

Response to action on AI, Survey outcome

Presented to CGMS-54 Working Group II session, agenda item 5

Background

Reference to CGMS-53 plenary discussions last year and response to action item 53.07:

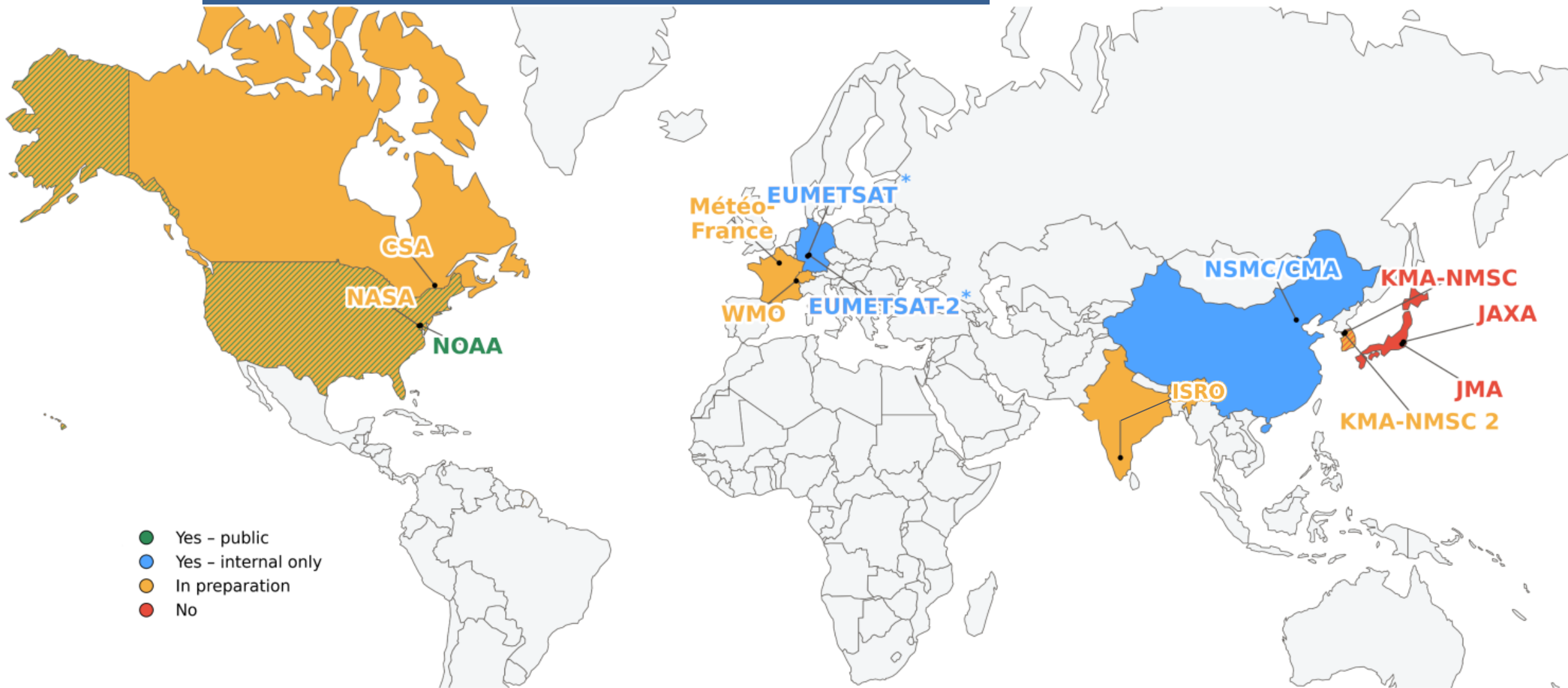
“On AI/ML: CGMS Secretariat to survey CGMS contributing agencies. WGII and WGIV to propose initial standards to be presented at CGMS-54”

