

Pētniecības datu pārvaldība dabas zinātņu (un inženierzinātņu) projekta ietvaros

10.04.2024.

Latvijas Zinātnes padomes seminārs

Īsi par mani



- Zooplanktona taksonomists un pelaģiskās vides ekologs. 2022.gadā iegūts doktora grāds (PhD) bioloģijā.
- Visvairāk interesē jūras pelaģiskā dzīvotne (t.i., ūdens) un procesi tajā. It īpaši zemāko trofisko līmeņu barības tīkli.
- Pieredze darbā ar datu apstrādi, analīzi, kvalitātes kontroli, starptautisko datu ziņošanu, projektu realizēšanu un datu pārvaldīšanu.
- Viena no atbildīgajiem Latmare datubāzes izveidošanā un attīstīšanā. Latmare ir brīvpieejas Latvijas jūras vides monitoringa datu portāls. <https://latmare.lhei.lv/>.

Astra Labuce

pētniece, Latvijas Hidroekoloģijas institūta hidrobioloģijas laboratorijas vadītāja

Datu pārvaldības cikls



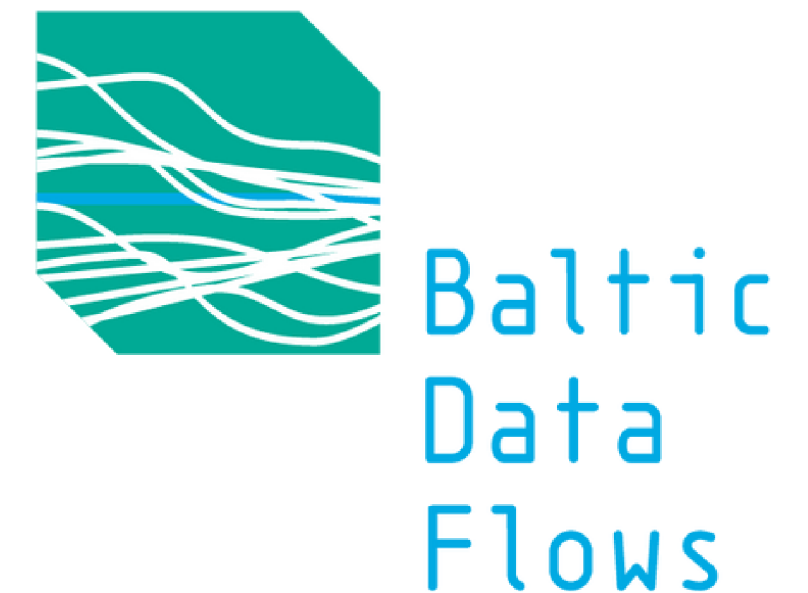
plānošana = DPP (*DMP*)

Piemērs I

Baltic Data Flows projekts CINEA finansēts

Contents

1. Introduction	3
2. Project summary	3
3. Data collection and generation	4
3.2. Eutrophication data	5
3.3. Hazardous substances data	7
3.4. Zooplankton, Phytoplankton and Zoobenthos data	8
4. FAIR Data Management	9
4.1. Making data findable	9
4.2. Making data accessible.....	10
4.3. Making data interoperable.....	10
4.4. Re-use.....	10
5. Data security	11
5.1 General.....	11
5.2 Protection and management of data.....	11
6. Ethical aspects	12



Co-financed by the
Connecting Europe Facility
of the European Union

DPP (*DMP*)

Piemērs I

Baltic Data Flows projekts CINEA finansēts



Baltic
Data
Flows



Co-financed by the
Connecting Europe Facility
of the European Union

Contents

1. Introduction	projektu aprakstošā daļa	3
2. Project summary		3
3. Data collection and generation	projektā izmantoto/iegūto datu aprakstošā daļa	4
3.2. Eutrophication data		5
3.3. Hazardous substances data		7
3.4. Zooplankton, Phytoplankton and Zoobenthos data		8
4. FAIR Data Management		9
4.1. Making data findable	metadati - padara datus mašīnrakstā atrodamus	9
4.2. Making data accessible		10
4.3. Making data interoperable		10
4.4. Re-use		10
5. Data security		11
5.1 General	cita būtiska informācija	11
5.2 Protection and management of data		11
6. Ethical aspects		12

DPP (*DMP*)

Piemērs II AquaINFRA projekts HORIZON finansēts

1	Executive Summary.....	5
2	Introduction	6
3	Data Summary	6
3.1	<i>Case studies.....</i>	<i>8</i>
3.2	<i>Final data products.....</i>	<i>18</i>
4	FAIR data	21
4.1	<i>Making data findable, including provisions for metadata.....</i>	<i>21</i>
4.2	<i>Making data openly accessible.....</i>	<i>23</i>
4.3	<i>Making data interoperable.....</i>	<i>23</i>
4.4	<i>Increase data re-use (through clarifying licenses).....</i>	<i>26</i>
5	Costs.....	26
6	Data Security.....	27
7	Ethical aspects.....	27
8	Other issues	27
9	Next Steps	27
	Appendix I: FAIR Data Management summary tables	28



DPP (*DMP*)

Piemērs II
AquaINFRA projekts
HORIZON finansēts

1	Executive Summary.....	5
2	Introduction	6
3	Data Summary	6
3.1	Case studies.....	8
3.2	Final data products.....	18
4	FAIR data	21
4.1	Making data findable, inc.....	21
4.2	Making data openly acces.....	23
4.3	Making data interoperabl.....	23
4.4	Increase data re-use (through clarifying licenses).....	26
5	Costs.....	26
6	Data Security.....	27
7	Ethical aspects.....	27
8	Other issues	27
9	Next Steps	27
	Appendix I: FAIR Data Management summary tables	28

projektu aprakstošā daļa

projektā izmantoto/iegūto
datu aprakstošā daļa

metadati - padara datus
mašīnrakstā atrodamus

cita būtiska informācija



DPP (*DMP*)

