



**Working Paper**

**SOPHIA**  
**ONTOLOGIA INIZIALE DEL PROGETTO**

**Dicembre 2021**



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 raccoglie e documenta i risultati raggiunti nella prima fase del progetto Sophia che attraverso la ricerca, analisi e organizzazione di diverse fonti dati ha portato a una prima ontologia di base del progetto.

Il progetto, finanziato dal PTA 2021-2023, 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.

Gruppo di lavoro: Chiara Carlucci; Valeria Cioccolo; Monia De Angelis; Giuseppina Di Iorio; Boris Sofronic

Sono autori del testo: Giuseppina Di Iorio (par.2), Monia De Angelis (par.2), Boris Sofronic (par. 1, 3, 4 e 5)

Testo chiuso a dicembre 2021

Pubblicato a marzo 2022

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

Alcuni diritti riservati [2021] [INAPP]

Quest'opera è rilasciata sotto i termini della licenza Creative Commons Attribuzione Non commerciale. Condividi allo stesso modo [4.0. Italia License](https://creativecommons.org/licenses/by-nc-sa/4.0/).



## Indice

1. Progetto Sophia: Real-time, User centric Multistakeholder Labour Market Intelligence System	p. 06
2. Vocabolari strutturati, ontologie, glossari, classificazioni, tassonomie e thesaurus: definizioni e distinzioni	p. 09
3. Ontologia di Sophia: verso un'ontologia supportata dell'intelligenza artificiale	p. 15
4. Principali classificazioni, tassonomie e sistemi confluiti nell'ontologia iniziale di Sophia	p. 19
5. Conclusioni	p. 30
Bibliografia	p. 32

## **ABSTRACT**

Il progetto Sophia è un Labour Market Intelligence System che intende fornire ai propri stakeholders - cittadini, comunità di ricerca e decision maker - le informazioni e le analisi relative allo skill mismatch in tempo reale.

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 e Natural language processing 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 e sistemi come ESCO, ISCO-08, CP-2011, NACE, ATECO, Atlante lavoro, SIP, ISCED, Classi di laurea ecc. Permette inoltre l'aggiunta di ulteriori termini e dizionari.

L'obiettivo del presente documento è quello di descrivere i processi di analisi e sviluppo dell'ontologia iniziale del Progetto.

## 1. PROGETTO SOPHIA: REAL-TIME, USER CENTRIC MULTISTAKEHOLDER LABOUR MARKET INTELLIGENCE SYSTEM

Quali sono le competenze, le professioni e le soft skills più richieste dal mercato del lavoro? Quali sono i settori economici più dinamici oggi in Europa? Come è cambiata la domanda di lavoro nell'ultimo periodo? Quali sono le tendenze del mercato del lavoro, quali nuove competenze richiederà e quali saranno le nuove professioni nel futuro a breve/medio termine? Ci sono le competenze richieste dal mercato e mancanti in un percorso di studi in cui sono coinvolto? Il mio CV sarà più spendibile completandolo con un'esperienza Erasmus o con un tirocinio? Quale lingua straniera studiare? Qual è la mia posizione attuale sul mercato del lavoro, quali competenze mi mancano e dove posso acquisirle?

Sophia si offre come una possibile risposta a questi e molti altri quesiti provenienti dal mondo della ricerca e dagli utenti in cerca di lavoro. Sophia risponde in tempo reale incrociando i dati delle offerte di lavoro, dei risultati di apprendimento dell'offerta formativa con i microdati del feedback degli utenti, valorizzando nel contempo gli strumenti dell'INAPP - come il Sistema informativo sulle professioni e l'Atlante del lavoro – e assicurando l'interoperabilità della soluzione proposta con gli strumenti della CE - EURES, Europass ed ESCO, come da Comunicazione della CE - [COM \(2017\) 134](#) "European Interoperability Framework (EIF) and Digital Single Market" e "[Linea di indirizzo sull'interoperabilità tecnica delle Pubbliche Amministrazioni](#)" dell'AGID.

Sophia si rivolge ai propri stakeholders mediante diversi front-end: un app mobile per i cittadini, una web app dedicata alla comunità di ricerca scientifica e una dashboard che si rivolge ai decision maker. Sono, inoltre, in fase di analisi ulteriori possibili implementazioni del progetto, es. "Sophia per le imprese" e "Sophia per gli enti di istruzione e formazione".

### **1.1. Applicazione mobile**

App mobile di Sophia permette agli utenti la consultazione dei profili professionali, titoli di studio, competenze, esperienze e lingue maggiormente richiesti dal 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, propensione alla mobilità ecc.).

