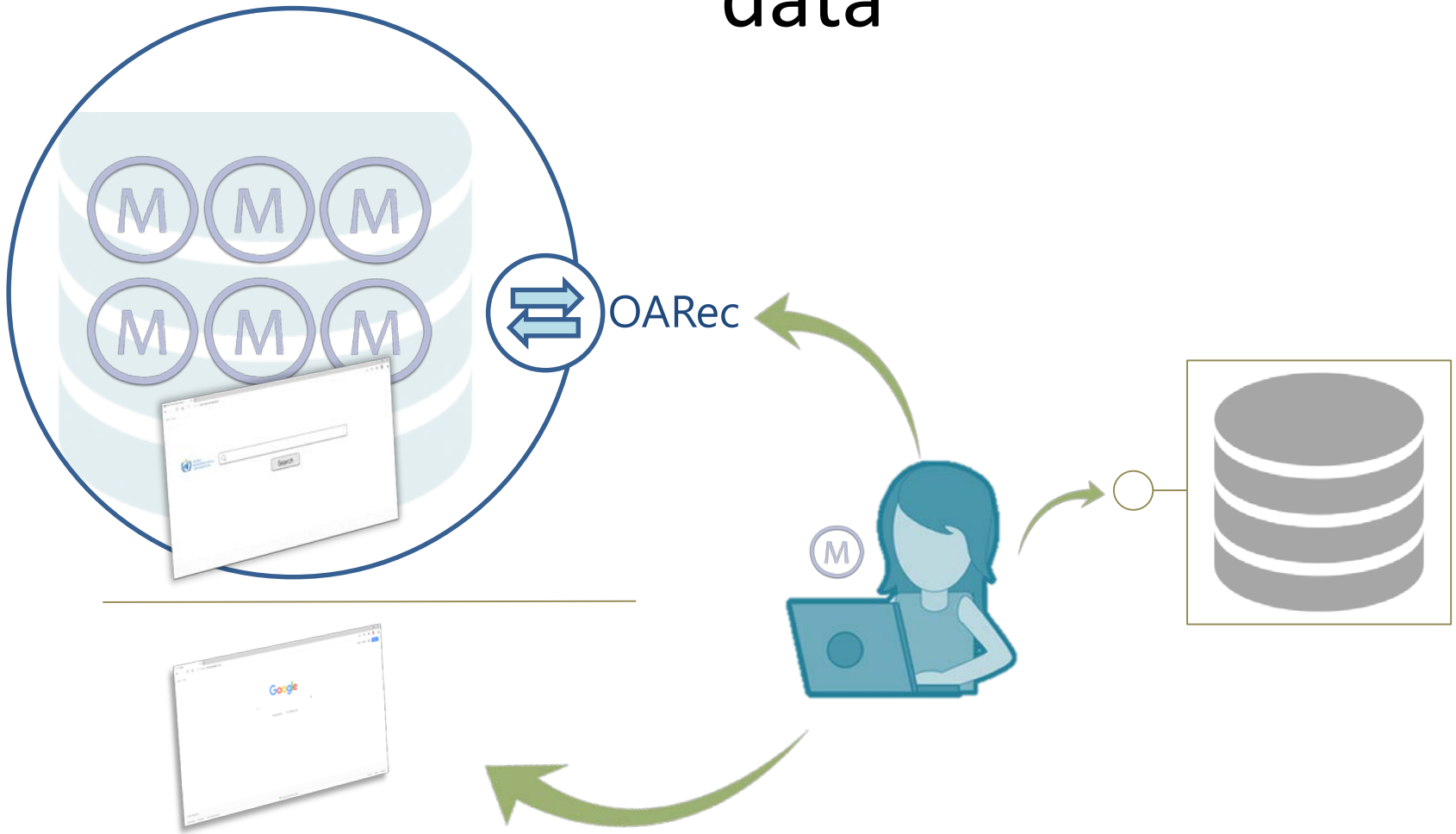
WMO OMM

How do I publish metadata?



