

Progetto NFFA-DI - PNRR Missione 4, "Istruzione e Ricerca" - Componente 2, "Dalla ricerca all'impresa" - Linea di investimento 3.1, "Fondo per la realizzazione di un sistema integrato di infrastrutture di ricerca e innovazione" - Azione 3.1.1, "Creazione di nuove IR o potenziamento di quelle esistenti che concorrono agli obiettivi di Eccellenza Scientifica di Horizon Europe e costituzione di reti" - CUP B53C22004310006.

# D2.5 - PROTOCOLS FOR REAL TIME ACCESS MONITORING, STATISTICS AND ANALYTICS

BIM 4

OI 2.1 - PROCEDURES FOR REAL-TIME  
MONITORING

WP 2 - NEW SCHEME FOR A DIGITAL  
OPERATIONAL WORKFLOW



Piano Nazionale di Ripresa e Resilienza

## Quadro riassuntivo rilasci documento

<i>Data</i>	<i>Stato documento</i>	<i>Realizzato da</i>	<i>Note</i>	<i>Supervisione</i>
07/08/2023	Draft	Stefano Cozzini	First draft	
29/09/2023	Draft	Mariarita de Luca, Federica Bazzocchi		
03/10/2023	Draft	Francesca Genuzio		Roberto Gotter
04/10/2023	Draft	Mariarita de Luca, Federica Bazzocchi		Stefano Cozzini
30/10/2023	Final	Francesca Genuzio, Federica Bazzocchi, Mariarita de Luca	Final draft	
17/11/2023	Final	Roberto Gotter	Final revision	

## Indice

1. Abstract .....	4
2. Risultati attesi .....	4
3. Obiettivi prefissati.....	5
3.1. Analisi dello stato attuale, struttura del DATABASE e elaborazione statistiche .....	6
3.2. Statistiche e sistemi di monitoraggio dell'andamento dell'infrastruttura .....	7
4. Considerazioni finali .....	9
4.1. Statistiche elaborate .....	10

## 1. ABSTRACT

In questo deliverable si propongono una serie di statistiche e di tools per monitorare l'accesso degli utenti, l'utilizzo della piattaforma di monitoraggio e il coordinamento digitale dell'attività dell'infrastruttura del progetto NFFA-DI che permetteranno di ottimizzare in itinere la programmazione degli esperimenti scientifici e i flussi di accesso all'infrastruttura e di valutare in modo quantitativo l'impatto del progetto. Le statistiche in esame permetteranno di osservare anche informazioni relative all'accesso all'infrastruttura tipo il genere del team di ricerca, il campo scientifico di provenienza, eventualmente in correlazione con la tipologia di tecniche analitiche utilizzate, o la localizzazione geografica degli enti coinvolti nei progetti, che potranno essere utilizzate in futuro per definire iniziative di promozione e diffusione dell'infrastruttura stessa.

L'attività di riferimento è la A2.2 (*Development of a new real-time monitoring scheme for integrated access provision*), coordinata da Area Science Park.

La piattaforma costituirà, da un lato, il portale digitale di accesso unico per gli utenti esterni (o SEP, single entry point) ai servizi dell'infrastruttura diffusa, e dall'altro lo strumento centralizzato a supporto della gestione delle attività del gruppo di coordinamento tecnico (Technical Liaison Network o TLNet) dell'infrastruttura. Infatti, gli utenti che intendono utilizzare le facilities messe a disposizione da NFFA-DI potranno creare un account sul portale di NFFA-DI attraverso il SEP (Single Entry Point), navigare il portale, consultare il catalogo delle tecniche sperimentali a disposizione, e infine sottomettere una richiesta di accesso all'infrastruttura, sulla base di un progetto scientifico (proposal) con modalità simili a quelle proposte sul portale del progetto NFFA Europe Pilot (NEP)

Contemporaneamente, gli amministratori del TLNet potranno agire in modo centralizzato sui progetti sottomessi, utilizzando specifici strumenti del SEP per: i) richiedere, formulare e associare le valutazioni tecniche al proposal, ii) coordinare la programmazione degli esperimenti e iii) gestire gli accessi ai singoli nodi locali/laboratori dell'infrastruttura iv) raccogliere tutte i dati necessari per l'analisi e la rendicontazione delle attività di accesso.

