eosc | Blue-Cloud2026

WP2 "**FAIR compliant Discovery and Access services for marine domains & beyond**"

Dick M.A. Schaap

MARIS

Funded by
the European Union

Facilitates users:

Federated search for discovering interesting data sets (currently more than 10 million) in a common way

Federated retrieval of identified data sets using a shopping basket mechanism

Download of data sets or push to Blue-Cloud VRE

Facilitates managers of Blue Data Infrastructures:

Wider outreach to potential users

Stay informed about data requests and users for their repository

Periodic reporting of downloads from their repository
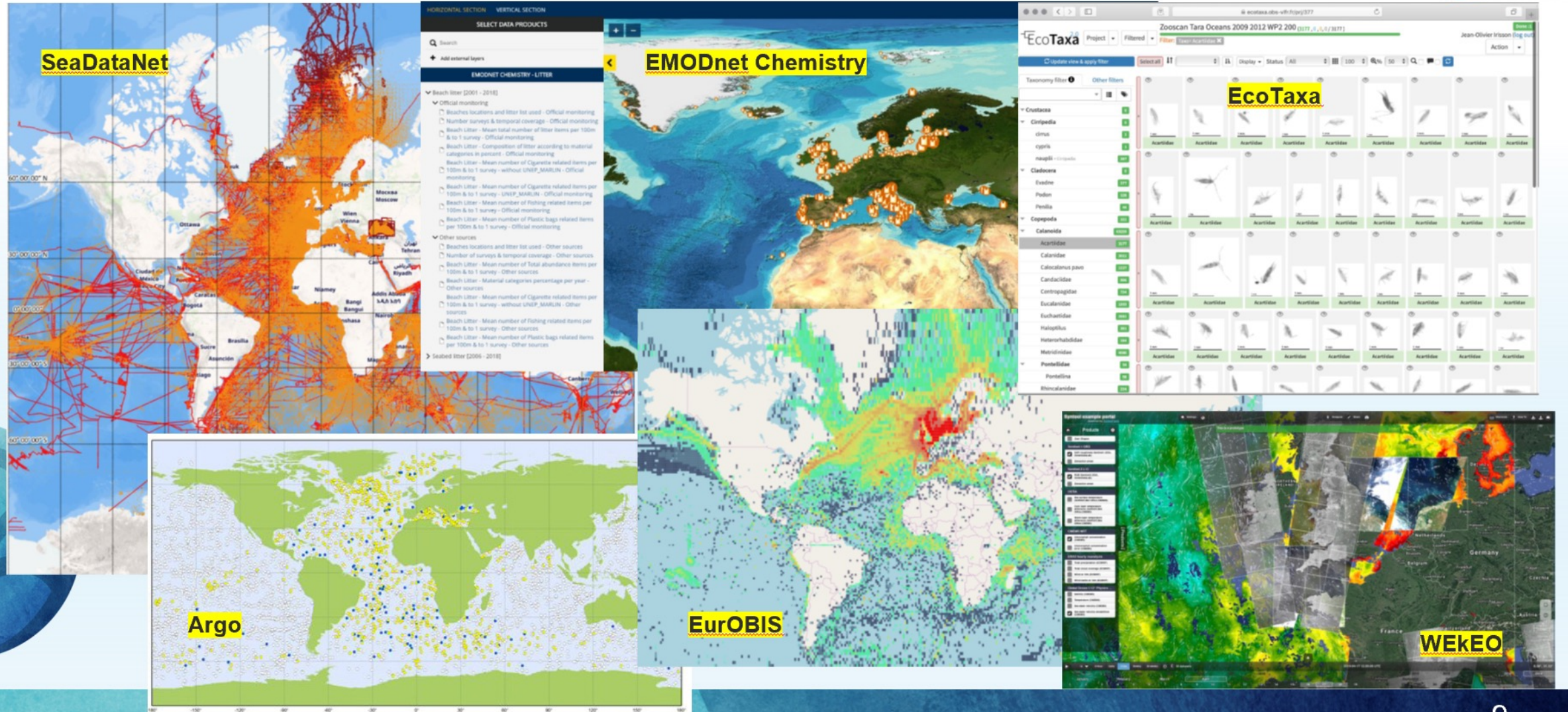
Federated discovery and retrieval of data sets and data products from the Blue Data Infrastructures

Concept of two-step search approach:

First step: identifying interesting data collections and products with few criteria

Second step: drilling down with more criteria to select specific data at granule level, where possible, otherwise at collection/products level