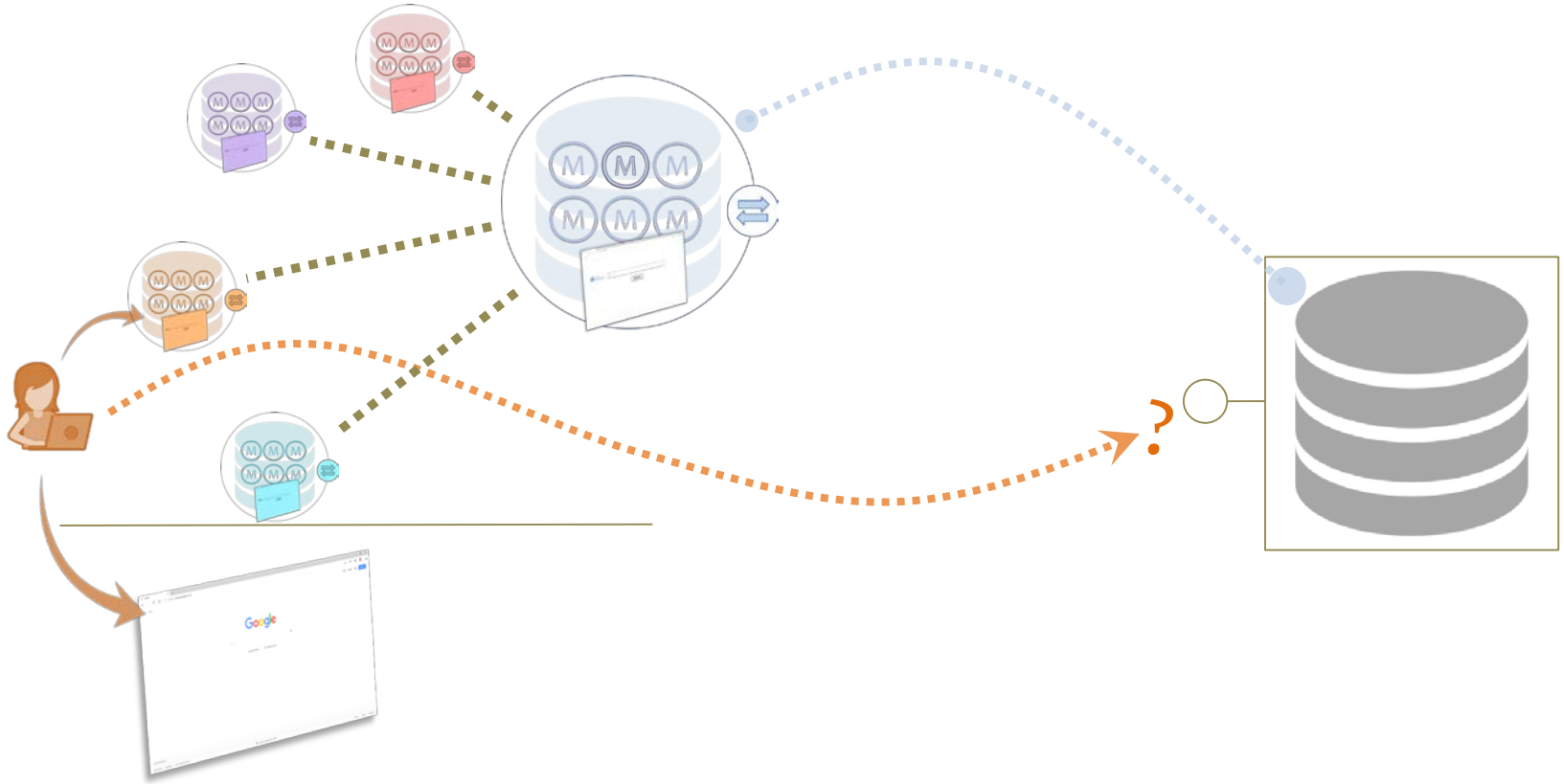
WMO OMM

Searching the WIS Catalogue, finding data



WMO OMM

Aggregation vs. Distributed Search?