Nell'implementazione del SEP di NFFA-DI si intendono sviluppare ulteriormente dei modelli di piattaforme esistenti e già operanti per altre infrastrutture diffuse. Nel caso specifico si è scelto di capitalizzare l'esperienza già sviluppata dai partner di progetto rispetto al portale dell'infrastruttura NFFA Europe Pilot (NEP), adottandone le metodologie e gli approcci ritenuti funzionali, analizzandone le criticità e utilizzando soluzioni alternative per quegli aspetti poco efficaci. Ai fini del presente report, sono state analizzate le caratteristiche del database del SEP di NEP, in modo da evidenziare quali caratteristiche possono essere mutate nel nuovo sistema e quali invece devono essere modificate per implementare funzionalità per ora non disponibili. Analogamente al sistema SEP sviluppato per NEP, la raccolta di tutte le informazioni relative agli utenti, ai proposal, alla gestione e alla rendicontazione degli accessi sarà gestita da un sistema front-end (ovvero l'interfaccia utente) e un sistema back-end che raccoglierà i dati e li organizzerà in un apposito database che potrà essere interrogato per estrarre e confrontare informazioni relative agli utenti e ai proposal sottomessi.

## 2. RISULTATI ATTESI

Il deliverable 2.5 si riferisce all'Attività 2.2 (*Development of a new real-time monitoring scheme for integrated access provision*), coordinata da Area Science Park, inclusa nel WP2, che prevede:

*The activity aims to implement an effective real-time monitoring of the access provided by the nodes integrated in a single digital multicompetence multi-site offer. Starting from data, dealers and actions of the access workflow identified in A2.1, schemes and related tools, will be developed and identified, respectively, in order to allow for monitoring the users access activity progress allowing*

*prompt reaction to the user requests and to implement real-time mitigations to possible emerging criticalities.*

*Datasets will be identified for specific monitoring needs, and tools will be implemented to be used by properly trained personnel to respond to queries. Configurable dashboards will be implemented to generate and report periodic statistics and analytics of the whole access experience. Datasets, experimental protocols, and tools will constitute a common platform for all the parties, with the challenge to fit it with the specific platforms used by the different partners, who will have thus the added value of an effective tool for a direct internal accounting.*

Questo deliverable è collegato all'obiettivo intermedio 2.5 "Procedures for real-time monitoring" legato al KPI 2.3 - # of compatibility tests on data of NFFA-Europe Pilot project che prevede la definizione di 3 statistiche di monitoraggio realizzabili su entrambi i portali NEP e NFFA-DI, come mostrato in Tabella 1.

In questo documento sono presentate alcune possibili statistiche da implementare sulla futura piattaforma del progetto NFFA-DI utili per monitorare la sottomissione dei proposal e la gestione dei relativi accessi e in particolare saranno individuate le tre statistiche di maggior rilievo prescelte realizzabili al momento anche sul portale NEP.

Tabella 1: KPI previsto corrispondente all'obiettivo intermedio 2.5.

Indicatore oggettivo	Valore	Note
KPI 2.3 - # of compatibility tests on data of NFFA-Europe Pilot project	3	Numero di statistiche pensate per fare un confronto con NEP

### 3. OBIETTIVI PREFISSATI

Alla base della progettazione del database del "single entry point" (SEP) di NFFA-DI, c'è la necessità di conciliare le funzioni operative del SEP in quanto mezzo per l'avanzamento del workflow (legato alla ricezione, valutazione e allocazione dei proposal sperimentali) con il suo ruolo di database dei dati relativi ai singoli passaggi dello stesso workflow.

