



**Nota tecnica**

# **SOPHIA – DOCUMENTAZIONE TECNICA DEL PROGETTO**

**Maggio 2022**



L'Istituto nazionale per l'analisi delle politiche pubbliche (INAPP) è un ente pubblico di ricerca che si occupa di analisi, monitoraggio e valutazione delle politiche del lavoro, delle politiche dell'istruzione e della formazione, delle politiche sociali e, in generale, di tutte le politiche economiche che hanno effetti sul mercato del lavoro. Nato il 1° dicembre 2016 a seguito della trasformazione dell'Isfol e vigilato dal Ministero del Lavoro e delle politiche sociali, l'Ente ha un ruolo strategico - stabilito dal decreto legislativo 14 settembre 2015, n. 150 - nel nuovo sistema di governance delle politiche sociali e del lavoro del Paese. L'Inapp fa parte del Sistema statistico nazionale (SISTAN) e collabora con le istituzioni europee. Da gennaio 2018 è Organismo Intermedio del PON Sistemi di Politiche Attive per l'Occupazione (SPA0) per svolgere attività di assistenza metodologica e scientifica per le azioni di sistema del Fondo sociale europeo ed è Agenzia nazionale del programma comunitario Erasmus+ per l'ambito istruzione e formazione professionale. È l'ente nazionale all'interno del consorzio europeo ERIC-ESS che conduce l'indagine European Social Survey.

Presidente: Sebastiano Fadda  
Direttore generale: Santo Darko Grillo

INAPP  
Corso d'Italia, 33  
00198 Roma  
Tel. + 39 06854471  
[www.inapp.org](http://www.inapp.org)

La pubblicazione rappresenta la documentazione tecnica sull'analisi funzionale, sui modelli cognitivo, descrittivo e prescrittivo e l'architettura della soluzione ICT del progetto Sophia.

Il progetto, finanziato dal PTA 2022-2024, è gestito nell'ambito delle attività del Gruppo *Analisi prospettica e comparata dei fabbisogni occupazionali* della Struttura Lavoro e Professioni dell'INAPP coordinata da Paolo Severati e si avvale delle competenze professionali trasversali all'istituto.

Autore del testo: Boris Sofronic

Testo chiuso a maggio 2022

Pubblicato a settembre 2022

Le opinioni espresse in questo lavoro impegnano la responsabilità degli autori e non necessariamente riflettono la posizione dell'Ente.

Alcuni diritti riservati [2022] [INAPP]

Quest'opera è rilasciata sotto i termini della licenza Creative Commons Attribuzione - Non commerciale

Condividi allo stesso modo 4.0. Italia License.

<http://creativecommons.org/licenses/by-nc-sa/4.0/>



# Sommario

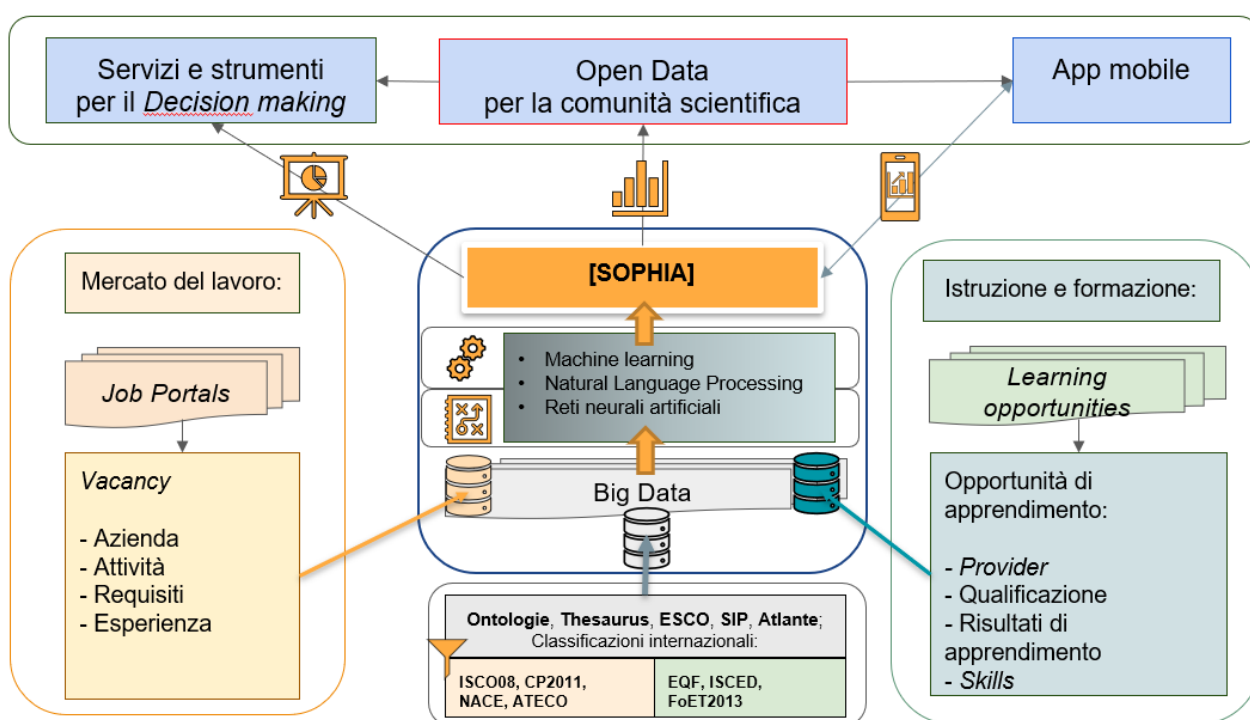
- 1. Introduzione..... 2
- 2. Obiettivi e contenuti ..... 2
- 3. Analisi funzionale..... 3
- 4. Modelli ..... 7
- 5. Architettura della soluzione..... 12

## 1. Introduzione

Il progetto Sophia è un sistema di Labour market intelligence (LMI) che in tempo reale fornisce ai propri Stakeholder - cittadini, imprese, comunità di ricerca e Policy makers - le informazioni e le analisi relative allo Skill mismatch.

In Sophia, il Mismatch viene individuato analizzando le Web job vacancies, i risultati di apprendimento dei percorsi formativi e i CV degli utenti. I dati vengono trattati con le tecnologie di intelligenza artificiale, come es. Machine learning (ML) e Natural Language processing (NLP) e catalogati mediante un'ontologia propria.

