

# La piattaforma IEMAP (Italian Energy Materials Acceleration Platform): Il ruolo dei dati aperti e condivisi

Claudio Ronchetti, Sergio Ferlito, Marco Puccini, Simone Giusepponi, Francesco Buonocore, Massimo Celino, Sara Marchio, Giovanni Ponti

ENEA - Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile  
TERIN - Dipartimento Tecnologie Energetiche e Fonti Rinnovabili ICT - Divisione per lo Sviluppo di Sistemi per l'Informatica e l'ICT

**Abstract.** La piattaforma IEMAP è costituita da una infrastruttura di supercalcolo e una rete di laboratori e ha lo scopo di condividere dati per accelerare la ricerca su nuovi materiali per l'energia. I dati sperimentali e computazionali prodotti vengono raccolti in un DB dedicato presso l'infrastruttura di supercalcolo. Mediante un approccio multidisciplinare che sfrutta le tecnologie Big Data e AI, la condivisione ed analisi dei dati potranno accelerare la progettazione dei materiali più adatti per una data tecnologia energetica. Materiali per batterie, elettrolizzatori e fotovoltaico (tre aree di ricerca fondamentali per la transizione energetica) sono i primi casi d'uso della piattaforma, il cui motore è il supercomputer CRESCO sul quale le tecnologie HPC sono implementate per la gestione dei dati e lo sviluppo/porting di software per il modelling molecolare. Alla piattaforma partecipano vari laboratori sperimentali e computazionali dei quattro enti di ricerca aderenti: ENEA, CNR, RSE e IIT.

**Keywords.** Interoperabilità e accesso a dati e servizi, Servizi ICT a supporto delle infrastrutture di ricerca e istruzione, Utilizzo e condivisione di dati in infrastrutture di rete e servizi.

## Introduzione

Mission Innovation (MI) è un'iniziativa di cooperazione multilaterale con lo scopo di accelerare i processi di innovazione delle tecnologie pulite attraverso l'impegno dei paesi aderenti, tra i quali l'Italia, a raddoppiare la quota pubblica degli investimenti per le attività di ricerca, sviluppo e innovazione delle tecnologie per la decarbonizzazione al fine di rendere l'energia pulita accessibile ai consumatori e di creare posti di lavoro verdi e opportunità commerciali. In abito italiano sono coinvolti i quattro enti di ricerca: ENEA, CNR, RSE e IIT. L'iniziativa si articola in tre aree tematiche Smart Grid, Idrogeno e Materiali avanzati per l'energia attraverso i progetti: MISSION (Microreti e sistemi smart, multivettore ed integrati per accelerare la transizione energetica), HYDROGEN DEMO VALLEY (infrastrutture polifunzionali per la sperimentazione e dimostrazione delle tecnologie dell'idrogeno) e IEMAP (Piattaforma Italiana Accelerata per i Materiali per l'Energia). In quest'ultimo

progetto, il Dipartimento TERIN con la divisione ICT è direttamente coinvolto.

Il progetto ha come finalità la realizzazione di una piattaforma digitale che attraverso la condivisione ed elaborazione di grandi quantità di dati sperimentali e computazionali favorisca la progettazione accelerata e la selezione dei materiali avanzati per l'energia. Piattaforma multidisciplinare che in modo automatico, combinando tecnologie BigData e di AI, possa accelerare il processo di analisi dei dati computazionali e sperimentali al fine di identificare i materiali più adatti per una data tecnologia energetica. Materiali per l'accumulo elettrochimico-batterie, per gli elettrolizzatori e per il fotovoltaico, rappresentano aree di ricerca centrali per favorire il processo di transizione energetica. Verranno perciò semplificati quei processi, complessi e costosi, necessari per il design di nuovi materiali che comprendono la simulazione, la sintesi e la caratterizzazione con l'esecuzione di numerosi test.

Il motore di questa piattaforma è l'infrastruttura ENEAGRID e in particolare il supercomputer CRESCO, installato presso il C. R. ENEA di Portici, su cui si implementeranno tecnologie HPC sia per la gestione dei dati sia per lo sviluppo ed implementazione di codici numerici per il modeling molecolare. Alla piattaforma partecipano vari laboratori sperimentali e computazionali dei quattro enti di ricerca aderenti al progetto.