metadati - padara datus
mašīnrakstā atrodamus



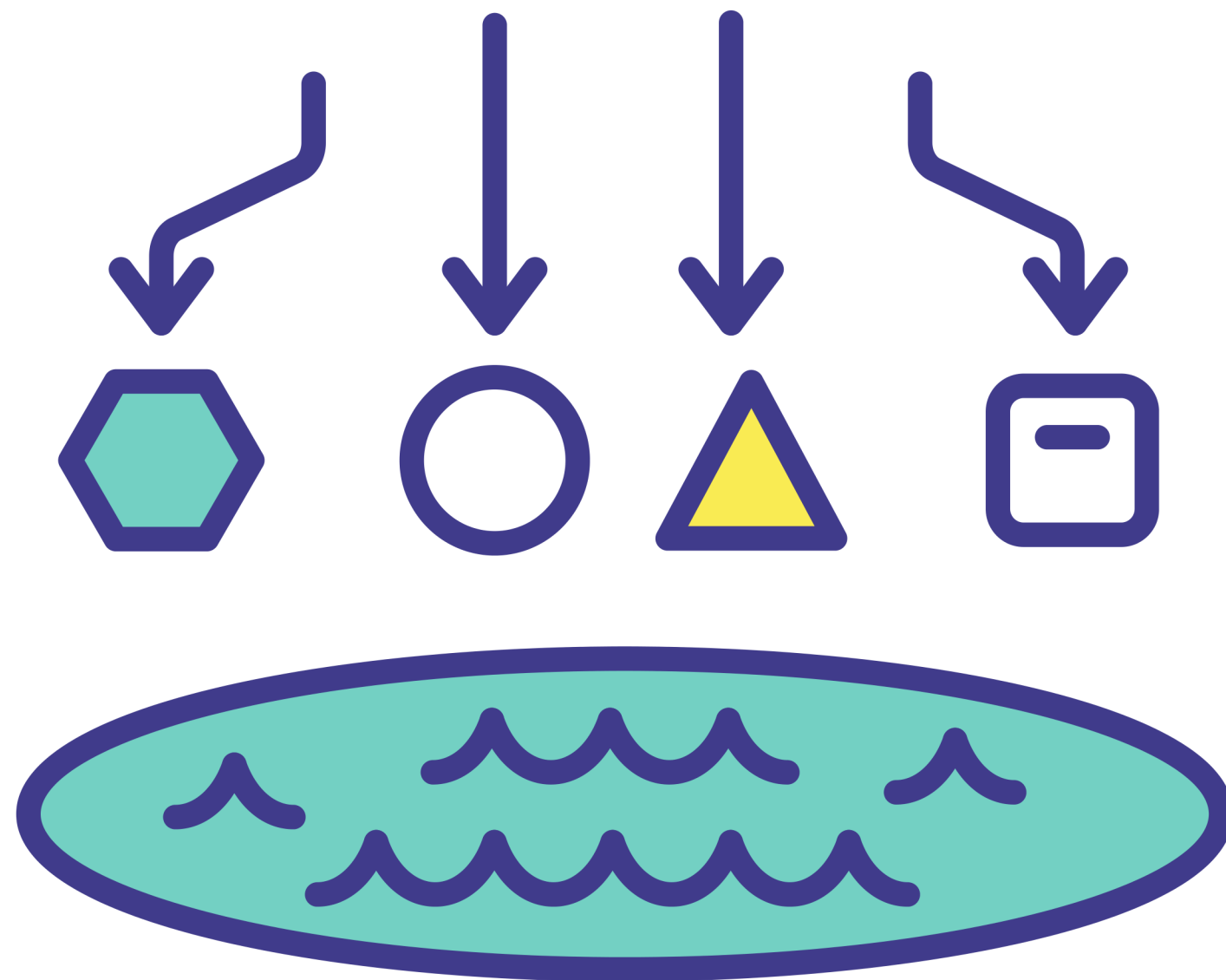
pieejamus

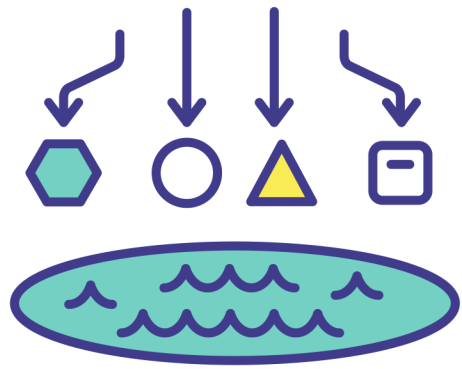


izmantojamus

FAIR dati

METADATI

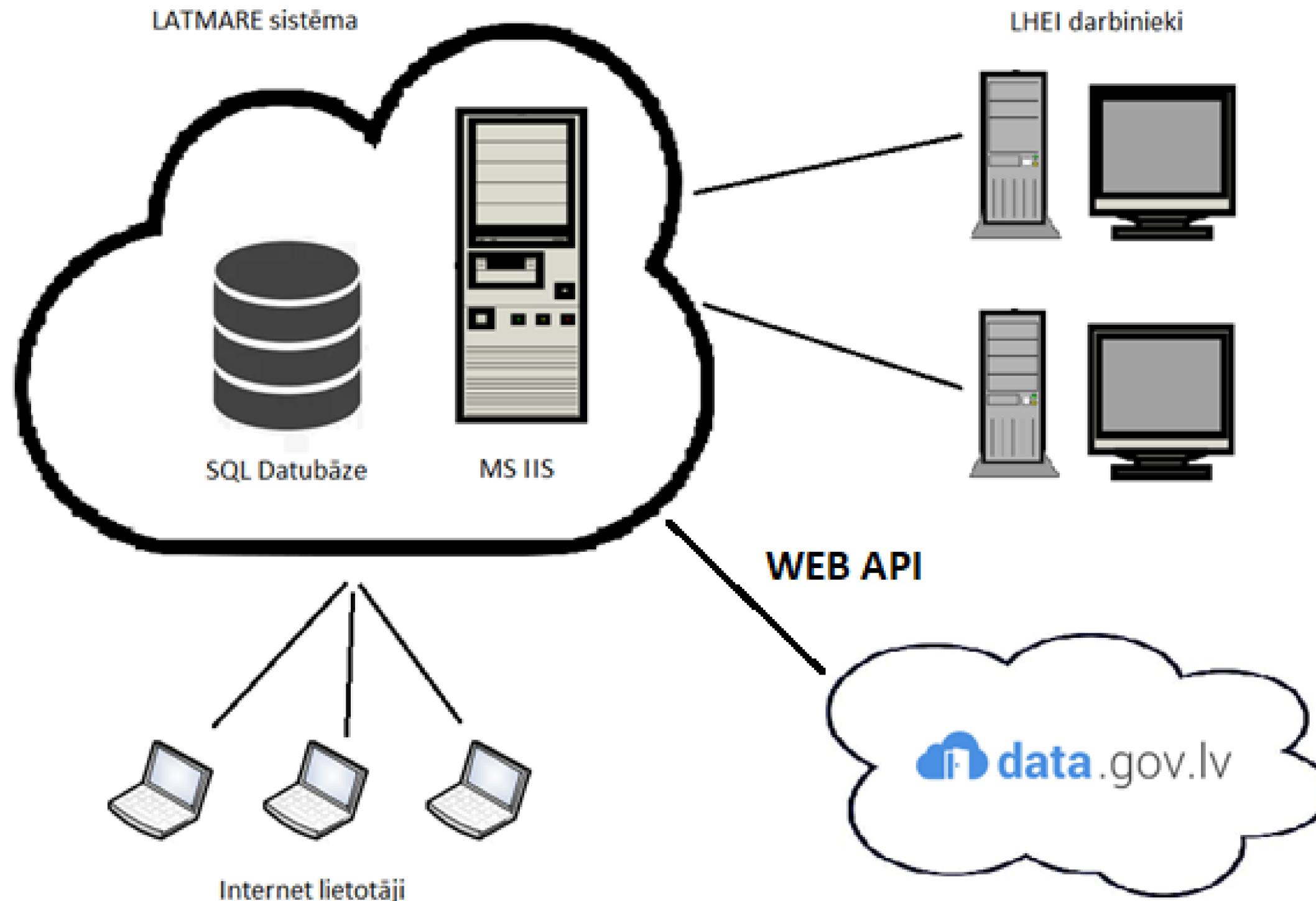


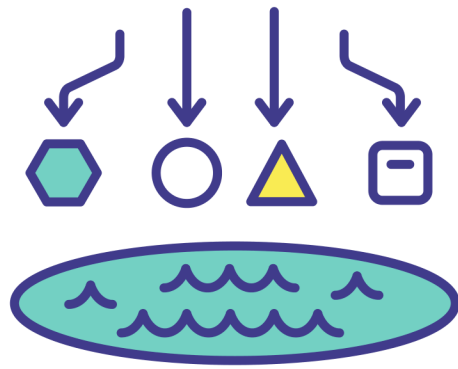


METADATI

DCAT-AP/INSPIRE

FAIR dati





