eSC Integration levels

How to choose the one that best meet the requirements of your infrastructure?

Prof Tamas Kiss

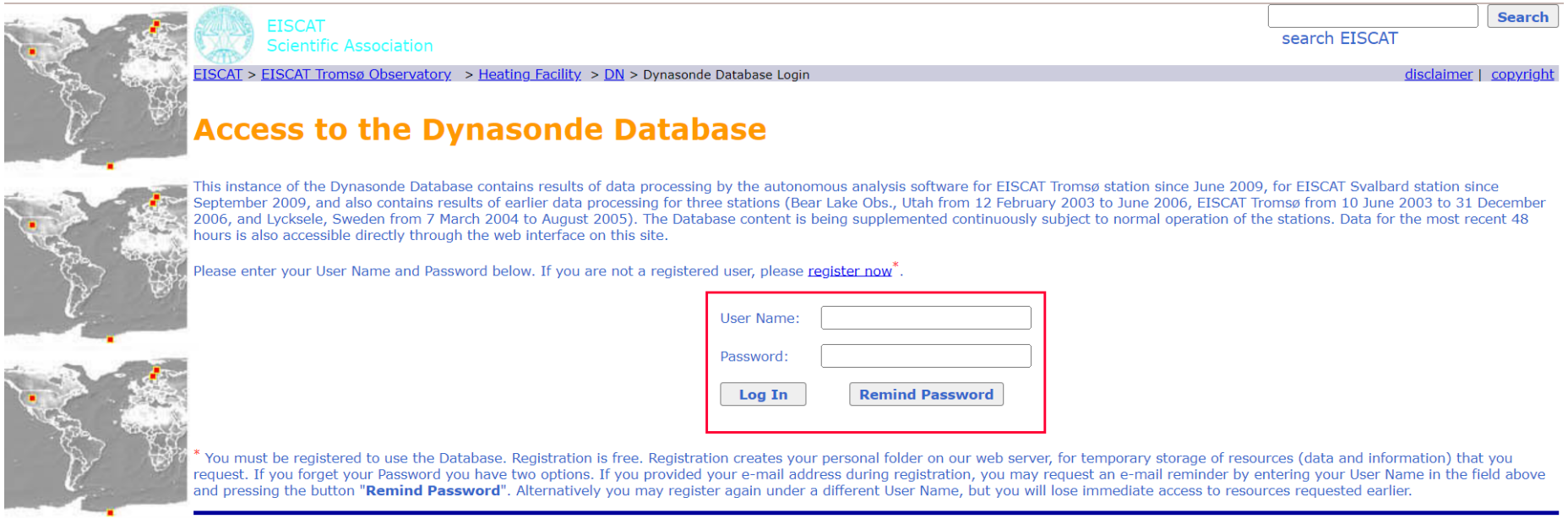
University of Westminster

TPW5 2023, London, UK

12-13 September 2023

What do we mean by integration level?

- Reminder: Data Collection = dataset or (executable) model
- When you publish a Data Collection (upload the XML files) you already integrate with the eSC
  - Your Data Collection can be accessed from the eSC
- However: This is simply a link taking the user outside the eSC



What do we mean by integration level?

- **We planned 4 integration levels with the eSC**
  1. Provide an external link
  2. Embed an automatically generated GUI into the eSC via API
  3. Dynamically deploy and execute model in the cloud
  4. Provide install package and instructions for local deployment and execution
- One Data Collection can be integrated in multiple ways

Which one to choose?

1. Provide an external link

- This is a natural part of the registration process – “linkage”
- Pros:
  - No additional effort
  - Data Collection can be found in the eSC
- Cons:
  - Not really integration – takes the user outside the eSC
  - Most of the Data Collections are currently integrated in this way

Whenever possible and justified, we must go beyond this

1. Provide an external link - examples

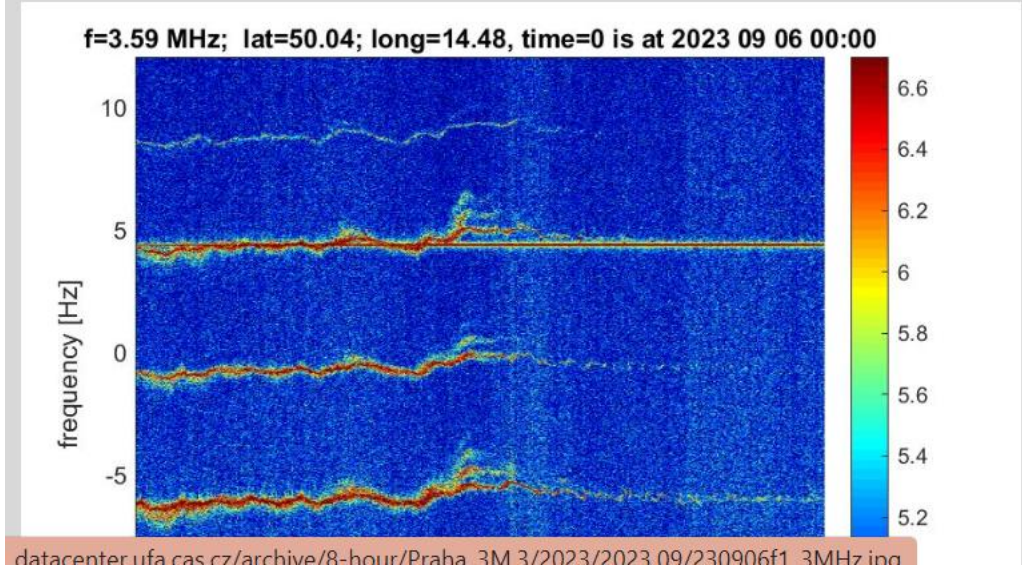
Doppler sounding of the Ionosphere - latest Doppler shift spectrograms

