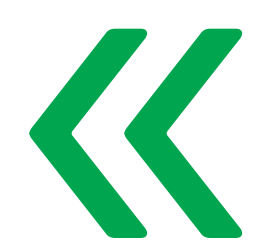


# Dissemination of Various Datasets via Mélodi and RMÉS



The Mélodi project aims to redesign our dissemination process in order to develop its open data supply. This process will result in significant improvements in terms of accessibility, by offering new open data services to users (including a catalogue, a data mining function to offer tables on demand, a mass downloading tool, options to print tables on demand and a major API covering all our disseminated databases). In view of this, producers will send their statistical data to Mélodi, which will then make them available to users via a wide range of services. All the datasets will be created by using the appropriate statistical metadata of our RMÉS repository, in a metadata-driven process approach. The complete standardisation of these metadata will enable better comparison of those results provided by different statistical operations.

Today, the dissemination of data involves...

**5 000**

downloadable files on insee.fr each year

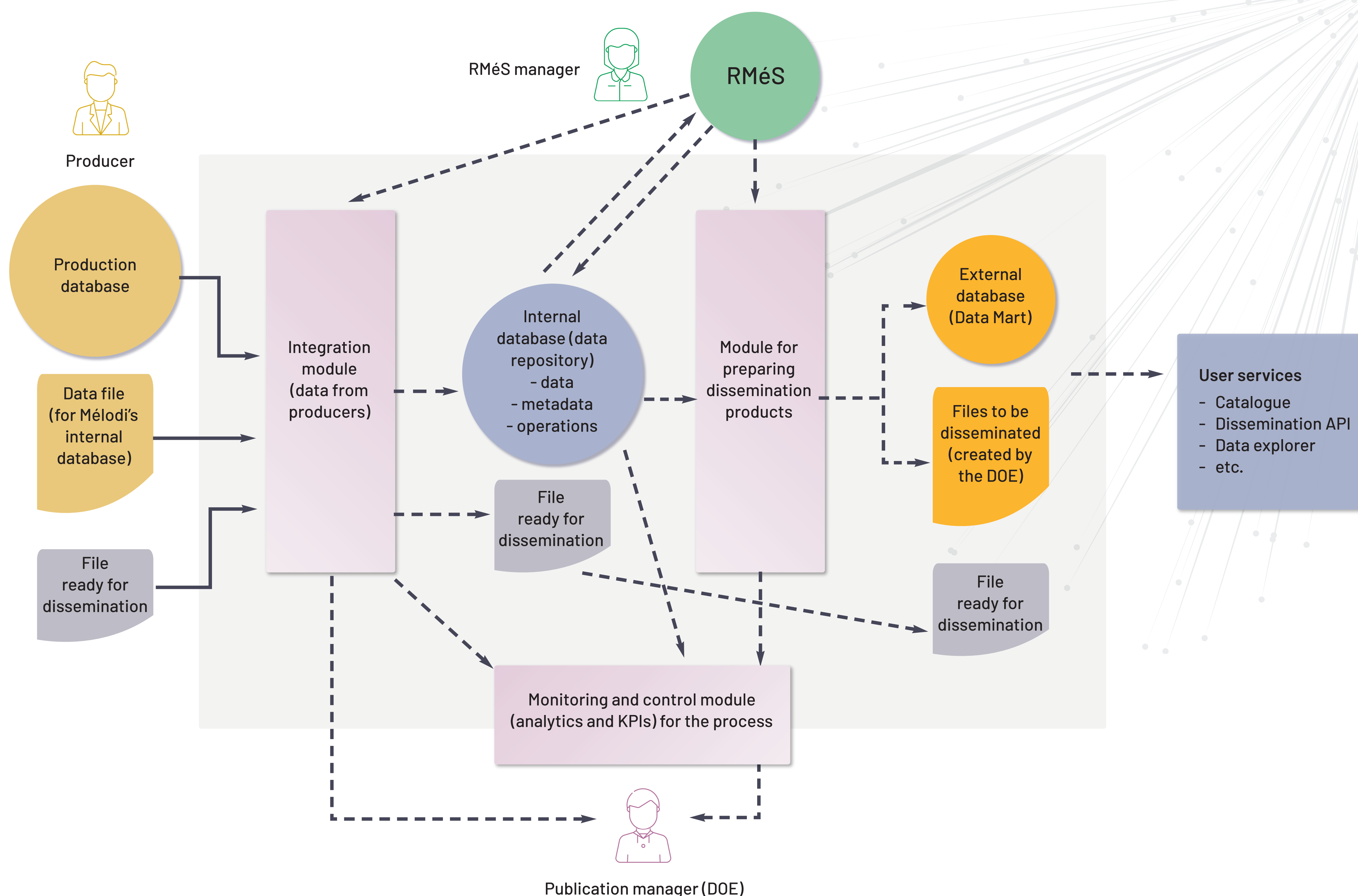
**+ 1 000**

files submitted to Eurostat each year

**150 000**

time series within the BDM

Jocelyne Mauguin, Head of the Mélodi statistics project



## The Mélodi project has two main objectives:

- Externally, to improve the service provided to users of our data by offering new Open Data services and by improving the quality of the data disseminated.
- Internally, to rationalise the dissemination of statistical data to achieve the external objective.

## For all users:

- the catalogue, which is a flagship product within the range and which draws the user's attention to the existence of dissemination datasets, while also improving natural referencing within search engines as part of the process;
- the data explorer to allow the user to extract the data sample of their choice with a view to preventing over-consumption of data and to allowing data mining of very large files;
- the statbot (which the project only plans to test) returning automatic answers to user questions concerning key figure requests, in order to make this information more immediately accessible.

## For the most technical users, with a view to reusing the data and positioning INSEE on upgraded dissemination channels:

- the roll-out of open data formats to facilitate the use of these files and to move up the international open data rankings (xlsx and csv formats rather than xls);
- a mass downloading tool, which avoids the user having to click repeatedly to access the site and which reduces the burden on the site by reducing the amount of time spent searching through it;
- a dissemination API hosted on the APIs' portal that promotes the reuse of data through a machine-to-machine dialogue (redissemination or integration into other ISs);
- finally, the provision of certain data as Linked Open Data for users who are experts in this field and to position INSEE in this niche.

And in order to feed into these services, a data repository in which the operations are driven by metadata from RMÉS.

**RMÉS**

Statistical Metadata Repository