



## Sviluppo portale – Implementazione sul portale SiMTE di SW con Intelligenza Artificiale per funzioni di lettura automatica, estrazione dati, assistenza utenti. (LA1.6)

E. Cosimi, F. Donato, S. Giammartini, G. Simbolotti (ENEA)

In collaborazione con il team Sistemi Informativi/IBM coinvolto con  
apposito contratto per la fornitura e l'addestramento del SW-AI

SVILUPPO PORTALE – IMPLEMENTAZIONE DI SW BASATI SU AI SUL PORTALE SIMTE PER FUNZIONI DI LETTURA AUTOMATICA, ESTRAZIONE DATI, ASSISTENZA UTENTI (LA 1.6)

Enrico Cosimi, Filippo Donato, Stefano Giammartini, Giorgio Simbolotti (ENEA)

In collaborazione con il team Sistemi Informativi/IBM coinvolto con apposito contratto per la fornitura e l'addestramento del SW-AI

Dicembre 2021

## Report Ricerca di Sistema Elettrico

Accordo di Programma Ministero della Transizione Ecologica - ENEA

Piano Triennale di Realizzazione 2019-2021 - III annualità 2021

Obiettivo: *Sistema Elettrico*

Progetto: 'Strumenti e modelli per scenari energetici ed elettrici adeguati all'evoluzione del sistema..., dei mercati e della regolazione'.

Linea di attività: *LA1.6 Sviluppo portale – Implementazione di SW basati su AI sul portale SIMTE per funzioni di lettura automatica, estrazione dati, assistenza utenti*

Responsabile del Progetto: Giorgio Simbolotti, ENEA

## Indice

SOMMARIO.....	4
1 INTRODUZIONE AL PROGETTO.....	6
1.1 IMPATTO DELLA PANDEMIA COVID-19 SULLE ATTIVITÀ PREVISTE PER L'ANNO 2020.....	9
1.2 BREVI INFORMAZIONI SULLA PIATTAFORMA SIMTE.....	10
2 DESCRIZIONE DELLE ATTIVITÀ SVOLTE E RISULTATI DELLA LA 1.6.....	12
2.1 INTRODUZIONE.....	12
2.2 ATTIVITÀ SVOLTE NELLA FASE 1.....	13
2.2.1 <i>Assistente Virtuale</i> .....	14
2.2.2 <i>Gestione dei documenti</i> .....	16
2.2.3 <i>Pre-annotazione e annotazione dei documenti di training e training del modello</i> .....	20
2.2.4 <i>Valutazione del modello</i> .....	21
2.2.5 <i>Pubblicazione del modello su Watson Discovery</i> .....	22
2.2.6 <i>Analisi e arricchimento cognitivo dei documenti con Watson Discovery</i> .....	22
2.2.7 <i>Training del SW cognitivo sulla struttura dei documenti e acquisizione delle tabelle dati</i> .....	22
2.3 RISULTATI DELLA FASE 1.....	24
2.4 ATTIVITÀ SVOLTE NELLA FASE 2.....	25
2.4.1 <i>Consolidamento del sistema relativamente alle tecnologie trattate nella Fase 1</i> .....	25
2.4.2 <i>Affinamento delle risposte dell'Assistente Virtuale e consolidamento della produzione di report</i> .....	26
2.4.3 <i>Addestramento del sistema su ulteriori cluster tecnologici</i> .....	26
2.4.4 <i>Interfaccia di preview</i> .....	27
2.4.5 <i>Interfaccia e integrazione del modello Scilab per il calcolo di prestazioni e costi delle tecnologie</i> .....	27
2.4.6 <i>Integrazione di funzioni statistiche a dati di prestazioni e costi derivati da cataloghi industriali</i> .....	36
2.5 RISULTATI DELLA FASE 2.....	37
2.6 TRAINING DEL PERSONALE ENEA SULLE TECNICHE DI ADDESTRAMENTO DEL SW AI.....	38
ALLEGATI.....	39

**Allegato 1:** RELAZIONE TECNICA - Appalto per fornitura di servizi di implementazione/addestramento e diritti di accesso (license & cloud) a SW-AI (SoftWare – Artificial Intelligence) per applicazione a piattaforma SIMTE-ENEA –RTI Sistemi Informativi/IBM e VAR Group -

**Allegato 2 :** RAPPORTO FINALE DI ATTIVITA per servizi di implementazione/addestramento e diritti di accesso (license & cloud) a SW-AI (SoftWare – Artificial Intelligence) per applicazione a piattaforma SIMTE-ENEA - Sistemi Informativi/IBM

**Allegato 3:** Specifiche Tecniche per l'acquisto di servizi di implementazione/addestramento e diritti di accesso (license & cloud) a SW-AI (Software - Artificial Intelligence) per applicazione alla piattaforma SIMTE (Sistema informativo e di Monitoraggio delle Tecnologie Energetiche), ENEA – Agosto 2020

## Sommario

Come previsto dal Piano Triennale Realizzativo (PTR) del progetto, nel corso dell'annualità 2021 sono stati svolti ed ultimati i lavori relativi alle linee di attività LA 1.5 e 1.6 che riguardano rispettivamente la prova di applicabilità e la successiva Implementazione di software Intelligenza Artificiale (SW-AI) sulla piattaforma SiMTE per svolgere le funzioni di: a) lettura automatica di documenti tecnico-economici con estrazione/gestione dei dati di interesse; b) assistenza agli utenti della piattaforma (assistente virtuale) nell'utilizzo dei servizi offerti (ricerca dati e documenti, interrogazione in linguaggio naturale, uso del modello per il calcolo di prestazioni e costi delle tecnologie energetiche ...).

L'introduzione dell'AI sulla piattaforma SiMTE è motivata dalla necessità di ridurre gli oneri (tempi e risorse) di aggiornamento della piattaforma (prestazioni e costi delle tecnologie) e di rendere più fruibili per gli utenti i servizi offerti. Essa costituisce inoltre la parte più **innovativa** del progetto

Nel lavori di introduzione dell'AI è anche confluita la LA 1.2 (Implementazione di funzioni statistiche per analisi di prestazioni e costi di tecnologie di larga diffusione commerciale) poiché le funzioni statistiche sono state introdotte con un SW aggiuntivo integrato nel sistema AI. Le tre LA 1.2, 1.5 e 1.6 danno luogo rispettivamente ai rapporti tecnici distinti 161, 163 e 071, tuttavia le LA 1.2 e 1.5 sono strettamente legate alla LA 1.6 e i rispettivi rapporti condividono parti comuni.

I ritardi riconducibili all'emergenza Covid e alcune difficoltà amministrative nell'assegnazione di contratti hanno prodotto lo slittamento di queste LA dal 2020 al 2021 (variante approvata) con inizio effettivo nell'Aprile 2021 e conseguente drastica riduzione dei tempi di lavoro (da 12 a circa 6 mesi). Cionondimeno, l'introduzione dell'AI ha richiesto un impegno di personale e di ore di lavoro superiore a quanto previsto a causa del laborioso processo di addestramento del SW che è stato necessario acquisire. Le LA 1.5 e 1.6 hanno quindi richiesto impegni superiori al preventivo iniziale, tuttavia compensati dal minor impegno richiesto per la LA 1.2 grazie alla sua integrazione nel sistema AI. Di questo si è tenuto conto in sede di rendicontazione.

Risultati - Pur ricorrendo all'uso di una versione beta del SW (IBM Watson), più avanzata e non ancora commercializzata, per svolgere le complesse funzioni richieste, i risultati della LA 1.5 (prova di applicabilità) sono stati soddisfacenti avendo raggiunto gli obiettivi fissati nelle specifiche tecniche e dimostrato la potenzialità dell'AI per i fini preposti. Nel corso della LA 1.6 il SW è stato quindi applicato ad un certo numero di tecnologie previo opportuno "addestramento" del SW stesso e contestuale training di personale ENEA sulle tecniche di addestramento per rendere autonoma l'Agenzia nell'uso del SW. Il processo di addestramento, inizialmente complesso e *time-consuming* sembra peraltro velocizzarsi con l'accrescimento del patrimonio cognitivo del SW (terminologia, entità, relazioni). L'addestramento<sup>1</sup> del SW è stato completato sui seguenti cluster tecnologici e sulle relative varianti:

- solare termico a concentrazione (Concentrating Solar Power, CSP);
- solare fotovoltaico (photovoltaic solar power);
- elettro-generazione da carbone (coal-fired power);
- pompe di calore aria-aria (air-to-air heat pumps);
- pompe di calore ad acqua e geotermiche (water and geothermal heat pumps);
- celle a combustibile (fuel cells, FC);

Relativamente a tali tecnologie sono stati raggiunti tutti gli obiettivi attesi. In particolare:

- Acquisire (leggere) i documenti tecnico-economici ed individuare i dati di interesse nelle rispettive unità di misura (es: efficienza energetica, fattore di capacità, vita operativa di impianti/dispositivi,

---

<sup>1</sup> Per velocizzare il processo, l'addestramento è stato eseguito in inglese: il SW-AI è infatti nativo in lingua inglese (pur potendo operare in altre lingue) e la terminologia utilizzata nel settore delle tecnologie energetiche è largamente diffusa e utilizzata in inglese. L'uso dell'inglese è stato possibile poiché tutti i documenti della piattaforma sono disponibili in doppia lingua inglese-italiano al fine di rendere accessibile il sistema anche in ambito internazionale.

emissioni, costi di investimento e di esercizio, costi specifici dei vettori energetici o prodotti finali, dati di mercato, ecc. e, ove disponibili, proiezioni di tali dati al 2030-50);

- trasferire i dati di interesse così individuati alla banca dati della piattaforma SiMTE integrando e aggiornando i dati già presenti;
- guidare l'utente - tramite l'assistente virtuale – nella formulazione di domande in linguaggio naturale, e rispondere alle domande poste con testi e tabelle (con congruenza dell'ordine del 70% rispetto alle domande formulate), nonché confrontare i dati relativi a tecnologie diverse;
- fornire a richiesta brevi report con testi e tabelle relativi alle domande formulate e, ovviamente, scaricare i documenti relativi ai vari cluster tecnologici presenti nella piattaforma SiMTE;
- guidare l'utente nell'uso del modello attraverso il quale la piattaforma consente (ad utenti esperti) di eseguire l'analisi tecnico-economica (prestazioni e costi, costi e benefici,) di una tecnologia e/o del relativo investimento, utilizzando dati derivati dalla banca dati SiMTE integrabili con dati immessi dall'utente stesso.

Il sistema è inoltre in grado di effettuare operazioni statistiche su prestazioni e costi di tecnologie di larga diffusione commerciale (es: veicoli, elettrodomestici) sulla base di dati derivati da cataloghi commerciali o forniti da operatori dei rispettivi mercati.

Alla fine della Fase 2 tutte le funzioni di cui sopra sono state provate con esiti soddisfacenti. Inutile forse precisare che il sistema, al pari di tutti i sistemi informativi, richiede una fase di ulteriore messa a punto che può essere ottenuta soltanto in base ad una più lunga esperienza di utilizzo.

## 1 Introduzione al progetto

Il **Piano Triennale RdS 2019-2021** per il **Tema di Ricerca 2.1** dal titolo *“Strumenti e modelli... per scenari energetici ed elettrici adeguati all’evoluzione del sistema..., dei mercati e della regolazione...”*, assegna ad Enea i seguenti sotto-temi di ricerca:

1. *“Sviluppo di un data-base con dati tecnico-economici sulle tecnologie energetiche (prestazioni, emissioni, costi, dati di mercato, proiezioni, etc.), utilizzabili nei modelli (energetici)...”*;
2. *“Sviluppo di un portale per l’accesso a documenti e dati... con funzioni di ricerca e analisi, e modelli ... per il calcolo di prestazioni e costi delle tecnologie...”*;
3. *“Aggiornamento dei **brief** tecnologici che per ogni tecnologia contestualizzano informazioni e dati tecnico-economici attuali e attesi ...”*;
4. *“Valutazioni previsionali energetiche regionali coerenti con scenari nazionali, finalizzate a un nuovo burden-sharing delle fonti rinnovabili e del potenziale regionale di efficienza energetica ...”*.

Nei punti 1), 2) e 3) di cui sopra si fa implicito riferimento (e.g. la parola *brief*) ad una attività pregressa conclusa nel 2018 (Convenzione MISE-ENEA denominata SIMTE) nell’ambito della quale, in ottemperanza ad alcuni dispositivi di legge<sup>2</sup>, l’Enea ha realizzato per conto del MISE un **Sistema di Monitoraggio dello stato e delle prospettive delle Tecnologie Energetiche (SIMTE)**.

Il sistema consiste in una piattaforma web ([www.simte.enea.it](http://www.simte.enea.it)) che offre all’utenza (prevalentemente operatori, analisti e decisori del settore energetico) documenti, informazioni, dati e strumenti di analisi sulle tecnologie energetiche commerciali e in via di sviluppo nei settori di offerta e domanda dell’energia: a) produzione e trasformazione di energia primaria; b) produzione e distribuzione di elettricità e calore; c) residenziale e servizi; d) industria; e) trasporti. Si tratta di un sistema informativo con focus sul mercato italiano e attenzione agli sviluppi internazionali, con orizzonte temporale 2018-2030 (2050) e aggiornamenti biennali (ove si renda disponibile una continuità di risorse).

Il sistema attualmente comprende:

- la piattaforma web (portale) aperta in via sperimentale all’indirizzo di cui sopra e destinata a decisori pubblici e privati, investitori e analisti del settore energia che possono accedere ai documenti e ai servizi offerti;
- circa 50 documenti tecnici (*brief*) redatti da esperti, aggiornati al Gennaio 2018, che contestualizzano dati attuali e attesi su prestazioni, emissioni, quote di mercato, normative, peculiarità nazionali, costi di investimento, di esercizio, di prodotto/servizio, etc. per altrettante tecnologie trattate;
- una banca-dati (derivata dai documenti di cui sopra e altre fonti) dotata di strumenti di ricerca, confronto e analisi dei dati;
- un modello per il calcolo di prestazioni e costi delle tecnologie in base a criteri omogenei e ad input modificabili dall’utente.

In base a quanto previsto nei punti 1-4 di cui sopra, il presente progetto si prefigge i seguenti obiettivi:

1. **Aggiornare/raddoppiare la base-dati e gli strumenti della banca-dati SIMTE**, in particolare,
  - a. aggiornare documenti e dati già presenti, il sistema di ricerca/analisi dati e il modello di calcolo prestazioni e costi delle tecnologie ad uso dell’utenza;
  - b. raddoppiare la base-dati con circa 50 nuovi documenti (*brief*) aggiuntivi su tecnologie ancora non presenti e/o tramite opportuna disaggregazione di quelle già presenti;
  - c. introdurre un database su misure di incentivazione per tecnologie rinnovabili ed efficienza energetica;

---

<sup>2</sup> DLgs 28 del 3.3.2011, c.7/art. 40, DM 6.7.2012 e DM 28.12.2012

- d. introdurre strumenti per valutazioni statistiche di prestazioni e costi di tecnologie di largo consumo (settori residenziale e trasporti), in base a dati di mercato.
2. **Conseguire un salto di qualità nella gestione della banca dati e del portale** (con sostanziale riduzione delle risorse necessarie) attraverso l'adozione di strumenti software (SW) basati su **Intelligenza Artificiale (AI)** al fine di rendere possibile:
    - a. lettura automatica di documenti tecnici con acquisizione e gestione dei dati di interesse;;
    - b. produzione automatica di report su richiesta dell'utente;
    - c. assistenza virtuale all'utenza per l'accesso ai servizi offerti dalla banca dati.
  3. **Attuare una adeguata strategia di diffusione** della piattaforma SIMTE e conseguire la piena operatività del portale verso l'utenza, i.e. operatori, decisori e investitori del settore energia, consumatori.
  4. **Produrre le valutazioni energetiche previsionali** a livello regionale coerenti con gli scenari nazionali (metodologia e applicazioni) come previsto nel sotto-tema 4.

Nella formulazione inizialmente approvata (successivamente oggetto di varianti nel 2020, vedi sez. 1.1 seguente) il Piano Triennale Realizzativo (PTR) del progetto si articola in **23 linee di attività (LA)** divise in **5 aree** e distribuite su **tre annualità**. Le 5 aree comprendono:

- Sviluppo Banca Dati, che comprende le LA 1.1, 1.2, 1.3 e 1.4;
- Sviluppo Portale, che comprende LA 1.5, 1.6 e 1.7 (con coinvolgimento di fornitore di SW-AI);
- Aggiornamento/Produzione Documenti-Brief, che comprende le LA da 1.8 a 1.16 (con coinvolgimento di partner universitari);
- Valutazioni Energetiche Regionali, che comprende le LA 1.17, 1.18, 1.19 e 1.20;
- Diffusione Risultati, che comprende le LA 1.21, 1.22 e 1.23 (con coinvolgimento di azienda specializzata in diffusione).

**La prima annualità 2019 comprende 4 LA** e una attività preparatoria auto-finanziata descritte di seguito:

- **LA1.1:** miglioramento del sistema di ricerca, analisi e presentazione dati; aggiornamento del modello deterministico per calcolo prestazioni/costi delle tecnologie da ambiente Excel/VB ad ambiente Scilab (open source) con inserimento di misure di incentivazione (e.g. super-ammortamento); preparazione casi-guida per analisi prestazioni/costi tecnologie;
- **LA1.8:** preparazione nuovo format per aggiornamento brief, raccolta dati e nuovi documenti;
- **LA1.17:** analisi energetiche regionali: messa a punto del metodo e applicazione a Regione Lazio;
- **LA1.21:** strategia diffusione portale (valutazioni di proposte fornite pro-bono da aziende di settore).