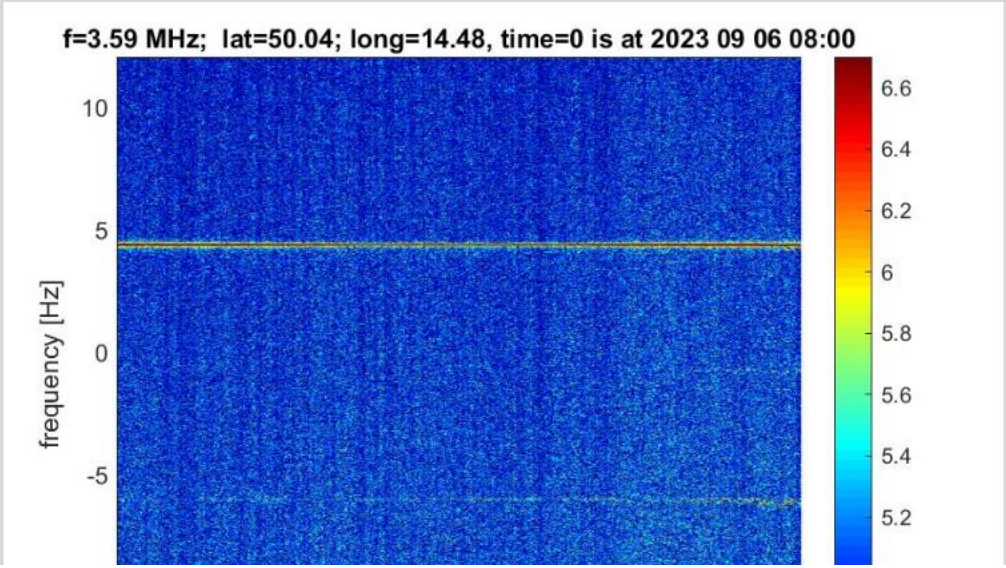
Spectrogram archive 8-hour spectrograms 2-hour spectrograms

Data from Tucuman were obtained in the collaboration with Universidad Nacional de Tucuman (Laboratorio de Telecomunicaciones).
Data from Hermanus were obtained in the collaboration with SANSA, Hermanus, South Africa.
Data from Taiwan (Zhongli,Puli) were obtained in the collaboration with National Central University, Jungli City, Taiwan.

Quick navigation: Prague 3.59 MHz Prague 4.65 MHz Prague 7.04 MHz Kasperk_3.59 MHz Kasperk_4.65 MHz Oberpfaffenhofen_4.65 MHz Tucuman 4.63 MHz Tucuman 6.80 MHz Taiwan-Puli Taiwan-Hsinchu Hermanus France Belgium Slovakia

3.59 MHz Prague





datacenter.ufa.cas.cz/archive/8-hour/Praha_3M.3/2023/2023.09/230906f1_3MHz.jpg

2. Embed an automatically generated GUI into the eSC via API

- You need to have an API for your Data Collection
  - Many of you have it
  - In limited cases UoW can help develop one (not scalable)
  - Your internal technical team can also help you to develop one as part of the project
- Data Collection needs to be accessible by API calls
  - It can be running on any server (your own, EGI, anywhere) until we can access the API endpoint
  - **BUT:** We have no expertise in your scientific software and for us it could be complicated to deploy it (e.g. MCM model)

2. Embed an automatically generated GUI into the eSC via API

- API-based solution works for both Datasets and Models
- API needs to be described in a JSON file – there are automated tools to do this, support and manual are available
- GUI will be generated automatically from this JSON
- Publication in eSC is simply providing a link to the JSON file

2. Embed an automatically generated GUI into the eSC via API

- Pros:
  - Much better integration with eSC – same look and feel for all Data Collections
  - Relatively simple if you already have an API
  - In the case of simple models, writing the API is not a big challenge
- Cons:
  - The automated tools available have limitations (e.g., automated filenames, problem with zip files, problem with complex structures)
  - Initial deployment of the Data Collection is your responsibility (EGI, local or anywhere)