Sebbene la struttura di dettaglio del workflow di NFFA-DI sia ancora da definirsi sulla base dell'analisi delle specifiche peculiarità dell'infrastruttura stessa nell'ambito dell'attività di progetto A2.1, ai fini del presente report ci limiteremo allo schema generale.

I passaggi principali che possono essere individuati nel "del ciclo di vita" di ogni proposal (workflow) sono:

- 1) Sottomissione; con informazioni associate relative a: anagrafica dell'utente (dati personali, dati dell'istituzione di provenienza, livello accademico...), contenuto del proposal: informazioni bibliografiche (titolo, acronimo, area scientifica di riferimento), tecnico-scientifiche (numero di step, installazioni richieste, tecniche richieste, unità richieste, interesse industriale, ...), informazioni sui campioni da analizzare (materiale, dimensioni, provenienza, ...) e informazioni relative alla tipologia di accesso (in presenza o da remoto).
- 2) Valutazioni; eleggibilità, fattibilità tecnica e punteggio nella valutazione scientifica del proposal.
- 3) Scheduling: informazioni relative alla programmazione temporale delle varie fasi dell'esperimento.

- 4) Esecuzione dell'esperimento, con la possibilità di introdurre modifiche a quanto precedentemente inserito.
- 5) Conclusione dell'esperimento: rendicontazione, rimborso, questionario di valutazione, report utenti e provider.

Pertanto, all'interno del SEP ogni proposal è associato alle informazioni relative alle fasi sopra elencate e tutte queste informazioni sono contestualmente salvate dentro un database.

### 3.1. Analisi dello stato attuale, struttura del DATABASE e elaborazione statistiche

Il database di NEP è al momento di tipo documentale DBMongo, organizzato in 15 "collections" riportate qui di seguito:

- counters
- erc\_sectors
- family
- feasibility
- files
- installation
- institution
- instrument
- management
- partner
- proposal
- scheduling
- site
- technique
- virtual\_access

In ogni "collection" le informazioni sono salvate con una struttura categorizzata su vari livelli gerarchici.

Allo stato attuale l'interazione fra front-end e database di NEP evidenzia che l'elaborazione delle statistiche risulta particolarmente laboriosa e non risulta implementato alcun tool di monitoraggio o analisi in tempo reale. Ad esempio, il monitoraggio disponibile sul sito NFFA.eu e riportato in Fig.1 è esclusivamente statico e aggiornato periodicamente da operatori. Inoltre, nella versione corrente non è possibile avere una valutazione in tempo reale dello stato di avanzamento del workflow, né dal lato dell'utente, (stato di avanzamento del suo proposal), né per il TLNEt centrale e i provider (visione globale dell'andamento delle valutazioni di fattibilità e dello scheduling).

transnational access statistics		
<b>24</b> calls for access	<b>824</b> proposals submitted	<b>63%</b> rate of acceptance
<b>29%</b> with Large Scale Facilities	<b>12%</b> with theory	<b>12%</b> with industry
<b>~3</b> average users per proposal	<b>63</b> countries applying	<b>2378</b> lab sessions

Figura 1: Monitoraggio periodico di alcuni indicatori relativi all'accesso al portale NEP (<https://www.nffa.eu/>)

Per superare queste limitazioni, nell'implementazione del sistema di NFFA-DI si propone di strutturare le interrogazioni del database a tre livelli:

- Un cruscotto visivo periodicamente aggiornato riportando le statistiche più generali, come ad esempio il numero di proposal sottomessi e/o accettati, proposal realizzati e programmati, le attività dei singoli provider, tecniche maggiormente richieste;
- Un tool di interrogazione filter-based in modo da interrogare il database in maniera più specifica rispetto alle statistiche riportate dal cruscotto;
- Interrogazione standard del database, per elaborazioni più specifiche e a richiesta. Quest'ultima funzionalità sarà concessa solo a dei super-utenti, quali gli amministratori di DB o delegati.