L'attività autofinanziata prevista per il 2019 comprende:

- **Analisi di SW basati su AI** (con la collaborazione di fornitori di SW coinvolti pro-bono) per: lettura automatica documenti, estrazione e gestione dati; redazione di brevi report su richiesta utente (meta-motori); assistenza virtuale all'utente della banca dati;

**La seconda annualità 2020 comprende 12 LA** e prevede il massimo impegno di risorse e investimenti con il coinvolgimento di due Università su temi di analisi e ricerca (Università di Roma la Sapienza – Dip. di Ingegneria Astronautica, Elettrica ed Energetica - DIAEE e Politecnico di Torino, Dip. Energia - DENERG), e di aziende private per la fornitura di consulenze e dei servizi di cui sopra:

- **LA1.2:** introduzione di funzioni statistiche per calcolo di prestazioni/costi di tecnologie a larga diffusione (residenziale, trasporti) alimentato da dati di mercato;
- **LA1.3:** introduzione di un nuovo DB su misure di incentivazione per le tecnologie rinnovabili e l'efficienza energetica;
- **LA1.5-1.6:** prove di applicazione e successiva implementazione/addestramento di SW basati su tecnologia AI per lettura automatica documenti e gestione dati, reporting, virtual assistant (in collaborazione con fornitore di SW-AI);

- **LA1.7:** ammodernamento HW del portale SiMTE (server, etc.);
- **LA1.9-1.12:** aggiornamento e produzione ex-novo di documenti e dati su tecnologie energetiche (settori: produzione, trasmissione e distribuzione di elettricità e calore; residenziale e servizi) in collaborazione con Politecnico di Torino e Università di Roma La Sapienza (circa 50 documenti/brief);
- **LA1.18-1.19:** valutazioni energetiche regionali: implementazione della metodologia sul portale SiMTE e applicazione alla Regione Basilicata;
- **LA1.22:** Predisposizione di una strategia di diffusione del portale tramite contratto assegnato ad una azienda di settore.

La terza annualità 2021 comprende 7 LA e prevede il completamento delle attività progettuali:

- **LA1.4:** applicazione dei modelli di calcolo prestazione e costi alle nuove tecnologie inserite nella banca dati nel 2020-21 (LA1.9-1.16) e adattamento dei modelli ai nuovi SW-AI (LA1.5-1.6);
- **LA1.13-1.16:** aggiornamento e produzione ex-novo di documenti e dati su tecnologie (settori: produzione e trasformazione di energia primaria; industria; trasporti) con Politecnico di Torino e Università di Roma La Sapienza (circa 45 doc./brief);
- **LA1.20:** Valutazioni energetiche regionali, applicazione alla Regione Piemonte;
- **LA1.23:** attuazione della strategia diffusione: lancio del portale SiMTE, organizzazione di eventi, workshop, etc.) in collaborazione con azienda di settore.

Il costo complessivo atteso del progetto è di € 2.316.797,94. La Figura 1 illustra la ripartizione del budget per voce di spesa (incluso spese generali), per anno, per obiettivo e per beneficiario (escluso spese generali). La Figura 2 illustra la ripartizione dei costi per voce di spesa nel triennio (escluso spese generali).

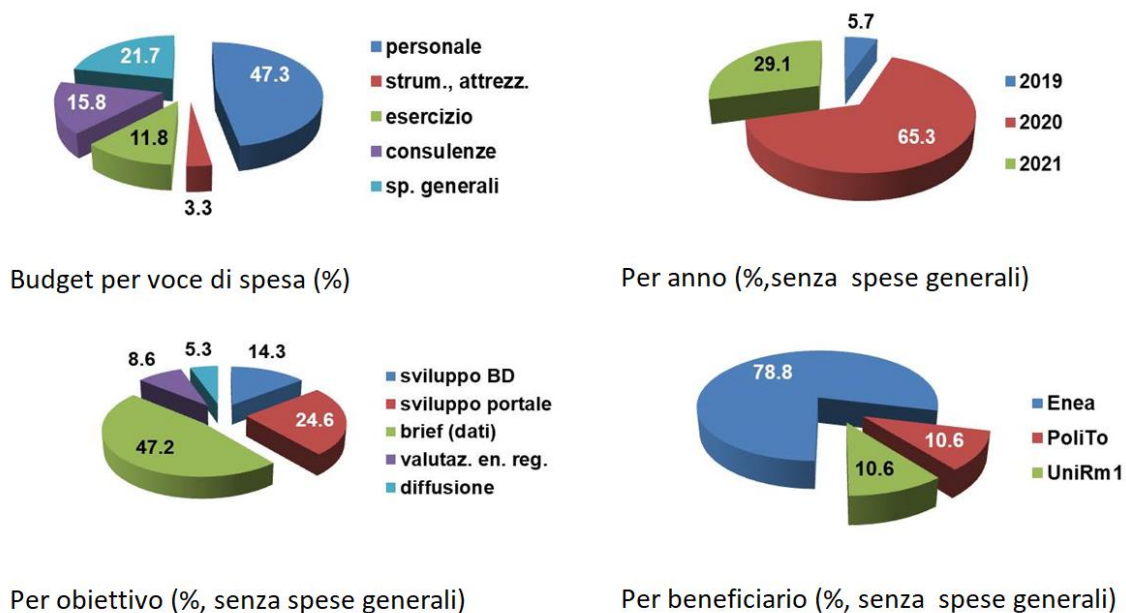


Figura 1 – Ripartizione del budget per voce di spesa, per anno, per obiettivo e per beneficiario.

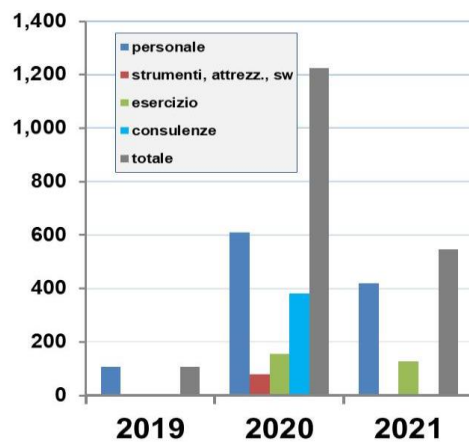


Figura 2 – Costi del progetto (k€) per voce di spesa nel triennio 2019-21 (senza spese generali)

### 1.1 Impatto della pandemia Covid-19 sulle attività previste per l'anno 2020

La pandemia Covid-19 che ha investito il paese all'inizio del 2020 ha causato ritardi nell'approvazione formale del progetto (Giugno 2020) e un conseguente ritardo nella disponibilità dei fondi. Successivamente, la riorganizzazione delle attività lavorative conseguente alla pandemia, unitamente ai tempi tecnici richiesti dalle procedure di gara per l'acquisto di forniture e consulenze esterne, non hanno consentito l'avvio e il regolare svolgimento di alcune linee di attività previste per il 2020. In particolare, non è stato possibile avviare la **LA 1.5** (Sviluppo portale: Valutazione e prova di applicabilità di strumenti SW basati su tecnologie AI), la **LA 1.6** (Sviluppo portale: Implementazione e addestramento di SW basati su AI). Conseguentemente, non è stato possibile avviare la **LA 1.7** (Sviluppo portale: ammodernamento HW), **LA 1.2** (Sviluppo data-base: Implementazione di funzioni statistiche per analisi prestazioni e costi) e la **LA 1.22** (Attività di diffusione 2020), queste ultime dipendenti per diverse ragioni dalle precedenti LA 1.5 e 1.6 e ad esse collegate. In particolare, la gara per la fornitura di SW-AI e per gli associati servizi di prova, implementazione e addestramento del SW stesso è stata aggiudicata in via provvisoria soltanto in data 16 Dicembre 2020 e le attività relative alle LA 1.5 e 1.6 sono ufficialmente iniziate nell'Aprile 2021 (verbale inizio lavori) e comprese nei restanti 8 mesi del 2021 a fronte di una durata inizialmente prevista per le LA 1.5 e 1.6 di complessivi 12 mesi (dal Gennaio al Dicembre 2020).

A tale proposito va rilevato che le attività di implementazione/addestramento del SW-AI non possono essere comprese in tempi sensibilmente più brevi di quelli previsti in ragione del loro carattere fortemente innovativo che riguarda l'addestramento del SW-AI alla lettura automatica di documenti tecnici e pubblicazioni tecnico-scientifiche, con riconoscimento ed estrazione di dati tecnico-economici di interesse (prestazioni, emissioni e costi delle tecnologie per la produzione e uso finale dell'energia) nelle rispettive unità di misura, e successiva gestione e restituzione di dati e informazioni su richiesta dell'utente della piattaforma SIMTE. Le interazioni con gli esperti della società vincitrice della gara (Sistemi Informativi/IBM) hanno confermato che l'applicazione in questione è assolutamente innovativa e di frontiera.

Allo sviluppo delle LA 1.5 e 1.6 sono peraltro collegate per le ragioni di seguito esposte la LA 1.7 (l'ammodernamento HW del portale dipende dall'implementazione del SW-AI), la LA 1.2 (le funzioni di analisi statistica hanno applicazione in presenza di una larga base acquisibile solo con il contributo di SW-AI), e la LA 1.22 (la promozione verso l'utenza dei servizi offerti dalla piattaforma SIMTE è fortemente caratterizzata dalla presenza di SW-AI e dell'associato *Assistente Virtuale*, soprattutto per quanto riguarda la facilità di fruizione dei servizi offerti da parte dell'utenza).

Per le motivazioni sopra esposte è stata proposta ed accettata una variante di progetto che posticipa all'annualità 2021 lo svolgimento delle LA 1.5, 1.6, 1.7, 1.2 e 1.22.

\*\*\*

Nel corso della terza annualità (periodo 01.01.2021 – 31.12.2021) sono state quindi svolte attività relative a 12 linee di attività ed in particolare **LA1.2, LA1.4, LA1.5, LA1.6, LA1.7, LA1.13, LA1.14, LA1.15, LA 1.16, LA1.20, LA1.22 e LA1.23**. Di queste:

- le LA 1.2, 1.5, 1.6 e 1.7 riguardano l'aggiornamento SW e HW del portale e del database della piattaforma SiMTE con introduzione di strumenti basati su intelligenza artificiale AI;
- le LA1.13-16 riguardano l'aggiornamento e/o la redazione ex novo di circa 50 documenti tecnici (relativi ad altrettante tecnologie energetiche) della piattaforma SiMTE. Si tratta di una attività divisa in 4 diverse LA per ragioni contabili e di rendicontazione in quanto affidata, oltre che ad Enea, a due diversi Co-beneficiari Universitari (Politecnico di Torino e Sapienza di Roma) e relativa a tre diversi macro-settori delle tecnologie energetiche (produzione e trasformazione dell'energia primaria, industria e trasporti);
- la LA 1.20 riguarda la valutazione energetica della regione Piemonte;
- le LA 1.22 e 1.23 riguardano le attività di diffusione

Le 12 LA di cui sopra hanno impegnato complessivamente risorse di personale per circa ore 20950 ore corrispondenti ad un impegno complessivo di circa **12.25 persona-anno** di personale professionale di diversi livelli di esperienza.

## 1.2 Brevi informazioni sulla piattaforma SiMTE

La banca dati e il portale SiMTE sono stati realizzati da Enea nel periodo Febbraio 2016-Gennaio 2018 per conto del MiSE. Si tratta, come previsto dai già citati dispositivi di legge, di un sistema informativo e di monitoraggio dello stato e delle prospettive delle tecnologie di offerta e domanda dell'energia a supporto di operatori e decisori pubblici e privati del settore energetico. Il sistema può essere inoltre di ausilio per affrontare le scelte di politica energetica necessarie per realizzare la rapida transizione energetica ed ecologica che dovrà caratterizzare il decennio 2020-30 al fine di far fronte alla questione climatica.

Nella configurazione 2018 il portale SiMTE ([www.simte.enea.it](http://www.simte.enea.it), Figura 3) era aperto in via *sperimentale* in attesa del necessario aggiornamento ed ampliamento. La banca dati offriva infatti documenti, informazioni e dati (prestazioni tecniche, consumi energetici, emissioni, normative, quote di mercato, peculiarità nazionali, costi di investimento, di esercizio, di prodotto/servizio, etc.) sulle tecnologie energetiche aggiornati al Gennaio 2018 e conteneva circa 50 tecnologie o cluster tecnologici divisi nei 5 settori già elencati in precedenza (produzione e trasformazione di energia primaria; produzione e distribuzione di elettricità e calore; residenziale e servizi; trasporti; industria). Al termine dell'attuale progetto è previsto il raddoppio delle tecnologie presenti sulla piattaforma.

Per ogni tecnologia o cluster di tecnologie (Figura 4) i dati e le informazioni vengono discussi e contestualizzati in un documento di base (brief) accessibile all'utenza, redatto da esperti e periodicamente aggiornato. I dati, trasferiti in banca dati, sono accessibili all'utenza mediante sistemi di ricerca e analisi che consentono anche il confronto tra varie tecnologie a richiesta dell'utente. I dati possono poi essere trasferiti ad un modello che ricalcola prestazioni e costi delle tecnologie in base a criteri omogenei. L'input del modello, derivato inizialmente dai dati presenti in banca-dati e reso pienamente trasparente per l'utente, può essere modificato con dati immessi dall'utente stesso. Il modello è applicabile a tutte le tecnologie di domanda/offerta. Di conseguenza, pur non sostituendosi alle accurate analisi costi-benefici a cura dell'investitore, il modello consente confronti costi/benefici (prestazioni, emissioni, costi) tra tecnologie anche afferenti a settori diversi. Esso è destinato prevalentemente ad utenti "esperti" (con conoscenze nel settore dell'analisi costi-benefici delle tecnologie energetiche) e consente anche l'analisi di

sistemi complessi (es: efficientamento nel settore residenziale, cogenerazione, impianti con *energy storage*). In tal caso l'utente potrà anche avvalersi di casi pilota già elaborati, disponibili sul portale.

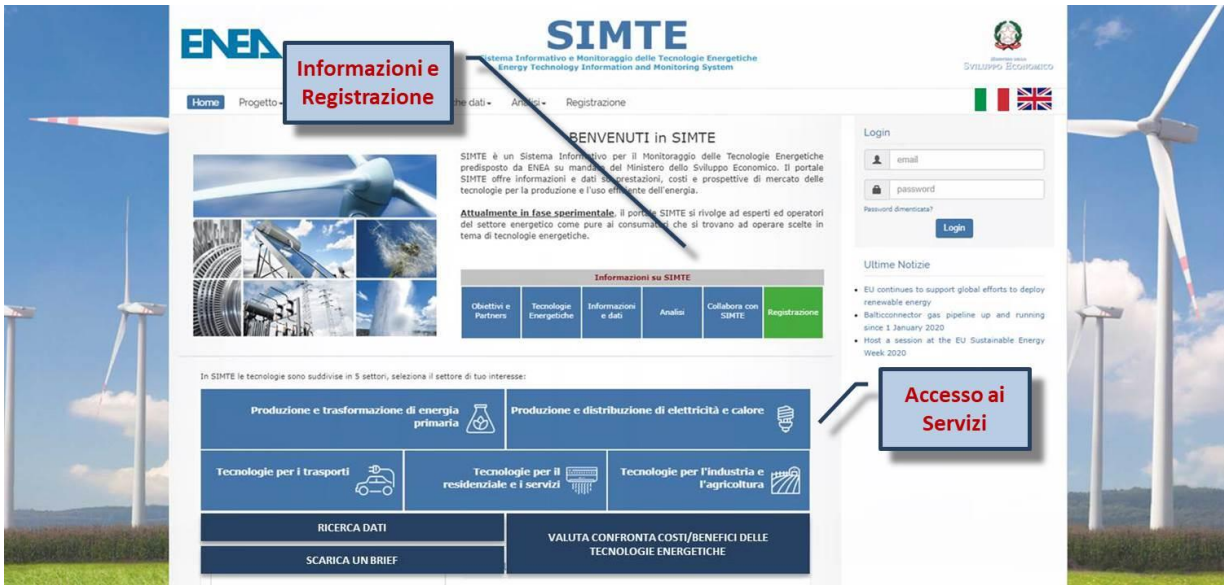


Figura 3 – Homepage della piattaforma SiMTE ([www.simte.enea.it](http://www.simte.enea.it) oppure google “simte”)

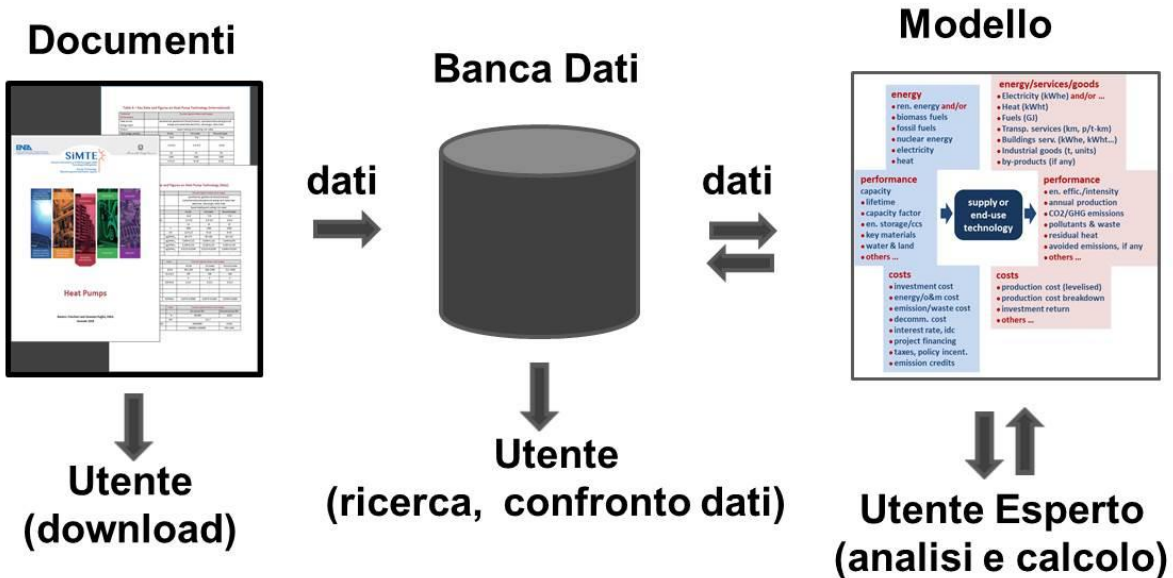


Figura 4 – Struttura, flusso dati e servizi della Banca-dati SiMTE

## 2 Descrizione delle attività svolte e risultati della LA 1.6

### 2.1 Introduzione

Come previsto dal Piano Triennale Realizzativo (PTR) del progetto, nel corso dell'annualità 2021 sono stati svolti ed ultimati i lavori relativi alla linea di attività LA 1.6 che riguarda l'implementazione sulla piattaforma SiMTE di un software (SW) cognitivo basato su intelligenza artificiale (AI, Artificial Intelligence) per svolgere le seguenti funzioni:

- a) lettura automatica di documenti tecnico-economici sulle tecnologie energetiche con riconoscimento ed estrazione dei dati di interesse (prestazioni, costi, aspetti di mercato) e trasferimento dei dati alla banca dati SiMTE;
- b) assistenza agli utenti della piattaforma SiMTE (mediante un assistente virtuale) nell'utilizzo dei servizi offerti tra i quali :
  - dati e informazioni sulle tecnologie energetiche (accesso alla banca dati), ivi incluso il download dei documenti presenti sulla piattaforma stessa;
  - modello di calcolo per l'analisi tecnico-economica delle tecnologie (analisi costi-benefici, con dati presenti nella banca dati o immessi dall'utente).