WMO OMM

Benefits: Broad Interoperability

- It’s just JSON; will work with any JSON tooling
- It’s GeoJSON; will work with any GIS tooling!
- Interoperability with Google and mass market search
- JSON-LD/schema.org

WMO OMM

Community Standardization Work

- WMO Expert Team on Metadata Standards
  - Task Team on WIS Metadata
  - WCMP 2.0
- OGC MetOceanDWG
  - MetOcean Best Practices for OGC API – Records
    - Extension to OGC API - Records

WMO OMM

WIS 2 Principles in the project

- 1 *WIS 2.0: adopts Web technologies and leverages industry best practices and open standards*

- HTTP, RESTful design patterns, as well as the evolving OGC API suite of standards

WMO OMM

WIS 2 Principles in the project

- 2 *WIS 2.0: uses Uniform Resource Locators (URL) to identify resources (i.e. Web pages, data, metadata, APIs)*

- Link relations in support of the hypermedia

WMO OMM

WIS 2 Principles in the project

- 3 *WIS 2.0: prioritizes use of public telecommunications networks (i.e. the Internet) when publishing digital resources*

- Leveraging the Web as the platform

WMO OMM

WIS 2 Principles in the project

- 4 *WIS 2.0: requires provision of Web service(s) to access or interact with digital resources (e.g. data, information, products) published using WIS*

- Discovery as a web service (API) as well as hypermedia controls to related actionable services/APIs
- **“Web services” mean either static resources or API machinery**

WMO OMM

WIS 2 Principles in the project

- 6 *WIS 2.0: will add open standard messaging protocols that use the publish-subscribe message pattern to the list of data exchange mechanisms approved for use within WIS and GTS*

- Actionable hypermedia controls to protocols and services put forth in Demonstration project 1 (Exploring the use of message querying protocols for GTS data exchange)

WMO OMM

WIS 2 Principles in the project

- 10 *WIS 2.0: will provide a catalogue containing metadata that describes both data and the service(s) provided to access that data*

- OGC API - Records as an approach for cataloguing WIS metadata for data and services

WMO OMM

WIS 2 Principles in the project

- 11 *WIS 2.0: encourages data providers to publish metadata describing their data and Web services in a way that can be indexed by commercial search engines*

- Enabling WIS metadata for SEO and mass market interoperability

WMO OMM

Project data standards

- Agnostic: the core focus of the pilot is on discovery metadata, providing a gateway to data
  - GRIB2, BUFR/CREX
  - GeoTIFF/CF-NetCDF
  - OGC Web Services/APIs
    - Maps (PNG/JPEG)
    - Coverages (GeoTIFF, CF-NetCDF)
    - Features (GeoJSON, CSV)

WMO OMM

Project metadata standards

Functions	Standard(s)
• Catalogue/discovery • Discovery metadata management (collections) • Harvesting/federation • Reporting	• OGC API – Records (OARec)
• metadata representation of various granularities (collections, items/assets)	• OARec record schema • SpatioTemporal Asset Catalog (STAC) • Schema.org/JSON-LD
• Content models baselined with GeoJSON	

WMO OMM

Data discovery

- Discovery is key to this pilot project
- The existence of a resource (data, services, processes, etc.)
- Filtering capabilities
  - Spatial (bbox=-152,42,-52,84)
  - Temporal (i.e. datetime=2000-11-11/2001-11-11)
  - Aspatial (title=foo)
  - Freetext (q=sea ice)
- Query capabilities
  - Sorting (sortby=-title,description)
  - Paging (limit=0&startindex=1000)
- Content negotiation
  - Schema (schema=iso19139)
  - Format (f=json, f=html)

WMO OMM

Data exchange

- Data exchange is facilitated indirectly from enabling users to “bind” to actionable links

WMO OMM

Input to WIS2

- Build out a topic hierarchies with Project 1
  - Expressed in metadata
  - Queryable via OARec API
- Build out a data identification scheme / granularity
- Metadata provisioning via a basic catalogue or API provisioning
  - OGC API – Records
  - STAC Items
- Key performance indicators
- Documentation / cookbooks for onboarding

WMO OMM

Resources / Outputs

- WIS 2.0 pilot report (in progress)
  - https://github.com/wmo-im/wis2-metadata-search
- OGC API – Records – Part 1: Core **draft**
  - http://docs.ogc.org/DRAFTS/20-004.html
- pygeoapi OGC API Python Server
  - https://pygeoapi.io

- OGC API – Records clients
  - OWSLib: https://geopython.github.io/OWSLib
  - QGIS MetaSearch: https://docs.qgis.org/latest/en/docs/user_manual/plugins/core_plugins/plugins_metasearch.html

WMO OMM

Demo

WMO OMM

WEATHER CLIMATE WATER
TEMPS CLIMAT EAU

Thank you
Merci

WMO OMM
World Meteorological Organization
Organisation météorologique mondiale