Un'altra innovazione prevista per il portale di NFFA-Di rispetto a quello di NFFA-NEP sarà l'aggiunta di alcuni "status di avanzamento proposal" e date inizio e fine validità degli stessi.

Con questo sistema, il singolo utente potrà visualizzare lo stato e il livello di avanzamento di ogni suo proposal sottomesso, mentre il gestore del portale (TLNET) avrà visione del cruscotto e potrà interrogare il DB col tool filter-based, in modo da avere un quadro generale delle operazioni della struttura dell'infrastruttura. Altri aspetti che potrebbero essere implementati sono ad esempio le tempistiche di risposta dei provider.

### 3.2. Statistiche e sistemi di monitoraggio dell'andamento dell'infrastruttura

La tabella sottostante riassume una lista non esaustiva di statistiche che si potrebbero realizzare nel contesto del monitoraggio dell'infrastruttura NFFA-DI, l'analisi dello status attuale per il database di NEP e le soluzioni implementative proposte:

tipologia	Statistiche di monitoraggio	Status attuale
<u>Statistiche su proposal</u>	status (draft/sottomesso/ritirato/feasible-not feasible/accettato-rejected/schedulato/running/abolito/finito/report)	Parzialmente implementata: alcuni status non sono definiti tra le attuali possibili opzioni. Inoltre, lo status corrente non è visibile all'utente. <b>La definizione più dettagliata ed eventualmente gerarchica dello status del proposal permetterà di</b>

		<b>ottenere le informazioni su altre statistiche in modo più semplice e schematico.</b>
<u>Statistiche su proposal</u>	pubblicazioni (flag da aggiungere a posteriori, DOI associato)?	Mancante: i dati non sono inseriti nel database
<u>Statistiche su proposal</u>	numero di Unità di accesso UoA del proposal. Per ogni step sperimentale: UoA proposte, valutate, assegnate e infine effettivamente eseguite, con la possibilità poi di raggruppare per installazione, tecnica, provider, strumento specifico utilizzato, ecc.;	Mancante: è richiesta un'elaborazione sostanziale dei dati così come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistiche su proposal</u>	valutazioni post-proposal, questionario non in database	Mancante: i dati non sono inseriti nel database
<u>Statistiche su proposal</u>	la presenza o meno di accesso alla teoria o a Large Scale Facilities	Parzialmente implementate: è richiesta un'elaborazione minore dei dati così come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistiche su proposal</u>	prosecuzione proposal precedente	Parzialmente implementate: è richiesta un'elaborazione minore dei dati come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistiche su proposal</u>	dominio scientifico di appartenenza	Parzialmente implementate: è richiesta un'elaborazione minore dei dati come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistiche su proposal</u>	countries affiliazione, (PI, maggioranza del team); UoA fatte per UE e non UE	Parzialmente implementate: è richiesta un'elaborazione minore dei dati come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistiche su installazioni e tecniche</u>	<b>combinazioni di installazioni utilizzate</b> (senza numero e ordine di esecuzione degli step; esempio. Structural+E/C+Structural e Structural+Structural+ E/C diventano entrambe sia: S+E/C	Mancante, è richiesta un'elaborazione sostanziale dei dati così come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistiche su installazioni e tecniche</u>	tecniche "rare"; definendo bene cosa si intendere per rare	Mancante: è richiesta un'elaborazione sostanziale dei dati così come forniti nell'estrazione attualmente realizzabile sul database e la realizzazione della lista di tecniche rare.
<u>Statistiche su installazioni e tecniche</u>	combinazioni di tecniche utilizzate; in caso scegliendo una tecnica (ad es. una tecnica rara) e vedendo come si combina; definire bene cosa si intende per rara	Mancante: è richiesta un'elaborazione sostanziale dei dati così come forniti nell'estrazione attualmente realizzabile sul database e la realizzazione della lista di tecniche rare.