L'introduzione dell'AI sulla piattaforma SiMTE costituisce la parte più **innovativa e sperimentale** del progetto. Essa è motivata dalla necessità di ridurre i notevoli oneri (in termini di tempo, risorse umane e costi) derivanti dal necessario continuo aggiornamento della piattaforma (per tenere il passo con l'evoluzione di innovazioni, prestazioni, costi e aspetti di mercato delle tecnologie energetiche) e di rendere più fruibili per l'utenza i servizi offerti. In assenza di SW cognitivo, l'aggiornamento dei dati e dei documenti presenti sulla piattaforma, come pure l'acquisizione di nuovi documenti e il trasferimento dei dati alla banca dati, sono processi che impegnano esperti dei vari settori tecnologici per tempi considerevoli. Inoltre, attraverso un *assistente virtuale*, l'AI è in grado di semplificare per l'utenza la fruizione dei servizi offerti dalla piattaforma guidando l'utente nelle varie operazioni per l'accesso alla banca dati e al modello.

L'implementazione di SW cognitivo richiede tuttavia non solo la fornitura del SW ma un vero e proprio processo di addestramento del SW stesso per le funzioni da svolgere, che richiede competenze specialistiche. Per tale ragione l'implementazione del SW ha richiesto l'assegnazione (tramite gara di appalto) di un contratto non solo per la fornitura (accesso al cloud ove il SW risiede) ma anche per l'addestramento del SW per svolgere le funzioni previste, oltre ad un training del personale ENEA per acquisire autonomia nella gestione e nell'utilizzo del SW (ivi compreso le competenze per l'addestramento).

La natura innovativa dell'applicazione e la ridotta conoscenza delle reali potenzialità dei nuovi SW cognitivi in relazione agli obiettivi preposti hanno indotto a dividere l'attività di implementazione in due fasi distinte e in due diverse linee di attività:

- LA 1.5 - Verifica di applicabilità dell'AI alle esigenze della piattaforma SiMTE (Fase 1)
- LA 1.6 - Implementazione dell'AI sulla piattaforma SiMTE (Fase 2, a valle di esito positivo della Fase 1).

La prima fase è stata dedicata in primo luogo allo scambio di competenze tra committente e fornitore (rispettivamente, tecnologie energetiche e tecnica di addestramento) al fine di pervenire ad una reciproca comprensione di esigenze ed obiettivi e, in secondo luogo, alla verifica della capacità del SW di raggiungere gli obiettivi prefissati;

La seconda fase ha riguardato l'effettiva implementazione del SW, l'addestramento e la messa a punto di tutte le funzionalità previste.

Una terza linea di attività, la LA 1.2, che prevede la implementazione sulla piattaforma SiMTE di funzioni statistiche per analisi di prestazioni e costi di tecnologie energetiche di larga diffusione commerciale (es: veicoli, elettrodomestici) è confluita in corso d'opera nella implementazione dell'AI poiché le funzioni statistiche, con un SW ed un lavoro aggiuntivo, sono state integrate nel sistema guidato dall'AI.

In ottemperanza alle regole di *reporting* della Ricerca di Sistema Elettrico, le tre LA di cui sopra (1.2, 1.5 e 1.6) sono oggetto di rapporti tecnici distinti (rispettivamente RdS/PTR2021/068, 070, 071) ma è evidente che sul piano sostanziale le LA 1.2 e 1.5 sono strettamente legate alla LA 1.6 e i rispettivi rapporti condividono parti del presente rapporto.

Vale la pena infine di ricordare che l'implementazione dell'AI era inizialmente prevista per l'annualità 2020 per un tempo di espletamento complessivo di 12 mesi (4 mesi per la Fase 1 e 8 mesi per la Fase 2). I ritardi conseguenti all'emergenza Covid unitamente ad alcune difficoltà dell'Amministrazione ENEA non hanno consentito tuttavia di procedere celermente all'assegnazione di un contratto di fornitura nell'ambito della cosiddetta "Convenzione CRUI", ambito nel quale, a suo tempo è stata espletata una gara internazionale (vinta da IBM) per la fornitura di SW-AI ad Università ed enti di ricerca. E' stato quindi necessario istruire una nuova gara di appalto internazionale i cui tempi di espletamento hanno portato ad iniziare i lavori soltanto nell'Aprile 2021 e alla necessità di comprimerli in poco più di 6 mesi. Oltre alla notevole compressione dei tempi di lavoro, è stato anche necessario richiedere l'approvazione di una variante di progetto per spostare al 2021 tutte le LA collegate alla introduzione dell'AI.

Nel seguito vengono descritti in sintesi le varie fasi del lavoro di implementazione e addestramento dell'AI e i risultati ottenuti, rimandando agli allegati tecnici (Allegati 1 e 2) redatti dalla società vincitrice della gara (Sistemi Informativi, gruppo IBM) per maggiori dettagli tecnici.

## 2.2 Attività svolte nella Fase 1

Come accennato nella introduzione, al fine di pervenire ad maggiore comprensione delle potenzialità (e dei limiti) del SW cognitivo da parte del Committente (ENEA) e degli obiettivi da raggiungere da parte del Fornitore del SW (Sistemi Informativi/IBM), nella parte iniziale della Fase 1 sono state messe in comune le conoscenze e la terminologia informatica del Fornitore con quelle del Committente in tema di caratterizzazione tecnico-economica delle tecnologie energetiche. Il Fornitore ha dovuto familiarizzare con termini quali *efficienza e intensità energetica, fattore di capacità, vita operativa, costi di investimento e di esercizio delle tecnologie* (e con le rispettive unità di misura) mentre il Committente ha dovuto familiarizzare con il significato attribuito a termini quali *intenti, entità e relazioni, indicatori*, utilizzati nella descrizione delle tecniche di addestramento dei SW cognitivi. Solo al termine di questa fase di *merging* delle competenze è stato possibile per il Fornitore comprendere con il necessario dettaglio gli obiettivi da raggiungere e le caratteristiche del sistema da realizzare e, per il Committente, comprendere le reali potenzialità e i limiti degli attuali SW cognitivi.

La fase di *merging* delle competenze ha evidenziato, tra l'altro, come l'assegnazione (tramite gara) di un contratto per la fornitura di servizi (imposta in base alla generica interpretazione della normativa vigente) mal si presti a realizzare una applicazione sperimentale del tipo in questione. Le procedure di gara non consentono infatti un approfondito dialogo preliminare tra i partecipanti alla gara e il committente per arrivare ad una precisa definizione di obiettivi raggiungibili. In pratica il committente non può conoscere in misura sufficiente i limiti e le potenzialità del SW cognitivo e i fornitori non possono conoscere in misura sufficiente l'ambito operativo e le esigenze del committente. Meglio sarebbe procedere con l'assegnazione diretta di un contratto di ricerca dopo una maggiore comprensione delle reciproche esigenze.

Al termine della fase di *merging* è stato quindi possibile iniziare il lavoro con la cosiddetta "**Raccolta dei requisiti funzionali**" (vedi Allegato 2).

Parallelamente alla raccolta requisiti funzionali, è stato preparato l'ambiente informatico operativo. Il SW cognitivo risiede infatti fisicamente su IBM Cloud. Sono quindi stati stabiliti i necessari collegamenti

informatici tra la piattaforma SiMTE residente su server ENEA e il Cloud IBM, inizializzando i servizi necessari per lo sviluppo del sistema e replicando sul Cloud la struttura del database SiMTE per il salvataggio dei dati estratti dai documenti.

Nella seconda parte della Fase 1 è stata verificata l'applicabilità del SW per gli obiettivi fissati mediante una applicazione dimostrativa ad alcune **tecnologie pilota**. In questo ambito sono state fatte alcune scelte relative alla lingua di addestramento del SW e alle tecnologie pilota e si è proceduto all'addestramento sulle tecnologie pilota per poi verificarne le funzionalità.

Per velocizzare il processo (dati i ristretti tempi a disposizione rispetto alla programmazione iniziale), è stato deciso di procedere all'addestramento del SW in lingua inglese per una duplice ragione: pur potendo operare in molte lingue, il SW è nativo in lingua inglese; inoltre la terminologia utilizzata nel settore delle tecnologie energetiche è largamente diffusa e utilizzata in lingua inglese. L'addestramento in lingua inglese è stato possibile poiché tutti i documenti sulle tecnologie energetiche presenti sulla piattaforma SiMTE sono stati prodotti in doppia lingua inglese-italiano, nella prospettiva di rendere accessibile il sistema anche in ambito internazionale.

Per quanto riguarda la scelta delle tecnologie pilota, sono state selezionate alcune delle più diffuse tecnologie del settore della elettro-generazione, in particolare, produzione di elettricità da combustibili fossili (COAL, coal-fired power), da solare fotovoltaico (PV, solar photovoltaics) e da solare a concentrazione (CSP, concentrating solar power) poiché la terminologia tecnico-economica utilizzata in tale settore è più largamente diffusa e standardizzata (es: *energy efficiency, capacity factor, lifetime, overnight investment cost, levelized cost of electricity, LCOE*).

Il SW utilizzato per l'applicazione è il sistema *IBM Watson* costituito da una famiglia di diversi SW cognitivi, ognuno dei quali preposto ad una specifica funzione (*Watson Assistant, Watson Discovery, Watson Knowledge Studio, Watson Natural Language Understanding*), coordinati da un SW cosiddetto "*Orchestratore*" che, in base ai requisiti funzionali e al progetto dell'architettura informatica del sistema, permette di integrare le funzioni dei vari componenti. Una descrizione delle funzioni dei vari SW, delle fasi di addestramento del SW cognitivo e della terminologia utilizzata (es: intenti, entità, relazioni) e viene fornita nel seguito (per descrizioni più dettagliate e in linguaggio più tecnico si rimanda all'Allegato 1 per i componenti del SW e all'Allegato 2 per le fasi di addestramento).

Le attività svolte in questa fase possono quindi essere riassunte come segue:

- analisi dei requisiti funzionali e definizione dell'architettura del sistema;
- implementazione del nucleo del SW cognitivo (orchestratore e componenti);
- integrazione con la base dati del sistema SiMTE, replicato su cloud IBM;
- sviluppo e integrazione dell'interfaccia per la funzione di **Assistente Virtuale**;
- *ingestion* dei documenti SiMTE sottoposti al SW cognitivo e implementazione delle funzioni di storage e download dei documenti.

Nella fase iniziale alcune sessioni di lavoro congiunto tra esperti ENEA e IBM/Sistemi sono state dedicate alla definizione di una sorta di **glossari** di termini (categorie, parametri indicatori, grandezze fisiche, tecniche ed economiche, con rispettivi sinonimi, varianti semantiche e unità di misura) che intervengono nella caratterizzazione tecnico-economica delle tecnologie energetiche. Tali glossari sono stati poi utilizzati nella varie fasi di addestramento del SW unitamente ai documenti relativi alle tecnologie pilota selezionate.

Di seguito vengono descritte le principali fasi di preparazione del sistema e di addestramento del SW cognitivo.

### 2.2.1 Assistente Virtuale

L'**Assistente Virtuale** è un componente fondamentale del sistema in quanto presiede a tutte le interazioni con l'utente della piattaforma SiMTE. Per lo sviluppo della funzione di Assistente Virtuale sono stati

sviluppati i componenti grafici di base dell'interfaccia utente (Figura 5) e i necessari collegamenti con l'Orchestratore e con i componenti Watson Assistant e Watson Discovery.

Il SW Watson Assistant consente di aggiungere alle applicazioni sviluppate la funzionalità di conversazione con l'utente in linguaggio naturale. È stato quindi creato l'assistente *ENEA AI SiMTE*, dotato di due "skills", corrispondenti a differenti capacità: una capacità di *Search*, in grado di richiamare le informazioni dal SW Watson Discovery che presiede alla gestione dei documenti, ed una capacità di *Dialog*, per permettere all'Assistente di comprendere e rispondere a domande e richieste dell'utente utilizzando tecniche di machine learning ed elaborazioni in linguaggio naturale. Le attività di addestramento dell'**Assistente Virtuale** possono essere suddivise in tre categorie relative agli **intenti**, alle **entità**, e al **dialogo**.

Per creare le conversazioni si parte dagli **intenti** che servono a categorizzare l'input dell'utente, "incanalando" le diverse tipologie di richieste. Per ogni intento l'assistente è stato addestrato con frasi di esempio per consentire l'identificazione dell'input dell'utente da parte dell'Assistente. Successivamente alla creazione e categorizzazione degli intenti, sono state create le **entità**, che rappresentano termini rilevanti per gli intenti gestiti e che forniscono quindi un contesto distintivo per un dato intento. In altri termini, le entità sono valori specifici (in questo caso, parole) da cercare nell'input dell'utente, con vari sinonimi per ciascun valore; esse rappresentano gli oggetti di interesse che aiutano a caratterizzare l'input (e quindi l'intenzione) dell'utente. Dopo aver definito intenti ed entità, è stato costruito il **dialogo**. Tale costruzione è stata fatta in maniera grafica (Figura 6), attraverso una serie di nodi collegati tra loro. Ogni nodo contiene condizioni che vengono valutate, al fine di accedere o meno ai relativi output. La valutazione dei nodi avviene dall'alto verso il basso.

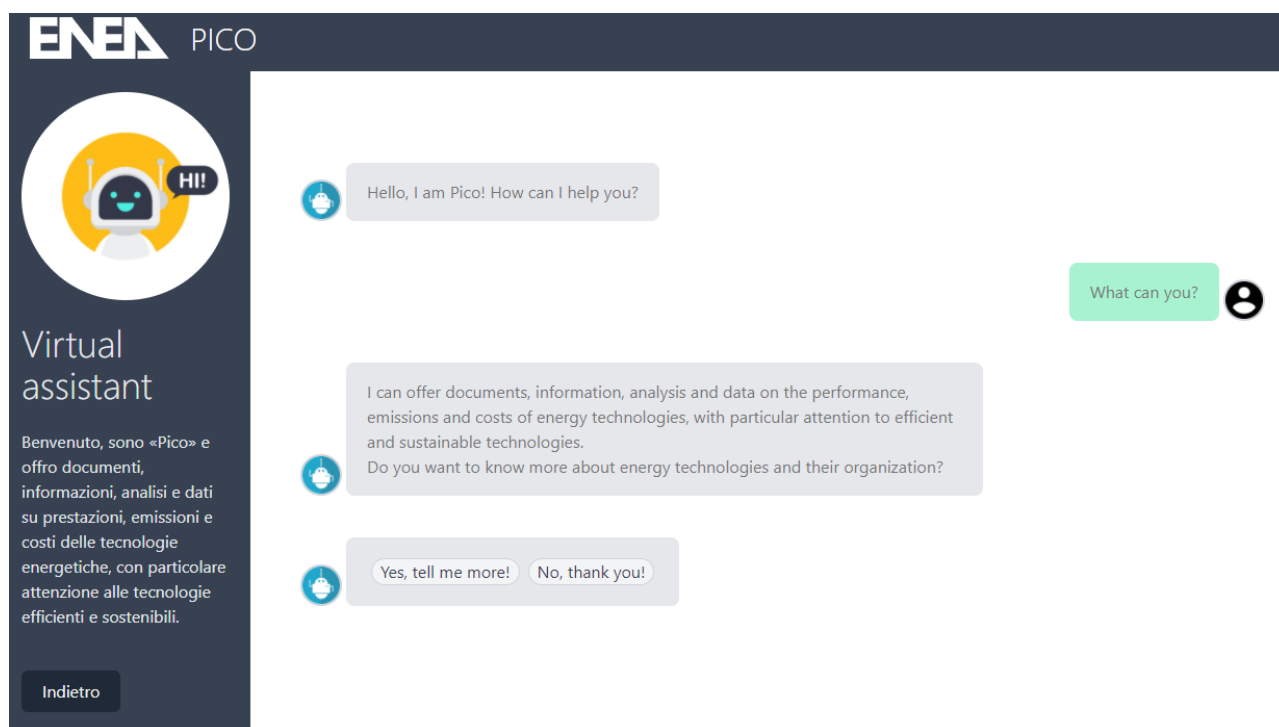


Figura 5 - Interfaccia Grafica dell'Assistente Virtuale "Pico"