L'ontologia del progetto è una rete neurale artificiale, comprendente sia il dominio di Istruzione e formazione sia quello delle Professioni e occupazione. Integra le relative tassonomie, vocabolari, classificazioni, sistemi e strumenti come ESCO, ISCO, CP, NACE, ATECO, Atlante lavoro, SIP, ISCED, Classi di laurea ecc. Permette inoltre l'aggiunta di ulteriori termini e dizionari.



Il back office offre strumenti e funzioni per l'amministrazione del sistema, gestione delle fonti dati e dell'ontologia, pubblicazione dei contenuti diversi sulle diverse istanze del front-end di Sofia - un'App mobile per i cittadini, un'applicazione web dedicata alla comunità di ricerca scientifica e le dashboard che si rivolgono alle imprese e ai Policy makers.

## 2. Obiettivi e contenuti

L'obiettivo della presente nota è quello di fornire la documentazione tecnica relativa allo stato di avanzamento lavori aggiornato al mese di maggio del 2022. Nel documento si specificano i risultati

dell'analisi funzionale, i modelli - cognitivo, descrittivo e prescrittivo - già implementati o in fase di implementazione nel progetto e l'architettura informatica della soluzione proposta.

### 3. Analisi funzionale

Nella fase di analisi funzionale sono stati individuati tre pilastri fondamentali del progetto:

1. Una ontologia, che raggruppa tutta una serie di tassonomie, classificazioni, termini ed entità trasversali ai domini del e affini alle attività e ai contenuti del progetto;
2. Una base dati, che raccoglie i dati provenienti dal MdL e sull'offerta formativa;
3. Un modello cognitivo, che classifica, analizza ed estrae informazioni utili per l'applicazione dai dati a disposizione.

#### **L'ontologia di Sophia:**

Di seguito vengono descritti brevemente i dataset che compongono l'ontologia della soluzione, suddivisi per tipo di fonte da cui tali dati provengono.

1. Fonti dati del dominio Employment:

Dato che la stragrande maggioranza delle aziende ormai affida la ricerca del personale ai Job portals privati e pubblici e/o alle Career pages dei propri siti web, è stato necessario svolgere un'analisi preliminare per individuare i principali Player del settore ed effettuare una mappatura prendendo in considerazione l'importanza del Player in diversi mercati e settori, il numero medio delle Job vacancies giornalmente pubblicate e il trend/popolarità di ognun Player in diversi territori. È emerso che tra i Job portals privati il principale Player è Indeed, seguito da Monster, tra le piattaforme vale la pena citare Burning Glass e LinkedIn, mentre tra i portali pubblici si è imposto EURES, il portale gestito dalla DG Employment della Commissione europea che raccoglie quotidianamente una raccolta delle Job vacancies proveniente dai Centri per l'impiego di tutti i Paesi membri dell'Unione europea. Per migliorare la precisione della classificazione dei settori economici è stato inoltre individuato il DB AIDA del Bureau Van Dijk, che raccoglie i nominativi delle aziende referenziati ai codici delle classificazioni NACE e ATECO.

2. Fonti dati del dominio E&T:

#### *Offerta formativa di competenza del MUR:*

Universitaly, il portale realizzato dal MIUR con la collaborazione della CRUI, del Cineca e di tutti gli atenei italiani, fornisce la più ampia descrizione di tutte le componenti del sistema universitario in Italia, compresi i titoli di studio, elenchi di Atenei, AFAM e ITS, nonché la descrizione dei percorsi delle singole opportunità formative utilizzando le schede SUA-CdS attraverso le quali è possibile risalire ai risultati di apprendimento. Questi ultimi sono

particolarmente utili al progetto in quanto i loro testi destrutturati si avvicinano molto ai testi destrutturati delle Web job vacancies.

*Offerta formativa di competenza delle Regioni (es. leFP, IFTS):*

A differenza dell'offerta di istruzione accademica e universitaria disponibile su un'unica fonte dati (University), il quadro della formazione professionale di competenza delle Regioni è estremamente frammentato. Sophia intende avvalersi della collaborazione di più progetti INAPP che da anni se ne occupano in maniera diretta:

- Atlante del lavoro e delle qualificazioni contenente il Repertorio nazionale delle qualificazioni di competenza delle Regioni;
- DB "Accreditamento" che censisce enti accreditati alla formazione regionale;
- Monitoraggi leFP e IFTS che, nel futuro, potrebbero includere i dati sui percorsi formativi.

*La formazione privata (o internazionale):*

Dalle analisi preliminari emerge che sono sempre più richieste dal Mercato del lavoro la formazione e le certificazioni conseguite da soggetti privati (es. Microsoft, IBM, Google, Facebook, RedHat, Cisco ecc.) e/o la conoscenza certificata delle lingue straniere (es. TOEFL ottenibile da soggetti privati accreditati negli USA e IELTS del British Council). La formazione in oggetto è spesso riconducibile alle Skill trasversali/Soft skills richieste dalle Web job vacancies come requisiti di ammissibilità o preferenziali/utili per lo Screening e la valutazione dei candidati. Dai primi studi di caso è altrettanto emerso che la stragrande maggioranza delle richieste del Mercato del lavoro relative alla formazione privata ricade in due soli settori: ICT e conoscenze linguistiche. Si è quindi preferito attendere ulteriori risultati di Machine learning prima di includere fonti dati relativi a questo tipo di formazione come fonte dati di Sophia.

*Opportunità formative in Europa:*

Dopo le esperienze dei progetti "Ploteus", EQF e "Learning Opportunities and Qualifications in Europe", la Commissione europea ha deciso di incorporare le opportunità formative nel nuovo portale Europass, il noto tool per la creazione dei CV in formato europeo. Il rilascio della sezione Europass "Study in EU" è prevista nel 2022.

3. Classificazioni e tassonomie del dominio Employment:

## ISCO

International Standard Classification of Occupations (ISCO), la classificazione internazionale standard delle professioni è una struttura di classificazione dell'Organizzazione internazionale del lavoro (ILO) per l'organizzazione delle informazioni sul lavoro e sui posti di lavoro.