### **1.2 Applicazione web**

Applicazione web mette a disposizione della comunità di ricerca scientifica tutto quanto già riportato per l'App, 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.

### **1.3 Dashboard**

La dashboard permette agli iscritti al servizio di individuare meglio le cause dello *skills mismatch*. Ai *decision 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 propri percorsi formativi. Offre un servizio di notifiche push personalizzate (es. le notifiche sulle nuove competenze rilevate sul MdL inerenti un percorso formativo/profilo professionale/qualificazione possono tornare utili a un preside/rettore per progettare meglio la propria offerta formativa).

Il back office dà 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 Sophia.

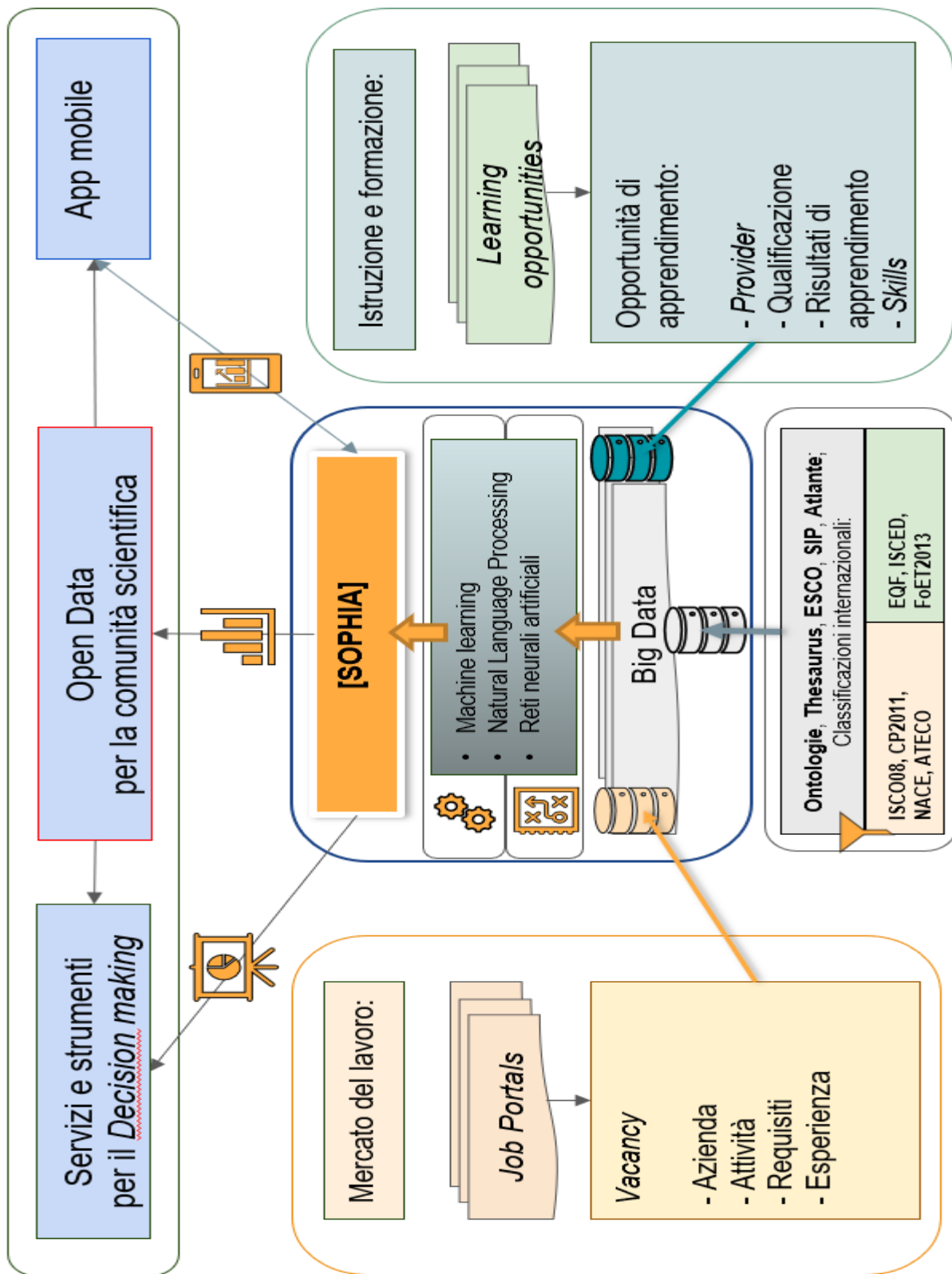


Figura 1. Sophia: schema dei componenti, funzioni e interconnessioni



## **2. VOCABOLARI STRUTTURATI, ONTOLOGIE, GLOSSARI, CLASSIFICAZIONI, TASSONOMIE E THESAURUS: DEFINIZIONI E DISTINZIONI**