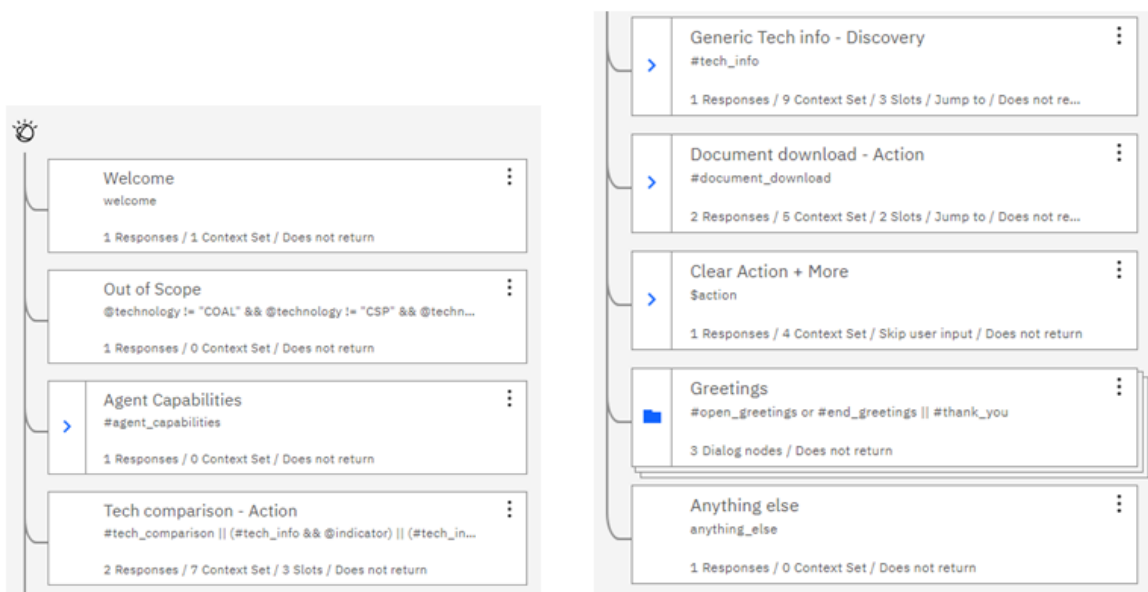


Figura 6 - Dialogo dell'Assistente Virtuale

### 2.2.2 Gestione dei documenti

Le attività relative alla gestione dei documenti coinvolgono i componenti Watson Knowledge Studio e Watson Discovery.

Il componente Watson Knowledge Studio genera modelli cognitivi personalizzati che identificano ed estraggono, da testi non strutturati, **entità e relazioni** univoche per il settore o dominio su cui il modello viene addestrato. In pratica il componente Watson Knowledge Studio presiede un flusso di lavoro (workflow) costituito da una serie di passi del tipo rappresentato in Figura 7.

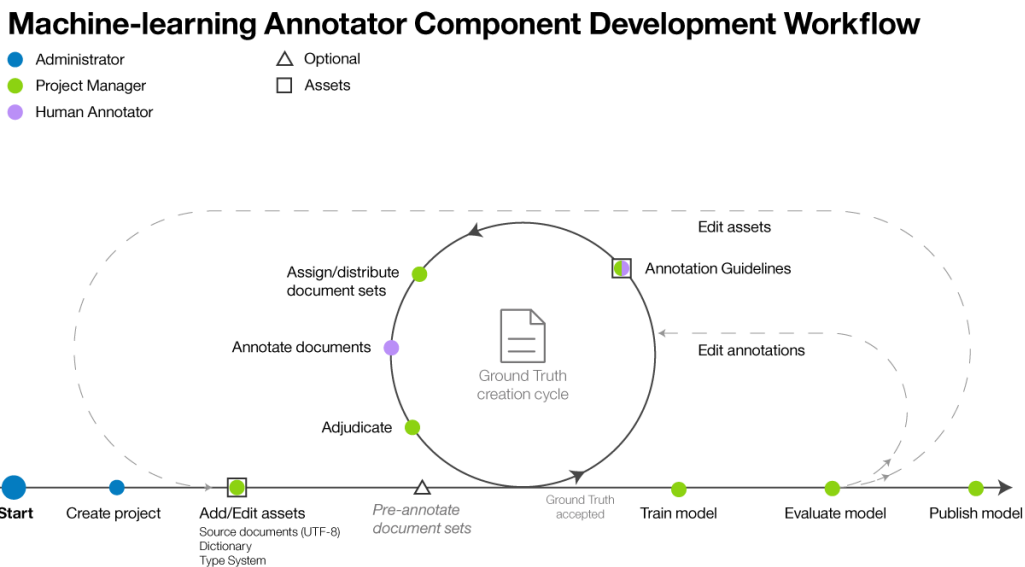


Figura 7 - Workflow di Watson Knowledge Studio per la creazione di un modello Machine Learning.

Il primo passo del workflow prevede la creazione del workspace dedicato al progetto, da generare nella lingua dei documenti che dovranno poi essere analizzati dal modello. Per il progetto SiMTE è stato creato il workspace "ENEA AI-SiMTE" in lingua inglese.

Il secondo passo del workflow prevede l'aggiunta degli asset, che si dividono in **type system**, **dizionari**, e **documenti di origine**.

Il **type system** (sistema tipo) definisce le **entità** e le **relazioni** tra le entità che si ritiene debbano essere estratte dai documenti. In particolare, una entità è una categorizzazione di una cosa reale, una "etichetta" da associare ad un particolare concetto; una relazione è un'associazione tra due entità, quindi un collegamento tra due concetti.

Il type system utilizzato in Fase 1 è stato definito in una serie di incontri tra Committente e Fornitore nell'ambito dei quali sono stati tra l'altro definiti i "**glossari**" precedentemente menzionati. Per la Fase 1 sono state aggiunte **106 entità e 5 "macro relazioni"**, che collegano fra loro più entità, per un totale di **1.218 relazioni effettive** (di seguito vengono riportati gli elenchi delle 106 entità utilizzate e delle 5 relazioni con relative descrizioni).

I **dizionari** raggruppano parole e/o frasi che devono essere trattate allo stesso modo dal modello, quindi parole o frasi che, pur non avendo uguale significato, identificano la stessa entità o la stessa relazione. Per la Fase 1 del progetto, sono stati aggiunti dizionari utili per la definizione delle entità, partendo dai termini stabiliti in fase di analisi, integrati con ulteriori definizioni estratte dai documenti oggetto di Fase 1.

I **documenti di origine** (source document) sono i documenti utilizzati per costruire il modello, per l'addestramento e per testare successivamente le funzionalità del modello stesso. Per la Fase 1 del progetto sono stati aggiunti otto documenti (tre per la tecnologia PV, tre per la tecnologia CSP, e due per la tecnologia COAL) suddivisi in otto set di training, e tre set di test. I set di training sono stati poi utilizzati per effettuare il training del modello, mentre quelli di test sono stati usati per testare le funzionalità del modello stesso.

Entità	Descrizione
CONVERSION_EFFICIENCY	L'entità fa riferimento all'indicatore <i>Annual Conversion Efficiency</i> .
AVAILABILITY_FACTOR	L'entità fa riferimento all'indicatore <i>Availability Factor</i> .
AVOIDED_CO2_EMISSIONS	L'entità fa riferimento all'indicatore <i>Avoided CO<sub>2</sub> Emissions</i> .
BOS_COST	L'entità fa riferimento all'indicatore <i>Balance Of System Cost</i> .
CAPACITY	L'entità fa riferimento all'indicatore <i>Capacity</i> .
CAPACITY_FACTOR	L'entità fa riferimento all'indicatore <i>Capacity Factor</i> .
CAPACITY_UNDER_CONSTRUCTION	L'entità fa riferimento all'indicatore <i>Capacity Under Construction</i> .
CELL_EFFICIENCY	L'entità fa riferimento all'indicatore <i>Cell Efficiency</i> .
CO2_EMISSIONS	L'entità fa riferimento all'indicatore <i>CO<sub>2</sub> Emissions</i> .
COAL	L'entità fa riferimento alla tecnologia del <i>Carbone (COAL)</i> .
COAL_IGCC	L'entità fa riferimento alla variante tecnologica <i>Integrated Gasification Combined Cycle (IGCC) Technology</i> del COAL.
COAL_PC	L'entità fa riferimento alla variante tecnologica <i>Pulverized Coal Technology</i> del COAL.
COAL_PC_AUSC	L'entità fa riferimento alla sotto variante tecnologica della <i>Pulverized Coal Technology</i> , la <i>Advanced Ultra-Super-Critical Cycle (AUSC) Technology</i> .
COAL_PC_SC	L'entità fa riferimento alla sotto variante tecnologica della <i>Pulverized Coal Technology</i> , la <i>Super-Critical Cycle (SC) Technology</i> .
COAL_PC_SUB	L'entità fa riferimento alla sotto variante tecnologica della <i>Pulverized Coal Technology</i> , la <i>Sub-Critical Cycle Technology</i> .
COAL_PC_USC	L'entità fa riferimento alla sotto variante tecnologica della <i>Pulverized Coal Technology</i> , la <i>Ultra-Super-Critical Cycle (USC) Technology</i> .
COMMERCIAL_MODULE_EFFICIENCY	L'entità fa riferimento all'indicatore <i>Commercial Module Efficiency</i> .
CONCENTRATION_RATIO	L'entità fa riferimento all'indicatore <i>Concentration Ratio</i> .
CONSTRUCTION_TIME	L'entità fa riferimento all'indicatore <i>Construction Time</i> .
COST_BREAKDOWN	L'entità fa riferimento all'indicatore <i>Cost Breakdown</i> .
CSP	L'entità fa riferimento alla tecnologia del <i>Concentrating Solar Power (CSP, o Solar Thermal Electricity)</i> .
CSP_VAR_CR	L'entità fa riferimento alla variante tecnologica <i>Central Receiver Technology</i> del CSP.
CSP_VAR_CR_SALT_WS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Tower HTF Salt With Storage Technology</i> .

CSP_VAR_CR_STEAM_NS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Tower HTF Water Without Storage Technology</i> .
CSP_VAR_DISH	L'entità fa riferimento alla variante tecnologica <i>Dish System Technology</i> del CSP.
CSP_VAR_LFR	L'entità fa riferimento alla variante tecnologica <i>Parabolic Trough Technology</i> del CSP.
CSP_VAR_LFR_STEAM_NS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Fresnel HTF Water Without Storage Technology</i> .
CSP_VAR_LFR_STEAM_WS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Fresnel HTF Water With Storage Technology</i> .
CSP_VAR_PT	L'entità fa riferimento alla variante tecnologica <i>Linear Fresnel Reflector Technology</i> del CSP.
CSP_VAR_PT_OIL_NS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Trough HTF Oil Without Storage Technology</i> .
CSP_VAR_PT_OIL_WS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Trough HTF Oil With Storage Technology</i> .
CSP_VAR_PT_OIL_WWS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Trough HTF Oil With/Without Storage Technology</i> .
CSP_VAR_PT_SALT_WS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Trough HTF Salt With Storage Technology</i> .
CSP_VAR_PT_STEAM_NS	L'entità fa riferimento alla sotto variante tecnologica della <i>Linear Fresnel Reflector Technology</i> , la <i>Trough HTF Water Without Storage Technology</i> .
CYCLE_PRESSURE	L'entità fa riferimento all'indicatore <i>Cycle Pressure</i> .
CYCLE_TEMPERATURE	L'entità fa riferimento all'indicatore <i>Cycle Temperature</i> .
CYCLE_TYPE	L'entità fa riferimento all'indicatore <i>Cycle Type</i> .
DECOMMISSIONING_COST	L'entità fa riferimento all'indicatore <i>Decommissioning Cost</i> .
ELECTRICITY_COST	L'entità fa riferimento all'indicatore <i>Electricity Cost</i> .
ELECTRICITY_SHARE	L'entità fa riferimento all'indicatore <i>Electricity Share</i> .
EMISSIONS	L'entità fa riferimento all'indicatore <i>Emissions</i> .
ENERGY_EFFICIENCY	L'entità fa riferimento all'indicatore <i>Energy Efficiency</i> .
ENERGY_INPUT_COST	L'entità fa riferimento all'indicatore <i>Energy Input Cost</i> .
ENERGY_PAYBACK_TIME	L'entità fa riferimento all'indicatore <i>Energy Payback Time</i> .
EXHAUSTED_HEAT_COOLING	L'entità fa riferimento all'indicatore <i>Exhausted Heat Cooling</i> .
EXTERNAL_COST	L'entità fa riferimento all'indicatore <i>External Cost</i> .
FINANCING_COST	L'entità fa riferimento all'indicatore <i>Financing Cost</i> .
FOM_COST	L'entità fa riferimento all'indicatore <i>Fixed Operation &amp; Maintenance Cost</i> .
FUEL_COST	L'entità fa riferimento all'indicatore <i>Fuel Cost</i> .
GHG_EMISSIONS_COST	L'entità fa riferimento all'indicatore <i>GHG Emissions Cost</i> .
HEAT_STORAGE_MEDIA	L'entità fa riferimento all'indicatore <i>Heat Storage Media</i> .
HTF_TEMPERATURE	L'entità fa riferimento all'indicatore <i>HTF Temperature</i> .
HTF_TYPE	L'entità fa riferimento all'indicatore <i>HTF Type</i> .
INSTALLED_CAPACITY	L'entità fa riferimento all'indicatore <i>Installed Capacity</i> .
LAND_USE	L'entità fa riferimento all'indicatore <i>Land Use</i> .
LIFETIME	L'entità fa riferimento all'indicatore <i>Lifetime</i> .
MARKET_SHARE	L'entità fa riferimento all'indicatore <i>Market Share</i> .
MARKET_SHARE_PROJECTION	L'entità fa riferimento all'indicatore <i>Market Share Projection</i> .
MATERIAL_USE	L'entità fa riferimento all'indicatore <i>Material Use</i> .
MEASUREMENT_UNIT	L'entità fa riferimento all'unità di misura relativa ai valori associati agli indicatori.
MODULE_COST	L'entità fa riferimento all'indicatore <i>Module Cost</i> .
MODULE_EFFICIENCY	L'entità fa riferimento all'indicatore <i>Module Efficiency</i> .
NOMINAL_EFFICIENCY	L'entità fa riferimento all'indicatore <i>Nominal Efficiency</i> .
NOX_EMISSIONS	L'entità fa riferimento all'indicatore <i>NOx Emissions</i> .
NUMBER_OF_UNITS	L'entità fa riferimento all'indicatore <i>Number Of Units</i> .
OM_COST	L'entità fa riferimento all'indicatore <i>Operation &amp; Maintenance Cost</i> .
PLANNED_CAPACITY	L'entità fa riferimento all'indicatore <i>Planned Capacity</i> .
PM_EMISSIONS	L'entità fa riferimento all'indicatore <i>PM Emissions</i> .
PRESSURE	L'entità fa riferimento all'indicatore <i>Pressure</i> .
PV	L'entità fa riferimento alla tecnologia del <i>Fotovoltaico (PV)</i> .
PV_APPLICATION_CS	L'entità fa riferimento al settore di applicazione del PV di tipo <i>Commercial</i> .
PV_APPLICATION_RS	L'entità fa riferimento al settore di applicazione del PV di tipo <i>Residential</i> .
PV_APPLICATION_US	L'entità fa riferimento al settore di applicazione del PV di tipo <i>Utility</i> .
PV_TYPE_BUILDING	L'entità fa riferimento all'installazione <i>Building Integrated PV System</i> , relativa alla tecnologiaPV.