## 1. Gestione dei dati

Nelle fasi iniziali del progetto, una sfida importante è stata quella di definire un modello di metadato comune per i domini sperimentale e computazionale provenienti dai differenti laboratori e simulazioni. A questo scopo si è scelto di basare la definizione del modello dati sull'ontologia riportata in (Andersen 2021). Il modello logico è rappresentato da quattro entità (Fig. 1) che costituiscono le informazioni del materiale preso in analisi; a) il processo utilizzato per analizzare il materiale, b) i parametri che incidono sul processo, c) le proprietà chimico-fisiche ricavate dal processo e d) la provenienza del dato generato. A questo possono essere legati file di dati grezzi generati dal processo.



Fig. 1  
Rappresentazione schematica del metadato con le informazioni del materiale considerato: processo, parametri, proprietà e provenienza. Enti di ricerca coinvolti nel progetto

## 2. Piattaforma digitale

Al fine di raccogliere, gestire e immagazzinare dati eterogenei provenienti da contesti sperimentali e numerici, avendo come riferimento i principi FAIR (Findable, Accessible, Interoperable, Reusable) (Wilkinson 2016) è stata realizzata un'infrastruttura digitale (Fig. 2) composta da quattro elementi principali, ovvero:

1. Interfaccia Web (front-end);
2. REST API (Application programming interface);
3. NoSQL DataBase;
4. File Storage S3.

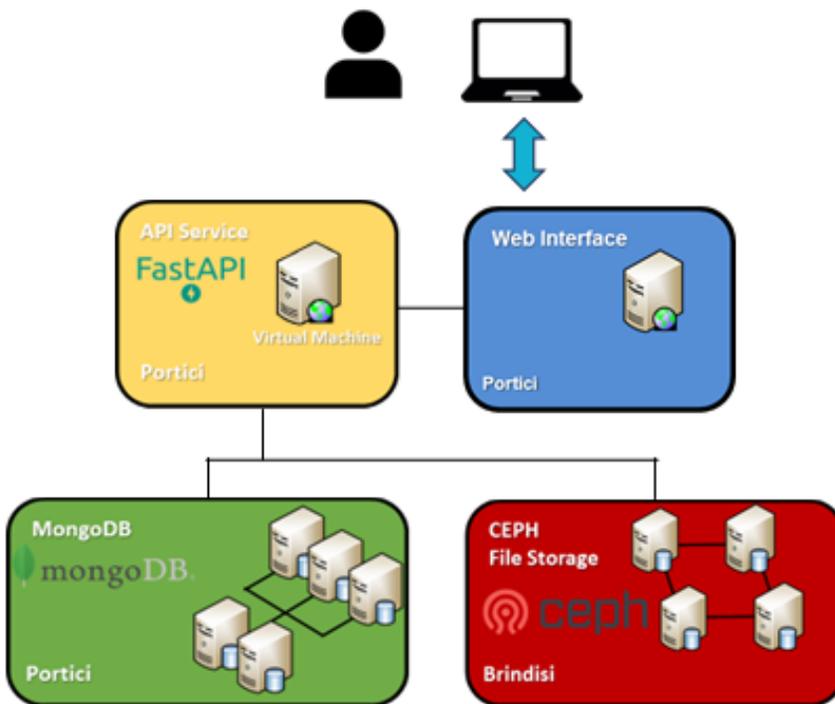


Fig. 2  
Rappresentazione schematica della piattaforma digitale: Web Interface, API service, MongoDB e CEPH File Storage

### 2.1 Interfaccia Web

L'interfaccia Web [iemap.enea.it] permette di rendere fruibili le funzionalità implementate nelle API agli utenti persone tenendo conto degli aspetti di usabilità e accessibilità. VUE 3 è la tecnologia utilizzata per l'interfaccia Web.

### 2.2 REST API

Le REST API sono indispensabili per la creazione e l'integrazione di software applicativi.

In questo caso forniscono diverse funzionalità legate agli accessi, al caricamento dei dati e alle interrogazioni sul DB. Le API rappresentano uno strato di sicurezza superiore per l'accesso ai dati e, inoltre, permettono interazioni sia con utenti macchina (direttamente) sia con utenti persone (indirettamente via interfaccia Web). Le REST API si basano su tecnologia FastAPI.

### 2.3 DataBase

Il Database NoSQL permette di immagazzinare i metadati (formato JSON) caricati dagli utenti partner della piattaforma e di effettuare query elaborate in modo efficiente. MongoDB è la soluzione adottata come database.

