P16

Sovereign digital commons for IA

Inria

FRANCE 2030

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Our mission

P16
Sovereign digital commons for IA

« Support the development and deployment of a **sovereign**, **open**, and **interoperable** platform of AI software libraries for French companies and scaling it up to the European level. » - *Measure No. 16 of the National Strategy for AI*

Inria

RÉPUBLIQUE FRANÇAISE
Liberté
Égalité
Fraternité

Our mission

P16
Sovereign digital commons for IA

Develop

Develop or support the development of software libraries

Integrate

Ensure seamless interoperability between libraries

Sustain

Maintain these commons at the state of the art in science and technical practice.

Disseminate

Promote adoption in academia and industry

Our setup

Inria

- **Identification** and assessment of libraries
- **Pre-industrialization** of libraries
- Development of **use cases** and business applications with **industry partners**
- **Scientific collaboration** and research outreach

:probabl.

- Mission-driven company for **Industrial and Digital Sovereignty**
- **Industrialization** of libraries
- Business opportunities
  - **Products**
  - **Professional services**
  - **Training** and **certification**

How we work with :Probabl.

P16
Sovereign digital commons for IA

:probabl.

Investigation unit
- Scientific and Technical Monitoring
- Ecosystem Manager
- Software Architect

Development unit
- OSS Engineers
- Software Architect

Use case unit
- Data scientists
- Ecosystem Manager

Community unit
- OSS engineers
- Ecosystem Manager
- Software Architect

Product

Ensure the optimal experience for data scientists in experimenting and building models.

Labs

Develop, maintain, sustain open source projects and communities in data science

Services

Provide training, certification and expert solutions for enterprise AI challenges and applications.

P16
Commons for AI

Inria

Our OSS commitment

P16
Sovereign digital commons for IA

Collective innovation

Build robust and scalable solutions together

Transparency

Develop mechanisms for traceability and compliance with ethical and regulatory standards

Customization

Provide flexible tools that can be adapted to users' needs

Training

Disseminate the knowledge necessary for tools use and development

Our technical actions

P16
Sovereign digital commons for IA

Data interoperability

- Standardization and distributed architecture prototyping
- Publication and exchange of data via open web standards
- Infrastructure to ensure the sustainability and independence of stakeholders

Data Wrangling

- Formatting data for AI, identified as a bottleneck by data scientists
- Development of a library for unified, flexible, and efficient data transformation, exploration, and quality measurement

Machine learning toolbox

- Continuous innovation and state-of-the-art relevance
- Support for different data formats (e.g., time series)
- Transparency and explainability of models
- Performance optimization

Our north star

scikit-learn is one of the most popular open-source libraries for machine learning, widely used in both academic research and industry.

Impact

- ✓ Trusted by thousands of developers & researchers
- ✓ Backbone of ML workflows in academia & industry
- ✓ Extensive contributions & continuous improvements

- **Stars:** 61.4k
- **Forks:** 25.7k
- **Contributors:** 2,900+
- **Commits:** 49,000+
- **Releases:** 40+
- **Dependent Projects:** Over 1 million
- **Dependent Packages:** Over 20,000

Features

- ❑ **Classification:** SVM, Random Forest, k-NN...
- ❑ **Regression:** Linear Regression, Ridge, Lasso...
- ❑ **Clustering:** k-Means, DBSCAN...
- ❑ **Dimensionality reduction:** PCA, LDA...
- ❑ **Model selection:** GridSearchCV, cross-validation…
- ❑ **Data pre-processing:** normalization, encoding, missing value imputation…