### **2.1. I vocabolari controllati: organizzazione e recupero semantico dell'informazione**

Nell'attuale contesto operativo, caratterizzato da un gran numero di risorse informative virtualmente disponibili, gli strumenti per l'organizzazione dell'informazione e della conoscenza devono essere in grado di produrre il risultato desiderato dall'utente nelle modalità meno dispendiose per tempi e costi.

Un'intelligente e corretta organizzazione dell'informazione è indispensabile non solo alla comunità della ricerca, ma all'intera società. Thesauri, tassonomie, schemi di classificazione, ontologie<sup>1</sup> sono gli strumenti utili non solo alla necessaria comprensione delle stesse aree tematiche o domini disciplinari, ma anche per dialogare con operatori di diverse aree linguistiche.

Gli strumenti di controllo terminologico non solo consentono all'utente, orientandolo, di recuperare le informazioni (*information retrieval*), ma soprattutto normalizzano la terminologia attraverso un metodo di analisi concettuale dell'area in oggetto, creando un metodo nell'analisi e nella gestione delle informazioni, in particolar modo nell'eterogenea e illimitata informazione accessibile via Web.

### **2.2. I vocabolari controllati**

I vocabolari controllati consentono di trovare un punto di incontro tra lessico usato dall'autore e quello usato da chi cerca un documento, stabiliscono una relazione biunivoca tra termine e concetto, così da creare univocità semantica, ossia un termine per ogni concetto e un concetto per ogni termine, determinando in tal modo un linguaggio artificiale.

Ciò è finalizzato a ridurre i problemi connessi all'utilizzo del linguaggio naturale, ricco di ridondanze, ambiguità, omonimie, ecc. che impattano sull'organizzazione dell'informazione.

Un vocabolario controllato assume un livello strutturale maggiore quando i termini sono inseriti all'interno di gerarchie. La relazione gerarchica non solo lega reciprocamente due termini

---

<sup>1</sup> Dal punto di vista strutturale, si differenziano gli uni dagli altri sostanzialmente per la presenza o meno di relazioni di tipo semantico tra i concetti in essi rappresentati e per il diverso grado di formalismo che caratterizza la modellizzazione dell'informazione.

(concetto/termine e viceversa), ma dà luogo ad un rapporto che va dal concetto più generico a quello più specifico.

Si hanno quindi:

- **Tassonomie**

In senso tradizionale, l'idea di tassonomia è legata alle discipline scientifiche e alla classificazione degli organismi nelle scienze biologiche, il termine viene usato spesso come sinonimo di sistematica. Nell'ambito della documentazione, invece, il termine acquista un'accezione più ampia, riferendosi all'organizzazione sistematica di un soggetto o dominio. Secondo la definizione fornita dallo standard americano ANSI/NISO Z39-19:2005<sup>2</sup>, infatti, per tassonomia si intende: «*a controlled vocabulary consisting of preferred terms, all of which are connected in a hierarchy or polyhierarchy*».

Liste di termini preferiti, in cui ogni termine appartiene almeno a una struttura gerarchica che ne determina la tipica organizzazione ad albero che rende visibili i rapporti tra termini sovra e sotto-ordinati. Spesso i termini tassonomia e thesaurus sono utilizzati indifferentemente, come se fossero sinonimi, ma tra le due tipologie esistono differenze significative: rispetto ad un thesaurus, infatti, una tassonomia non prevede le relazioni di equivalenza, né quelle associative per l'esplicitazione di rapporti semantici diversi da quelli gerarchici. Generalmente, le relazioni ontologiche non vengono esplicitate attraverso una notazione, ossia non sono codificate all'interno di un sistema di simboli (cifre, lettere o altro) che rimandi al tipo di relazione e, in definitiva, all'ordine prestabilito all'interno della struttura classificatoria.

Le tassonomie vengono impiegate in modo massiccio nell'organizzazione dell'informazione in ambienti digitali.

---

<sup>2</sup> ANSI/NISO Z39-19:2005, *Guidelines for the construction, format, and management of monolingual controlled vocabularies*, p. 18.