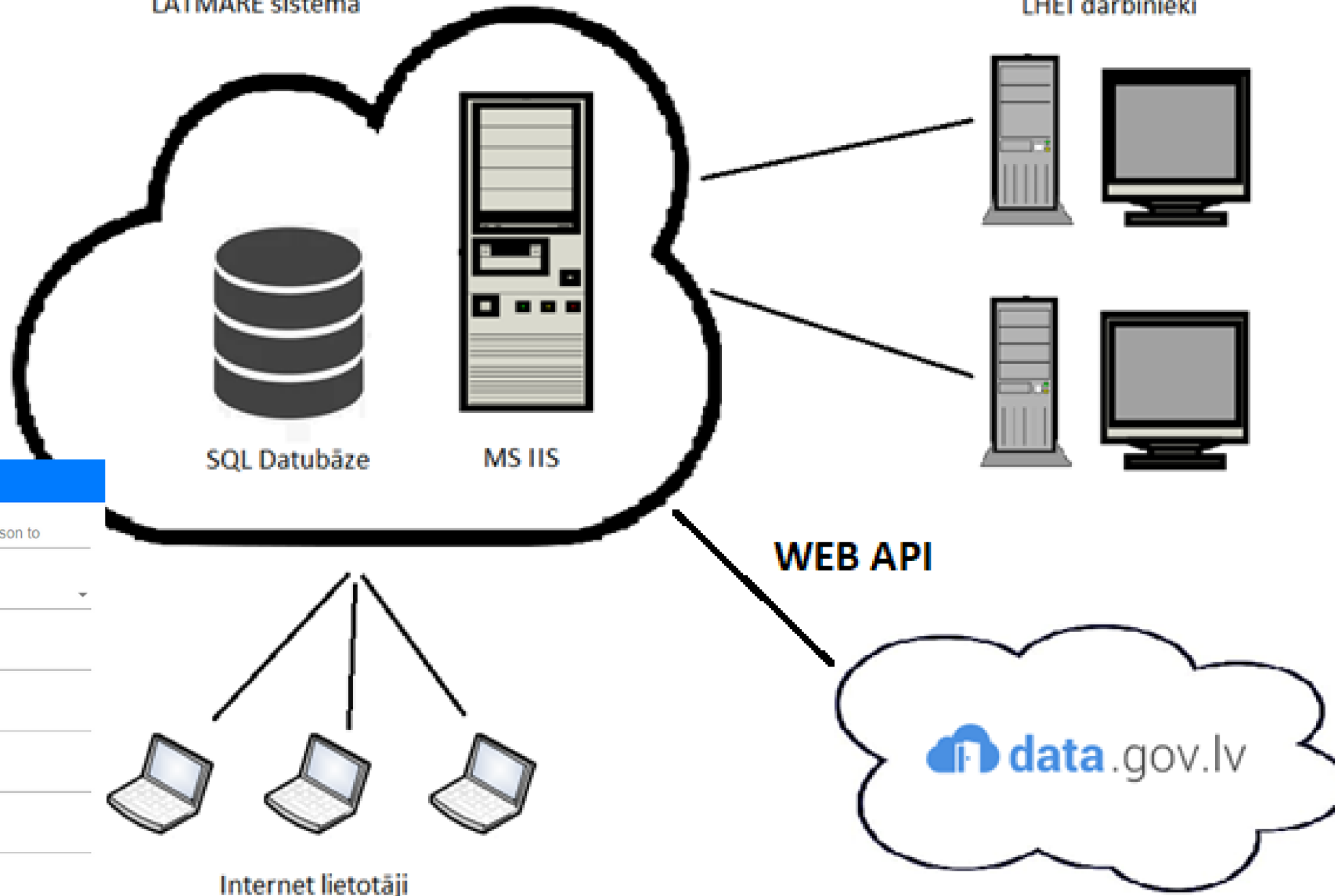
METADATI

DCAT-AP/INSPIRE

FAIR dati

LATMARE sistēma

LHEI darbinieki



<https://latmare.lhei.lv>

LATMARE Observations

Monitoring period * Season from Season to

Group of parameters

Parameters *

Cruise

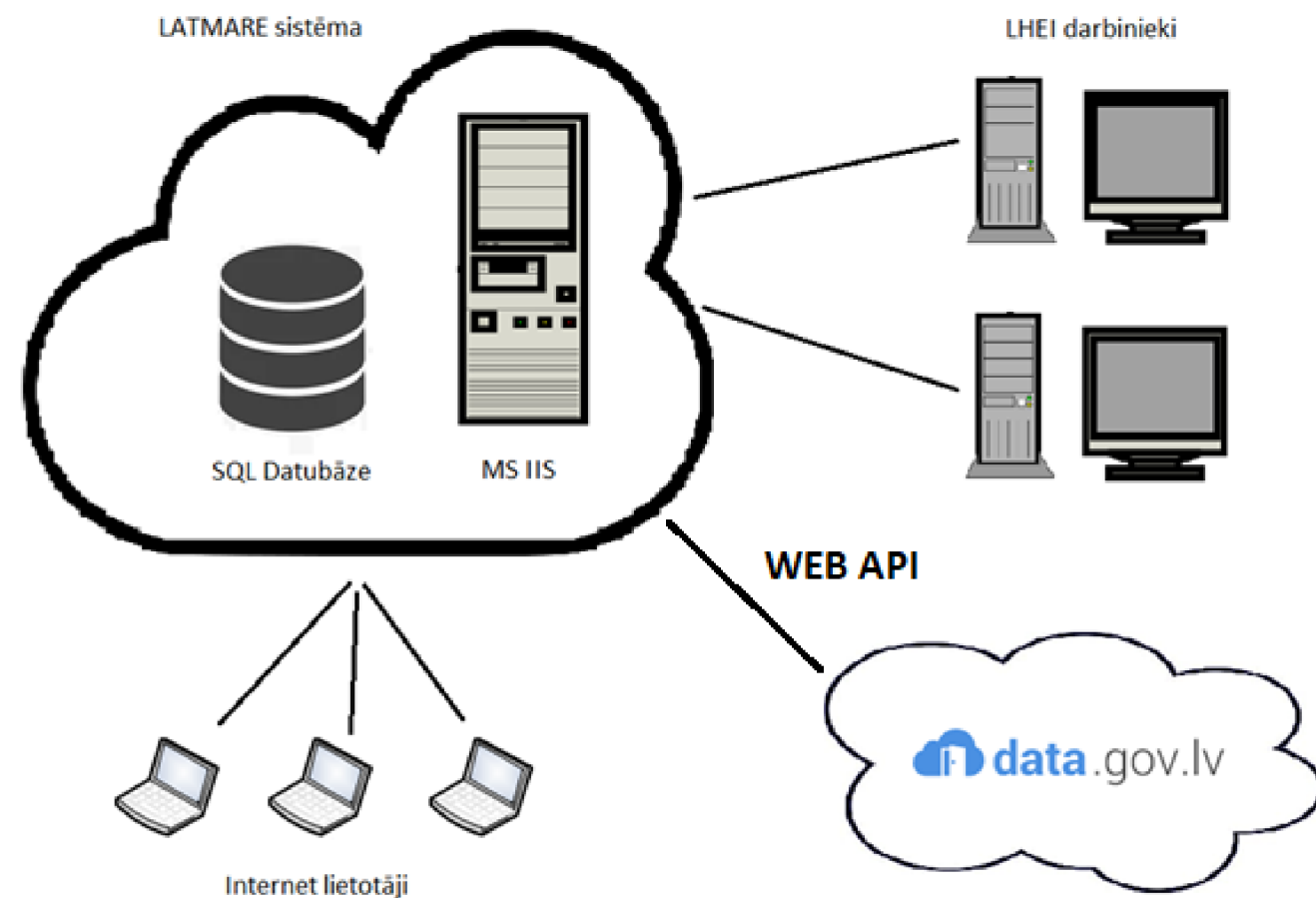
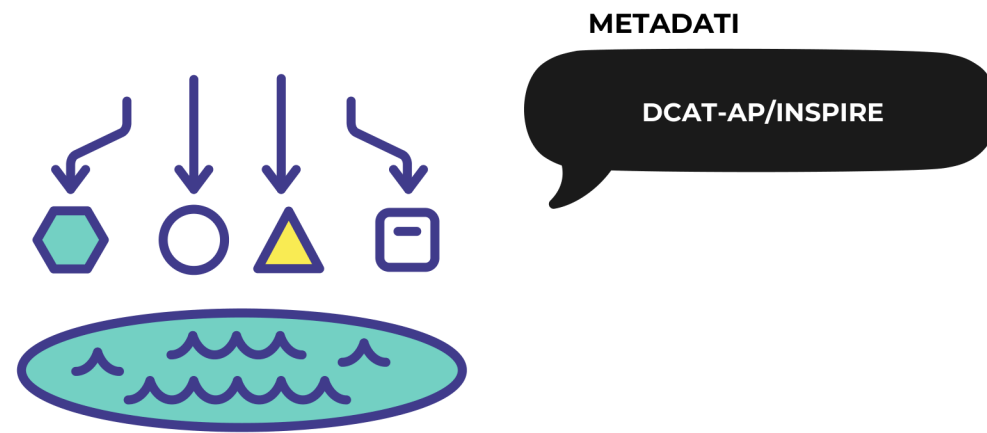
Station name