## ESCO

ESCO è la classificazione multilingue delle professioni e competenze in Europa. Individua e classifica le professioni e le competenze rilevanti per il mercato del lavoro dell'Ue e le relazioni tra i diversi concetti.

Il pilastro delle professioni organizza i concetti relativi all'occupazione. Ogni concetto contiene un termine preferenziale, un numero di termini non preferenziali e termini in ciascuna delle lingue di ESCO. Ciascuna professione è accompagnata da un profilo professionale. I profili sono corredati da una definizione, una descrizione e una nota operativa. Elencano inoltre le conoscenze, le abilità e le competenze considerati rilevanti da un punto di vista terminologico a livello europeo.

Ogni professione è abbinata in modo univoco ad un codice ISCO-08 che fornisce i primi quattro livelli al pilastro delle occupazioni. Le professioni ESCO sono reperibili nel livello 5 e successivi.

## CP2011

La classificazione CP2011 fornisce uno strumento per ricondurre le professioni esistenti nel mercato del lavoro all'interno di un numero limitato di raggruppamenti professionali, da utilizzare per comunicare, diffondere e scambiare dati statistici e amministrativi sulle professioni, comparabili a livello internazionale.

## ATECO

La classificazione ATECO è la declinazione italiana del NACE dell'Eurostat, che a sua volta deriva dalla classificazione ISIC dell'ONU.

## NACE

La classificazione statistica delle attività economiche nella Comunità europea è un sistema di classificazione generale utilizzato per sistematizzare ed uniformare le definizioni delle attività economico/industriali nei diversi Stati membri.

## Comunicazioni Obbligatorie

Il Sistema Informatico per le Comunicazioni Obbligatorie (CO) costituisce il punto di accesso unico per l'invio on-line delle comunicazioni di instaurazione, proroga, trasformazione, cessazione di un rapporto di lavoro, secondo i modelli unificati definiti dal Ministero del Lavoro e delle Politiche Sociali, da parte di tutti i soggetti obbligati e abilitati.



#### 4. Classificazioni e tassonomie del dominio E & T

##### ISCED

La classificazione internazionale standard dell'istruzione (ISCED) è il sistema standard internazionale dell'UNESCO di classificazione dei corsi di studio e dei relativi titoli.

##### EQF

L'EQF è un quadro basato sui risultati dell'apprendimento articolato su 8 livelli per tutti i tipi di qualificazioni, che funge da strumento di "traduzione" tra i diversi quadri nazionali delle qualificazioni. Tale quadro contribuisce a migliorare la trasparenza, la comparabilità e la portabilità delle qualificazioni delle persone e consente di confrontare le qualifiche di paesi e istituzioni diversi.

#### 5. Progetti e strumenti INAPP

Il portale Professioni e occupazione e il Sistema informativo integrato sulle professioni (SIP) Il progetto costituisce una risorsa conoscitiva per tutti gli attori istituzionali, economici e sociali interessati a comprendere la natura e le evoluzioni, in atto o tendenziali, delle professioni e del mondo del lavoro al fine di definire più mirate ed efficaci politiche del lavoro e della formazione.

##### Atlante Lavoro

L'Atlante Lavoro descrive i contenuti del lavoro attraverso 23 settori economico – professionali (SEP) derivati dalle classificazioni ISTAT delle attività economiche (ATECO) e delle professioni (CP) e un settore trasversale, "Area comune", il quale descrive i processi di supporto alla produzione di beni e servizi. I SEP e l'Area comune sono descritti per processi, sequenze di processo e aree di attività (ADA). Le ADA sono le principali unità dell'Atlante. Contengono le descrizioni delle singole attività che la costituiscono e i riferimenti ai codici delle classificazioni ISTAT CP 2011 e ATECO 2007.

##### Archivio Nazionale delle Strutture Formative Accreditate (ANSFA)

ANSFA raccoglie le principali informazioni su tutti i sistemi di accreditamento di Regioni e Province Autonome e sui soggetti accreditati per la formazione in tutta Italia.

##### Stage4eu

Stage4eu è un'App mobile e un sito web che, oltre a un set completo di informazioni e riferimenti utili, seleziona quotidianamente le più interessanti offerte di stage nei diversi Paesi europei.

In Sophia, Stage4eu può avere una duplice implementazione:

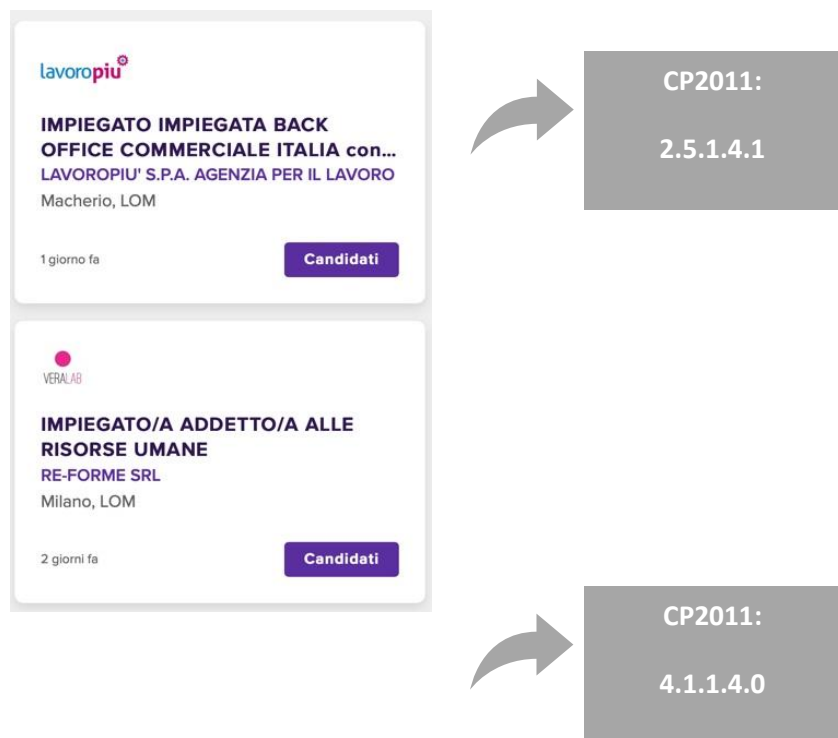
- Come fonte dati delle Web job vacancies nel settore pubblico in Europa;
- Come strumento di orientamento all'esito del CV check dell'utente nei casi in cui l'Engine di intelligenza artificiale di Sophia rileva lo Skill mismatch relativo

all'autonomia, intraprendenza, responsabilità, capacità di adattamento ai contesti multiculturali, internazionali o comunque non abituali, conoscenza insufficiente di una lingua straniera ecc.