PV_TYPE_GRID	L'entità fa riferimento all'installazione <i>Grid Connected PV System</i> , relativa alla tecnologia PV.
PV_TYPE_GROUND	L'entità fa riferimento all'installazione <i>Ground Based PV System</i> , relativa alla tecnologia PV.
PV_TYPE_OFFGRID	L'entità fa riferimento al tipo di installazione <i>Off-Grid PV System</i> , relativa alla tecnologia PV.
PV_TYPE_ROOF	L'entità fa riferimento al tipo di installazione <i>Rooftop PV System</i> , relativa alla tecnologia PV.
PV_VAR_CSI	L'entità fa riferimento alla variante tecnologica <i>Wafer-Based Crystalline Silicon Technology</i> PV.
PV_VAR_CSI_MC	L'entità fa riferimento alla sotto variante tecnologica del <i>Wafer-Based Crystalline Silicon Technology</i> , la <i>Single-Crystalline Silicon Technology</i> .
PV_VAR_CSI_SC	L'entità fa riferimento alla sotto variante tecnologica del <i>Wafer-Based Crystalline Silicon Technology</i> , la <i>Multi-Crystalline Silicon Technology</i> .
PV_VAR_NOVEL	L'entità fa riferimento alla variante tecnologica <i>Emerging And Novel PV Technology</i> del PV.
PV_VAR_NOVEL_AITF	L'entità fa riferimento alla sotto variante tecnologica della <i>Emerging And Novel PV Technology</i> , la <i>Advanced Inorganic Thin-Films Technology</i> .
PV_VAR_NOVEL_CPV	L'entità fa riferimento alla sotto variante tecnologica della <i>Emerging And Novel PV Technology</i> , la <i>Concentrating Photovoltaics Technology</i> .
PV_VAR_NOVEL_OSC	L'entità fa riferimento alla sotto variante tecnologica della <i>Emerging And Novel PV Technology</i> , la <i>Organic Solar Cells Technology</i> .
PV_VAR_NOVEL_OSC_DSSC	L'entità fa riferimento alla sotto variante tecnologica della <i>Organic Solar Cells Technology</i> della <i>Emerging And Novel PV Technology</i> , la <i>Hybrid Dye-sensitised Solar Cells Technology</i> .
PV_VAR_NOVEL_OSC_OPV	L'entità fa riferimento alla sotto variante tecnologica della <i>Organic Solar Cells Technology</i> della <i>Emerging And Novel PV Technology</i> , la <i>Organic Photovoltaic Cells Technology</i> .
PV_VAR_TF	L'entità fa riferimento alla variante tecnologica <i>Thin-Film Technology</i> del PV.
PV_VAR_TF_AS	L'entità fa riferimento alla sotto variante tecnologica del <i>Thin-Film Technology</i> , la <i>Amorphous Silicon Technology</i> .
PV_VAR_TF_CIGS	L'entità fa riferimento alla sotto variante tecnologica del <i>Thin-Film Technology</i> , la <i>Copper-indium-[gallium]-[di]selenide-[di]sulphide Film Technology</i> .
PV_VAR_TF_CT	L'entità fa riferimento alla sotto variante tecnologica del <i>Thin-Film Technology</i> , la <i>Cadmium-telluride Film Technology</i> .
PV_VAR_TF_MJS	L'entità fa riferimento alla sotto variante tecnologica del <i>Thin-Film Technology</i> , la <i>Multi-Junction Silicon Technology</i> .
SO2_EMISSIONS	L'entità fa riferimento all'indicatore <i>SO<sub>2</sub> Emissions</i> .
SOLAR_MULTIPLE	L'entità fa riferimento all'indicatore <i>Solar Multiple</i> .
STORAGE	L'entità fa riferimento all'indicatore <i>Storage</i> .
STORAGE_CAPACITY	L'entità fa riferimento all'indicatore <i>Storage Capacity</i> .
SYSTEM_COST	L'entità fa riferimento all'indicatore <i>System Cost</i> .
TEMPERATURE	L'entità fa riferimento all'indicatore <i>Temperature</i> .
TOTAL_CUMULATIVE_CAPACITY	L'entità fa riferimento all'indicatore <i>Total Cumulative Capacity</i> .
TYPICAL_SIZE	L'entità fa riferimento all'indicatore <i>Typical Size</i> .
UNITS_IN_OPERATION	L'entità fa riferimento all'indicatore <i>Units In Operation</i> .
VALUE	L'entità fa riferimento al valore numerico, tipicamente associato agli indicatori.
VALUE_MAX	L'entità fa riferimento al massimo valore numerico, tipicamente associato agli indicatori.
VALUE_MIN	L'entità fa riferimento al minimo valore numerico, tipicamente associato agli indicatori.
VOM_COST	L'entità fa riferimento all'indicatore <i>Variable Operation &amp; Maintenance Cost</i> .
WAFER_THICKNESS	L'entità fa riferimento all'indicatore <i>Wafer Thickness</i> .

Relazione 1	Prima entità	Seconda entità
INDICATORE_DI_COAL  Collega gli indicatori alla tecnologia del COAL o alle sue varianti	FOM_COST, ENERGY_INPUT_COST, INSTALLED_CAPACITY, LIFETIME, OM_COST, EMISSIONS, TEMPERATURE, FUEL_COST, DECOMMISSIONING_COST, PM_EMISSIONS, VOM_COST, ELECTRICITY_COST, CO2_EMISSIONS, MARKET_SHARE, SO2_EMISSIONS, EXTERNAL_COST, NUMBER_OF_UNITS, ENERGY_EFFICIENCY, GHG_EMISSIONS_COST, PRESSURE, CAPACITY_FACTOR, SYSTEM_COST, ELECTRICITY_SHARE, TYPICAL_SIZE, NOX_EMISSIONS	COAL_PC_AUSC, COAL_PC_SC, COAL_IGCC, COAL_PC, COAL_PC_SUB, COAL, COAL_PC_USC

Relazione 2	Prima entità	Seconda entità
INDICATORE_DI_CSP  Collega gli indicatori alla tecnologia del CSP o alle sue	HEAT_STORAGE_MEDIA, ELECTRICITY_SHARE, OM_COST, SOLAR_MULTIPLE, ELECTRICITY_COST, CAPACITY, FINANCING_COST, MARKET_SHARE, CAPACITY_FACTOR, CONSTRUCTION_TIME, INSTALLED_CAPACITY, CYCLE_PRESSURE, HTF_TEMPERATURE, TOTAL_CUMULATIVE_CAPACITY,	CSP_VAR_LFR_STEAM_WS, CSP_VAR_CR_SALT_WS, CSP_VAR_CR, CSP_VAR_CR_STEAM_NS, CSP_VAR_PT, CSP, CSP_VAR_PT_OIL_WS, CSP_VAR_LFR,

varianti	STORAGE_CAPACITY, TEMPERATURE, LIFETIME, NOMINAL_EFFICIENCY, ANNUAL_CONVERSION_EFFICIENCY, ENERGY_EFFICIENCY, CYCLE_TEMPERATURE, CONCENTRATION_RATIO, CAPACITY_UNDER_CONSTRUCTION, SYSTEM_COST, LAND_USE, HTF_TYPE, STORAGE, TYPICAL_SIZE, CYCLE_TYPE, NUMBER_OF_UNITS, PLANNED_CAPACITY	CSP_VAR_LFR_STEAM_NS, CSP_VAR_PT_OIL_WWS, CSP_VAR_PT_SALT_WS, CSP_VAR_PT_STEAM_NS, CSP_VAR_DISH, CSP_VAR_PT_OIL_NS
----------	--	--

Relazione 3	Prima entità	Seconda entità
INDICATORE_DI_PV  collega gli indicatori alla tecnologia del PV o alle sue varianti	MARKET_SHARE_PROJECTION, WAFER_THICKNESS, CELL_EFFICIENCY, ANNUAL_CONVERSION_EFFICIENCY, MODULE_EFFICIENCY, AVOIDED_CO2_EMISSIONS, COMMERCIAL_MODULE_EFFICIENCY, CAPACITY, EMISSIONS, MARKET_SHARE, BOS_COST, NOMINAL_EFFICIENCY, INSTALLED_CAPACITY, COST_BREAKDOWN, CO2_EMISSIONS, OM_COST, MATERIAL_USE, CAPACITY_FACTOR, MODULE_COST, LIFETIME, ENERGY_PAYBACK_TIME, ELECTRICITY_COST, SYSTEM_COST, AVAILABILITY_FACTOR, TYPICAL_SIZE, ENERGY_EFFICIENCY, ELECTRICITY_SHARE, TOTAL_CUMULATIVE_CAPACITY, LAND_USE	PV_VAR_NOVEL_OSC_DSSC, PV_VAR_NOVEL_OSC_OPV, PV_VAR_NOVEL_CPV, PV_VAR_TF_CIGS, PV_VAR_NOVEL_OSC, PV_VAR_NOVEL_AITF, PV_VAR_TF_AS, PV_VAR_CSI, PV_VAR_NOVEL_PV_VAR_TF_MJS, PV, PV_VAR_TF, PV_VAR_CSI_MC, PV_VAR_TF_CT, PV_VAR_CSI_SC

Relazione 4	Prima entità	Seconda entità
UNITA_DI  collega l'unità di misura al relativo valore	MEASUREMENT_UNIT	VALUE, VALUE_MIN, VALUE_MAX

Relazione 5	Prima entità	Seconda entità
VALORE_DI  collega il valore al relativo indicatore	VALUE, VALUE_MAX, VALUE_MIN	MARKET_SHARE, HTF_TEMPERATURE, AVAILABILITY_FACTOR, LAND_USE, CAPACITY_UNDER_CONSTRUCTION, CAPACITY, ELECTRICITY_SHARE, PLANNED_CAPACITY, ENERGY_PAYBACK_TIME, MODULE_EFFICIENCY, NOMINAL_EFFICIENCY, OM_COST, FINANCING_COST, LIFETIME, ENERGY_INPUT_COST, BOS_COST, INSTALLED_CAPACITY, WAFER_THICKNESS, SO2_EMISSIONS, MARKET_SHARE_PROJECTION, VOM_COST, TYPICAL_SIZE, EMISSIONS, PM_EMISSIONS, TOTAL_CUMULATIVE_CAPACITY, ANNUAL_CONVERSION_EFFICIENCY, STORAGE_CAPACITY, TEMPERATURE, DECOMMISSIONING_COST, ELECTRICITY_COST, GHG_EMISSIONS_COST, COMMERCIAL_MODULE_EFFICIENCY, SYSTEM_COST, HEAT_STORAGE_MEDIA, EXTERNAL_COST, CONCENTRATION_RATIO, CONSTRUCTION_TIME, FOM_COST, AVOIDED_CO2_EMISSIONS, NOX_EMISSIONS, CYCLE_TEMPERATURE, MODULE_COST, FUEL_COST, STORAGE, CYCLE_TYPE, SOLAR_MULTIPLE, MATERIAL_USE, COST_BREAKDOWN, CELL_EFFICIENCY, CO2_EMISSIONS, CYCLE_PRESSURE, HTF_TYPE, PRESSURE, NUMBER_OF_UNITS, UNITS_IN_OPERATION, ENERGY_EFFICIENCY, CAPACITY_FACTOR

### 2.2.3 Pre-annotazione e annotazione dei documenti di training e training del modello

Una volta definiti e aggiunti gli asset, si è proceduto alla pre-annotazione dei documenti di training. Nella Fase 1 la pre-annotazione è stata effettuata partendo dai dizionari e da un modello Rule Based appositamente creato per definire le entità VALUE, VALUE\_MAX e VALUE\_MIN, necessarie per categorizzare rispettivamente i valori tipici, massimi e minimi delle entità indicatori (i parametri tecnici ed economici che caratterizzano la tecnologia). Terminata la pre-annotazione, per ogni set di documenti di training è stato creato il relativo Annotation Set, una copia dei set pre-annotati destinati ad essere lavorati dagli annotatori umani. In Fase 1 del progetto l'annotazione degli otto Annotation Set di training è stata suddivisa fra due annotatori identificati nelle persone di due Senior Expert IBM. L'annotazione da parte dell'annotatore umano prevede l'eventuale correzione e/o integrazione delle entità e relazioni annotate automaticamente da Watson Knowledge Studio tramite la pre-annotazione, oltre all'eventuale annotazione della co-reference che permette al modello di capire se le entità all'interno del documento si riferiscono allo stesso concetto. Nella Figura 8 sono osservabili una serie di entità estratte da un documento di training

della Fase 1. Dopo aver completato e confermato l'annotazione degli Annotation Set di training, è stato eseguito il primo training del modello sugli otto Annotation Set di training.

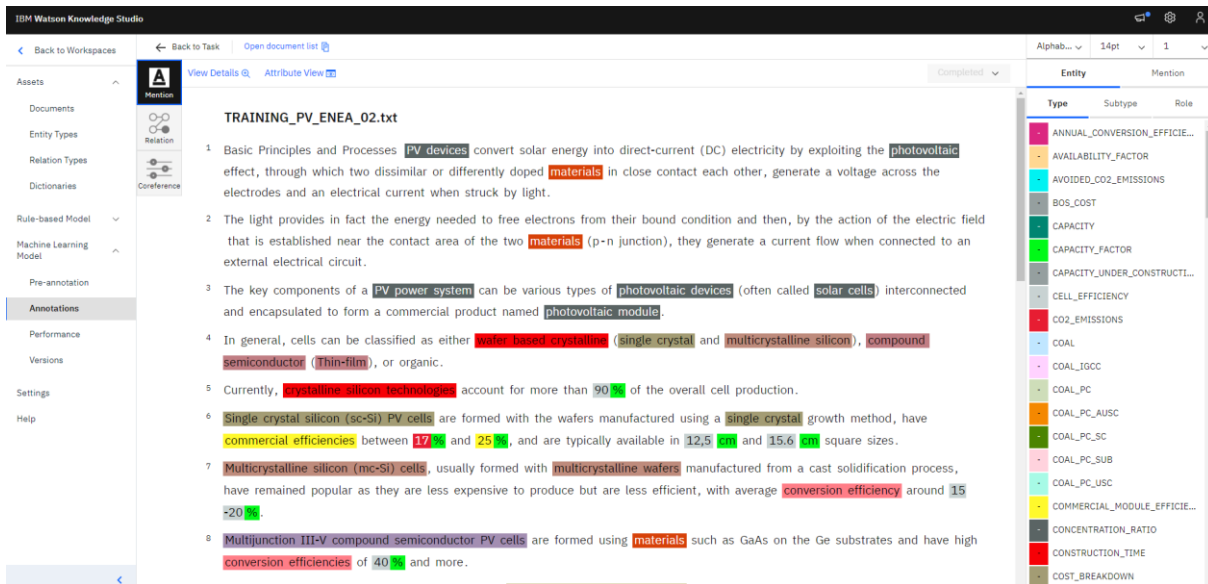


Figura 8 - Annotazione di un documento di training

## 2.2.4 Valutazione del modello

Terminato il training del modello è stata eseguita una valutazione delle sue prestazioni tramite degli appositi test. In questa fase sono stati utilizzati alcuni dei documenti (documenti di origine) forniti dal Committente nella fase iniziale, non utilizzati per l'addestramento e appositamente accantonati per essere utilizzati nella fase di valutazione (tre **set di test**). La valutazione prevede in sostanza la ripetizione sui set di test della procedura appena descritta. In particolare:

- la pre-annotazione dei set di test con il modello Machine Learning appena creato;
- la creazione degli Annotation Set di test, associati ai set di test pre-annotati;
- l'assegnazione degli Annotation Set di test agli annotatori umani;
- l'annotazione degli Annotation Set di test da parte degli annotatori umani;
- la valutazione del modello su tali Annotation Set di test.

La Figura 9 mostra le performance del modello creato nella Fase 1 del progetto. In generale, ove i risultati ottenuti non fossero soddisfacenti, si può procedere ad un eventuale ulteriore training del modello per ottenere risultati migliori.

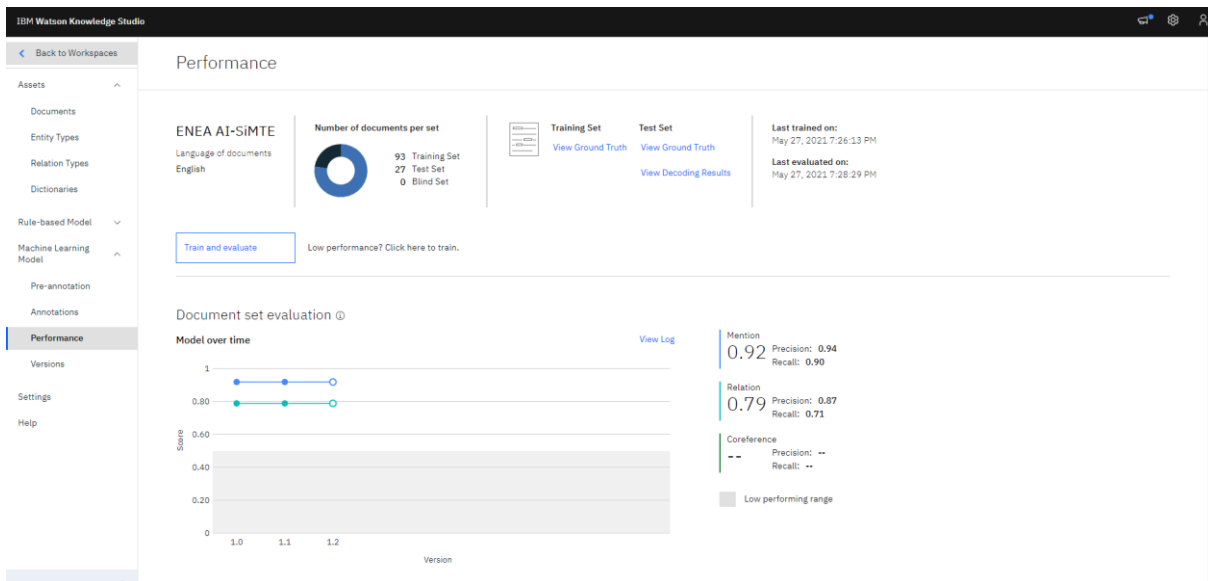


Figura 9 - Performance del modello cognitivo

### 2.2.5 Pubblicazione del modello su Watson Discovery

Nella Fase 1 del progetto, il modello generato dal componente Watson Knowledge Studio è stato “pubblicato” nel componente Watson Discovery, al fine di arricchire i documenti caricati in Watson Discovery con le entità e le relazioni sulle quali il modello è stato trainato. Il componente Watson Discovery è uno strumento di ricerca ed analisi basato sull’intelligenza artificiale, che estrae informazioni da ogni tipo di documento. In aggiunta a ciò, Watson Discovery conserva i documenti e i relativi arricchimenti come sentiment e concetti chiave, frutto delle analisi cognitive effettuate, e permette di effettuare query immediate in linguaggio naturale o linguaggio specifico Discovery, direttamente sui documenti. Il componente Watson Discovery è stato collegato anche al componente Watson Assistant e, tramite quest’ultimo, è stato utilizzato per permettere la lettura dei documenti in linguaggio naturale e in linguaggio Discovery.