Project

environmental observations



FAIR dati

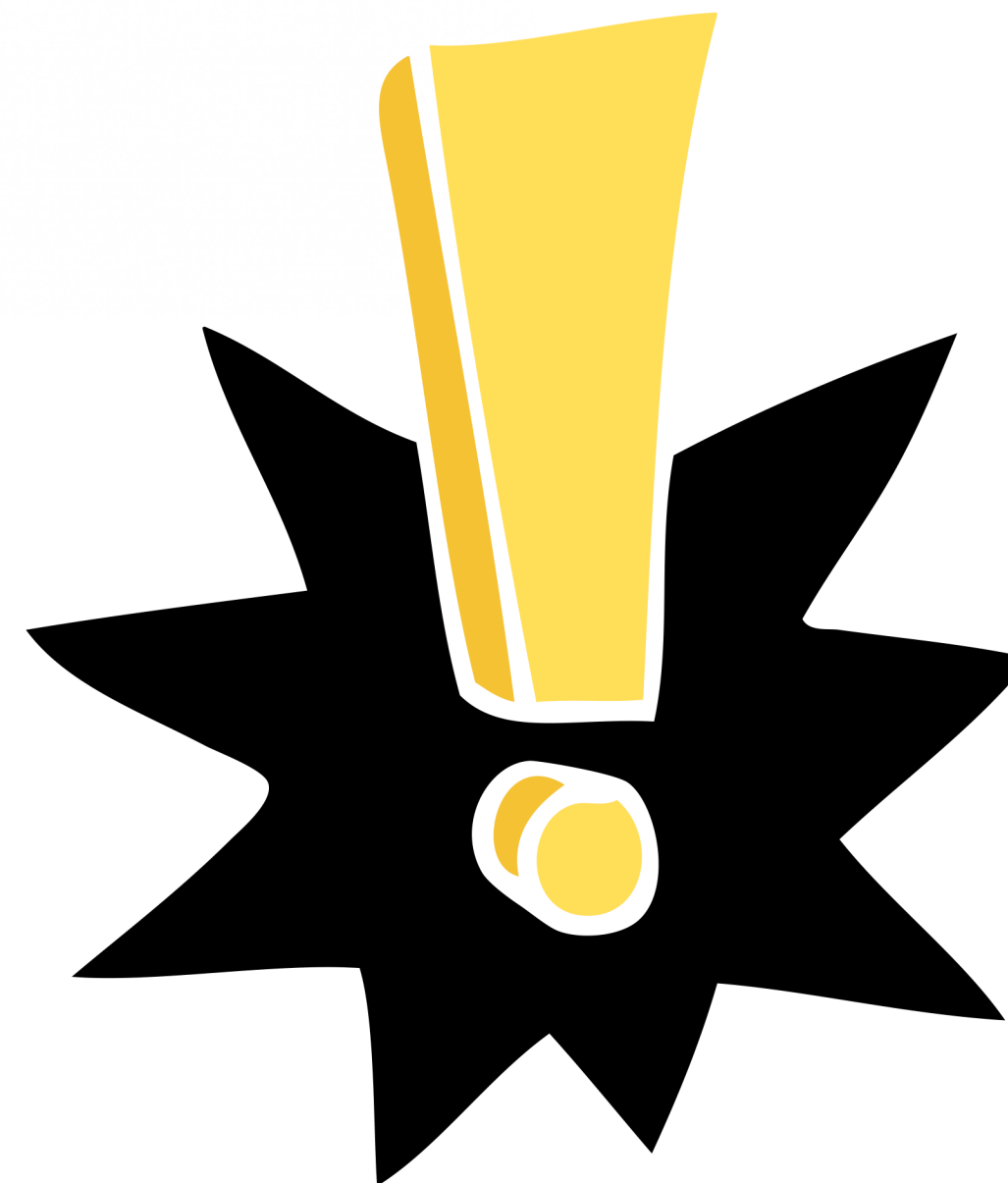
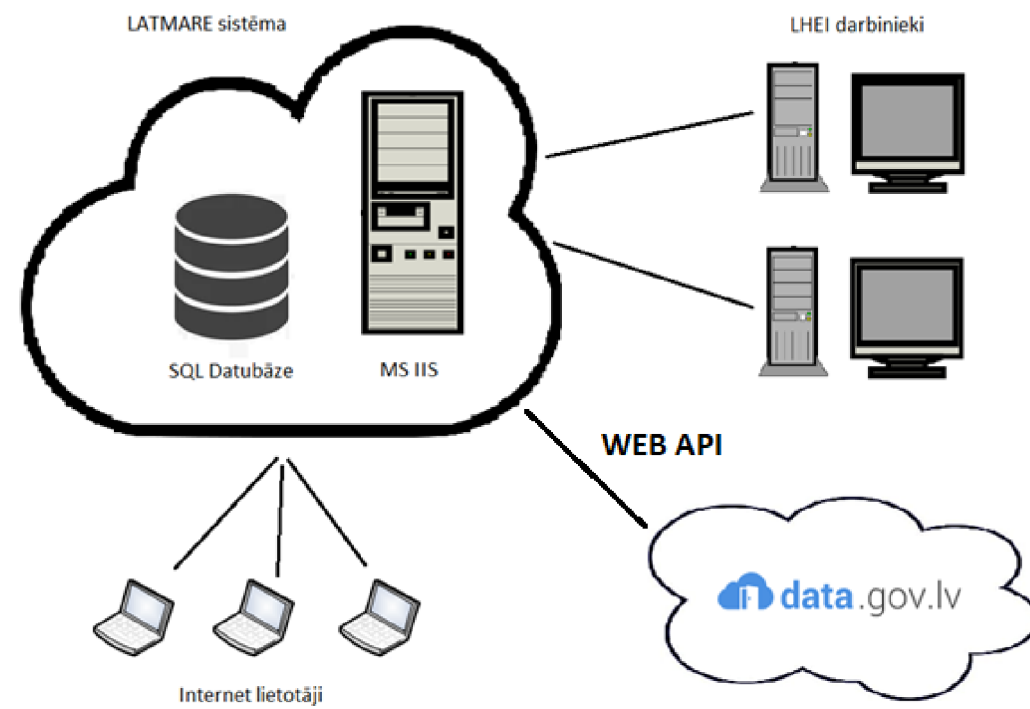
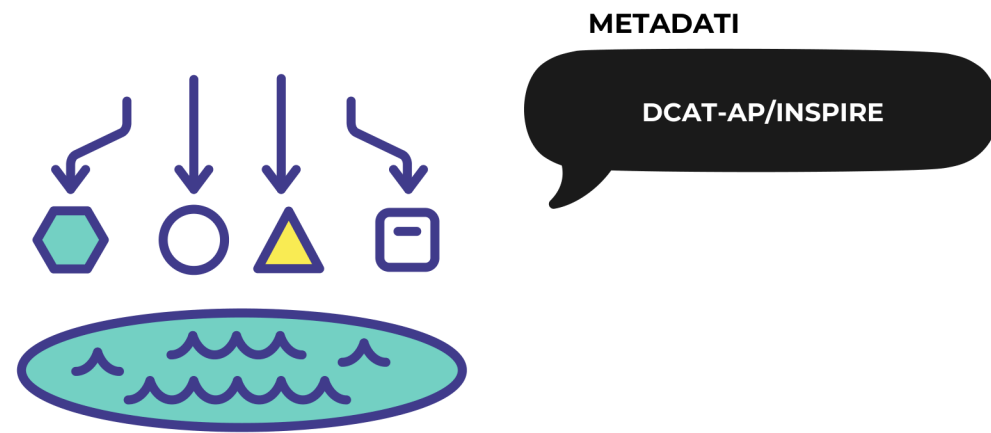