## 4. Modelli

### Il modello cognitivo di classificazione

Il modello cognitivo classifica offerte lavorative secondo il quinto livello della classificazione CP2011. Questa classificazione viene eseguita da Watson NLU (Natural Language Understanding), uno dei servizi Watson in grado di associare la corretta unità professionale CP2011 al testo che gli si sottomette:



I risultati ottenuti fino ad ora attraverso il training del modello cognitivo di classificazione risultano incoraggianti. Infatti, come è possibile vedere nel test sotto, Watson NLU riesce a classificare correttamente l'88.72% delle posizioni lavorative nei suoi primi tre risultati:

File Name: C:\\$user\Cienti\NAPP\Test\_Classificatore\test\_\_0.0.2.csv  
 NLU Model ID: c16c981c-67aa-4e5d-b0a4-8550939c58e0  
 Total Contents: 195 (Right Count: 173 - Wrong Count: 22)  
 Right Classes %: 88.72 %  
 Wrong Classes %: 11.28 %

All Classes  Right Classes Only  Wrong Classes Only

**Statistics by Class**

Expected Class	Actual First 3 Classes	Total Count	Right Count (%)	Wrong Count (%)
↕ 3.1.2.4.0	3.1.2.4.0   2.1.1.5.3   2.5.1.5.2	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.2.5.0	2.1.1.5.1   3.1.2.5.0   3.1.2.6.1	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.2.6.1	3.1.2.6.1   3.1.3.4.0   6.2.4.4.0	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.3.1.0	3.1.3.1.0   2.2.1.1.1   3.1.3.3.0	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.3.2.3	3.1.3.1.0   2.2.1.1.1   3.1.3.2.3	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.3.3.0	3.1.3.3.0   3.1.1.2.0   3.1.4.1.2	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.3.4.0	3.1.3.3.0   3.1.7.3.0   2.2.1.1.4	1	0 (0.00 %)	1 (100.00 %)
↕ 3.1.3.5.0	3.1.3.5.0   3.1.4.2.2   4.2.1.3.0	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.3.6.0	3.1.3.6.0   5.2.2.2.3   6.2.2.3.1	1	1 (100.00 %)	0 (0.00 %)
↕ 3.1.3.7.1	2.2.1.3.0   4.2.2.4.0   3.1.3.3.0	1	0 (0.00 %)	1 (100.00 %)
↕ 3.1.3.7.2	2.3.1.3.0   3.1.3.7.2   7.2.6.1.0	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.2.8	3.2.1.2.8   5.4.3.3.0   3.2.1.2.6	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.3.2	3.2.1.3.2   2.3.1.1.1   2.2.1.8.0	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.3.3	2.4.1.6.0   3.2.1.2.4   3.2.1.3.3	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.3.4	3.2.1.3.4   3.2.1.2.5   2.3.1.1.1	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.4.1	3.2.1.4.1   6.2.4.2.0   5.1.3.2.0	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.4.2	3.2.1.4.2   3.1.8.3.1   6.3.1.6.1	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.4.3	3.2.1.4.3   2.4.1.5.0   3.2.1.6.2	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.4.5	3.2.1.4.5   2.4.1.7.1   2.3.1.1.4	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.5.1	3.2.1.5.1   3.1.8.2.0   2.5.1.3.1	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.5.2	3.2.1.5.2   2.4.1.1.0   3.2.1.2.8	1	1 (100.00 %)	0 (0.00 %)
↕ 3.2.1.6.1	3.2.1.6.1   6.3.1.5.1   5.1.2.5.1	1	1 (100.00 %)	0 (0.00 %)

La percentuale di classificazioni corrette da parte del modello può essere aumentata fornendo ulteriori esempi di offerte di lavoro, avendo cura che la qualità degli esempi raccolti sia comunque alta e l’offerta lavorativa racchiuda diverse caratteristiche univoche o quasi, al fine di estrapolare i punti principali di divergenza dalle altre unità CP2011.

**Il modello cognitivo per l’estrazione di entità**

Il modello cognitivo per l’estrazione di entità analizza le offerte di lavoro, i CV degli utenti e i risultati di apprendimento dei percorsi di istruzione e formazione.

La finalità è quella di estrapolare ulteriori informazioni, come ad esempio:

- Competenze;
- Professioni;
- Aziende;
- Sedi operative;
- Corsi di laurea;
- Livelli di conoscenza delle lingue;
- Requisiti richiesti e skill possedute.

Questa analisi arricchisce la base dati di Sophia, permettendo di registrare, per ogni unità professionale, tutte le caratteristiche che la differenziano dalle altre.

Inoltre, riuscendo ad estrapolare i dati appena descritti da più offerte di lavoro associate ad una stessa unità professionale, è possibile portare a termine una profilatura sempre più completa ed aggiornata delle qualifiche, competenze ed esperienze richieste da parte dei datori di lavoro per ogni specifica professione. Tutto ciò è replicabile anche lato qualificazioni, competenze ed esperienze descritte nei CV degli utenti.

### Esempio analisi cognitiva di un'offerta di lavoro

La figura seguente rappresenta un esempio dell'output del modello cognitivo di Sophia, relativo all'analisi di una offerta di lavoro. L'immagine ritrae l'ambiente di training del modello di estrazione entità Watson Knowledge Studio (WKS).

Il modello all'interno di WKS, una volta ultimato, viene inserito all'interno di Watson Natural Language Understanding (NLU) e reso disponibile per l'analisi Runtime dei vari tipi di documento. Le parti di testo colorate corrispondono alle entità estratte dal modello cognitivo. Ad esempio, i testi evidenziati in azzurro corrispondono alle competenze ESCO (es. market analysis o supporto alla pianificazione strategica), quelli in verde chiaro ai titoli di studio (es. laurea magistrale), mentre quelle in arancione ai comuni (es. Prato o Milano):

#### JV13

"**Orienta** SPA Agenzia per il Lavoro ricerca, per azienda cliente specializzata nella realizzazione di progetti di **ricerca e sviluppo** e di trasferimento tecnologico a livello nazionale ed europeo,

UN/A INGEGNERE **MATEMATICO**.