Metadata and Data Brokerage services interacting **Machine-to-Machine** with web services and APIs as provided and operated by the Blue Data Infrastructures
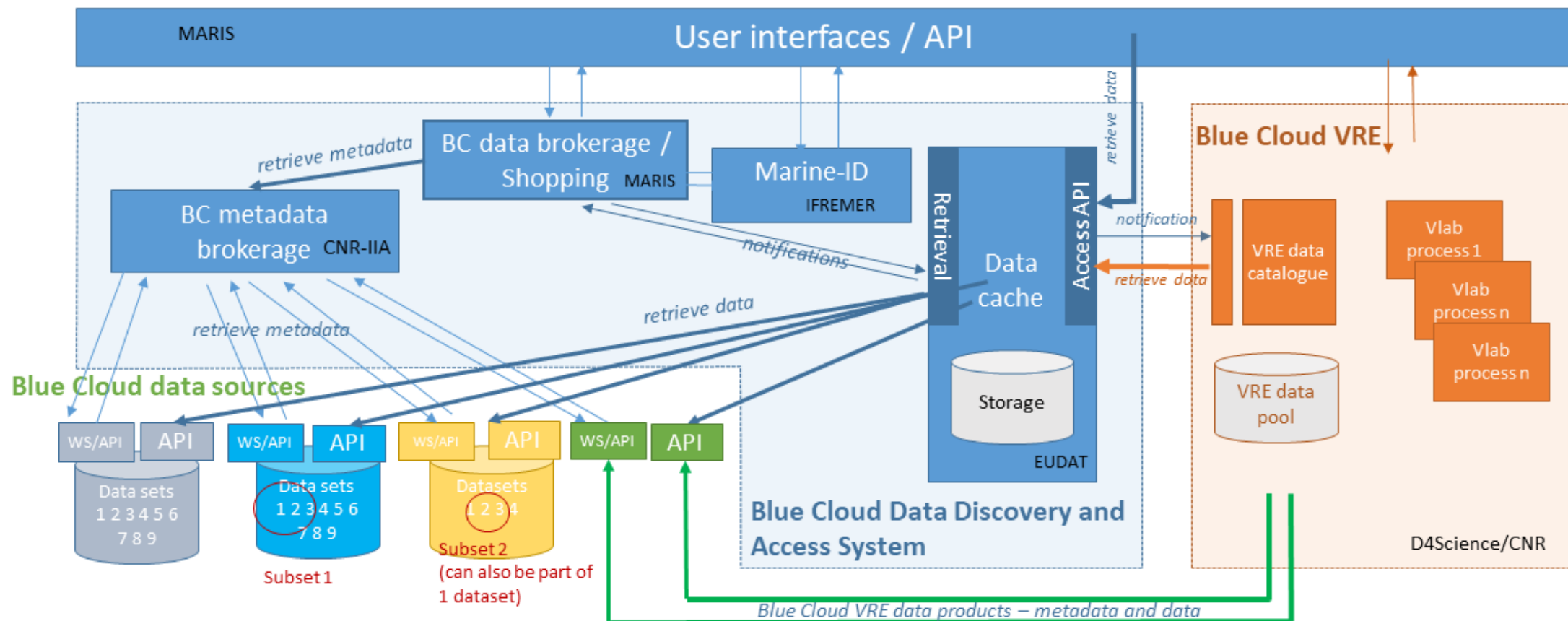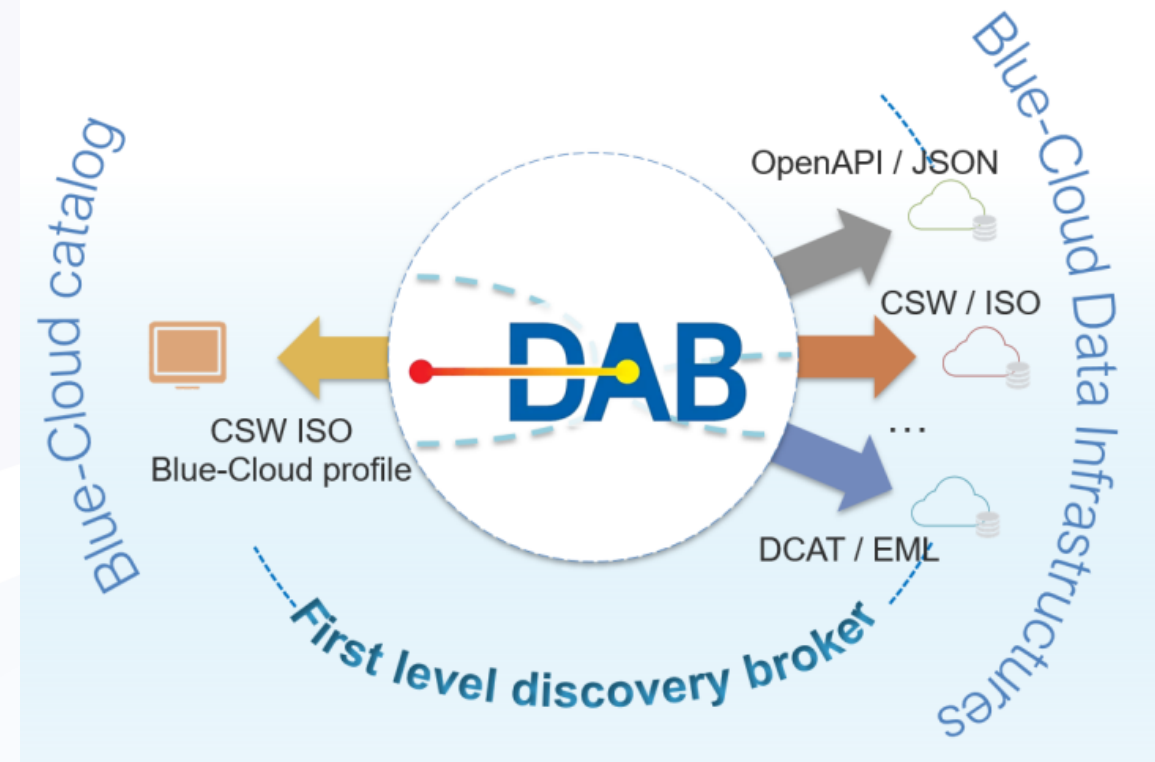
The common Blue-Cloud metadata elements are:
• IDENTIFIER: Blue-Cloud unique and persistent code for the metadata record
• TITLE: a characteristic, and often unique, name by which the collection is known
• ABSTRACT: a short description of the collection
• KEYWORD: a commonly used word, formalised word or phrase used to describe the subject
• BOUNDING_BOX: extent of the resource in the geographic space given as a bounding box
• TEMPORAL_EXTENT: time period covered by the content of the collection
• PARAMETER: name of the attribute described by the measurement value
D2.4 Blue-Cloud Data Discovery and Access service
15
• INSTRUMENT: measuring instrument used to acquire the data
• PLATFORM: platform from which the data were taken
• ORGANIZATION: organization associated with the collection
• DATESTAMP: the latest update date of the metadata description
• REVISION_DATE: the latest update date of the data
• RESOURCE_LINKS: download links where available and useful