<u>Statistica su utenti:</u>	<b>nuovi utenti</b> (numero di nuovi utenti per ciascuna call, sia PI che team members)	Parzialmente implementate: è richiesta un'elaborazione sostanziale dei dati così come forniti nell'estrazione attualmente realizzabile sul database. Non attivo il riconoscimento nome + cognome o cognome + nome
<u>Statistica su utenti:</u>	<b>composizione di genere</b>	Parzialmente implementate: è richiesta un'elaborazione minore dei dati così come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistica su utenti:</u>	tipologia ente di provenienza (università/ente ricerca/facility)	Parzialmente implementate: è richiesta un'elaborazione minore dei dati così come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistica su utenti:</u>	titolo di formazione dei partecipanti (dottorandi/post doc/...) (si può vedere come si compongono i gruppi che pubblicano di più, ad esempio...)	Parzialmente implementate: è richiesta un'elaborazione minore dei dati così come forniti nell'estrazione attualmente realizzabile sul database.
<u>Statistiche su operazioni gestite dai tools</u>	adempimento della valutazione di feasibility da parte dei nodi locali della TLNet a seguito di richiesta da parte del nodo centrale TLNet	Mancante: è richiesta un'elaborazione sostanziale dei dati così come forniti dal database e la realizzazione della lista di tecniche rare. Si potrà ottenere in futuro con una successiva implementazione delle estrazioni realizzabili sul database.
<u>Statistiche su operazioni gestite dai tools</u>	scheduling: quanti, dopo quanto tempo dalla submission (prima notifica)	Parzialmente implementate: è richiesta un'elaborazione minore dei dati così come forniti nell'estrazione attualmente realizzabile sul database

#### 4. CONSIDERAZIONI FINALI

Grazie a una collaborazione molto attiva con alcuni partner del progetto NEP è stato possibile esaminare il work-flow digitale delle procedure del portale [www.nffa.eu](http://www.nffa.eu) e proporre delle possibili statistiche e tecniche di monitoraggio in tempo reale che potranno essere implementate nel portale di NFFA-DI.

Sono state realizzate le tre seguenti statistiche, attualmente implementabili sia sul portale NEP che su quello di NFFA-DI:

- **Installazioni e tecniche:**
  1. Combinazioni di installazioni utilizzate (senza numero e ordine di installazione degli step)
- **Utenti:**
  2. numero dei nuovi utenti per ciascuna call (sia PI che team members)

### 3. composizione di genere (2A E 2B)

In particolare nello studio per la redazione del presente report è risultato essenziale implementare nell'ambito di NFFA-DI un monitoraggio automatico dei seguenti indicatori:

-pubblici: quelli già presenti sul sito e rappresentati in Figura 1; numero nuovi utenti per call  
-ad uso interno (TLNET e providers): adempimento della valutazione di feasibility dal parte dei nodi locali della TLNet a seguito di richiesta da parte del nodo centrale TLNet; tempistiche dello scheduling ad esempio dopo quanto tempo dalla submission (prima notifica).

#### 4.1. Statistiche elaborate

- 1- Statistica delle combinazioni di installazioni richieste nei proposal passati alla fase di scheduling in Figura 2;
- 2- Distribuzione per call del numero di utenti coinvolti nei proposal sottomessi (num\_Proposal), evidenziando gli utenti esclusivamente nuovi, ovvero con nessuna presenza né come PI né come team member in proposal precedenti (num\_Proposal\_utenti\_nuovi) in Figura 3;
- 3- Percentuale di presenza femminile nei proposal sottomessi in Figura 4.