La figura ricercata si occuperà della risoluzione di problematiche complesse, che richiedono un'approfondita indagine di tipo modellistico-numerico e di tipo probabilistico-statistico.

E' richiesta **laurea magistrale** in ingegneria **matematica**.

Completano il profilo l'attitudine al **problem solving** e la **flessibilità** nella **gestione** delle tecnologie dell'ingegneria.

Luogo di lavoro: **Prato**

Orario di lavoro: **Full Time** "

"Per la nostra sede di **Milano** siamo alla ricerca di una risorsa da inserire nel team di **Market Analysis** & Price Forecasting che si occupa dell'**analisi** e previsioni di fondamentali di mercato e dei prezzi sui mercati energetici di interesse come input alla strategia di **gestione** del portafoglio energetico di **A2A**.

La struttura si occupa inoltre della **produzione** di scenari energetici di lungo periodo, elemento di **supporto alla pianificazione strategica** per la Società.

Attività

- **Analisi** su fondamentali e delle dinamiche dei prezzi per i mercati energetici di interesse per il Gruppo, con particolare riferimento al mercato elettrico in Italia ed Europa, ivi compreso il contesto regolatorio e gli scenari climatici.
- **Elaborazione, manutenzione** e misura delle performance di modelli di previsione dei prezzi sui mercati dell'energia elettrica, con orizzonte di breve-medio periodo, a **supporto decisionale** per l'operatività dell'area Energy Management.
- **Elaborazione**, in **collaborazione** con il team, di scenari su fondamentali e prezzi relativi al mercato dell'elettricità in Europa nel più lungo periodo, a **supporto** dell'attività di Scenario **Planning**.

Requisiti

- **Laurea Specialistica** in Ingegneria o equivalenti, con preferenza per indirizzi accademici in ambito quantitativo.
- Precedente esperienza di almeno tre anni su mercati dell'energia **italiano/europei**.

Esperienze in ambito **market analysis** o trading saranno valutate positivamente.

- Eccellenti **competenze analitiche** e di **presentazione**, ivi compreso l'uso **avanzato** di software come Microsoft **Excel** e **Powerpoint**.
- La conoscenza pregressa di linguaggio **SQL** e di linguaggi di programmazione (**R**, **Python**, **Matlab** o equiparabili) sarà requisito preferenziale.
- **Ottima conoscenza** della lingua **Inglese**.
- Attitudine al **lavoro in team** e capacità di **gestire** efficacemente tempo e scadenze.

## Esempio analisi cognitiva di un CV

L'immagine che segue fa riferimento all'analisi cognitiva di un CV. Anche in questo caso, le parti di testo colorate corrispondono alle entità estratte dal modello cognitivo. I testi evidenziati in azzurro corrispondono alle competenze (es. tecniche di conservazione e restauro del materiale cartaceo), quelli in verde olivastro alle lingue (es. italiano, inglese o spagnolo), mentre quelle in rosa ai livelli di conoscenze linguistiche (es. lingua madre, B2 o buona conoscenza):

### COMPETENZE PERSONALI

Lingua madre

italiano

Altre lingue

COMPRESIONE, PARLATO, PRODUZIONE SCRITTA

Ascolto, Lettura, Interazione, Produzione orale

Inglese: B2, C1, B2, B2, B2

Spagnolo: B1, C1, B2, C2, B2

Competenze comunicative

Ottime capacità comunicative e organizzative.

Buona attitudine alla multiculturalità.

Competenze organizzative e gestionali

Ben organizzata e in grado di gestire situazioni di stress.

Esperienza nel lavoro di equipe e pianificazione di progetti strutturati.

Ben organizzata nel problem solving grazie all'esperienza decennale derivata dal ruolo di capitano nella squadra agonistica di pallavolo della Prometeo Monteverde

Competenze professionali:

Buona conoscenza sulle tecniche di conservazione e restauro del materiale cartaceo e pergameneo: operazioni per via umida, reintegro delle lacune e reintegrazione cromatica.

Conoscenza di alcune tecniche diagnostiche: Imaging multispettrale, XRF, Spettroscopia Raman, Termografia IR.

Riordino archivistico digitale e metadati.

Conoscenza Photoshoph

Altre competenze

Conoscenza dei software Microsoft (Word, Excel, Power Point) e degli ultimi dispositivi (iPhone, tablet).

Buona conoscenza dei sistemi operativi (IOS, Windows).

Patente di guida: B

## Esempio analisi cognitiva di un'offerta formativa

La figura si riferisce all'analisi cognitiva di un'opportunità formativa. Le parti di testo colorate corrispondono alle entità estratte dal modello cognitivo. I testi evidenziati in azzurro corrispondono alle competenze (es. abilità informatiche di base), mentre quelli in giallo alle soft skill (es. comunicazione o abilità di inserimento in attività di gruppo):

### OF1

Il Corso di laurea in SERVIZIO SOCIALE dà diritto ad accedere, previo Esame di Stato, alla professione di Assistente Sociale (ai sensi della legge 84/93 e del DPR 328/2001).

Sulla base dell'Ordinamento didattico, e dei relativi obiettivi qualificanti, il nuovo progetto formativo si basa sulla coniugazione delle tre aree fondanti la professione di assistente sociale: 1.

i valori declinati nella professione del servizio sociale (il perché); 2.

le conoscenze teorico-professionali (il che cosa); 3.

le abilità professionali o skills (il come).