### 2.2.6 Analisi e arricchimento cognitivo dei documenti con Watson Discovery

I tre documenti su Photovoltaics, Concentrating Solar Power e Coal, selezionati come documenti di riferimento per la Fase 1, cioè come documenti costituenti la fonte da cui reperire le informazioni da fornire all’utente finale tramite l’assistente virtuale, sono stati caricati tramite l’interfaccia grafica del sistema in una *collection* di progetto, appositamente creata in lingua inglese. Watson Discovery ha quindi effettuato, in autonomia, analisi cognitive sui documenti, e li ha arricchiti con degli insight di default, fra cui, ad esempio, *keyword*, *sentiment*, e *category classification*. Alla collection di progetto è poi stato associato il modello Machine Learning generato da Watson Knowledge Studio, al fine di arricchire i documenti con due ulteriori tipologie di insight: le entità e le relazioni. Terminata l’associazione del modello, Watson Discovery ha infatti effettuato un’ulteriore analisi cognitiva per estrarre dai testi dei documenti le entità e le relazioni definite dal suddetto modello Machine Learning.

### 2.2.7 Training del SW cognitivo sulla struttura dei documenti e acquisizione delle tabelle dati

Una volta concluse le varie fasi fin qui descritte, è stato effettuato un apposito training del SW cognitivo con l’obiettivo di far comprendere a Watson Discovery la struttura dei documenti SiMTE. I documenti disponibili sulla piattaforma SiMTE che caratterizzano le varie tecnologie energetiche hanno infatti un format specifico costituito da:

- un Abstract (Highlights) con sezioni caratterizzate da titoli standard (Process, Technologies and Market Aspects, Performance and Costs, Potential and Barriers);
- un testo con figure e tabelle diviso in capitoli e paragrafi, che in molti casi riproducono in forma più estesa la struttura dell'Abstract;
- una tabella finale che riassume per la tecnologia in esame e per le sue varianti i dati di maggior rilievo riguardanti prestazioni, costi e aspetti di mercato;
- una bibliografia.

Tale struttura dei documenti, pensata non specificatamente in vista dell'addestramento del SW cognitivo, ha favorito tuttavia l'addestramento in misura significativa, consentendo di ottenere risultati soddisfacenti anche nei ridotti tempi a disposizione per l'esecuzione delle attività. In particolare, come evidenziato in Figura 10, Watson Discovery è stato addestrato con l'obiettivo di identificare, all'interno dei documenti, parti dei documenti stessi, quali il titolo, l'autore, la data in cui il documento è stato redatto, i titoli dei capitoli, quelli dei paragrafi, il footer, le immagini, le tabelle, le note a piè di pagina, e le caption di tabelle e immagini. La corretta definizione della struttura dei documenti ha permesso la successiva suddivisione degli stessi per sottotitoli.

La struttura formattata dei documenti SiMTE è stata di aiuto anche nella fase di affinamento della risposta del sistema a interrogazioni (query) in linguaggio naturale. È stato infatti possibile istruire il sistema ad associare alla specifica query solo gli estratti dei documenti rilevanti e più in linea con la query stessa (come osservabile nella Figura 11).

Per quanto riguarda infine l'estrazione delle tabelle riassuntive finali dei documenti SiMTE (e dei dati in esse contenuti che vanno a costituire parte della banca dati SiMTE), trattandosi di tabelle non strutturate che possono contenere sia dati numerici che brevi testi e note di commento e precisazione, è stato necessario ricorrere ad una funzionalità Beta (non ancora commercializzata) di Watson Discovery. Tale funzionalità sfrutta un modello di Machine Learning pre-trained incluso in Watson Discovery per trasformare le tabelle presenti nei documenti in file in formato JSON, che vengono poi lavorati dall'orchestratore. Questo meccanismo consente l'acquisizione delle tabelle e il trasferimento dei dati alla banca dati SiMTE previa verifica di correttezza operata da un supervisore umano.

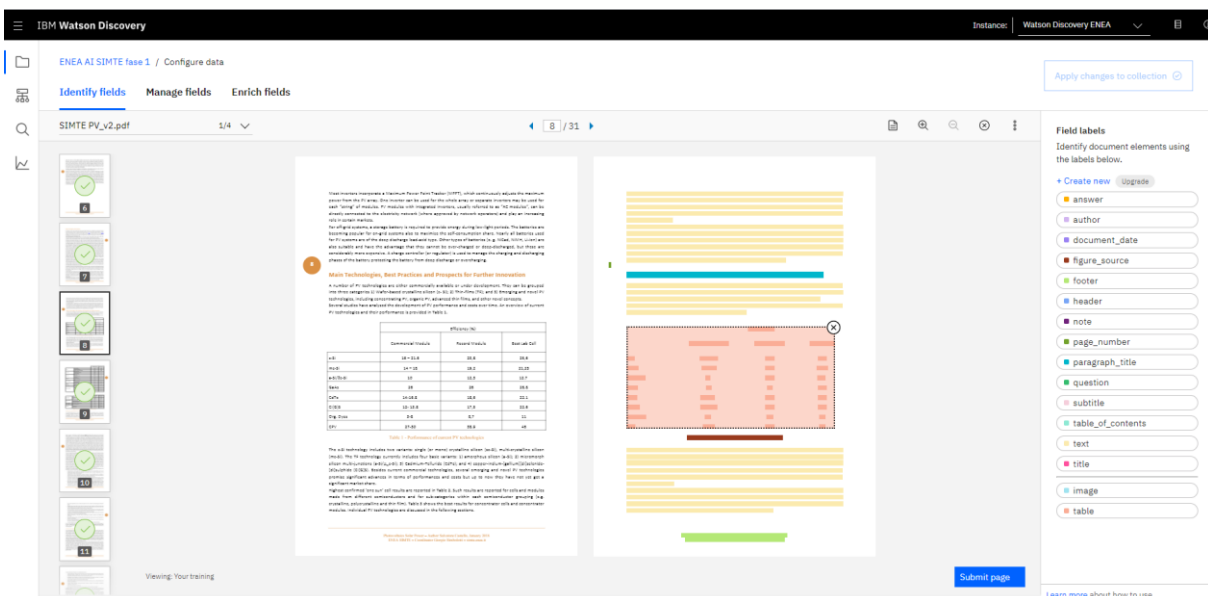


Figura 10 - Training strutturale dei documenti.

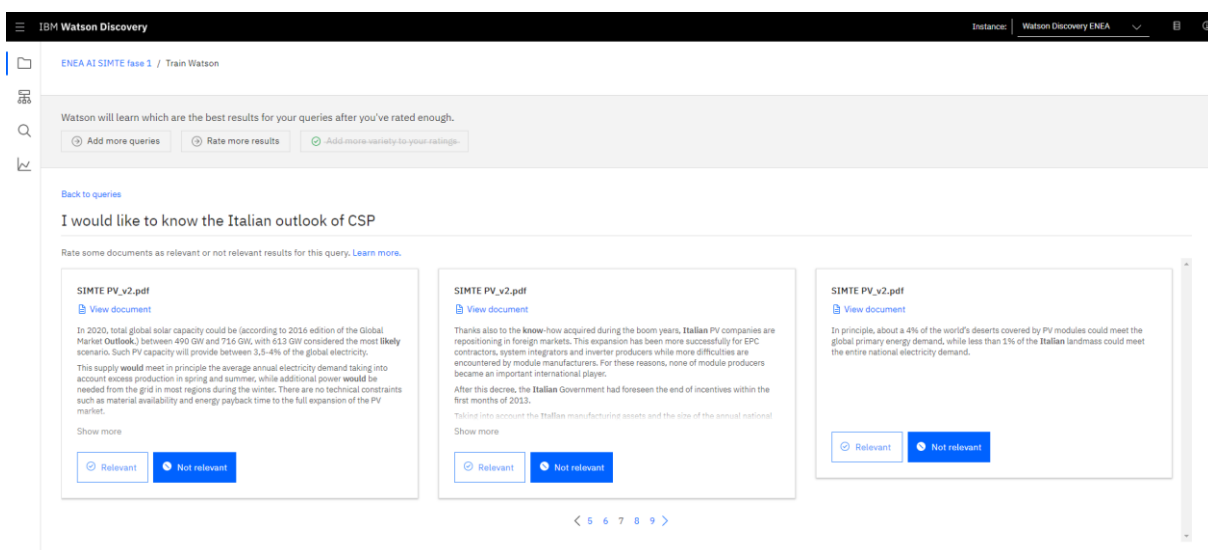


Figura 11 - Training query in linguaggio naturale.

### 2.3 Risultati della Fase 1

Al termine della Fase 1, il Fornitore Sistemi Informativi/IBM ha fornito al Committente ENEA l'accesso ai servizi sviluppati su IBM Cloud e ai vari componenti che costituiscono il sistema e una dimostrazione delle varie funzionalità del SW cognitivo relativamente ai tre cluster tecnologici oggetto di addestramento.

I test di funzionalità sono stati eseguiti congiuntamente dal Fornitore e dal Committente in data 17/06/2021 secondo i criteri stabiliti in fase di raccolta requisiti e previsti per la fine della Fase 1 nelle Specifiche Tecniche del contratto stipulato (Allegato 3, pagg. 2 e 4). I test hanno riguardato in particolare:

- la capacità di acquisizione dei documenti (lettura automatica ed estrazione dei dati di interesse, con trasferimento alla banca dati SiMTE, relativamente alle tecnologie oggetto di addestramento;
- la capacità dell'assistente virtuale di rispondere in modo congruo (con congruenza almeno nel 60% dei casi) a domande poste dall'utente in linguaggio naturale sulle tecnologie oggetto di addestramento;
- la capacità di fornire, a richiesta dell'utente, brevi report con testo, dati e tabelle estratti dai documenti presenti sulla piattaforma, in base alle domande formulate dall'utente stesso.

Infine, in data 22/06/2021, è stato svolto un training per il personale ENEA dedicato alla presentazione in dettaglio delle componenti cognitive della soluzione, Watson Assistant, Watson Discovery e Watson Knowledge Studio e ad una prima introduzione alle tecniche di addestramento.

I test effettuati in data 17/06/2021 hanno fornito esiti soddisfacenti e incoraggianti ai fini del passaggio alla prevista Fase 2 del contratto, dimostrando la potenzialità del sistema (ancorché creato in soli due mesi) alle di rispondere alle esigenze della piattaforma SiMTE, pur nella necessità di una ulteriore messa a punto. È risultato altresì evidente che i tempi disponibili e le risorse umane ed economiche allocate per l'attività contrattuale (sia dal Committente che dal Fornitore) non avrebbero permesso di applicare il sistema a tutte le circa 100 tecnologie energetiche contenute nella piattaforma SiMTE (obiettivo peraltro non previsto nelle Specifiche Tecniche). È stato pertanto deciso (in ottemperanza agli accordi contrattuali) di concentrare i restanti 4 mesi di lavoro della Fase 2 sugli obiettivi previsti (Allegato 3, pagg 2 e 5):

- applicazione del SW cognitivo ad un numero limitato ulteriori cluster tecnologici, con messa a punto di tutte le funzioni previste;
- Integrazione nel sistema delle funzioni statistiche da applicare su dati di prestazioni e costi derivati da cataloghi industriali relativi a tecnologie di largo uso commerciale (es: veicoli, elettrodomestici)

- Realizzazione di una interfaccia con il modello per il calcolo di prestazioni e costi delle tecnologie e relativa integrazione nel sistema.

A tali obiettivi si è aggiunta la necessità di messa punto e affinamento delle funzionalità del sistema per i cluster tecnologici già trattati nella Fase 1.

Naturalmente, al termine della Fase 1, è stato possibile stimare con buona approssimazione i tempi di applicazione del SW cognitivo ai nuovi cluster tecnologici e definire quindi il loro numero in relazione al tempo a disposizione. Molto più incerta era invece la stima dei tempi richiesti per la realizzazione dell'interfaccia con il modello (formulato in linguaggio Scilab).

La realizzazione di tale interfaccia è motivata dalla necessità di poter guidare l'utente esterno nell'uso del modello e in particolare nella preparazione dell'input. Per realizzare l'interfaccia è stata necessaria da parte del Fornitore una piena comprensione delle logiche e della struttura del modello e, da parte del Committente, una piena comprensione delle modalità di creazione dell'interfaccia. La realizzazione dell'interfaccia e l'integrazione del modello hanno costituito una delle parti più impegnative dell'intero lavoro di costruzione del Sistema sia sul piano concettuale che sul piano informatico. In particolare, hanno assorbito circa 2 mesi di lavoro congiunto di un esperto Senior del Committente (il creatore del modello Scilab) e 2-3 esperti del Fornitore (a seconda delle fasi di lavoro).

## 2.4 Attività Svolte nella Fase 2

All'inizio della Fase 2, nella consapevolezza del ridotto tempo di lavoro, gli obiettivi previsti sono stati riarticolati come segue:

- Consolidamento del sistema relativamente alle tecnologie trattate nella Fase 1 con affinamento delle risposte dell'Assistente Virtuale e della produzione di report;
- Estensione del sistema ad altri cluster tecnologici;
- Interfaccia utente di preview;
- Interfaccia con il modello Scilab e integrazione nel sistema;
- Integrazione delle funzioni statistiche su dati di prestazioni e costi derivati da cataloghi industriali relativi a tecnologie di largo uso commerciale (veicoli, elettrodomestici).

Per ogni obiettivo sono stati analizzati i requisiti funzionali e definita una pianificazione comprendente vincoli e date di verifica delle applicazioni realizzate. Di seguito sono riportate le attività svolte.

### 2.4.1 Consolidamento del sistema relativamente alle tecnologie trattate nella Fase 1

In questa fase si è proceduto ad affinare le modalità di riconoscimento dei contenuti dei documenti tecnici. La struttura delle tabelle finali dei documenti è stata variata per consentire di avere un algoritmo di riconoscimento unico valido per tutte le tecnologie; sono state inoltre stabilite regole da rispettare anche per le successive revisioni delle tabelle finali dei documenti. Questa logica ha consentito di riconoscere correttamente:

- gli indicatori (i parametri tecnico-economici della tecnologia);
- i valori relativi agli indicatori, con riconoscimento di eventuali valori tipici e intervalli minimo-massimo;
- le unità di misura;
- le varianti e le sottovarianti della tecnologia trattata e a cui i dati si riferiscono (ciò che compare sulla colonna o sull'insieme di colonne sopra il dato).

Il sistema, in fase di acquisizione dei documenti, memorizza tutte le informazioni di cui sopra in un database, consentendo quindi la ricerca puntuale dei dati. La funzione di acquisizione e memorizzazione dei dati è stata progettata per essere indipendente dalla specifica tecnologia, l'unico vincolo da rispettare è la correttezza della struttura delle tabelle da cui estrarre indicatori e relativi valori. Eventuali brevi testi

presenti in tabella (commenti e note) sono trattati separatamente come tali in modo da non perdere l'informazione in essi contenuta.

#### 2.4.2 Affinamento delle risposte dell'Assistente Virtuale e consolidamento della produzione di report

Un lavoro congiunto tra Committente e Fornitore ha permesso di eseguire test di interrogazione dell'Assistente con domande poste da esperti delle varie tecnologie riportato l'esito di tali test al fornitore con suggerimenti su come migliorare la congruenza della risposta rispetto alla domanda. In tal modo è stato possibile affinare le risposte dell'Assistente definendo percorsi di dialogo più articolati di quelli realizzati in Fase 1. In particolare, sono state stabilite le seguenti "linee guida":

- per ogni richiesta dell'utente, il sistema esegue sia la ricerca sul testo del documento, laddove riconosca il nome di uno o più indicatori, che la ricerca sulla base dati relazionale contenente gli indicatori e i relativi valori;
- i risultati della ricerca sul testo sono presentati con una finestra di dialogo modale che contiene i 3 paragrafi più rilevanti; i risultati della ricerca sugli indicatori sono presentati con una finestra di dialogo modale che contiene una tabella
- se la domanda su una tecnologia è generica, il sistema chiede se si vogliono visualizzare i contenuti degli Highlights o se si vuole eseguire una ricerca con le parole della domanda posta;
- se la ricerca non da alcun esito (perché il sistema non è ancora addestrato per la tecnologia richiesta trattata, l'Assistente Virtuale informa l'utente che il sistema è ancora in fase di apprendimento per la tecnologia richiesta;
- se la richiesta dell'utente è relativa ad una tecnologia per la quale non sia stato eseguito l'addestramento del sistema, ma sia stato memorizzato il documento tecnico, il sistema propone la possibilità di scaricare il documento.

Il lavoro di miglioramento sulle risposte dell'Assistente Virtuale è continuato per tutta la durata del progetto ed è ulteriormente migliorabile con una più estesa esperienza di uso. Analogamente, il tempo di addestramento del sistema per una determinata tecnologia si riduce al crescere del numero di tecnologie sulle quali il sistema viene addestrato. Infatti, la disponibilità di un lessico (intenti, entità, relazioni) sempre più ampio favorisce la rapidità di addestramento. Sia stima che, superata la fase di startup e messa appunto del sistema, un addestratore mediamente esperto sia in grado di addestrare il sistema su una nuova tecnologia in un tempo di circa 2 settimane.

Per quanto riguarda il consolidamento della funzione di report, oltre a poter fornire un report con testo e tabelle relativo alla avvenuta interazione con l'utente (al termine dell'interazione il sistema chiede all'utente se desidera avere un report contenente le informazioni scambiate), ove l'utente chiedesse un report relativo ad una specifica tecnologia, il sistema propone in una finestra di dialogo i contenuti degli highlights e le tabelle finali del documento. È possibile in ogni caso chiedere il download in formato pdf dell'intero documento disponibile sulla piattaforma SiMTE per la tecnologia esaminata.

#### 2.4.3 Addestramento del sistema su ulteriori cluster tecnologici

Dopo aver valutato i tempi a disposizione per le attività della Fase 2 si è proceduto ad addestrare il sistema su ulteriori 3 cluster tecnologici ripercorrendo tutte le fasi di addestramento descritte in precedenza per la Fase 1 nel presente rapporto. Le tre tecnologie selezionate con rispettive varianti tecnologiche sono le pompe di calore aria-aria, le pompe di calore con sorgente geotermica e idrica e la tecnologia delle celle combustibile con relative varianti (Air-to-air Heat Pumps, Water and Geothermal Heat Pumps, Low-Temperature Fuel Cells). La scelta di tali tecnologie è motivata dalla necessità di addestrare il sistema su tecnologie diverse dalle quelle di elettro-generazione considerate nella Fase 1 costruendo in parte su un lessico già acquisito, allargandone tuttavia i confini a tecnologie in grado di produrre calore (oltre che energia elettrica) e a tecnologie in uso nel settore residenziale (pompe di calore) e, in futuro, nel settore dei

trasporti (celle a combustibile a bassa temperatura). Per ognuna delle tecnologie ENEA ha fornito i documenti tecnici con tabelle finali strutturate secondo le modalità di cui sopra per un più facile addestramento del sistema. Il lavoro svolto in questa fase è ovviamente analogo a quanto già descritto per la Fase 1.

#### 2.4.4 Interfaccia di preview

La necessità di sviluppare tale interfaccia è stata evidenziata in corso d'opera al fine di facilitare il compito di chi deve eseguire il caricamento nel sistema di un documento tecnico relativo ad una nuova tecnologia e l'acquisizione dei dati inerenti agli indicatori tecnico-economici contenuti nelle tabelle finali del documento. Per tale scopo è stata realizzata una funzionalità ad-hoc costituita da una schermata del sistema divisa in 2 sezioni affiancate: la parte sinistra serve all'operatore che esegue il caricamento; la parte destra visualizza il documento da acquisire. La funzione richiede le seguenti fasi:

- l'operatore seleziona il documento che si intende caricare e il relativo file;
- il sistema esegue un controllo formale della struttura delle tabelle finali e fornisce una preview dei dati riconosciuti, evidenziando con colori diversi le celle che contengono valori non congruenti.
- L'operatore verifica le eventuali incongruenze presentate dal sistema e decide se confermare l'acquisizione e la scrittura dei dati sul database o annullare l'operazione, per applicare eventuali correzioni al file originale e rieseguire la funzione.

#### 2.4.5 Interfaccia e integrazione del modello Scilab per il calcolo di prestazioni e costi delle tecnologie

Come accennato in precedenza, la piattaforma SiMTE dispone di un modello (realizzato in linguaggio Scilab) per la valutazione tecnico-economica (analisi costi-benefici) di prestazioni e costi delle tecnologie energetiche sulla base di dati di input che possono essere derivati dalla banca dati SiMTE o immessi dall'utente.

Il modello è applicabile a tutte le tecnologie di domanda/offerta. Di conseguenza, pur non sostituendosi a più accurate analisi costi-benefici eseguibili dagli analisti o dagli investitori del settore energetico, consente confronti costi/benefici (prestazioni, emissioni, costi) tra tecnologie anche afferenti a settori diversi. Esso è destinato prevalentemente ad utenti "esperti" (con conoscenze nel settore dell'analisi costi-benefici delle tecnologie energetiche) e consente anche l'analisi di sistemi complessi (es: cogenerazione, impianti con *energy storage*, efficientamento di edifici nel settore residenziale). In tali casi l'utente può anche avvalersi di casi pilota già elaborati, disponibili sul portale. Il modello può simulare una tecnologia (un impianto o un dispositivo energetico) come un unico componente (caso semplice) dotato di vettori di input (es: input energetico) e di output (es: prodotto o servizio energetico reso, emissioni, ecc.) oppure come un sistema a più componenti (casi più complessi). L'input del modello richiede l'immissione di dati che riguardano aspetti tecnici ed economici.

Per facilitare l'uso del modello da parte dell'utente è stata realizzata una interfaccia per guidare l'utente nella preparazione dell'input. L'interfaccia è stata realizzata creando un contenitore cloud dell'applicativo che espone le funzionalità del modello in una un'interfaccia web integrata nel portale SiMTE. Per tale integrazione è stato configurato un server (scritto su stack nodeJS e deployato) su cloud IBM.

Attraverso l'interfaccia l'utente ha la possibilità di creare ex-novo modelli personali e personalizzati, avendo un feedback visivo immediato della costruzione in corso tramite un grafo (diagramma a blocchi) costruito in base alle relazioni create nella definizione del caso trattato. L'utente può ovviamente anche salvare il modello creato costituendo una base di partenza per nuovi calcoli o per creare una diversa simulazione di impianto o dispositivo.

Nel seguito vengono descritte le varie fasi di preparazione di un caso di calcolo e le varie funzionalità dell'interfaccia.

**Flusso dei dati** - La peculiare conformazione del modello ha fatto emergere la necessità di una struttura dati condivisa scritta su file per scambiare informazioni tra server e interfaccia. Per tale motivo è stata definita una struttura JSON da cui il modello legge e verso cui il server scrive che mantiene tutti i collegamenti necessari al funzionamento del modello creato dall'utente. I principali campi e componenti di tale struttura sono riportati di seguito:

- **vectors:** Lista dei vettori definiti dall'utente.
- **components:** Lista dei component definiti dall'utente.
- **processes:** Lista dei processi definiti dall'utente.
- **componentsParameters:** Lista dei parametri in uscita dal modello e modificabili dall'utente.
- **constraintVectors:** Lista dei vincoli sui vettori definita dall'utente.
- **constraintProcesses:** Lista dei vincoli sui processi definita dall'utente.
- **financialData:** Insieme dei dati finanziari relativi al modello.

Durante l'interazione con l'interfaccia, la struttura dati si aggiorna sulla base delle azioni intraprese dall'utente e, nel momento in cui tutte le condizioni di validità vengano rispettate, è possibile richiedere l'output dell'analisi.

Per arrivare ad ottenere un output, l'interazione è divisibile in due macro fasi:

- Costruzione del modello tramite le componenti messe a disposizione dall'interfaccia
- Definizione di dati finanziari e vincoli relativi al modello di interesse

In prima istanza l'utente è chiamato a popolare il modello congruamente, inserendo specifiche entità con propri specifici vincoli. Per poter costruire il modello, si ha la possibilità di inserire tre tipologie di entità logiche: *Componenti*, *Processi* e *Vettori*. Un componente (di un impianto o dispositivo) può essere sede di più processi. Ogni entità ha il suo sottoinsieme di dati specifici e caratterizzanti che permettono l'aggancio logico con le altre entità.

Dopo aver completato la definizione del modello, l'utente è chiamato a inputare dati relativi ai parametri finanziari e a definire i vincoli necessari per l'esecuzione, tra le entità precedentemente create.

**Interazione con l'interfaccia e guida per l'utente** – Trattandosi di un modello molto avanzato, è stato necessario definire una modalità per guidare l'utente nella compilazione del medesimo. Per tale motivo, l'interfaccia sul portale SIMTE si configura come un area di lavoro con una banda azzurra di destra ove compaiono dei messaggi contestuali rispetto al punto dell'interazione in cui si trova l'utente, per guidarlo nel corretto inserimento dei dati (Figura 12).

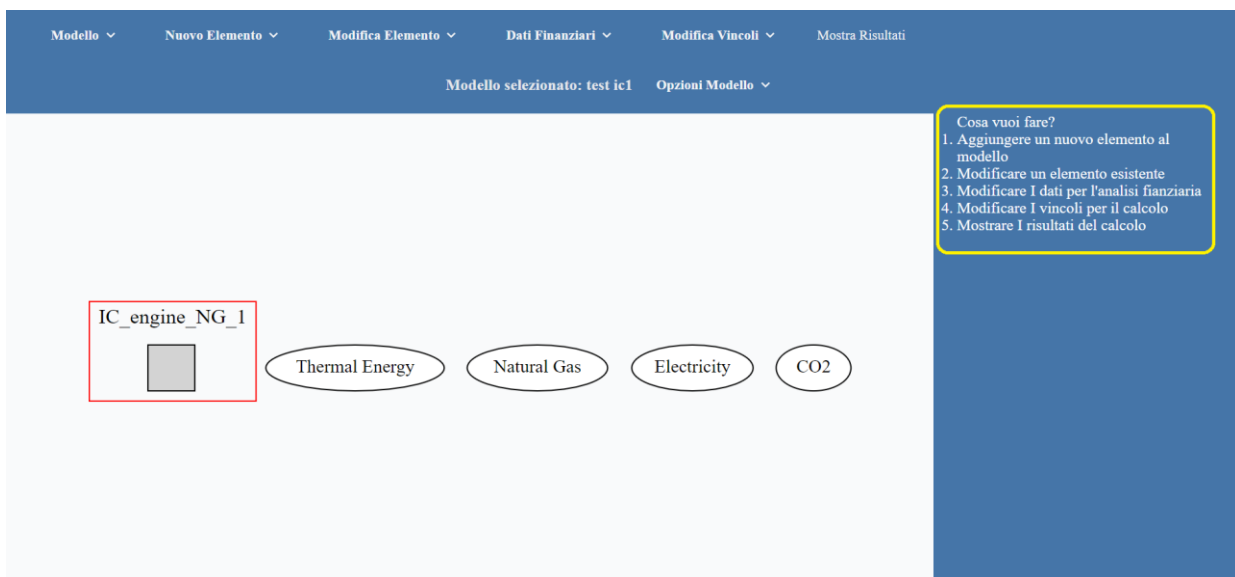


Figura 12 - Guida per l'utente

**Scelta Modello** - Accedendo all'interfaccia, è possibile eseguire la scelta iniziale: scegliere se creare un nuovo modello o partire da uno già creato (Figura 13).

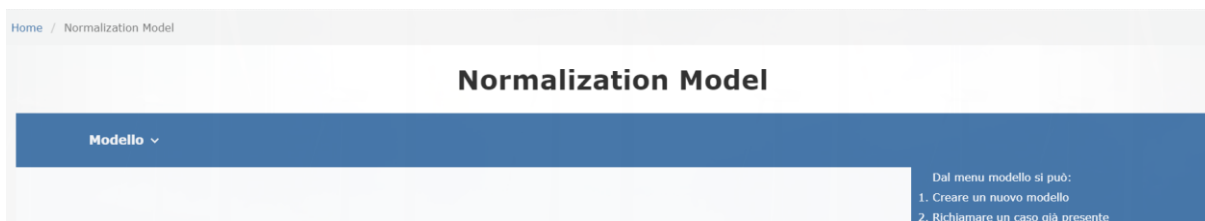


Figura 13 - Scelta modello

Nel caso in cui l'utente scelga di riprendere un modello esistente, questo viene caricato nell'interfaccia. Se invece l'utente procede con la creazione, dovrà assegnare un nuovo nome al modello insieme a poche altre informazioni di carattere globale. In entrambi i casi, si arriverà all'interno del menu contestuale al modello dove poter iniziare ad interagire (Figura 14).

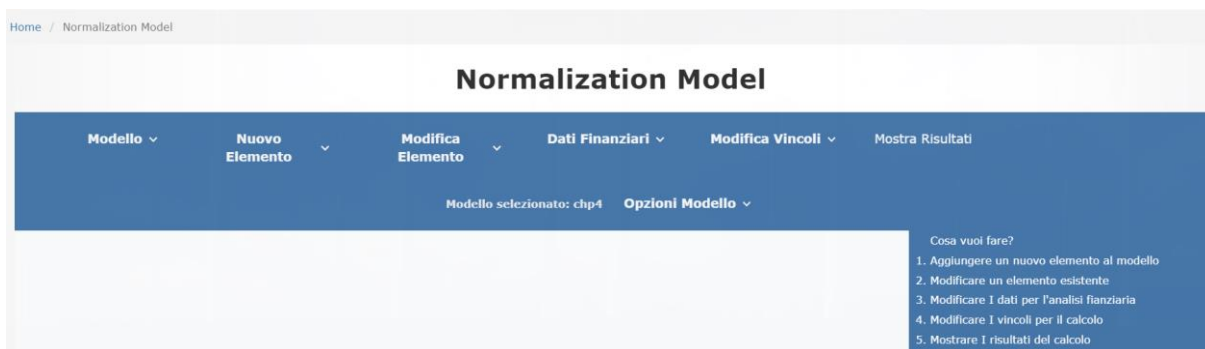


Figura 14 - Menu contestuale modello

**Creazione degli Elementi** - Attraverso la voce di menu “Nuovo Elemento” l’utente può creare una delle sopracitate entità: *Componente*, *Vettore*, *Processo*. A seconda della tipologia della entità, il grafo di preview mostra un nodo adeguato (Figura 15) :

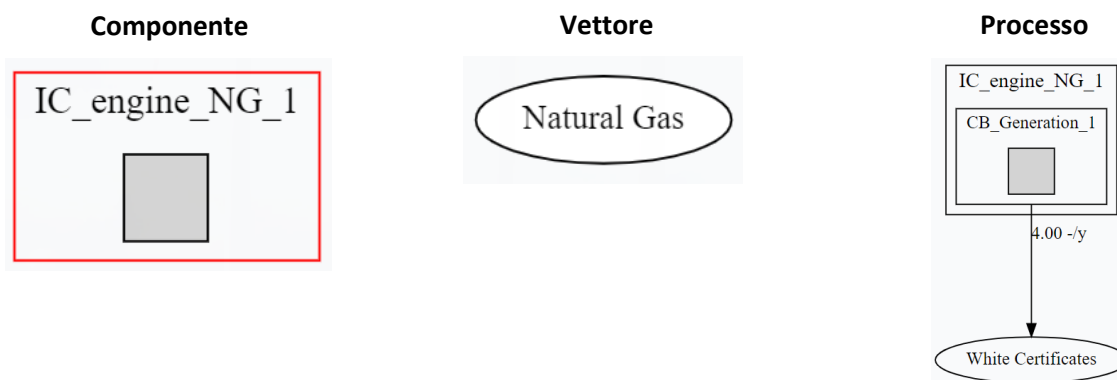


Figura 15 - Tipologie di entità

Selezionando la voce *Componente* l’interfaccia richiede di scegliere il nuovo componente da instanziare tramite dropdown (Figura 16).

The screenshot shows the "Nuovo Componente" dialog box. It has a title bar with a close button (X). Below the title, there are two dropdown menus:

- "Seleziona un settore" with a downward arrow.
- "Seleziona un componente" with the text "Nessuna opzione." and a downward arrow.

At the bottom left is a red button labeled "Chiudi". At the bottom right is a grey button labeled "Conferma".

The screenshot shows the "Nuovo Componente" dialog box with the "Seleziona un settore" dropdown menu open. The list of options is:

- Tecnologie per i trasporti
- Tecnologie per l'industria e l'agricoltura
- Tecnologie per il residenziale e i servizi
- Produzione e trasformazione di energia primaria
- Produzione e distribuzione di elettricità e calore

The "Chiudi" button is visible at the bottom left.

**Nuovo Componente** ×

Seleziona un settore

Seleziona un componente

Figura 16 - Selezione settore e componente

Ricevuta la conferma per il componente da instanziare, è necessario inputare dati relativi allo specifico componente e vincolati dalle regole dello stesso (Figura 17).

**Nuovo Componente IC\_engine\_NG**

year0

Valore non valido o anno minore di Project Year

Descrizione	Valore	Unità di misura
Power output at nominal conditions	<input type="text"/>	<input type="text" value="v"/>
Thermal Power Output at nominal conditions	<input type="text"/>	<input type="text" value="v"/>
Total efficiency	<input type="text"/>	<input type="text" value="v"/>
White Certificates production per year	<input type="text"/>	<input type="text" value="v"/>

Figura 17 - Compilazione dati per un nuovo componente

Non appena compilato e salvato il nuovo componente, l'interfaccia aggiunge il corrispettivo nodo al grafo di preview. In modo analogo, si procede con la scelta del vettore e la relativa compilazione dati. Specificatamente per i vettori, è stato predisposto un pool di vettori precompilati così che, nel caso in cui l'utente scelga un vettore per cui è disponibile un vettore precompilato, avrà la possibilità di scegliere se usare il vettore precompilato o iniziare a compilarlo da capo (Figura 18).

### Nuovo Vettore ×

Seleziona un settore Tecnologie per l'industria e l'agricoltura ▾

Seleziona un componente IC\_engine\_NG ▾

Seleziona un processo Generation ▾

Seleziona la tipologia del vettore In ▾

Seleziona un vettore Natural Gas ▾

Chiudi
Conferma

### Nuovo Vettore Natural Gas

Variabile	Valore	Unità di misura
Name	<input style="width: 90%;" type="text" value="Natural Gas"/>	
UOM	<span style="border: 1px solid gray; padding: 2px 5px;">kg</span> ▾	
LHV	<input style="width: 90%;" type="text" value="0"/>	J/kg
Pr_En	<input style="width: 90%;" type="text" value="0"/>	J/kg
CEF	<input style="width: 90%;" type="text" value="0"/>	kgCO2eq/kg
UCEF	<input style="width: 90%;" type="text" value="0"/>	kgCO2eq/kg
P	<input style="width: 90%;" type="text" value="0"/>	Pa
T	<input style="width: 90%;" type="text" value="0"/>	K
RCMV	<input style="width: 40%;" type="text" value="2020"/> <input style="width: 40%;" type="text" value="0"/> <span style="margin-left: 5px;">⊕ ⊗</span>	€/kg
	<input style="width: 40%;" type="text" value="2120"/> <input style="width: 40%;" type="text" value="0"/> <span style="margin-left: 5px;">⊕ ⊗</span>	

Chiudi
Salva

Figura 18 - Finestre di dialogo Nuovo Vettore

La creazione di un processo si articola diversamente rispetto a vettore e processo. Dopo una prima scrematura tramite dropdown per individuare il giusto processo da inserire, viene chiesto all'utente di scegliere tra i vettori presenti un vettore da legare logicamente al processo, nel caso non esistano, viene lasciata la possibilità di crearli direttamente dall'interfaccia processo tramite il pulsante "Crea". Inserendo i dati necessari, è quindi possibile salvare il processo e vederlo in preview (Figura 19).

**Nuovo Processo**

Seleziona un componente IC\_engine\_NG-1 ▾

Seleziona un processo Generation ▾

Vettori in ingresso	Selezionare vettore
Natural Gas	Natural Gas ▾ Natural Gas
Vettori in uscita	Selezionare vettore
Electricity	Electricity ▾
Thermal Energy	Thermal Energy ▾
CO2	Nessun vettore <span style="float: right;">Crea</span>
Fattore di parzializzazione <input style="width: 100%;" type="text"/>	

Chiudi
Salva

Figura 19 - Finestra di dialogo Nuovo Processo

**Modifica degli Elementi** - Accedendo alla voce di menù “Modifica Elemento”, si ha un campo dropdown analogo a quello disponibile nel caso della creazione.

In questa fase, nonostante l’insieme dei dati modificabili relativi ai vettori restino gli stessi usati in fase di creazione, l’insieme dei dati per componenti e processi si espande aggiungendo i parametri calcolati e mandati in output dopo una prima compilazione del modello.

Proprio per tale ragione, è stato necessario definire una logica per gestire la popolazione delle tabelle con dati di default provenienti dal modello insieme alla loro eventuale sovrascrittura in fase di compilazione, qualora questi siano stati aggiornati da parte dell’utente.

Inoltre, essendo il campo “year0” e i campi della prima tabella della finestra di modifica “oc\_params” funzionali all’output del modello, si è deciso di procedere sovrascrivendo eventuali modifiche avvenute sulla seconda tabella qualora durante l’operazione di salvataggio vengano modificati il campo “year0” o i campi “oc\_params” ma non i dati della seconda tabella.

In particolare, nella finestra di modifica del componente (Figura 20) saranno possibili operazioni molteplici quali:

- Eliminazione del componente
- Impostazione del componente come **core** del modello
- Modifica dei parametri inputati in fase di creazione
- Modifica dei parametri ricevuti in output dal modello

### Modifica Componente IC\_engine\_NG-1

Elimina
Imposta come componente core

year0

Descrizione	Valore	Unità di misura
Power output at nominal conditions	<input type="text" value="1"/>	kW <span style="font-size: small;">▼</span>
Thermal Power Output at nominal conditions	<input type="text" value="2"/>	kWt <span style="font-size: small;">▼</span>
Total efficiency	<input type="text" value="3"/>	- <span style="font-size: small;">▼</span>
White Certificates production per year	<input type="text" value="4"/>	- <span style="font-size: small;">▼</span>

Variabile	Valore	Unità di misura
TOCC	<input type="text" value="2000"/>	€2020
comp_CT	<input type="text" value="0"/>	€2020
FOMC	<input type="text" value="127.4"/>	€2020
REVC	<input type="text" value="0"/>	€2020
CCFRV	<input type="text" value="0"/>	€2020
TLT	<input type="text" value="8"/>	€2020
DC	<input type="text" value="0"/>	€2020
DCT	<input type="text" value="2"/>	€2020
AVF	<input type="text" value="1"/>	€2020
CCE	<input type="text" value="0"/>	kgCO2eq

Chiudi
Salva

Figura 20 - Modifica componente

Analogamente, in fase modifica processo (Figura 21), sarà possibile eseguire operazioni come:

- Eliminazione del processo
- Modifica dei parametri inputati in fase di creazione
- Modifica dei parametri ricevuti in output dal modello

### Componente IC\_engine\_NG-1 Modifica Processo Generation\_1

Elimina

Variabile	Valore	Unità di misura	
Cons_vect_proc_0	Natural Gas <span style="font-size: small;">▼</span>	<input type="text" value="107.774598"/>	kg/8760 h
	Elettricità autoprodotta <span style="font-size: small;">▼</span>	<input type="text" value="0.6"/>	MWhe/8760 h
TNC_0	calore <span style="font-size: small;">▼</span>	<input type="text" value="0.643"/>	MWht/8760 h
	CO2 <span style="font-size: small;">▼</span>	<input type="text" value="296.380143"/>	kgCO2e/8760 h
SVOMC	<input type="text" value="0"/>	€2020/(8760 h)	
ER0_in	Natural Gas	1	1/(8760 h)
	Elettricità autoprodotta	1	1/(8760 h)
	calore	1	1/(8760 h)
ER0_out	CO2	1	1/(8760 h)

Chiudi
Salva

Figura 20 - Modifica processo

Per i vettori, attraverso un'interfaccia identica a quella di creazione è invece possibile:

- Eliminare il vettore
- Modificare i parametri inputati in fase di creazione

**Dati Finanziari** - La struttura dei dati finanziari è stata definita sulla base dell'input accettato dal modello e viene assemblata strada facendo, mentre l'utente popola le tabelle relative alle singole voci di menu sotto la voce "Dati Finanziari". In questa fase, risulta necessario compilare tutte le tabelle di tutte le sotto-voci del menù, per poter eseguire correttamente il run del modello.

**Vincoli** - In questa sezione si definiscono i vincoli che regolano il funzionamento del modello. Alla voce "Modifica Vincoli" si hanno due possibilità: inserire e/o modificare vincoli per processi e vettori.

In modifica vincoli sui vettori (Figura 21) secondo le regole ed i vincoli derivati dai calcoli intermedi del modello, è necessario inserire valori relativi ai singoli vettori selezionabili dal dropdown nella prima colonna. Nel caso in cui si vogliono aggiungere vincoli, tramite l'apposito pulsante è possibile continuare a inserire vincoli fintanto che ci sono vettori non vincolati.

Vettori	1	2	3	4	5	6	7	CT+LT	9	AT
Natural Gas	0	0	0	0	0	0	0	0	0	0

Figura 21 - Modifica vincoli sui vettori

In modifica vincoli sui processi (Figura 22), si può invece definire un range temporale entro cui far valere i vincoli. Il campo "Anni di validità" è predisposto per ricevere range temporali, sia in formato numerico che letterale, secondo lo standard concordato.

Essendo i vincoli temporali derivati da dati intermedi calcolati dal modello, l'interfaccia è predisposta per controllare la correttezza dell'input, onde evitare incoerenza tra i dati. Per agevolare l'inputazione, è accettata l'inputazione tramite valori numerici e letterali secondo i seguenti vincoli:

- I valori letterali riconosciuti sono: AT, CT, LT
- I range sono esprimibili tramite la formula: <anno\_apertura> : <anno\_chiusura>
- I valori puntuali sono espressi concatenando valori con la virgola: anno\_1, anno\_2
- I valori letterali possono essere sommati o sottratti tra loro e con numeri interi
- Range e valori puntuali possono essere concatenati

**Modifica vincoli sui processi**  
year0: 2019 AT: 2028

N° Vincolo	Generation_1	CB_Generation_1	Generation_2	Production_1	RHS	Anni di validità	Azioni
1	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="1"/>	2019:2026	<input type="button" value="⊕"/> <input type="button" value="⊗"/>
1	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	<input type="text" value="0"/>	2027:2028	<input type="button" value="⊕"/> <input type="button" value="⊗"/>

Figura 22 - Modifica vincoli sui processi

**Output e Risultati del Modello** - Nell'apposito menù di funzionalità disponibile nell'interfaccia, è possibile richiedere l'esecuzione del modello tramite la voce "Lancia" ottenendo l'output finale. Una volta lanciato il modello con successo, spostandosi nella voce di menù "Mostra Risultati" è possibile visionare l'output della computazione relativa al modello creato (Figura 23).

**Risultati**

Indicatori globali							
Indicatori tecnici		Indicatori economici			Indicatori Tecno-Economici		
Tot_En_Cons	251055.814 GJ	EOLAPV	€ 2090353	UEI		120.102 GJ/(1000 €)	
Ult_Pr_En_Cons	251055.814 GJ	IRROE	1.007	UCI		7.126 tCO2e/(1000 €)	
Tot_Carb_Em	14473.868 tCO2e	DPP	2.25 y				
Ult_Carb_Em	14894.926 tCO2e						

Indicatori per vettore							
	Natural Gas	Elettricità autoprodotta	calore	CO2	White Certificates	Product	elettricità acquistata
Net_Prod	-5263224.602277 kg	0.000000 MWhe	0.000000 MWht	14473867.656261 kgCO2e	4064.000000 -	80000.000000 t	6692.810202 MWhe
R_LPC	0.852409 €/kg	0.000000 €/MWhe	-66285377942272432.000000 €/MWht	-0.176607 €/kgCO2e	-593.368111 €/-	19.994600 €/t	-178.804256 €/MWhe
N_LPC	1.374924 €/kg	0.000000 €/MWhe	-1069173695104731008.000000 €/MWht	-0.284865 €/kgCO2e	-957.094302 €/-	32.251005 €/t	-288.408716 €/MWhe

Figura 23 - Risultati

#### 2.4.6 Integrazione di funzioni statistiche da applicare a dati di prestazioni e costi derivati da cataloghi industriali

A scopo esclusivamente dimostrativo, le funzioni statistiche sono state implementate ed applicate ai dati di prestazioni e costo di autoveicoli derivati dai cataloghi commerciali presenti sulle più note riviste automobilistiche. I dati relativi potenza (CV, kW), consumi energetici (km/L, L/100km), emissioni (gCO2/km) costo di mercato e di esercizio (euro, euro/km) sono stati acquisiti e riallocati in base al tipo di

alimentazione (es. benzina, diesel, metano, gpl e dual fuel, full hybrid, mild hybrid, hybrid plug-in, electric vehicle) e alla classificazione dei segmenti dei veicoli in uso nel mercato nazionale ed europeo (minicar/citycar; small car/utilitarie/sub compatte; medium car/lower medium/compatte; large car/berline&SW medio-grandi; executive car/upper medium/grandi berline&SW; luxury car/berline lussuose; J-small/B-SUV/piccoli SUV da città; J-compact/C-SUV/SUV compatti; J-medium/D-SUV/SUV medi; j-large/E-SUV/grandi SUV; multipurpose car/van/monovolumi passeggeri e commerciali; sport/coupè/roadster; convertible; supercar/hypercar ...ecc. ) con distinzione all'interno di ogni segmento delle categorie premium e commerciali.

Una volta acquisiti i dati presenti nei cataloghi, le funzioni statistiche permettono ad esempio di calcolare e richiedere all'Assistente Virtuale per uno specifico segmento i valori medi, massimi e minimi relativi ai vari parametri (dimensione, pesi, potenze, consumi, emissioni, prezzi di acquisto e di esercizio, tassazione, etc.);

Le funzioni statistiche possono essere ovviamente applicate a tutte le tecnologie di larga diffusione commerciale presenti sul mercato con grandi varietà di modelli diversi in termini di prestazioni energetiche e costi (es: elettrodomestici) in supporto delle scelte di operatori e consumatori.

## 2.5 Risultati della Fase 2

Al termine della Fase 2, le funzionalità del sistema sono state collaudate con esito positivo in più sedute di collaudo preventivamente concordate per verificare il progressivo raggiungimento degli obiettivi fissati entro i termini previsti. Un ulteriore collaudo è stato effettuato nell'ambito della *valutazione di conformità degli obiettivi raggiunti* (prevista dalla vigente normativa) da parte del RUP del contratto di appalto assegnato al fornitore. Anche in tale occasione sono state richiamate con esito positivo tutte le principali funzionalità implementate sulla piattaforma SiMTE.

Le funzioni dell'Assistente Virtuale per tutte le tecnologie oggetto di addestramento del SW sono state verificate ponendo una serie di domande riportate nella matrice in Figura 23. Ogni domanda ha dato luogo non solo ad una singola risposta ma ad una interazione articolata con l'Assistente Virtuale nella quale sono stati forniti testi, dati e tabelle in risposta alle domande poste fornendo poi, alla fine di ogni interazione, un report scaricabile in pdf sui temi della interrogazione avvenuta. La congruenza delle risposte fornite dall'Assistente Virtuale rispetto alle domande poste è risultata prossima al 75%. Nel corso di tale prova è stata anche implicitamente verificata la corretta estrazione dei dati delle tecnologie dai documenti oggetto di addestramento e il corretto riversamento nella banca dati SiMTE. Occorre infine notare che nonostante l'addestramento sia stato eseguito in lingua inglese per le ragioni precedentemente descritte, il sistema risponde correntemente anche ad interrogazioni in Italiano.

	Questions: please, tell me ...	Technology				
		CSP, STE	PV	Coal Power	HP A-A and W-G	FC (LTFC)
1	something on ...	CSP; STE; concentr. solar power; solar th. electricity	PV; PV power; solar power;	Coal power; coal-fired power; coal technology	Heat pumps, HP; HP devices	FC, Fuel cells Low temp. FC; High temp. FC;
2	something on ...	Th. energy storage; Heat storage	PV thin film; Crystalline silicon PV; c-Si PV	Emissions from coal power plants	HP process; HP market	FC process; FC market
3	the technology variants of ...	CSP technology	PV technology	Coal power	HP	FC
4	the energy efficiency of ...	CSP plants	PV power plants; PV modules; PV cells	Coal power plants		FC, Fuel cells; Low Temp. FC; High Temp. FC;
5	the COP of ... the coeff. of performance of ... the energy efficiency of ...				HP; Air to air HP; Geothermal HP;	
6	the capacity factor of the CF of .. the lifetime of ...	CSP plants	PV power plants	Coal power plants	HP	FC
7	the investment cost of ... the capital cost of ... the overnight capital cost of ...	CSP plants	PV power plants	Coal power plants	HP	FC
8	the electricity cost of ... the el. production cost of ... the LCOE of ...	CSP plants	PV power plants	Coal power plants		FC
9	the O&M costs of ...	CSP plants	PV power plants	Coal power plants	HP	FC
10	the heat cost for ...				HP	FC
11	something on the future of ... the potential & barriers for ...	CSP plants	PV power plants	Coal power plants	HP	FC
12	the projected/future costs of electricity from ...	CSP plants	PV power plants	Coal power plants	HP	FC
13	the GHG emissions from ...	CSP plants	PV power plants	Coal power plants	HP	FC
14	something on the XX process ...	CSP	PV	Coal power	HP	FC
15	st on the Italian/global market of	CSP	PV	Coal power	HP	FC
	Please, compare ... energy efficiency, emissions, capacity factor and investment cost of ...	CSP, PV, coal power,			HP, FC	
	Do the XX plants produce heat?	CSP plants	PV power plants	Coal power plants	HP	FC

Figura 23 – Quesiti di collaudo per l'Assistente Virtuale

E' stata altresì eseguita con esito positivo una applicazione del modello di calcolo al caso di un cogeneratore di elettricità e calore con contestuale verifica della funzionalità dell'interfaccia modello precedentemente descritta. Il buon esito della prova ha dimostrato implicitamente anche la buona funzionalità di tutti i collegamenti informatici stabiliti tra il cloud IBM e la piattaforma SiMTE.

In definitiva, i collaudi effettuati confermano il raggiungimento di tutti gli obiettivi previsti nel contratto con il fornitore e di tutti gli obiettivi previsti nel progetto in merito alla implementazione di Intelligenza Artificiale nella piattaforma SiMTE.

## 2.6 Training del personale ENEA sulle tecniche di addestramento del SW AI

Nell'ambito della Fase 2 era anche previsto un ulteriore training del personale ENEA sulle tecniche di addestramento del SW-AI, più approfondito di quello eseguito nella prima fase e allargato ad un maggior numero di partecipanti. Tale training era necessario al fine di rendere autonoma l'Agenzia nell'uso del SW anche dopo il termine delle attività contrattuali con il Fornitore. Il previsto training di Fase 2 è stato realizzato in due sedute in modalità remota per la durata complessiva di circa 6h in data **3 e 10 Dicembre 2021** con la partecipazione di 14 ricercatori e tecnici ENEA (lista di seguito). Naturalmente, la piena acquisizione delle tecniche di addestramento è conseguibile soltanto con una certa esperienza di uso del SW. Tuttavia il contratto stipulato con il Fornitore prevede assistenza all'uso del sistema e manutenzione del sistema realizzato per la durata di 1 anno dopo il termine delle attività contrattuali, oltre al diritto di accesso al cloud IBM e all'uso del SW per la durata di 3 anni.

**Personale ENEA che ha partecipato al Training su Tecniche di Addestramento AI:**

**Enrico Cosimi,**

**Filippo Donato,**

**Giorgio Simbolotti,**

**Flavio Fontana,**

**Amedeo Buonanno,**

**Silvia Orchi,**

**Alessandro Giocoli,**

**Angelo Mariano,**

**Vittoria Cozza,**

**Fabio Cignini,**

**Simona De Iuliis,**

**Martina Botticelli,**

**Carlo Petrovich, I .**

**Maria Salvato.**

## **Allegati**

**Allegato 1:** RELAZIONE TECNICA - Appalto per fornitura di servizi di implementazione/addestramento e diritti di accesso (license & cloud) a SW-AI (SoftWare – Artificial Intelligence) per applicazione a piattaforma SiMTE-ENEA –RTI Sistemi Informativi/IBM e VAR Group -

**Allegato 2 :** RAPPORTO FINALE DI ATTIVITA per servizi di implementazione/addestramento e diritti di accesso (license & cloud) a SW-AI (SoftWare – Artificial Intelligence) per applicazione a piattaforma SiMTE-ENEA - Sistemi Informativi/IBM

**Allegato 3:** Specifiche Tecniche per l'acquisto di servizi di implementazione/addestramento e diritti di accesso (license & cloud) a SW-AI (Software - Artificial Intelligence) per applicazione alla piattaforma SiMTE (Sistema informativo e di Monitoraggio delle Tecnologie Energetiche), ENEA – Agosto 2020