EUROPEAN
DATA PORTAL



The DCAT Application profile for data portals in Europe (DCAT-AP) is a specification based on the Data Catalogue vocabulary (DCAT) for describing public sector datasets in Europe.

FAIR dati



Pieejamība

<u>Lejupielādes URL</u>	0%	<u>Biežākās piekļuves URL...</u>	200	<u>Biežākās lejupielādes U...</u>	n/a
-------------------------	----	----------------------------------	-----	-----------------------------------	-----

Atkārtota izmantojamība

<u>Piekļuves ierobežojumi</u>	false	<u>Licences informācija</u>	0%	<u>Piekļuves ierobežojumu v...</u>	{}	<u>Kontaktpunkts</u>	true
-------------------------------	-------	-----------------------------	----	------------------------------------	----	----------------------	------

<u>Publicētājs</u>	true
--------------------	------

Kontekstualitāte

<u>Faila lielums</u>	50%	<u>Tiesības</u>	0%	<u>Modificēšanas datums</u>	true	<u>Modificēšanas datums</u>	0%
----------------------	-----	-----------------	----	-----------------------------	------	-----------------------------	----

Atrodamība

<u>Atslēgvārdu izmantošana</u>	true	<u>Kategorijas</u>	false	<u>Ģeogrāfiskā meklēšana</u>	false	<u>Uz laiku balstīta meklē...</u>	false
--------------------------------	------	--------------------	-------	------------------------------	-------	-----------------------------------	-------

Savietojamība

<u>DCAT-AP atbilstība</u>	true	<u>Formāts</u>	100%	<u>Multivides veids</u>	50%	<u>Formāts/multivides vei...</u>	50%
---------------------------	------	----------------	------	-------------------------	-----	----------------------------------	-----

FAIR pārbaude





Gaignard, A., Rosnet, T., de Lamotte, F., Lefort, V., & Devignes, M. (2023). **FAIR-Checker: supporting digital resource findability and reuse with Knowledge Graphs and Semantic Web standards.** *Journal of Biomedical Semantics*, 14. <https://doi.org/10.1186/s13326-023-00289-5>



Rosnet, Thomas, Lefort, Vincent, Devignes, Marie-Dominique, & Gaignard, Alban. (2021). **FAIR-Checker, a web tool to support the findability and reusability of digital life science resources.** *JOBIM (JOBIM)*, Paris. <https://doi.org/10.5281/zenodo.5914307>



Gaignard, Alban, Rosnet, Thomas, de Lamotte, Frédéric, & Devignes, Marie-Dominique. (2021, June 4). **Automatic evaluation of FAIR metrics.** *Elixir All-Hands meeting*. <https://doi.org/10.5281/zenodo.5914367>

Latvijas atvērtās zinātnes stratēģija

2021.-2027. gadam

Zinātniskajām institūcijām jāveido drošus, FAIR principi un OAI-PMH protokolam⁶⁵ atbilstošus pētniecības datu repozitorijus. VPC kā organizācijai ar tehnisku kompetenci un ciešu sadarbību ar zinātniskajām institūcijām, ar Izglītības un zinātnes ministrijas atbalstu jāizveido vispārējo pētniecības datu repozitoriju tīkls *DataverseLV*⁶⁶, kas nodrošinās iespēju jebkuram interesentam atrast un piekļūt Latvijas pētnieku pētniecības datiem.

Datu atvēršana: piemērs I

kopā ar publikāciju



Dataset on microplastic concentrations, characteristics, and chemical composition in the marine surface waters of Latvia – the Eastern Gotland basin and the Gulf of Riga

Published: 6 February 2023 | Version 2 | DOI: [10.17632/x9ptrn83sz.2](https://doi.org/10.17632/x9ptrn83sz.2)

Contributors: [Marta Barone](#), [Natalija Suhareva](#), [Juris Aigars](#), [Ieva Putna](#), [Inta Dimante-Deimantovica](#)

Description

This dataset provides information on composition and spatial distribution of spectroscopically verified microplastics of two Baltic Sea sub-basins – the semi-enclosed Gulf of Riga and the Eastern Gotland Basin – as well as meteorological conditions during sampling events for further studies and research activities. The data can be used e.g., for modelling spatial distribution or pinpointing sources and transport pathways of spectroscopically verified microplastics in marine environment and for the calibration of existing models.