DAB Service endpoint (global): https://blue-cloud.geodab.eu/gsservice/services/essi/view/blue-cloud/csw
- o Service endpoint (Euro Argo - Argo): https://blue-cloud.geodab.eu/gsservice/services/essi/view/argo/csw
- o Service endpoint (ELIXIR-ENA): https://blue-cloud.geodab.eu/gsservice/services/essi/view/elixir-ena/csw
- o Service endpoint (EMODnet Chemistry): https://blue-cloud.geodab.eu/gsservice/services/essi/view/emodnet-chemistry/csw
- o Service endpoint (EurOBIS): https://blue-cloud.geodab.eu/gsservice/services/essi/view/eurobis/csw
- o Service endpoint (ICOS Data Portal): https://blue-cloud.geodab.eu/gsservice/services/essi/view/icos-data-portal/csw
- o Service endpoint (ICOS SOCAT): https://blue-cloud.geodab.eu/gsservice/services/essi/view/icos-socat/csw
- o Service endpoint (SeaDataNet CDI service): https://blue-cloud.geodab.eu/gsservice/services/essi/view/seadatanet-open/csw
- o Service endpoint (SeaDataNet Products): https://blue-cloud.geodab.eu/gsservice/services/essi/view/seadatanet-products/csw
- Service interface: CSW ISO v. 2.0.2
- Data model: Blue-Cloud ISO 19115 profile

- https://dabreporting.s3.amazonaws.com/BlueCloud/BlueCloudReport_brief.html for completeness reporting

SeaDataNet	Dedicated API
SeaDataNet Products	OGC CSW service
EMODnet Chemistry	OGC CSW service
EuroArgo - Argo	Dedicated API
EurOBIS – EMODnet Biology	DCAT service
Ecotaxa	Dedicated API
ELIXIR - ENA	Dedicated API
ICOS Marine	SPARQL service
SOCAT	ERDDAP service

https://data.blue-cloud.org

eosc | Blue-Cloud2026

WP2 will expand & optimise the **Blue-Cloud Data Discovery & Access service** (DD&AS) and its FAIRness by:

- harmonising and expanding functionality of web services as operated by each BDI for discovery and access of managed data resources, and as used in DD&AS, following FAIRness review