Executive summary of the WP

- Strong community interest in making EO data and services more AI-ready, but **maturity remains uneven across agencies**: Few agencies are already advancing with roadmaps, pilot services, Spatio-Temporal Asset Catalogues (STAC) adoption, cloud-native formats, and AI-oriented access approaches, while others remain in planning or early exploration phases
- Agencies consistently highlight that users need an **usable access for an efficient implementation into ML workflows** (efficient data filtering and extraction, chunking, machine-readable metadata, provenance, and practical tooling) remain central expectations
- There is a **demand for a shared AI/ML catalogue**, mainly to avoid duplication, accelerate uptake, improve benchmarking, and strengthen collaboration
- **Challenges** reported by users remain concentrated around data access speed, preprocessing burden, interoperability across sources, standards maturity, and platform or usage constraints
- There is a **clear interest in a community coordination on best “AI-ready” EO practices under a CGMS –WMO framework**, with respondents indicating that joint guidance, shared practices, and lightweight coordination mechanisms would be useful next steps



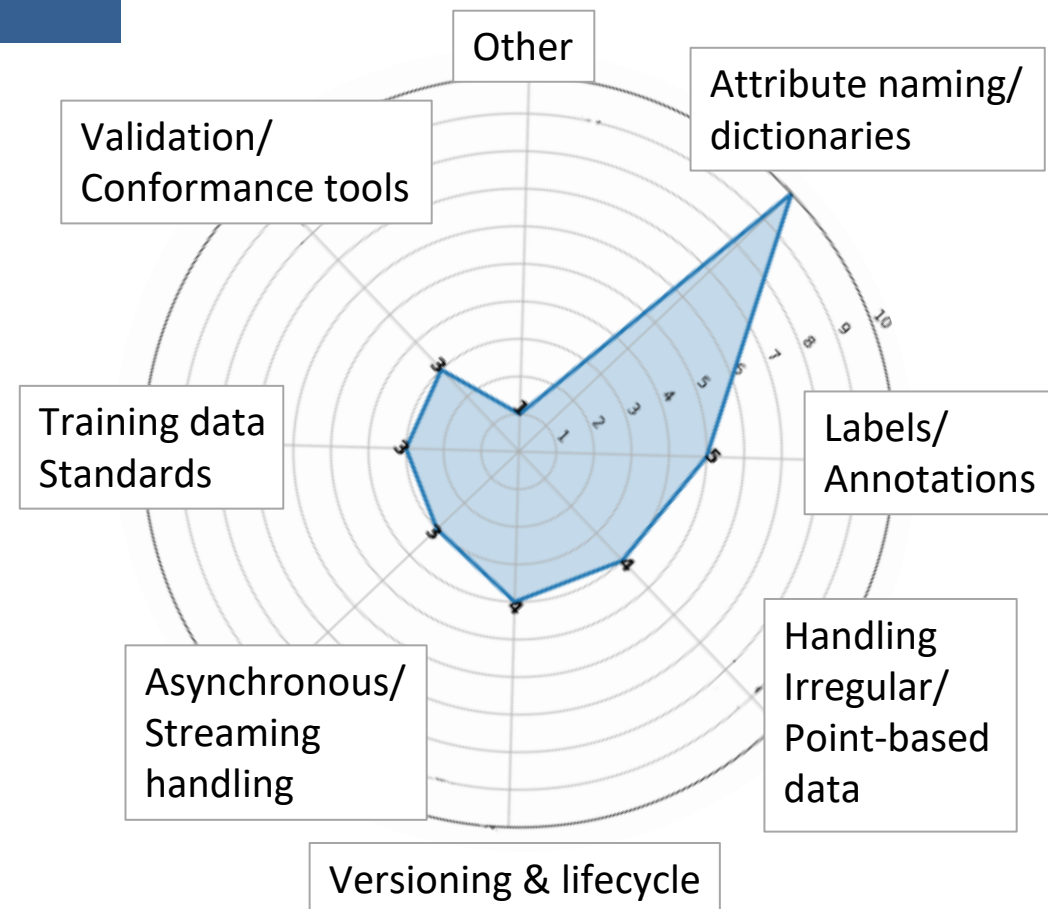
CGMS – WMO Survey on AI-Readiness



- Yes - public
- Yes - internal only
- In preparation
- No

1. Metadata & Catalogues

- Spatio Temporal Asset Catalogues (STAC): Catalog standard for organizing, describing and easily discovering geospatial datasets to be used consistently across platforms and tools
- STAC adoption is progressing: 5 respondents are currently operationally using STAC for EO discovery or at a pilot stage. 3 are planning using STAC and 5 are not planning it
- Semantic metadata commonly used: Climate and Forecast conventions as well as in-house dictionaries
- Clear demand for attribute naming /dictionaries (10) followed by label/annotations (5) being harmonised across agencies



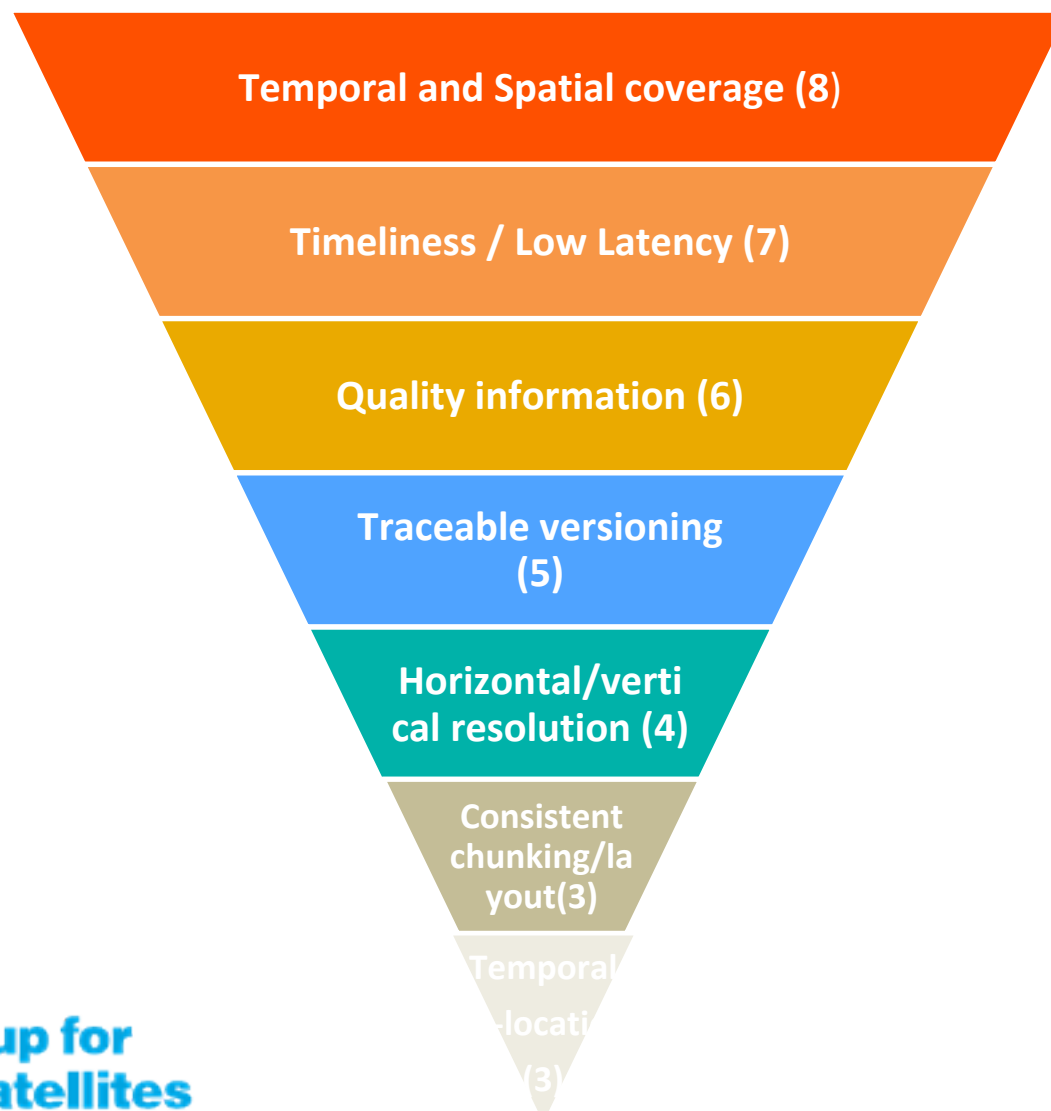
Spatio Temporal Asset Catalogues standardisation priorities



