



Contribution ID: 150

Type: **Presentation**

Develop data-centric web apps in Jupyter with Voilà and VOIS

Tuesday 12 March 2024 10:20 (20 minutes)

After more than four years of experience in developing dashboards with Jupyter and Voilà [1] and the development of a library that simplifies the creation of user interfaces for compelling interactive visualization, we can say that yes: it is possible to use Jupyter as an advanced development environment for the creation of complex web applications, centred on data and equipped with a simple but modern user interface and perfectly capable of supporting interactive and integrated exploration on multiple datasets.

What we have done in this period at the Joint Research Centre of the European Commission, within the on-premise cloud infrastructure called **BDAP** [2] [3] (Big Data Analytics Platform), was to support our fellow scientists and researchers with advanced storage services for big data, parallel and GPU processing, data analysis and Machine Learning/Deep Learning.

In the final phase of many of the collaborations we have established with the scientific units of the JRC, the communication of research results has gradually taken a more important role. Whether it is to inform colleagues and collaborators, as well as to communicate results to the political DGs of the Commission, to publish in scientific journals or to present research work to an audience of non-experts, over time the creation of interactive tools that allow autonomous exploration of the data and results generated by research have met with growing success.

We started from simple dashboards where basic graphs allowed us to summarize the concepts being researched, and then gradually added geo-spatial visualizations with the aim of geographically framing the phenomena studied, to finally arrive at complex applications that can be used via the web for interactive communication and the exploration of multiple and interconnected datasets. Geographical, textual and numerical data integrate with customized charts and visualizations created ad-hoc to best describe the object of study and the results obtained. Very often, these applications accompany and complete the publication of the research in scientific journals, providing a tool, open to all, which allows for the interactive and autonomous verification and exploration of the study results. Moreover, we recognise the importance of Jupyter in this process. As it is the standard environment vastly used by data scientists and researchers from all over the world for the interactive exploration and analysis of data, being able to use the same framework also for the communication of research results, it undoubtedly guarantees enormous advantages in terms of continuity and simplicity. Researchers, data scientists and software developers can freely move from the focus of the research to its final communication while always remaining within the same work environment.

The creation of Voilà in 2019 enabled Jupyter notebooks to be automatically transformed into interactive dashboards, opening up the possibility of using Jupyter as an application development environment. The availability of widget libraries (ipywidgets [4] initially, then followed by ipyvuetify [5] and others) made it possible to insert graphic elements into dashboards that were directly connected to the data and allowed guided exploration. In this context, with the experiences made in BDAP, we began to collect code snippets that could be generalized and reused, until they became the core of a Python library. This is how the **VOIS** (VOilà Simplification library) library was born, which a few months ago we managed to make Open Source and publish in the code.europa.eu repository of the European Commission [6] [7].

Through the VOIS library, we aim to simplify the development of data visualization applications and support the rapid creation of the application's visual infrastructure by making available a very vast catalogue of reusable and customizable graphic widgets. The library has been enriched over time with tools to:

- manage responsiveness in a simplified way (i.e. to automatically adapt the display to the different graphic resolutions of the screens, targeting smartphones, tablets and desktop screens of any resolution),

- create multi-page applications (where the user can navigate through different pages that alternate on the screen and present distinct and complementary functions),
- add modal dialog boxes that allow forms and graphic views to be superimposed on the background page allowing for a modern interaction similar to desktop applications.

All this is accompanied by specific modules to manage the subdivision of space on the page, functions that allow for the download/upload of data between the user's PC and the application that runs on the cloud, a gallery of interactive charts created in SVG [8] to fill any gaps present in the open source charting libraries, a centralized management of colours and themes that allows for easy control of the graphic coherence in all parts of an application.

All this has allowed us to move from the creation of dashboards to the creation of real web applications, with multiple windows and complex integrated functions, to provide the user with advanced visualizations and highly interactive experiences.

It is also interesting to note that the recently introduced VaaS service (Voilà as a Service) enables BDAP users to autonomously create and deploy their dashboard in production, through an automated procedure based on Gitlab repositories. This new service, together with the intensive training on the usage of the VOIS library, is contributing to the spreading of Voilà dashboards usage by many research groups in the JRC.

With this presentation, we intend to provide a review of the main applications that we have developed in recent years using Jupyter, Voilà and the VOIS library and which touch on a great variety of different scientific fields, from citizen science to earth observation, from the analysis of European agriculture to air quality monitoring. Moreover, in showing these applications, we also want to talk about the difficulties encountered, the obstacles overcome and the lessons learned, providing ideas on how to best use Jupyter as a development environment for complex web applications. Jupyter was not born with a focus on software development, but on data analysis. Nevertheless, with the experiences that we have had and the guidelines we have derived from it, we want to demonstrate that it also performs reasonably well as a development and testing environment for data-centric web applications.

The development of the VOIS library was partially funded by the JRC's participation in the Horizon2020 project called **CS3MESH4EOSC** and led by CERN which ended in 2023 [9].

[1] <https://voila.readthedocs.io/>

[2] P. Soille, A. Burger, D. De Marchi, P. Kempeneers, D. Rodriguez, V. Syrris, and V. Vasilev. "A Versatile Data-Intensive Computing Platform for Information Retrieval from Big Geospatial Data". *Future Generation Computer Systems* 81.4 (Apr. 2018), pp. 30-40.
<https://doi.org/10.1016/j.future.2017.11.007>.

[3] D. De Marchi, A. Burger, P. Kempeneers, and P. Soille. "Interactive visualisation and analysis of geospatial data with Jupyter". In: *Proc. of the BiDS'17*. 2017, pp. 71-74.
<https://zenodo.org/record/3248741#.XeDvSuhKg2w>.

[4] <https://ipywidgets.readthedocs.io/en/latest/>

[5] <https://ipyvuetify.readthedocs.io/en/latest/>

[6] <https://vois.readthedocs.io/>

[7] <https://code.europa.eu/jrc-bdap/vois>

[8] <https://developer.mozilla.org/en-US/docs/Web/SVG>

[9] <https://cs3mesh4eosc.eu/>

Primary author: DE MARCHI, Davide (European Commission - Joint Research Centre)

Co-authors: BURGER, Armin (European Commission - Joint Research Centre); SOILLE, Pierre (European Commission - Joint Research Centre); KEMPENEERS, Pieter (European Commission - Joint Research Centre)

Presenter: DE MARCHI, Davide (European Commission - Joint Research Centre)

Session Classification: Collaborative Data Science and Visualisation

Track Classification: Main sessions: User Voice: Innovative Applications, Data Science Environments & Open Data