- developing and deploying semantic brokering as part of DD&AS interface

- federating additional BDIs into the DD&AS (**EMSO, EMBRC, SIOS, EMODnet Physics, EuroBioImaging)**

- reviewing, and if missing, developing and deploying data sub-setting and extracting services, operated by each BDI, for feeding Blue-Cloud 'raw data' Data Lakes,

- developing and deploying Blue-Cloud Data Lakes for storing and maintaining 1) 'raw data' extracted from BDIs and 2) harmonised and validated data collections for selected data types, as resulting from the WP3 EOV Work Benches

- tuning Data Lakes developments with Digital Twin of the Ocean (DTO) developments, in particular EDITO-Infra.

Task 2.1 [M1-M42] "Optimising the functioning of the Data Discovery & Access Service".
Lead: MARIS; Partners: CNR, IFREMER, CINECA, GRNET, SU, VLIZ, EBI, NOC-BODC, UVA, FORTH, EMSO, EMBRC, SIOS, ETT, IOPAN.

Task 2.2 [M1-M42] "Horizontal expansion: federation of additional Blue Data Infrastructures".
Lead: MARIS; Partners: CNR (IIA), CINECA, ETT, EMSO, EMBRC, SIOS, SOCIB, SMHI, NERSC, IOPAN.

Task 2.3 [M1-M42] "Vertical expansion: developing and deploying sub-setting services at BDIs and Blue-Cloud Data Lakes".
Lead: MARIS; Partners: CNR, MOI, IFREMER, NOC-BODC, VLIZ, SU, EBI, SIOS, NERSC, IOPAN, EMBRC, ETT, CINECA, GRNET, UVA, EMSO, EGI, Hub Ocean.

Task 2.4 [M1-M42] "DTO Taskforce for tuning Blue-Cloud data lakes with DTO developments".
Lead: MOI; Partners: MARIS, CNR, IFREMER, VLIZ, EGI, SSBE.

Partner	Efforts (PM)	Partner	Efforts (PM)
MARIS	33	IRD	2
CNR	13	NOC-BODC	14
Ifremer	6	ETT	12
SSBE	2	EMSO-ERIC	12
MOI	10	EGI	4
VLIZ	12	SOCIB	5
EMBL	11	SMHI	3
CINECA	20	EMBRC-ERIC	12
GRNET	4	SIOS	5
SU	8	IOPAN	4
UVA	6	NERSC	4
FORTH	2	Hub Ocean	7
TOTAL	211	PM	20.40%

D2.1: Existing DD&AS and Blue Data Infrastructures – Review and Specifications for Optimisation Report: Report with FAIRness review of existing BDI web services for discovery and access and existing DD&AS central services **(M8 – MARIS)**

D2.2: New Blue Data Infrastructures – Service Analysis Report: Detailed descriptions of the new to be connected BDIs (EMSO, EMBRC, SIOS, EMODnet Physics) with details on their local discovery and access mechanisms, types of data, metadata format, data formats, use of vocabularies, possible restrictions, and existing web services **(M10 – MARIS**)

D2.3: Optimised and expanded Blue Cloud Data Discovery and Access Service – Documentation Report: Report documenting the new release of the Blue Cloud Data Discovery and Access Service with optimised services, semantic interoperability, new BDIs connected, and new or adapted discovery and access BDI web services **(M24 – MARIS)**

D2.4: BDI sub-setting APIs and Data Lakes – Concept and Specifications Report: Report with descriptions and analyses of existing sub-setting services at each of the BDIs, formulation of common requirements for BDI sub-setting APIs, and how BDIs might adapt existing or develop new sub-setting APIs **(M14 – MARIS)**

D2.5: Established BDI sub-setting APIs and Data Lakes – Documentation Report: Report documenting the release of the Blue Cloud Data Lakes, their architecture and deployment, and the established sub-setting APIs at each BDI **(M26 – MARIS)**

D2.6: Tuning between Blue-Cloud Data Lakes and DTO development, 1st report: Progress report about the tuning process and the interfacing between the Blue-Cloud Data Lakes development and the DTO development **(M14 – MOI)**

D2.7: Tuning between Blue-Cloud Data Lakes and DTO development, 2nd report: Progress report about the tuning process and the interfacing between the Blue-Cloud Data Lakes development and the DTO development **(M26 – MOI)**

Discussion

Presentation of new Blue Data Infrastructures:

- **EMSO**

- **EMBRC**

- **SIOS**

- **EMODnet Physics**

- **EuroBioImaging**

TSC meeting in Amsterdam, 28 – 29 March 2023