CGMS

Add CGMS agency logo here (in the slide master)

2. Data formats and cloud-native access



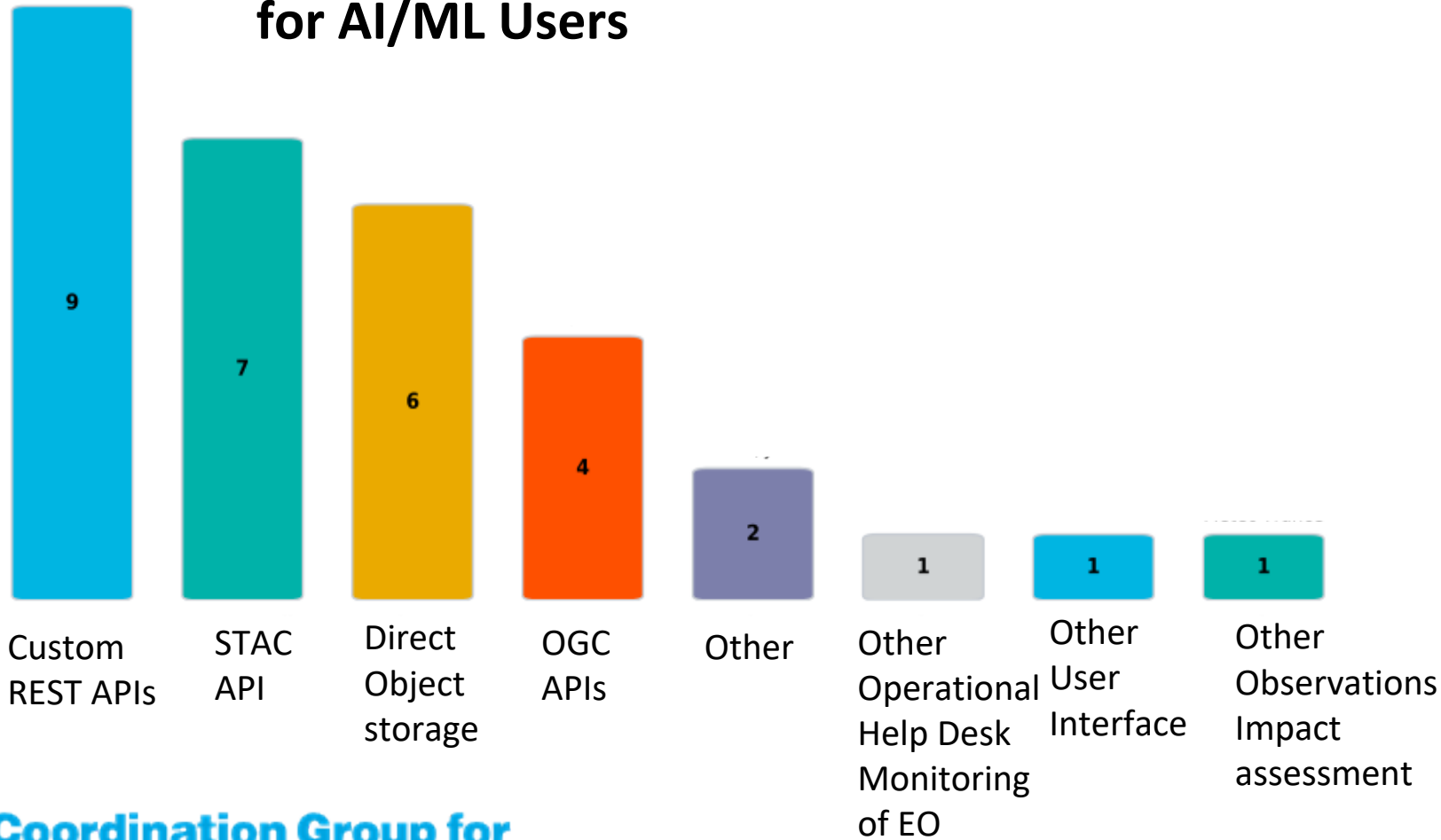
Critical data characteristics for Users in AI-workflows