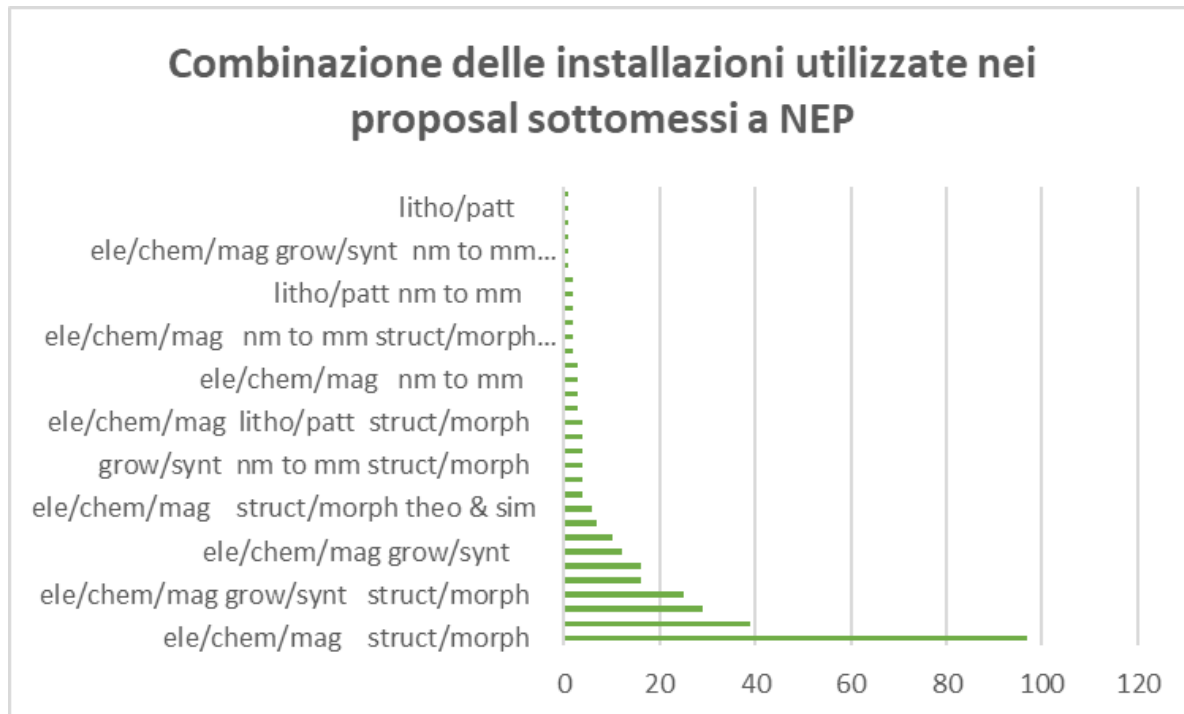


Figura 2 - Combinazioni di installazioni utilizzate (senza numero e ordine di installazione degli step)