2. Current state of the process

- One API is running and demonstrated several times (DTM2020)
- Several APIs are deployed as first prototypes and under improvement
- Several APIs are under specification and development

2. Current state of the process

Inst.	Onwer	Data Collection	In eSC	Hosted by	API Status	Comments
AERONOMIE	Edith Botek	BPIM	NO	UOW at EGI	development in progress	BPIM is deployed. API design has been agreed and is under development.
CNES	Sean Bruisma	DTM2020-operational	YES	UOW at EGI	production version deployed	Plotting requires smoothing. Workaround has been applied for UI limitations.
CNES	Sean Bruisma	MCM	NO	UOW at EGI	development in progress	Issues with MCM libraries. API is ready but not tested.
INGV	Emanuele Pica	wsstation	NO	INVG	development in progress	Use of a built-in API has been explored. A custom API will be provided.
NOA	Anna Belehaki	TechTIDE LSTID activity index	YES	NOA	development in progress	Issues with the current API have been resolved. An API will be created for the eSC.
NOA	Anna Belehaki	NOA Athens Digisonde (AT138) Data	YES	NOA	prototype deployed	Issues with the current API have been resolved. An API is created for the eSC.
NOA	Anna Belehaki	SWIF Model	YES	NOA	prototype deployed	Issues with the current API have been resolved. An API is created for the eSC.

2. Current state of the process

Inst.	Onwer	DataCollection	In eSC	Hosted by	API Status	Comments
OBSEBRE	Toni Segarra	hmF2_qModel	YES	OBSEBRE	prototype deployed	UI limitations with plots and zip files. Fixes will be applied.
OBSEBRE	Toni Segarra	B0B1_qModel	YES	OBSEBRE	prototype deployed	UI limitations with plots and zip files. Fixes will be applied.
OBSEBRE	Toni Segarra	EPB_detectionTool	YES	OBSEBRE	prototype deployed	UI limitations with plots and zip files. Fixes will be applied.
DIAS	Shane Maloney	SolarMonitor2	-	DIAS	prototype deployed	API is being finalised and planned for registration.

3. Dynamically deploy and execute a model in the cloud

- This is a really cool and modern way of doing it
- Model is not running permanently only when it is needed
- User can deploy the model by one click, run it and then free up the resources when done
- Newly developed models are likely to fit the requirements of this solution well (but not old Fortran models)

This solution is not available yet

3. Dynamically deploy and execute model in the cloud

- Pros:
  - Full integration with the eSC and EGI
  - Model deployed this way becomes easily portable and executable anywhere
  - We only use resources that are really needed
  - Scalability – resources can easily scale up or down
  - We can parallelise runs of a model in multiple instances
  - We can have personal running instances of a model

3. Dynamically deploy and execute model in the cloud

- Cons:
  - This requires the most effort on both sides (eSC developers and providers)
- Requirements on provider side:
  - Model needs to run in a Docker container – so far, we have not met a model that fulfils this
- Requirements on eSC side:
  - We have some development to do – but this is the core focus as our group anyway

We already prototyped the STIM model this way and it is feasible

4. Install-package for local deployment

- This is for relatively simple models that can be deployed and run on a local computer
- Provider uploads an install package and detailed instructions to the eSC
- User downloads it and installs on local machine
- Pros:
  - Very simple to realise
- Cons:
  - Limited integration with eSC
  - Limited possible application

4. Install-package for local deployment

- Is this a potential candidate?

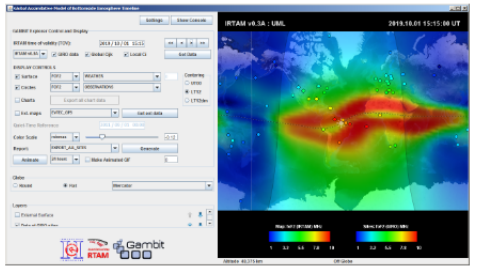
GIRO | Ionogram Data | Plasma Drift Data | TID Data | GAMBIT Weather Maps | Radio Link Evaluation | Examples | Info

Open-Data Guest Access to GAMBIT Database

Open data access to retrospective ionospheric weather data in display format:

Download : GAMBIT Explorer UserApp 1.3E

Please use versions 1.3 and later (MySQL 8 database, "best available" search capability).



- GAMBIT-X UserApp Installation Instructions
- GAMBIT-X User Guide version 1.0
- No need to contact us to obtain a guest account
- Open academic-use access for *organizations*
  - Send a message to Ivan Galkin to apply for your organization account.

Integration of GAMBIT data in research applications

Download IRTAM coefficients from LGDC to use in your investigation:

- Manual download: giro.uml.edu/rix/gambit-coeffs/
- URL-based scripted download:
  - Populate URL with date-time (ISO format) and name:
    - https://lgdc.uml.edu/rix/gambit-coeffs?time=yyyy.mm.ddThh:mm&charName=nnnn

Thank you for your attention!

WEB: https://www.pithia-nrf.eu

The PITHIA-NRF project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 101007599