- Temporal & spatial coverage as well as Timeliness/low latency and quality information are the most critical to users for AI-workflows



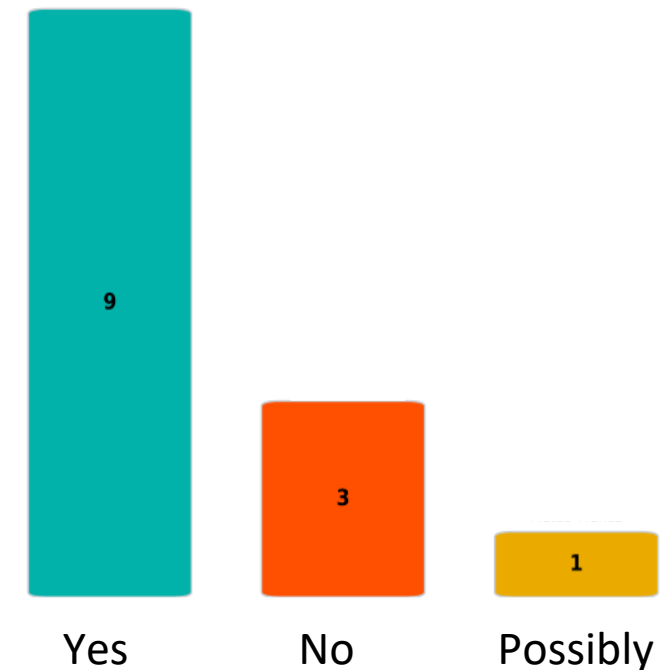
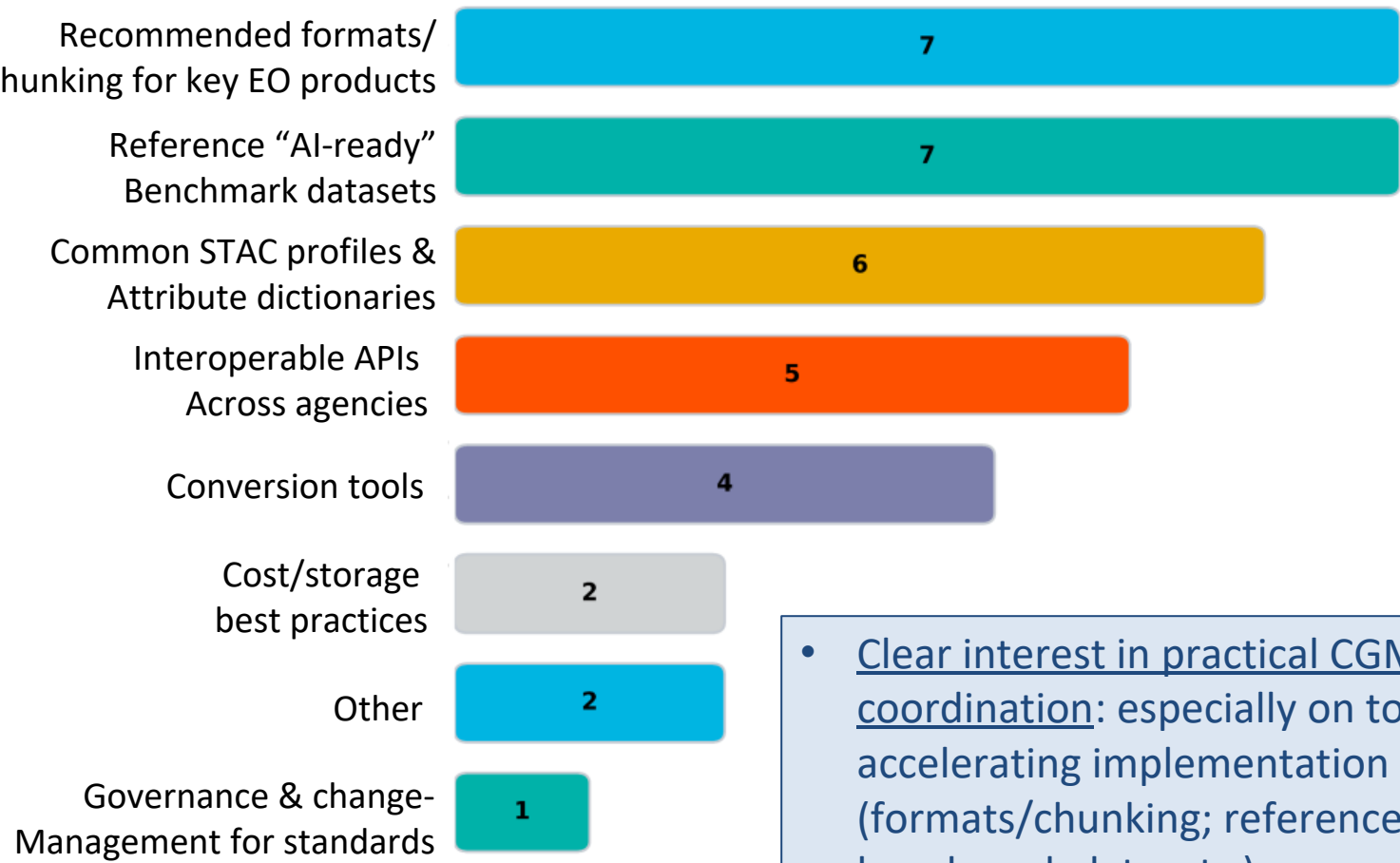
3. APIs, tools & users workflows