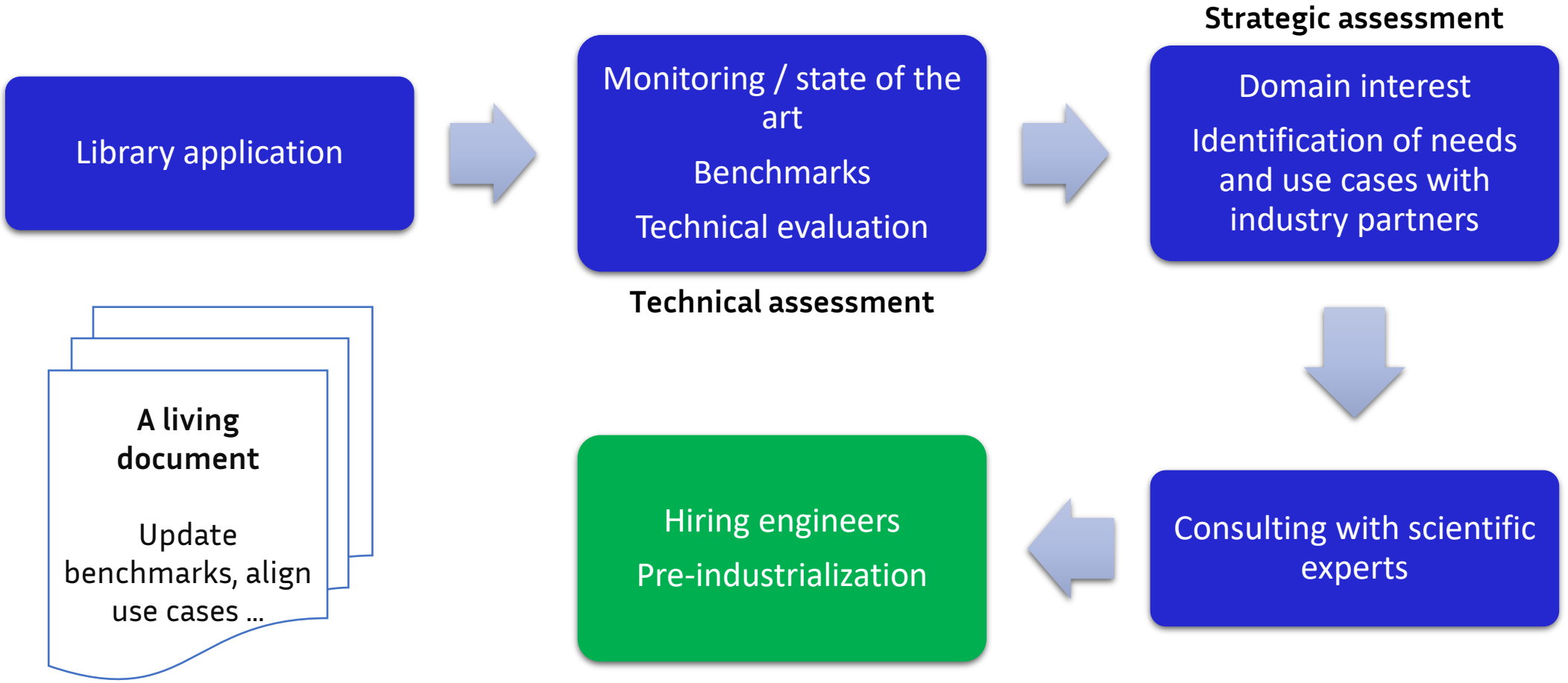
Our call for interest

P16
Sovereign digital commons for IA

Integration into the data cycle	Interest for digital sovereignty	Matching the needs of the ecosystem
demonstrate how the library contributes to the overall vision of P16 and can integrate or extend one of the technical actions (interoperability, preparation, learning)	demonstrate how the library contributes to strengthening the national and European position in the field of AI (competitive landscape, governance ...)	demonstrate how the developed solutions can meet specific use cases, targeting identified industrial needs.

+ **Generic use**

Libraries selection



[Application processing worflow]

Life of supported libraries

skrub	Data cleaning and preparation
CORESE	Data interoperability and knowledge graphs
tslearn	Time series clustering and classification
MAPIE	Uncertainty quantification
aeon	Time series toolbox
OpenHosta	Machine Learning / LLM Hybridization
shapash	Machine learning explainability

Technical Development

- Dedicated support to structure, industrialize and enrich the library.
- Support for open-source community management.

Dissemination

- Participation in events to increase visibility.
- Organization of workshops and hackathons to bring together users and contributors.

Training

- Setting up training to meet the needs of potential users.
- Development of pedagogical modules in collaboration with initiatives such as Inria Academy.

Ecosystem collaboration

P16
Sovereign digital commons for IA

1

Encourage adoption and contribution

- Participate in workshops, webinars, and hands-on sessions
- Continuous education program development with Inria Academy

2

Identify AI needs

- Tailor our work to the needs of our partners
- Large ecosystem consultations

3

Co-develop use cases

- Demonstrate, collectively, the relevance of P16 libraries

EDF · Capgemini · SAFRAN AEROSPACE · DEFENCE · SECURITY · framatome · GROUPE SEB · life.augmented · GROUPE VYV · MAIF

Inria

Thank you for your attention!

Any questions?

1 **Contact**

programme-ia.p16@inria.fr

2 **Website**

p16.inria.fr

P16
Sovereign digital commons for IA

Inria