Datu atvēršana: piemērs I

kopā ar publikāciju

Steps to reproduce

Marine surface water microplastic samples were collected from 44 transects in two Baltic Sea sub-basins – the semi-enclosed Gulf of Riga and the Eastern Gotland Basin. Sampling was performed by using Manta trawl with a mesh size of 300 µm. This was followed by organic material digestion using sodium hydroxide, hydrogen peroxide and enzymes. After that the samples were filtered on glass fiber filters and analyzed visually, registering the shape, size, and color of each item. Chemical composition of particles was determined using Attenuated Total Reflection Fourier Transform Infrared (ATR-FTIR) spectroscopy method.

Institutions

Latvijas Hidroekoloģijas instituta, Daugavpils Universitāte

Categories

Spectroscopy, Pollution, Fourier Transform Infrared Spectroscopy, Baltic Sea, Microplastics

Funding

European Social Fund

Project No. 8.2.2.0/20/I/003

Ministry of Environmental Protection and Regional Development project “Improvement of knowledge of the state of the marine environment in the marine waters under the jurisdiction of Latvia”

Project No. IL/106/2017

Related Links

Article

<https://doi.org/10.1016/j.marpolbul.2021.112860>

is related to this dataset

Licence

CC BY 4.0

[Learn more](#)



Dataset on microplastic concentrations, characteristics, and chemical composition in the marine surface waters of Latvia – the Eastern Gotland basin and the Gulf of Riga

Published: 6 February 2023 | Version 2 | DOI: 10.17632/x9ptrn83sz.2

Contributors: Marta Barone, Natalija Suhareva, Juris Aigars, Ieva Putna, Inga Dimante-Deimantovica

Description

This dataset provides information on composition and spatial distribution of spectroscopically verified microplastics of two Baltic Sea sub-basins – the semi-enclosed Gulf of Riga and the Eastern Gotland Basin – as well as meteorological conditions during sampling events for further studies and research activities. The data can be used e.g., for modelling spatial distribution or pinpointing sources and transport pathways of spectroscopically verified microplastics in marine environment and for the calibration of existing models.





Datu atvēršana: piemērs I

kopā ar publikāciju









runMSTS

Labuce, Astra¹ ; Gorokhova, Elena² 

This release provides a script-based workflow applying widely-used R-packages to compute the core indicator Zooplankton Mean Size and Total Stock (MSTS) for assessing the ecological status of the pelagic habitat in the Baltic Sea. The indicator MSTS was developed as a part of HELCOM (Helsinki Commission; The Baltic Marine Environment Protection Commission) activities (Gorokhova et. al., 2016) and used in the holistic environmental assessment in the Baltic Sea (HELCOM, 2018).

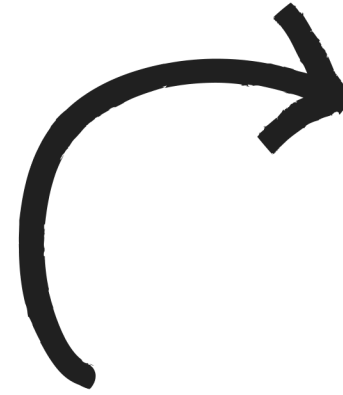
runMSTS was developed within the "Baltic Data Flows" and "HELCOM BLUES" projects funded by the Climate Infrastructure and Environment Executive Agency (CINEA) of the European Union and the Directorate-General for Environment of the European Commission, respectively. Complementary funding was provided by The Swedish Agency for Marine and Water Management (SwAM).

Files

runMSTS.zip		▼
 runMSTS.zip		
 README.txt		825 Bytes
 input		
 addInfo.txt		156 Bytes
 nonaggregated.txt		495.7 kB
 yearlymeans.txt		739 Bytes
 runMSTS.Rmd		71.8 kB



METADATI



runMSTS

Labuce, Astra¹ ; Gorokhova, Elena²

Show affiliations

This release provides a script-based workflow applying widely-used R-packages to compute the core indicator Zooplankton Mean Size and Total Stock (MSTS) for assessing the ecological status of the pelagic habitat in the Baltic Sea. The indicator MSTS was developed as a part of HELCOM (Helsinki Commission; The Baltic Marine Environment Protection Commission) activities (Gorokhova et. al., 2016) and used in the holistic environmental assessment in the Baltic Sea (HELCOM, 2018).

runMSTS was developed within the "Baltic Data Flows" and "HELCOM BLUES" projects funded by the Climate Infrastructure and Environment Executive Agency (CINEA) of the European Union and the Directorate-General for Environment of the European Commission, respectively. Complementary funding was provided by The Swedish Agency for Marine and Water Management (SwAM).

Files

runMSTS.zip	
runMSTS.zip	
README.txt	825 Bytes
input	
addInfo.txt	156 Bytes
nonaggregated.txt	495.7 kB
yearlymeans.txt	739 Bytes
runMSTS.Rmd	71.8 kB

Details

DOI

DOI [10.5281/zenodo.7467616](https://doi.org/10.5281/zenodo.7467616)

Resource type

Software

Publisher

Zenodo

Languages

English

Rights

GNU General Public License v2.0 only

Citation

Labuce, A., & Gorokhova, E. (2022). runMSTS (1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.7467616>

Style APA



Datu atvēršana

nespecifiskie repozitoriji



Datu atvēršana: piemērs II platforma

The screenshot shows the website for 'WAVE CLIMATOLOGY DATA FOR THE BALTIC SEA'. The page features a navigation bar with 'HOME', 'THE PROJECT', 'THE DATA', 'THE SCIENCE', and 'SUPPORT'. A central image shows a laptop displaying a map of the Baltic Sea with a color-coded wave climatology overlay. To the right of the laptop, the text reads 'WAVE CLIMATOLOGY DATA FOR THE BALTIC SEA' and 'Available for download'. Below this, there is a list of features: 'High Resolution Wave Data', 'Free for non-commercial use', '20th Numerical Model', and 'Extensively validated dataset'. At the bottom of the page, there are three icons with text: 'Wave Model' (using SWAN wave model), 'Grids' (three-level nested scheme), and 'Forcing Wind' (ERA5 reanalysis by ECMWF and ERA5+ reanalysis from ERA5-40). The footer contains the text 'TalTech - Wave Engineering Laboratory' and 'TalTech's high-resolution wave data repository'. The URL 'taltech High Resolution Wave Climatology Data for the Baltic Sea /' is visible at the bottom left.