Access or planned interfaces for AI/ML Users



- Agencies are moving toward more AI-oriented access services, but implementation is still uneven across the community (Use of Custom REST APIs, STAC API, Direct object storage., or other interface)
- For most respondents: agencies do not yet provide tools to help users re-chunk or reformat data for AI, but planned to do so
- Main pain points reported by users when using data for AI/ML:
 - Data access/transfer speed
 - Data format/ preprocessing burden
 - Standards catalogue immaturity

4. Coordination & next steps



Participation in task team on "AI-ready EO practices"

- Clear interest in practical CGMS-WMO coordination: especially on topics for accelerating implementation (formats/chunking; reference "AI-ready" benchmark datasets)
- 9 positive answers to participate in a small task team on "AI-ready EO practices"

CGMS – WMO Survey on AI-Readiness

Conclusions

- The survey shows a community that is ready to **move from experimentation toward coordinated implementation of AI-ready EO data practices**
- Across agencies, the strongest common needs are: **better discoverability, more consistent metadata, usable cloud-native access, and practical support for ML workflows**
- There is **clear interest in a shared AI/ML catalogue**, but only if the catalogue becomes an operational resource, not just a static inventory. It should provide: structured, quality-controlled, and supported by documentation, standards, code, and examples
- Agencies favour a catalogue **coordinated through a joint CGMS–WMO governance model**, with peer-reviewed contributions

Discussion points:

- CGMS members to take note of the survey result
- CGMS Secretariat recommends to organise annual CGMS-WMO workshops, with a first focused on priority topics: metadata harmonisation, catalogue structure and minimum metadata fields, fast and scalable data access, cloud-native formats and chunking strategies, traceability
- CGMS Secretariat recommends to promote practical standards and good practices to improve interoperability across agencies and help reduce discrepancies between the community in AI-ready maturity
- For the latter point, the implementation of pilot activities is recommended to test feasibility of proposed practices and refine the approach before broader implementation across the community