Pertanto, i laureati nel Corso di laurea di Servizio Sociale della classe L-39 devono acquisire: • conoscenze di tipo interdisciplinare di base in ambito sociologico, antropologico, politologico, giuridico e psicologico al fine di comprendere le caratteristiche delle società moderne e di collaborare alla costruzione di progetti di intervento individuale, di gruppo e di comunità; • un'adeguata sensibilizzazione al metodo di ricerca sociale; • conoscenze adeguate delle teorie, dei metodi e delle tecniche proprie del Servizio sociale; • competenze professionali relative al rilevamento, alla progettazione, all'intervento e alla prevenzione di situazioni di disagio sociale, riferite sia a singoli sia a gruppi e comunità; • competenze utili alla gestione e alla progettazione della comunicazione e dell'informazione in generale, anche tramite abilità informatiche di base, e specificamente per quanto attiene ai diritti di cittadinanza e all'accompagnamento dei soggetti in difficoltà; • capacità e abilità di inserimento in attività di gruppo mono e pluri-professionale; • conoscenze per utilizzare efficacemente, in forma scritta e orale, almeno una lingua dell'Unione Europea, oltre l'italiano, nell'ambito specifico di competenza e per lo scambio di informazioni generali; • conoscenze, competenze, e capacità utili ad intervenire nell'ambito delle dinamiche migratorie; • competenze relative alla qualità del servizio, con riferimento sia alla qualità oggettiva delle prestazioni erogate e dei processi attivati, sia alla qualità percepita; • competenze per la costruzione di reti fra i servizi, gli enti e le organizzazioni coinvolte nel rapporto con l'utente/cliente/cittadino; • competenze interdisciplinari utili a ricomporre le diverse logiche progettuali presenti nell'intervento globale in favore delle persone; • conoscenze, competenze, atteggiamenti specifici della professione nel raccordo tra teorie e pratiche di intervento, sia attraverso laboratori didattici sia attraverso la guida al tirocinio e il tirocinio in organizzazioni pubbliche e private, adeguate e convenzionate con l'Ateneo e il Dipartimento, per un totale di 18 CFU.

## Modello descrittivo

Il modello descrittivo viene costruito in tempo reale, partendo dalla base dati iniziale di Sophia e arricchito con le informazioni estratte dal modello cognitivo. Ha la funzione di cercare informazioni specifiche e produrre statistiche sulle loro tendenze ordinabili in base alla professione, al settore, o al tipo di competenza richiesta.

Il modello prevede quattro tipi di statistiche principali:

1. Una prima statistica consiste nella classifica "Top 10", la classifica delle competenze, abilità, soft skill, qualificazioni ed esperienze lavorative maggiormente richieste in un certo periodo di tempo in un territorio di riferimento.
2. Un'altra statistica prevede la composizione del mercato del lavoro in percentuale. Tale indicatore offrirà un quadro della composizione relativa del mercato del lavoro, riferita alle singole professioni o alle qualificazioni ed esperienze (richieste o trovate sui CV) in una specifica data. La formula che permette il calcolo di tale statistica è rappresentata dal rapporto tra il numero di osservazioni di ciò che ci interessa rispetto alla somma del numero di osservazioni totali, quindi:

$$Freq. Rel. X = \frac{frequenzaX_t}{frequenzaTotale_t}$$

3. Una terza statistica prevede la ricerca di una o più competenze nei risultati di apprendimento e fornisce indicazioni sul dove poterle acquisire.
4. L'ultima statistica generata dal modello descrittivo è quella relativa alla tendenza del mercato del lavoro, ricavato attraverso il confronto tra dati registrati in periodi diversi. Anche in questo caso sarà possibile impostare il periodo temporale di interesse. La tendenza per singola unità professionale verrà calcolata attraverso una percentuale ottenuta secondo il rapporto

$$trend = \frac{frequenza_t}{frequenza_{t-1}}$$

Il trend potrà essere associato ad un grafico settimanale e/o mensile, per una più chiara visualizzazione dell'andamento dell'informazione di interesse.

## Modello prescrittivo

Il modello prescrittivo, l'ultimo dei modelli previsti per l'applicazione Sophia, analizza le offerte di lavoro e i CV, andando a formulare consigli di orientamento su quale percorso di studio sarebbe più adatto al fine di massimizzare le possibilità di assunzione del candidato.

Le raccomandazioni sui corsi di formazione saranno composte grazie ai risultati di apprendimento definiti per ciascun percorso di istruzione e formazione. È infatti possibile,

attraverso i dati a disposizione, avere un collegamento diretto tra formazione e sbocchi professionali, rendendo solido il processo prescrittivo.

Il modello prescrittivo prevede un sistema “a punteggio” che supporta Sophia nella fase di fornire raccomandazioni al candidato:

Nel dettaglio, dopo il caricamento del CV all’interno dell’applicativo, i modelli cognitivi lo classificano e ne estrapolano le componenti fondamentali, quali esperienze lavorative, qualificazioni e conoscenze linguistiche. Il sistema assegna, quindi, un punteggio ad ognuna delle offerte lavorative. Questo punteggio viene personalizzato in relazione al CV di ogni utente e rappresenta quanto una specifica offerta di lavoro sia attinente al profilo del candidato.

#### Risultati attesi del modello prescrittivo

Il modello prescrittivo sarà in grado di fornire consigli riguardanti la professione più in linea con il profilo del candidato, suggerire corsi da frequentare o qualificazioni da conseguire, al fine di essere in linea con un certo tipo di carriera.

In particolare, al candidato sarà data la possibilità di filtrare le offerte di lavoro secondo filtri come la professione o le qualifiche richieste e, di conseguenza, di visualizzare a quanto ammonta il punteggio di attinenza delle offerte lavorative che più lo appassionano, ma che potrebbero non essere ricopribili nell’immediato a causa di una sua mancanza, totale o parziale, dei requisiti richiesti.

Trovata la professione di interesse, il modello prescrittivo evidenzierà le cause che impediscono al candidato di essere competitivo per il processo di selezione e gli fornirà consigli di orientamento sui percorsi di istruzione e formazione più adatti alla professione desiderata individuando anche i compiti, attività specifiche della professione, conoscenze, competenze, attitudini e altre caratteristiche richieste ma mancanti nel su CV.

