الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
المدرسة العليا للإعلام الآلي 80 ماي 5491 - سيدي بلعباس

People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Higher School of Computer Science – May 8, 1945 – Sidi Bel Abbès

Graduation Thesis

In partial fulfillment of the requirements for the degree of State Engineer in Computer Science

Specialization: Information Systems and Web

Topic

Modernizing Project Management in EarthPortal

A Semantic and Federated Approach to Metadata Integration and User Interface Design

---

Presented by

Hakim ALLEM

Defended in: **10 March 2025**

In front of the jury composed of:

Mr. Mohamed Yacine Kazi Tani	President of the Jury
Mr. Mohamed Bedjaoui (ESI-SBA)	Academic Supervisor
Mrs. Christelle Pierkot (Data Terra)	Professional Supervisor
Mrs. Khayra Benchrif	Examiner

Abstract

Modern scientific data platforms face the challenge of managing large volumes of heterogeneous project metadata while preserving interoperability and semantic consistency. *Problem statement:* how can we reliably integrate and semantically enrich project information coming from diverse, structurally incompatible sources within a single platform?

Methodology. We (1) studied and selected relevant external sources; (2) designed dedicated connectors to fetch and normalize project records; and (3) mapped key fields to a unified project model aligned with Linked Data principles.

Within EarthPortal, two external data sources, **CORDIS** and **ANR**, were integrated through these connectors, enabling ingestion of funding data, project duration, keywords, and participating organizations. The redesigned project model supports ontology references, funding agency identifiers, and additional metadata, while the enhanced UI facilitates project creation, filtering, and categorization by domain and funding.

These contributions strengthen the platform's capacity to deliver structured, reusable, and searchable information on funded research projects across Europe and beyond.

Keywords— Project integration, Research metadata, Linked Data, Semantic enrichment, CORDIS, ANR, EarthPortal

Résumé

Les plateformes scientifiques modernes doivent gérer de grands volumes de métadonnées de projets hétérogènes tout en garantissant l'interopérabilité et la cohérence sémantique. *Problématique :* comment intégrer de manière fiable et enrichir sémantiquement des informations de projets issues de sources diverses et structurellement incompatibles au sein d'une même plateforme ?

Méthodologie. Nous avons (1) étudié et sélectionné des sources externes pertinentes ; (2) conçu des connecteurs dédiés pour extraire et normaliser les enregistrements ; et (3) mappé les champs clés vers un modèle unifié aligné sur les principes du Web de données.

Dans EarthPortal, deux sources externes, **CORDIS** et **ANR**, ont été intégrées via ces connecteurs, permettant l'ingestion des informations de financement, de la durée, des mots-clés et des organisations impliquées. Le modèle de projet repensé intègre des références aux ontologies, des identifiants de financeurs et d'autres métadonnées, tandis que l'interface utilisateur facilitée prend en charge la création, le filtrage et la catégorisation par domaine et financement.

Ces contributions renforcent la capacité de la plateforme à fournir des informations structurées, réutilisables et interrogeables sur les projets de recherche financés en Europe et au-delà.

Mots-clés— Intégration de projets, Métadonnées de recherche, Web de données, Enrichissement sémantique, CORDIS, ANR, EarthPortal

Arabic abstract

تواجه المنصات العلمية الحديثة تحدّي إدارة كميات كبيرة من بيانات المشاريع غير المتجانسة، مع الحفاظ على قابلية التشغيل البيني والتناسق الدلالي. إشكالية البحث: كيف يمكن دمج معلومات المشاريع القادمة من مصادر متعددة ومختلفة البنية وإثراؤها دلالياً داخل منصة واحدة بشكل موثوق؟

المنهجية. اعتمدنا على: (1) دراسة واختيار مصادر خارجية مناسبة؛ (2) تصميم موصلات خاصة لاستخراج سجلات المشاريع وتوحيد بنيتها؛ و(3) ربط الحقول الأساسية بنموذج موحّد للمشاريع متوافق مع مبادئ بيانات الويب المترابطة.

ضمن منصة **إيرث بورتال (EarthPortal)**، تم دمج مصدرين خارجيين هما **CORDIS** و**الوكالة الوطنية للبحث الفرنسية (ANR)** عبر هذه الموصلات، مما أتاح استيراد معلومات التمويل، مدة المشروع، الكلمات المفتاحية، والجهات المشاركة. كما تم إعادة تصميم نموذج المشروع لدعم الإحالات إلى الأنطولوجيات ومعرّفات جهات التمويل وبيانات وصفية إضافية، مع تحسين واجهة المستخدم لتسهيل إنشاء المشاريع وترشيحها وتصنيفها حسب المجال والتمويل.

تُعزّز هذه المساهمات قدرة المنصة على تقديم معلومات منظّمة وقابلة لإعادة الاستخدام والبحث حول المشاريع المموّلة في أوروبا وخارجها.

الكلمات المفتاحية— تكامل المشاريع، البيانات الوصفية للبحث، بيانات الويب المترابطة، الإثراء الدلالي، CORDIS، ANR، EarthPortal