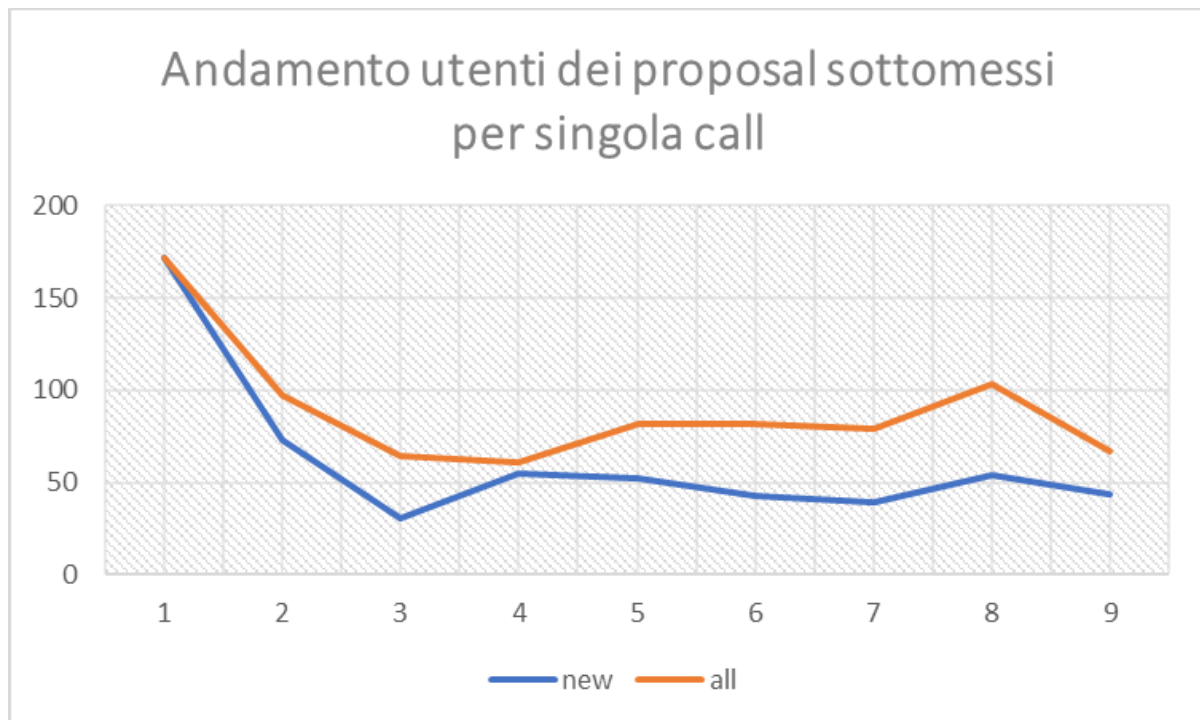


Figura 3 - Grafico dell'andamento del numero totale e dei nuovi utenti per ciascuna call.