## 5. Architettura della soluzione

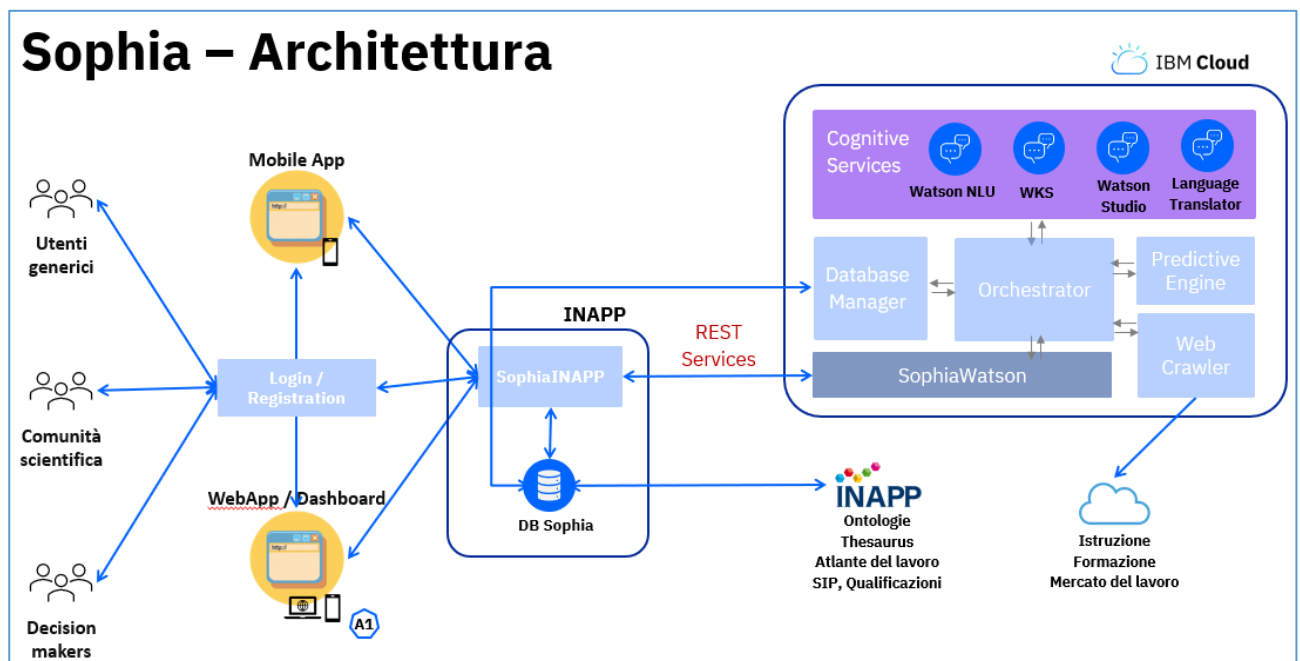
### IBM Watson Cognitive Services

L’applicazione Sophia si avvale di un insieme di servizi IBM Watson, basati su algoritmi di intelligenza artificiale, per analizzare il testo dei documenti ed estrarne le informazioni ritenute rilevanti ai fini applicativi. I servizi utilizzati sono i seguenti:

- **Watson Knowledge Studio (WKS):** con questo servizio è possibile creare dei modelli cognitivi in grado di riconoscere specifici pattern all’interno del testo ed estrarre quindi entità e/o relazioni tra due entità
- **Watson Natural Language Understanding (NLU):** questo servizio, tramite uno o più modelli cognitivi costruiti con WKS, analizza il testo ed estrae in Runtime le entità e le

relazioni secondo cui sono stati addestrati i modelli; inoltre è in grado di esaminare la sintassi del testo, estrarre le Part of Speech presenti in esso (nomi, verbi, aggettivi, ...) e fornirne il lemma.

- **Watson Language Translator:** traduce da una lingua all'altra il testo che gli viene sottomesso; nel caso dell'applicazione Sophia, tradurrà in italiano i testi delle offerte di lavoro o delle offerte formative eventualmente scritti in inglese.
- **Watson Studio:** mette a disposizione un ambiente specializzato per la creazione di modelli predittivi; il modello risultante dal training svolto sarà poi utilizzato dal modulo Predictive Engine per estrapolare andamenti e tendenze dei dati forniti in input.



Esigenze del progetto:	Soluzioni implementate:
AI, Analytics predittivi e prescrittivi; Open data e Visual tools per data scientists; Open source (Python, R, PyTorch, TensorFlow, scikit-learn)	IBM Watson® Studio consente a data scientist, sviluppatori e analisti di costruire e gestire i modelli AI e di ottimizzare le decisioni ovunque su IBM Cloud Pak® for Data. Automatizza i cicli di vita dell'AI e accelera il time to value su un'architettura multicloud open source. Raggruppa i framework open source come PyTorch, TensorFlow e scikit-learn con IBM e i relativi strumenti dell'ecosistema per una data science basata sul codice e visiva. Lavora in linguaggi come Python, R e Scala.
Usare dati disponibili in rete invece di replicarli; Big data; Crawler/spider, APIs, web scraping	IBM Watson® Discovery è una pluripremiata piattaforma di ricerca e analytics del testo basata sull'AI che elimina i silos di dati e recupera le informazioni che si nascondono nei dati aziendali. La piattaforma utilizza un'innovativa NLP (natural language processing) leader del settore per scoprire insight di business significativi da documenti, pagine web e big data.



Mettere a sistema diverse classificazioni Inapp, nazionali e internazionali	IBM Watson® Knowledge Catalog , basato su IBM Cloud Pak™ for Data, è un catalogo dati strettamente integrato con una piattaforma di governance dei dati aziendali. Un catalogo dati consente ai data citizen di trovare, organizzare e comprendere facilmente i dati di cui hanno bisogno.
Tradurre vacancies, learning outcomes e CVs	Watson Language Translator traduce il testo da una lingua all'altra.
Natural Language Processing;  Capire e mettere in relazione job vacancies, learning outcomes e CV degli utenti usando diverse classificazioni	Watson Natural Language Understanding analizza il testo per estrarre i metadati dal contenuto, quali concetti, entità, parole chiave, categorie, sentimenti, emozioni, relazioni e ruoli semantici utilizzando la comprensione del linguaggio naturale.
Normalizzare dati prima di offrirli in pasto agli analytics e alle applicazioni front-end web e mobile	Lo strumento Data Refinery, disponibile tramite Watson Studio e Watson Knowledge Catalog, consente di risparmiare tempo nella preparazione dei dati attraverso la trasformazione rapida di enormi quantità di dati non elaborati in informazioni di qualità, utilizzabili, pronte per l'analytics.

L'architettura della soluzione si basa su un modulo centrale di back-end che gira sui sistemi INAPP e che prende in carico le richieste dell'utente finale (utente generico, comunità scientifica, impresa, Decision maker), richiama un modulo su IBM Cloud (Orchestrator) per l'analisi dell'eventuale documento sottomesso dall'utente (CV) e, in base alla tipologia dell'utente, fornisce informazioni specifiche.