### 2.4 File Storage

Il File Storage S3 permette di conservare i file grezzi ottenuti dai processi che sono strettamente legati al metadato. La tecnologia per l'immagazzinamento dei dati grezzi è quella fornita da CEPH.

## 3. Caso d'uso

Il nostro gruppo sta svolgendo ricerca sui materiali per l'accumulo elettrochimico. Lo scopo è quello di trovare alternative alle batterie a base di ioni di Litio (LIB) con altre a base di ioni di Sodio (NIB) vista la sempre più limitata disponibilità di Litio (Li) a fronte della crescente richiesta di mercato. Invece, il Sodio (Na), essendo il 6° elemento chimico per disponibilità, è poco costoso e ampiamente distribuito rendendo le NIB una valida alternativa alle LIB.

Le NIB sono batterie ricaricabili analoghe alle LIB. Il catodo è un materiale allo stato solido contenente gli ioni alcalini, l'anodo è un materiale allo stato solido non contenente necessariamente gli ioni alcalini e l'elettrolita è una soluzione contenente gli ioni alcalini. In entrambi i catodi viene utilizzato il meccanismo di reazione di intercalazione. A causa dei loro elevati potenziali operativi e delle elevate capacità, i catodi a base di sodio hanno ricevuto crescente attenzione. Perciò, buone prestazioni e costi inferiori, insieme al fatto di poter utilizzare la stessa tecnologia e produzione già sviluppata per le LIB senza costi aggiuntivi, hanno suscitato un rinnovato interesse per questi sistemi.

Gli svantaggi al momento riguardano la densità di energia e la ciclabilità: nelle NIB la densità di energia è ancora inferiore rispetto alle LIB e la struttura geometrica può subire una distorsione Jahn Teller durante il ciclo di scarica/ricarica. Molti studi sono dedicati al miglioramento delle prestazioni delle NIB.

In particolare, si è visto che l'ossido metallico lamellare  $\text{NaMnO}_2$  in geometria P2 opportunamente drogato, mostra buone proprietà di ciclicità e capacità (Lee 2013, Wang 2013, Prosini 2019). In questo ambito si è definito un protocollo di calcolo ab-initio delle energie di formazione e dei potenziali redox (la prima è legata alla stabilità del composto e quindi alla ciclabilità, il secondo è legato alla densità di energia) e, partendo dal cristallo base, è stato predisposto un workflow per la generazione automatica di modelli atomistici di  $\text{NaMnO}_2$  drogato cui si sono variati l'elemento (Ni e Ti) e il livello di drogaggio. Viste

le numerose combinazioni possibili e nonostante le sempre maggiori capacità di calcolo, un approccio basato sui soli codici quantistici risulta proibitivo. Si è fatto perciò ricorso a tecniche di IA per fare delle previsioni sui tipi di strutture più promettenti andandone a prevedere l'energia di formazione e il potenziale redox. Dall'analisi si è visto che queste sono quelle che hanno concentrazioni simili dei tre elementi (Mn, Ti, e Ni) e con distribuzioni spaziali il più possibile uniformi (Ronchetti 2022). Indicazioni confermate da calcoli con codici quantistici a livello atomistico.

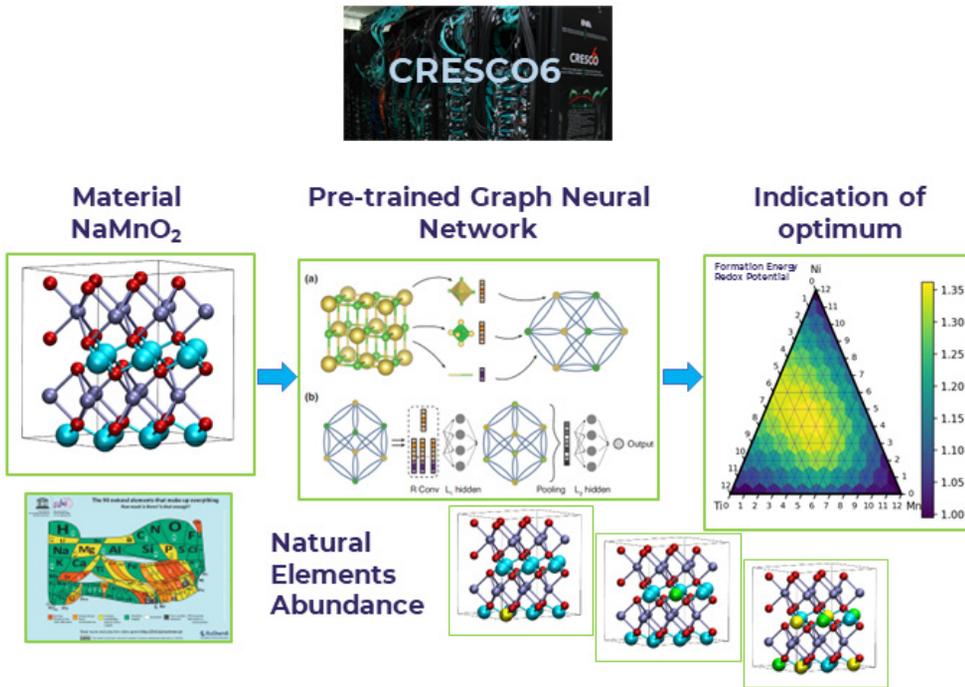


Fig. 3  
Rappresentazione del workflow che dal materiale di partenza  $\text{NaMnO}_2$ , attraverso tecniche di IA predice le concentrazioni droganti ottimali di Ni e Ti.

#### 4. Conclusioni

La piattaforma IEMAP è costituita da una infrastruttura di supercalcolo e una rete di laboratori afferenti ai quattro enti partecipanti all'iniziativa MI con lo scopo di condividere dati per accelerare la ricerca su nuovi materiali per l'energia. La piattaforma digitale, costituita da quattro componenti principali che sono l'interfaccia Web, le API, il DB e il File Storage, permette la condivisione dei dati sperimentali e computazionali e, mediante un approccio multidisciplinare potrà accelerare la progettazione di materiali avanzati per le batterie, gli elettrolizzatori e il fotovoltaico.

#### Ringraziamenti

Questo lavoro fa parte del progetto IEMAP finanziato dal MASE (Ministero dell'Ambiente e della Sicurezza Energetica) nell'ambito dell'iniziativa internazionale MI (Aspuru-Guzik 2018).

## Riferimenti bibliografici

Andersen C. W. et al. (2021), OPTIMADE, an API for exchanging materials data, *Sci. Data*, (8), 217.

Aspuru-Guzik A. et al. (2018), Materials Acceleration Platform: Accelerating Advanced Energy Materials Discovery by Integrating High-Throughput Methods and Artificial Intelligence, *Mission Innovation, Innovation Challenge 6*.

Lee D. H. et al. (2013), An advanced cathode for Na-ion batteries with high rate and excellent structural stability, *Phys. Chem. Chem. Phys.*, (15), pp. 3304-3312.

Prosini P. P. et al. (2019), Tin-Decorated Reduced Graphene Oxide and NaLi<sub>0.2</sub>Ni<sub>0.25</sub>Mn<sub>0.75</sub>O as Electrode Materials for Sodium-Ion Batteries, *Materials*, (12), 1074.

Ronchetti C. et al. (2022), Machine Learning Techniques for Data Analysis in Materials Science. 2022 AEIT International Annual Conference (AEIT), Rome, Italy, 2022, pp. 1-6.

Wang, H. et al. (2013), Electrochemical properties of P2-Na<sub>2/3</sub>[Ni<sub>1/3</sub>Mn<sub>2/3</sub>]O<sub>2</sub> cathode material for sodium ion batteries when cycled in different voltage ranges, *Electrochim. Acta*, (113), pp. 200-204.

Wilkinson M. D. et al. (2016), The FAIR Guiding Principles for scientific data management and stewardship, *Sci. Data*, (3) 160018.

## Autori



**Claudio Ronchetti** [claudio.ronchetti3@gmail.com](mailto:claudio.ronchetti3@gmail.com)

Claudio Ronchetti ha conseguito la laurea in Ingegneria Informatica presso l'Università di Roma Tre. Nel corso del suo percorso professionale ha applicato soluzioni di AI all'avanguardia in ambito pubblico e privato, spaziando in vari domini come, ad esempio, il rilevamento delle frodi, la scienza dei materiali, la trascrizione di documenti storici e digitali, la computer vision e l'elaborazione del linguaggio naturale.

**Sergio Ferlito** [sergio.ferlito@enea.it](mailto:sergio.ferlito@enea.it)

Sergio Ferlito è nato a Napoli, Italia nel 1969. Laureato in Ingegneria Elettronica presso l'Università degli Studi di Napoli "Federico II" nel 1998.

Dal 1999 lavora presso l'ENEA, Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile. I suoi interessi di ricerca includono modelli statistici e di machine learning applicati ai settori delle energie rinnovabili e dell'IoT, sviluppo software web (front/back-end), framework di machine learning online e DevOps.



**Marco Puccini** [marcopuccini@pm.me](mailto:marcopuccini@pm.me)

Laureato in Fisica all'Università di Roma "La Sapienza", ha lavorato con l'INM del CNR, il

DIMA ed il Dipartimento di Medicina Molecolare de "La Sapienza" come data scientist. Come Assegnista di Ricerca presso la Divisione ICT del Dipartimento TERIN di ENEA si è occupato principalmente di data management, sviluppo di soluzioni software cloud in diversi progetti come Mission Innovation IEMAP, DTECH, ECODIGIT, RAFAEL.



**Simone Giusepponi** [simone.giusepponi@enea.it](mailto:simone.giusepponi@enea.it)

Ricercatore presso la Divisione ICT del Dipartimento TERIN dell'ENEA. Dottorato in Fisica con una tesi nel campo delle simulazioni numeriche applicate ai sistemi non lineari. I suoi interessi di ricerca riguardano le applicazioni del calcolo parallelo e ad alte prestazioni con esperienze di ricerca nell'area della scienza dei materiali. Coinvolto in progetti di ricerca italiani ed europei come Mission Innovation IEMAP, VIPERLAB, StoRIES, ICSC, EoCoE-II, I-NEST, ROMA-TECHNOPOLE.

**Francesco Buonocore** [francesco.buonocore@enea.it](mailto:francesco.buonocore@enea.it)

Ricercatore ENEA dal 2012, lavora nella Divisione ICT del Dipartimento TERIN. Fisico specializzato in teoria della materia condensata, ha conseguito il dottorato presso l'Università di Napoli Federico II nel 2001. Le sue ricerche si concentrano su calcoli ab initio per nanomateriali e interfacce. È impegnato in progetti di ricerca italiani ed europei quali Mission Innovation IEMAP, StoRIES, ICSC, EoCoE e I-NEST ed è co-inventore di diversi brevetti e autore di numerose pubblicazioni.



**Massimo Celino** [massimo.celino@enea.it](mailto:massimo.celino@enea.it)

biografia: Massimo Celino, PhD in Fisica della Materia (Università di Strasburgo, Francia), laureato in Fisica (Università "La Sapienza" di Roma). Interessi: modellistica molecolare, calcolo ad alte prestazioni, gestione dei dati. Coinvolto in progetti europei (EoCoE, EERA-data), vicecoordinatore JP EERA "Digitalization for Energy", coordinatore progetto EuroHPC "Textarossa" per tecnologie exascale. 90+ pubblicazioni scientifiche.

**Sara Marchio** [sara.marchio@enea.it](mailto:sara.marchio@enea.it)

biografia: Ricercatrice ENEA presso la Divisione ICT del Dipartimento TERIN. Laurea in Fisica nel 2015 e PhD in Meccanica Teorica e Applicata nel 2019 (Università di Roma "La Sapienza") con una tesi nel campo delle simulazioni atomistiche combinate a tecniche per eventi rari con applicazioni nell'ambito della fluidodinamica. I suoi interessi attuali sono relativi alla scienza dei materiali ed al calcolo ad alte prestazioni. È coinvolta nei progetti ICSC e ROME-TECHNOPOLE.



**Giovanni Ponti** [giovanni.ponti@enea.it](mailto:giovanni.ponti@enea.it)

biografia: Direttore della Divisione ICT dell'ENEA e membro del CTS del GARR. PhD in Ingegneria Informatica nel 2010 (Università della Calabria) concentrandosi sulla modellazione dei dati e gli algoritmi di data mining. Ricercatore ENEA per 12 anni con attività di ricerca nell'HPC, Data Science, Cloud Computing, Big Data e Data Analytics. Campi nei quali è stato autore di prestigiose pubblicazioni scientifiche, membro del comitato di programma di conferenze e revisore per importanti riviste e convegni.