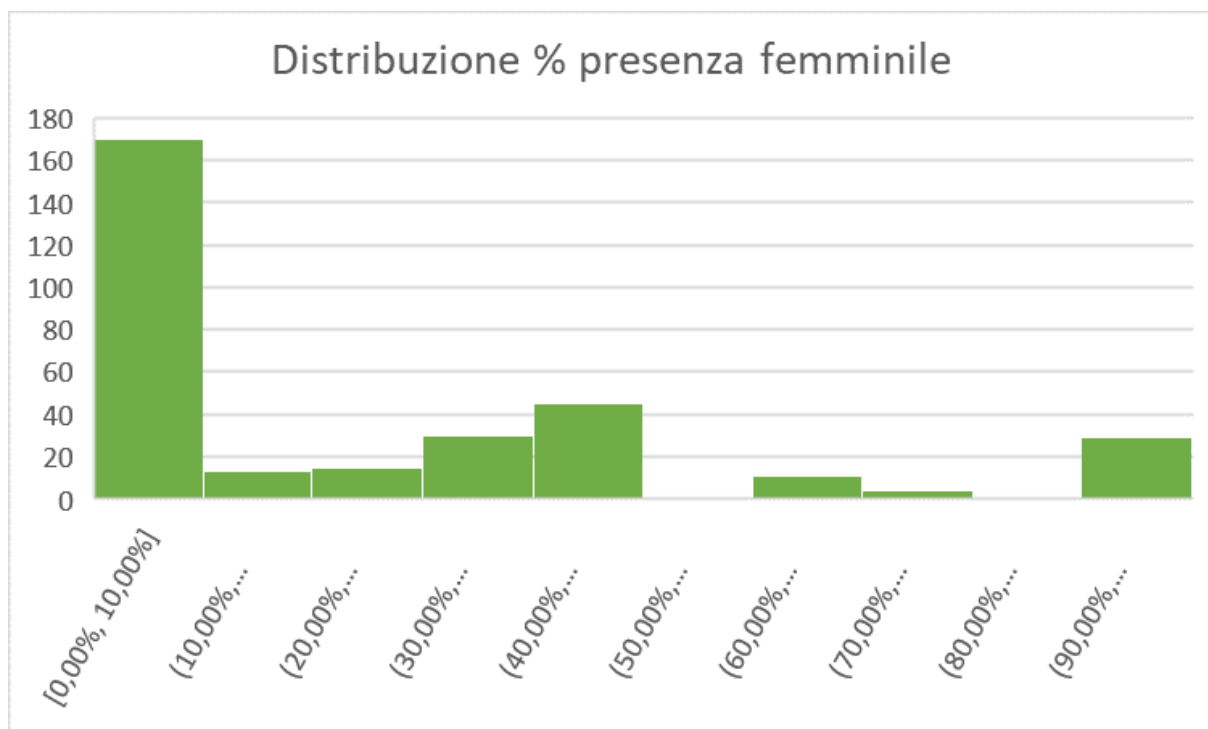


Figura 4 - Grafico della distribuzione del numero di proposal sottomessi in funzione della percentuale di presenza femminile. La media di presenza femminile si attesta attorno al 25%.

Le tre statistiche realizzate hanno permesso di valutare quantitativamente alcuni dati relativi alla richiesta combinata di diverse installazioni e agli utenti che derivano semplicemente dall'utilizzo della piattaforma da parte degli utenti stessi al momento della sottomissione dei proposal.

Anche se non richiesto nell'ambito del presente Deliverable, vale la pena fare alcune osservazioni qualitative.

La statistica sulla combinazione delle installazioni utilizzate evidenzia una presenza predominante della caratterizzazione (combinata "struct/morph" ed "ele/chem/mag"), eventualmente a sua volta associata con la sintesi "grow/synt", evidenziando quindi una carenza nelle richieste di fabbricazione ("lith/patt") e scale-up ("nm to mm") dell'offerta di NFFA.

Per quanto riguarda il monitoraggio degli utenti mostrato in Figura 3, se si interpreta l'alto valore per la prima call come dovuto ad un prolungato tempo di attesa (superiore rispetto ai soliti tre mesi che intercorrono tra una call e la successiva) da parte della comunità scientifica già fidelizzata col precedente progetto NFFA-EUROPE, si evidenzia un andamento abbastanza costante di nuovi e vecchi utenti.

In una recente valutazione sulla percentuale di donne nelle STEM<sup>1</sup> è stato osservato che solo il 34% dei laureati in discipline STEM è di sesso femminile e che nei percorsi post-universitari, le donne rappresentano solo il 22,4% del totale in ambito ICT e il 29,4% del totale in ambito ingegneristico. La percentuale di presenza femminile tra i ricercatori che hanno sottomesso i proposal al portale NEP si attesta attorno al 25% e può essere quindi considerata allineata con le altre percentuali di presenza femminile nelle discipline in ambito scientifico-tecnologico. Tale media si attesta al 23% se ci limitiamo ai proposal accettati. E' doveroso però chiarire che tale andamento è fortemente disomogeneo rispetto ai diversi settori scientifici.

In conclusione, per migliorare il monitoraggio del portale di NFFA-DI ci si propone di realizzare a breve l'adempimento della valutazione di feasibility dal parte dei nodi locali della TLNet a seguito di richiesta da parte del nodo centrale TLNet, al fine di poter gestire al meglio l'attribuzione dei proposal ai diversi laboratori, avendo la possibilità di monitorare la risposta (o assenza di risposta) da parte dei laboratori stessi. Nel contesto di questo deliverable, non è stato possibile elaborare questa statistica sulla base dei dati estraibili dalla attuale versione del database. Pertanto, questo aspetto si configura come uno degli elementi miglioratori per la nuova versione del SEP.

---

<sup>1</sup> Dossier "La condizione delle bambine e delle ragazze nel mondo", 12° edizione, 2023.