Il modulo Orchestrator su IBM Cloud ha anche il compito di scaricare da specifici siti web informazioni sulle job vacancy e sulle offerte formative di università, enti pubblici e aziende, analizzare tramite i servizi IBM Watson questi documenti e popolare il database con le informazioni strutturate ricavate dall'analisi.

In base alle informazioni ricavate dalle suddette analisi cognitive, uno specifico modulo chiamato Predictive Engine ha il compito di estrapolare statistiche e tendenze future del MdL.

## Front-end:

Sophia prevede diverse istanze front-end in funzione delle diverse esigenze dei suoi molteplici Stakeholder che richiedono diverse esperienze utente (UX), Graphic User Interface (GUI) e servizi personalizzati (es. notifiche “push”).

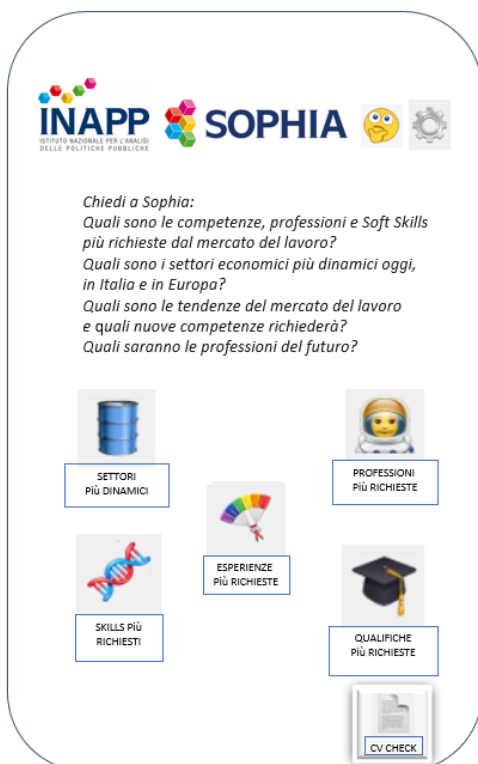
Sono state progettate diverse interfacce utente, che al momento della chiusura del presente documento si trovano in fase di sviluppo e graduale implementazione:

### 1. APP

L’App mobile di Sophia permette agli utenti la consultazione dei profili professionali, titoli di studio, competenze, esperienze, Soft skills e lingue maggiormente richieste dal Mercato del lavoro (MdL) nazionale/europeo in tempo reale e l’analisi predittiva delle tendenze del MdL a medio termine.

Previo consenso al trattamento dati, offre anche servizi di confronto del proprio CV con i requisiti delle job vacancies inerenti il profilo dell’utente e servizi di orientamento su dove acquisire le competenze mancanti. Invia, inoltre, le notifiche push relative agli aggiornamenti sull’andamento del MdL e/o analisi prescrittive sulle nuove opportunità formative.

Sempre previo consenso da parte dell’utente, l’app alimenta il core di Sophia con i microdati utente utili alla ricerca scientifica (geolocalizzazione, sesso, età, titolo di studio, profilo professionale, esperienze, lingue ecc.):



## 2. APPLICAZIONE WEB

L'applicazione web mette a disposizione della comunità di ricerca scientifica i datasets di Open data disponibili per download e per ulteriori elaborazioni statistiche, servizi di online data meshing, dati in formato Machine-readable e strumenti visuali per la creazione delle infografiche e animazioni:

The screenshot displays the SOPHIA web application interface. At the top, there are logos for INAPP (Istituto Nazionale per l'Analisi delle Politiche Pubbliche) and SOPHIA, followed by a search bar and a 'CERCA' button. A user profile area on the right shows 'Area riservata' and a smiley face icon.

Navigation links include: | Cittadini | **Ricercatori** | Servizi e strumenti avanzati | | **Open Data** | Strumenti visuali |

Filter options are provided: [  In tempo reale |  Tendenze ] | Area geografica: | dal: \_\_/\_\_/\_\_ al: \_\_/\_\_/\_\_ ]

On the left, a sidebar titled 'Aggiungi Dataset:' contains a search bar and a 'CERCA' button. Below it, a list of categories is shown: 'Vedi tutti/Scegli uno o più Dataset:', 'NACE', 'Professioni' (with '[ISCO|CP2021]'), 'Skills' (with 'Competenze', 'Conoscenze', 'Abilità', 'Soft skills'), 'Qualificazioni' (with 'Istruzione terziaria', 'Classi Laurea', 'Istruzione secondaria'), 'Esperienze', and 'Indagini INAPP:'.

The main content area features a list of agricultural codes (A1 to A130) and their descriptions, such as 'Agriculture, forestry and fishing', 'Crop and animal production, hunting and related service activities', 'Growing of non-perennial crops', etc.

Visualizations include a map of Italy with location markers, two 'Calendario 2021' (calendar) widgets, and a table titled 'Classificazione delle Classi di Laurea secondo la FOET\_2013'. The table has columns: Numero, Denominazione, Ordinamento, and three columns for ICFIS (I, F, S) with values ranging from 05 to 09.

At the bottom, there are options to 'Scarica:' (Download) using CSV or RDF formats, and 'Crea:' (Create) using 'Infografica' or 'animated data'. A 'footer' link is also present.

### 3. DASHBOARD

Le Dashboard di Sophia permettono agli iscritti al servizio – imprese, progettisti della formazione e Policy makers – servizi utili a individuare le cause dello Skill mismatch. Ai Policy makers delle ALMP offre i dati utili per progettare e analizzare le politiche attive e a quelli dell'istruzione e formazione offre spunti di riflessione sull'adeguatezza dei percorsi formativi di propria competenza. Offre un servizio di notifiche push personalizzate (es. le notifiche sulle nuove competenze rilevate sul Mdl inerenti a un percorso formativo, un profilo professionale o una qualificazione). Alle imprese Sophia può offrire informazioni sui trend del settore economico di propria appartenenza, sulla presenza di specifiche competenze nel territorio d'interesse, sui corsi di formazione professionale relativi al proprio settore economico ed erogati nei territori d'interesse e le analisi predittive sui settori economici, professioni, competenze e qualificazioni emergenti:

