

Sí — **hay papers y recursos que pueden servir como “roadmap académico” o guía sólida para un proyecto de meta-learning**, aunque no existe una única fuente que cubra *todo* de manera completa (desde extracción de datos hasta recomendación de algoritmos con código listo). Lo que sí hay son **estudios, datasets, código y repositorios que puedes usar como referencias concretas para cada parte del roadmap**.

Aquí te dejo una lista útil de **papers, recursos y repositorios** relevantes:

Papers académicos que funcionan como roadmap parcial


Meta-features y meta-learning con OpenML

1. **Bilalli et al. (2017): On the predictive power of meta-features in OpenML** – Estudia cómo los meta-features extraídos de OpenML ayudan a predecir el rendimiento de algoritmos. – Analiza la relación entre meta-features y performance de algoritmos y usa ese conocimiento para meta-learning. ([Paradigm](#))

 Útil para entender cómo seleccionar y evaluar meta-features para tu proyecto.


Meta-learning aplicado a AutoML dinámico

2. **Learning meta-features for AutoML (ICLR 2022)** – Propone nuevos meta-features aprendidos y los usa para mejorar sistemas AutoML. – Usa el benchmark OpenML CC-18 como caso de estudio. ([paperswithcode.com](#))

 Te da estrategias modernas de cómo ampliar tu meta-feature set además de usar los convencionales.

Meta-datasets reales extraídos de OpenML

3. **OpenML Study 7 — meta-datasets (Mendeley / OSF)** – Conjunto de meta-datasets reales creados a partir de OpenML. – Incluye evaluaciones de muchos algoritmos en muchos datasets y descriptores básicos. ([data.mendeley.com](#))

 **Documento casi perfecto como ejemplo de cómo construir meta-datasets** — incluso puedes descargarlos directamente y reproducir su construcción.


Otras aplicaciones de meta-learning

(útiles para inspiración aunque no directamente con OpenML)

- **Meta-QSAR**: aplicación a diseño de fármacos usando meta-learning con recursos públicos de OpenML. ([arXiv](#))


Código y repositorios relevantes

GitHub con papers y enlaces de meta-learning


 **Meta-Learning-Papers-with-Code** Repositorio con listas de **papers + código** sobre meta-learning y meta-RL. Puede ayudarte a encontrar implementaciones concretas de métodos de meta-learning. ([GitHub](#))


➡ No está centrado en AutoML/OpenML, pero es útil para entender los enfoques que existen y ver código de referencia.

OpenML repositorio oficial

 **openml/OpenML (GitHub)** Código de OpenML mismo (el backend y APIs). Útil si quieres entender cómo se estructuran los datos, runs, evaluations y meta-data en OpenML. ([GitHub](#))

Otras herramientas útiles

 **AutoML Toolkit (amltk)** Incluye utilidades para meta-learning como extracción de meta-features, portfolios, selección de configuraciones, etc. ([automl.github.io](#))

 **SmartML (R)** Un paquete de meta-learning con base de conocimiento (knowledge base) que va recolectando información para seleccionar algoritmos automáticamente. ([rdrr.io](#))

Cómo usar todo esto como “roadmap académico”

Puedes estructurarlo así:

♦ 1) Bases teóricas / literatura

- Lee la colección de papers de meta-learning con código en el repositorio de Meta-Learning-Papers-with-Code. ([GitHub](#))
- Estudia cómo los meta-features (OpenML) realmente predicen performance (p. ej., Bilalli et al.). ([Paradigm](#))
- Revisa trabajos más recientes como *Learning meta-features for AutoML*. ([paperswithcode.com](#))

♦ 2) Datos y benchmarks

- Descarga y explora meta-datasets reales del estudio OpenML Study 7. ([data.mendeley.com](#))

♦ 3) Construcción de meta-dataset

- Basarte en la descripción de cómo esos meta-datasets fueron formados (dataset descriptors + evaluaciones). ([osf.io](#))

♦ 4) Modelado y evaluación

- Implementar modelos de ranking o predicción de performance (kNN, RF, XGBoost).
- Comparar con trabajos académicos para validar.

♦ 5) Código de producción

- Integrar extracción de meta-features (amltk, OpenML API) con un pipeline que produzca meta-dataset final.

- Opcional: usar repositorios con implementaciones de meta-learning si expandes hacia aprendizaje few-shot o RL.

Recomendación final

No existe aún un **único repositorio que contenga TODO el pipeline académico + código listo para meta-learning con OpenML**, pero **sí existe literatura + herramientas + datasets que juntos forman un roadmap académico válido** para tu proyecto.

Si quieres, puedo darte:

♦ una **lista de papers más citados en meta-learning y meta-feature extraction**, ♦ una **plantilla de código reproducible** para construir un meta-dataset como hacen en esos estudios, ♦ o ayudarte a **comparar diferentes enfoques de meta-learning** side-by-side.

¿Cuál prefieres primero?