Snoop is a **framework** dedicated to **large-scale content-based image retrieval**. Co-developed by Ina and Inria, it is the visual search engine used by **PlantNet** applications.

The GallicaSnoop project studies the application of Snoop to digitized cultural heritage collections.

---

**SNOOP’S MAIN FEATURES**

- the extraction and **efficient indexing of visual features** (hand-crafted or learned through deep learning)
- the search of similar images through **approximate k-nearest neighbors**
- the supervised recognition of **trained visual concepts**.

---

**CHALLENGES**

- Allow the user to **interactively define visual classes** (human in the loop)
- Search in large scale collections with **compact representation** (hash) of descriptors and **high performances index and search** (query > 20 ms)
- Quality control on the **visual similarity parameter**
- **Plugins architecture** to adapt to any use case

---

**RESSOURCES**

- Images dataset from **Gallica (1.2 million)** harvested with **IIIF protocol**
- 2 **case studies**:
  - Recognition of reproduction of agency photos in the press
  - Human-in-the-loop approach

---

**WHAT’S NEXT ?**

- GallicaSnoop as a service at the BnF DataLab (2022)
- Snoop as a shared digital system within cultural institutions

---

Funded by the framework agreement between the Ministry of Culture and INRIA (2019-2020)

Pour tester GallicaSnoop : [https://snoop.inria.fr/bnf/](https://snoop.inria.fr/bnf/)

Utilisateur : Mot de passe : QR Code