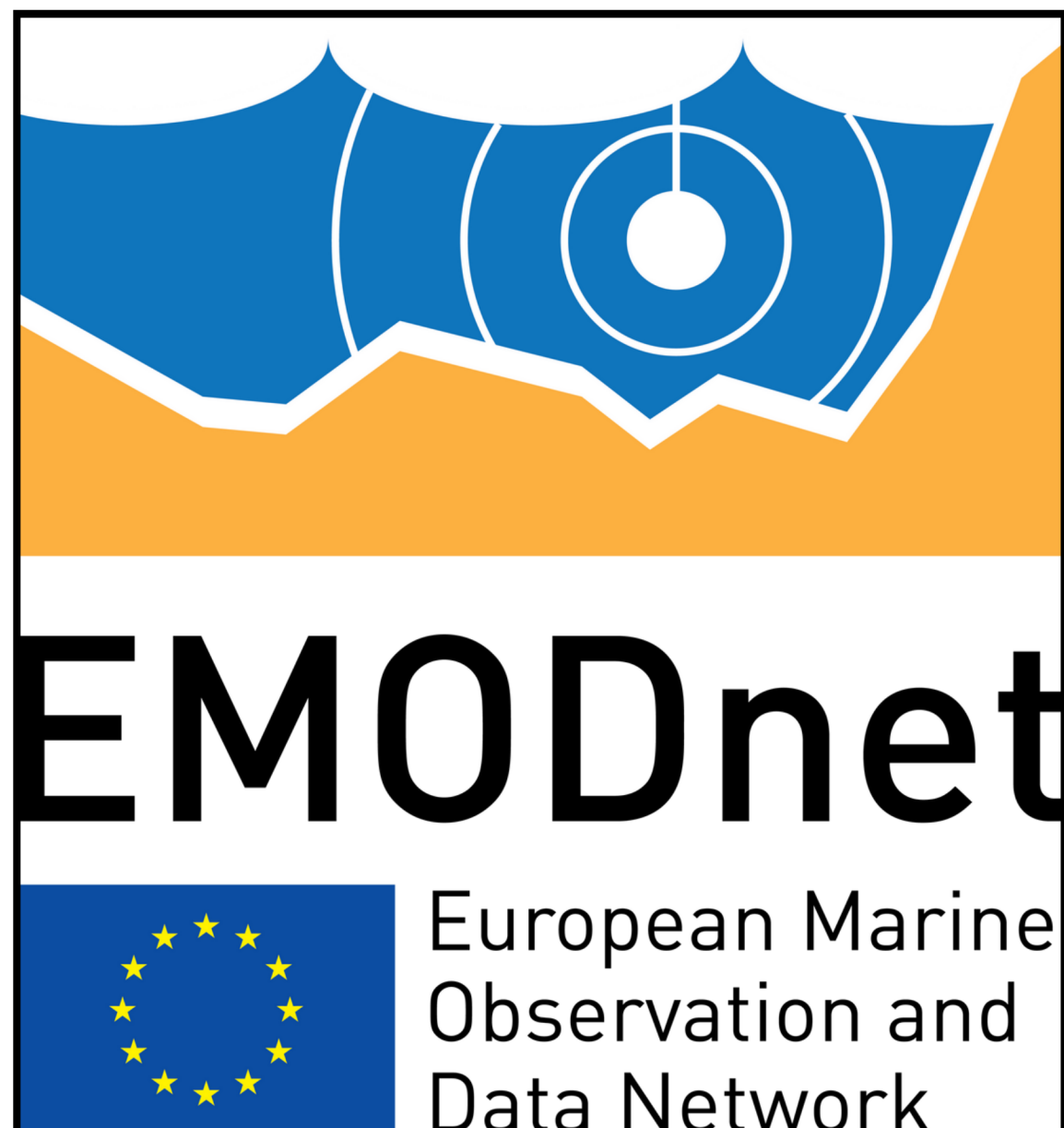
FAIR ?

Datu atvēršana

nozares repozitoriji
institūcijas repozitoriji

Datu atvēršana: piemērs III

startautisks nozares repozitorijs



konkrēts publicēšanas formāts

SEANOE SEA SCIENTIFIC OPEN DATA PUBLICATION

Search

Search everywhere

Reset filters

1406 result(s)

Open access 2023-06

Spotter buoy data f
Houghton Isabel , Arduin F

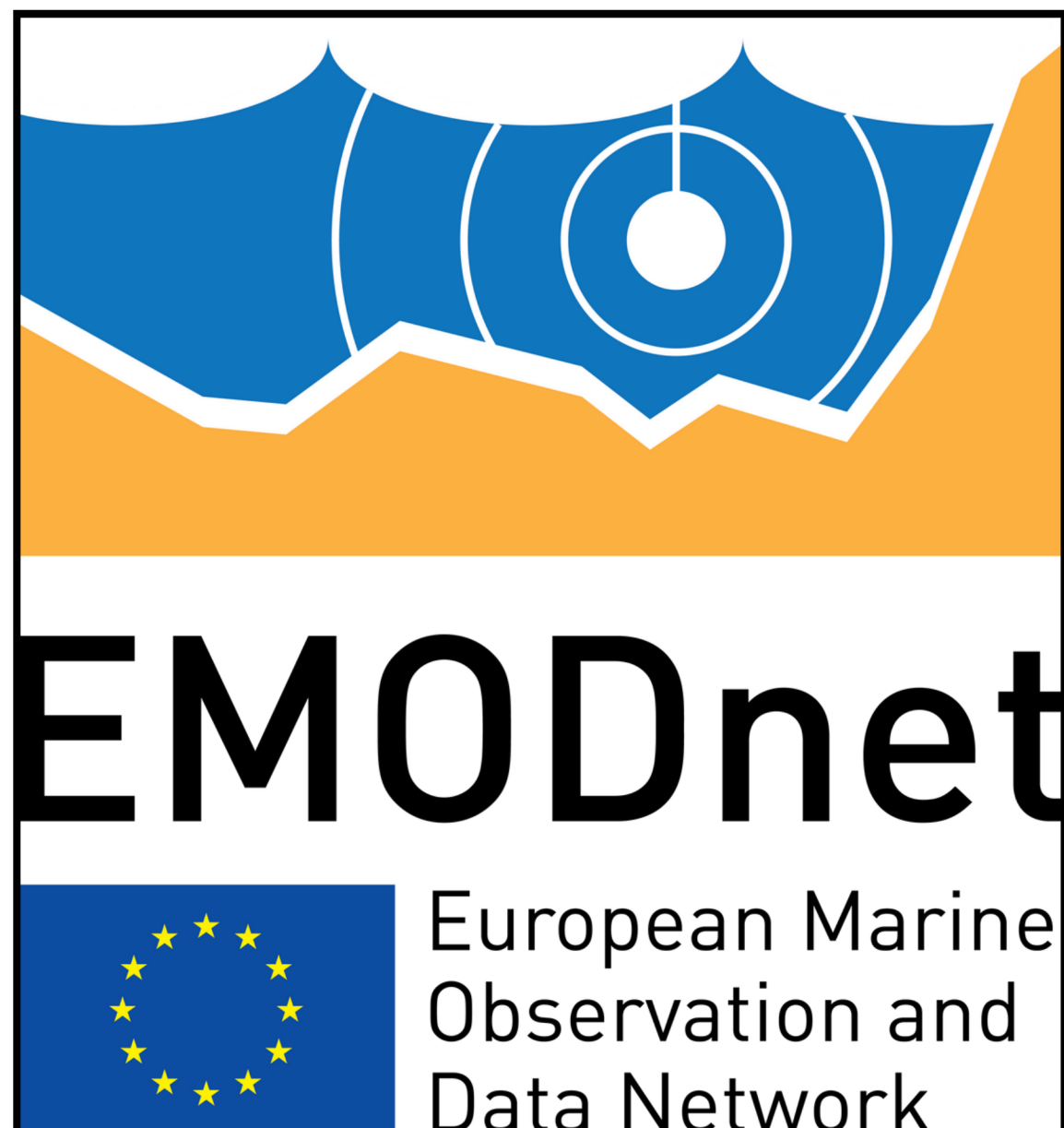
This dataset contains Spo storm "Fabrice". This datas SWOT altimetry data: reve with SWOT data and nume

PUBLICATION YEAR	
<input type="checkbox"/> 2024	78
<input type="checkbox"/> 2023	230
<input type="checkbox"/> 2022	176
<input type="checkbox"/> 2021	168

DISCIPLINE

Datu atvēršana: piemērs III

startautisks nozares repozitorijs



!Datu kvalitāte!

attīstās un virzās arī uz liela
apjomu datu uzturēšanu, piem,
datu plūsma no flowCam

SEANOE SEA SCIENTIFIC OPEN DATA PUBLICATION

Search

Search everywhere

Reset filters

1406 result(s)

Open access 2023-06

Spotter buoy data f
Houghton Isabel , Arduin F

This dataset contains Spo
storm "Fabrice". This datas
SWOT altimetry data: reve
with SWOT data and nume

PUBLICATION YEAR	
<input type="checkbox"/> 2024	78
<input type="checkbox"/> 2023	230
<input type="checkbox"/> 2022	176
<input type="checkbox"/> 2021	168

DISCIPLINE

Kādā formātā dati jāglabā?

- **iekļaut tikai datus**
 - izvairīties no grafiskus attēlojumus (ko mašīna nevar interpretēt)

Kā tos sagatavot?

Kādā formātā dati jāglabā?

- iekļaut tikai datus
- **labāk mazāk, bet lielākus failus**
 - atkārtotai izmantošanai vieglāk ir paņemt daļu no lielas datu kopas nekā apvienot daudz mazas

Kā tos sagatavot?

Kādā formātā dati jāglabā?

- iekļaut tikai datus
- labāk mazāk, bet lielākus failus
- **izvairīties no “saspiešanas”**
 - lai gan atsevišķi (parasti nespecifiskie) repozitoriji to pieļauj

Kā tos sagatavot?

Kādā formātā dati jāglabā?

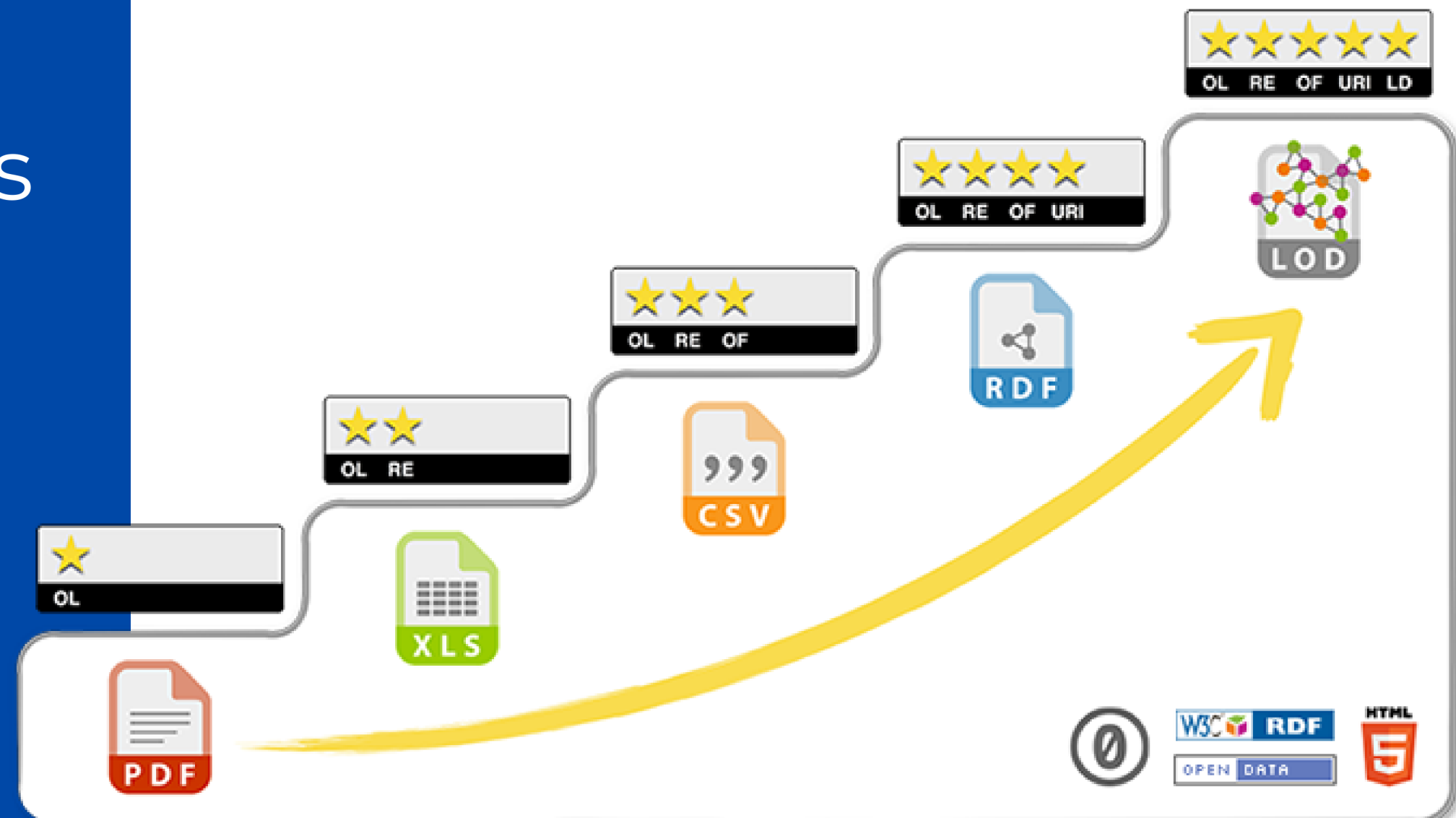
- iekļaut tikai datus
- labāk mazāk, bet lielākus failus
- izvairīties no “saspiešanas”
- ja konkrētais repozitorijs neprasa noteiktu datu formātu, **pēc iespējas .txt vai .csv faili tabulāriem datiem**

Kā tos sagatavot?

Kādā formātā dati jāglabā?

- iekļaut tikai datus
- labāk mazāk, bet lielākus failus
- izvairīties no “saspiešanas”
- ja konkrētais repozitorijs neprasa noteiktu datu formātu, **pēc iespējas .txt vai .csv faili tabulāriem datiem**

Kā tos sagatavot?



- Datu pārvaldīšanai, atvēršanai, metadatu sagatavošana ir jāparedz laiks un cilvēkstundas!!
- Dati pirms atvēršanas vienmēr ir jāpārbauda.
- Metadati ir bagātīgi un precīzi jāapraksta

Kā tos sagatavot?

- organizēt kolonās un rindās
- katrs mērījums/datu punkts ar unikālu identifikatoru
- katrai rindai vienāds kolonu skaits un katrai kolonai vienāds rindu skaits (NA aizpildīt pēc repazitorija noteikumiem vai brīvi aprakstot metadatos kā)



**Paldies par
uzmanību!**

Astra Labuce
astra.labuce@lhei.lv

<http://www.lhei.lv>