## - ***Schemi di classificazione***

Costituiscono strumenti fondamentali per la ricerca in campi disciplinari circoscritti: la costruzione di banche dati specialistiche e la loro modalità di fruizione hanno posto l'esigenza di rappresentare i diversi settori conoscitivi attraverso sistemi in grado di descrivere gli elementi identificativi dei documenti, e allo stesso tempo di cogliere le strutture semantiche riconducibili ai documenti stessi, consentendone perciò sia l'archiviazione che l'accesso per il recupero delle informazioni<sup>3</sup>.

Le classificazioni semantiche si concentrano infatti su un'ideale divisione della conoscenza e del sapere in "contenitori" chiamate "classi", a cui attribuire i documenti che le materializzano, assegnando a ciascuna classe un simbolo (notazione) che serva a distinguerla e ordinarla rispetto alle altre. L'esempio più conosciuto e diffuso in ambito biblioteconomico è la Classificazione Decimale Dewey (CDD). Ideata (1873) dal bibliotecario americano Melvil Dewey – pubblicata per la prima volta nel 1876 e ormai giunta alla 23. edizione attraverso un processo continuo di revisione ed aggiornamento cooperativo internazionale – la *Decimal Dewey Classification* (DDC) costituisce il sistema di classificazione più usato al mondo per l'organizzazione sistematica della conoscenza (indipendentemente dai supporti e dalle forme in cui questa si presenta) nonché per l'accesso all'informazione nel web. La DDC è costruita su solidi principi che la rendono ideale come strumento generale di sistematizzazione: notazione espressa mediante numeri arabi, categorie ben definite, gerarchie ben sviluppate e un'ampia rete di relazioni tra i soggetti. La classificazione è ordinata per discipline: la collocazione di un soggetto all'interno di una disciplina ha carattere gerarchico e sottende ad un processo di costruzione notazionale (espresso appunto in numeri arabi) che va dal generale al particolare. Qualunque soggetto può essere trattato secondo il punto di vista di più di una disciplina; può quindi comparire in più luoghi della classificazione. La gerarchia classificatoria si estrinseca in 10 classi principali, 100 divisioni (10 per ogni classe) e 1000 sezioni (10 per ogni divisione).

Per contestualizzare il discorso classificatorio sulle materie di rilevanza istituzionale Inapp, la classe di riferimento nel sistema documentale d'Istituto è la 300 - SCIENZE SOCIALI, SOCIOLOGIA E ANTROPOLOGIA nel cui ambito risultano declinabili tutti i contenuti semantici afferenti al mondo del lavoro, alla formazione ed ai sistemi sociali ed economici connessi.

---

<sup>3</sup> Chiara Carlucci (a cura di), *I termini della formazione: il controllo terminologico come strumento per la ricerca*, Roma, Isfol, 2002 (I libri del Fondo sociale europeo; 1).

## - *Thesauri*

All'apice della complessità degli schemi di classificazione troviamo i Thesauri che hanno tutte le proprietà dei precedenti vocabolari controllati e in più altri tipi di relazione, che sono definiti al momento della costruzione del thesaurus stesso.

Possiamo definire il Thesaurus<sup>4</sup> come un vocabolario di un linguaggio di indicizzazione controllato, organizzato in maniera formalizzata in modo che le relazioni a priori tra i concetti siano rese esplicite. Tale struttura si snoda secondo un percorso ad albero gerarchico, cui si affianca il reticolo di nessi associativi.

Le caratteristiche principali dei thesauri sono:

1) Una struttura classificatoria basata su tre tipi di relazioni<sup>5</sup> esplicite e formalizzate dall'uso di simboli prescritti da standard internazionali:

relazioni di equivalenza = USE (*usa*) per rinviare dal termine non preferito o non descrittore al preferito o descrittore; UF (*used for, usato per*) per rinviare dal termine descrittore al non descrittore (casa-abitazione);

relazioni gerarchiche = BT (*broader term, termine superiore*) per rinviare al termine sovraordinato, più generale; NT (*narrower term, termine inferiore*) per rinviare al termine subordinato, più specifico (gatto-mammifero);

relazioni associative = RT (*related term, termine correlato*) per correlare due o più termini caratterizzati da una forte implicazione reciproca e/o da una frequente ricorrenza nello stesso contesto (gatto- topo);

2) organizzazione delle forme scelte e presentazione dei termini in ordine alfabetico, sistematico e grafico;

---

<sup>4</sup> [...] mezzo per il controllo terminologico usato per tradurre dal linguaggio naturale in un linguaggio sistematico in modo tale da poter riconvertire il linguaggio sistematico in naturale; [...] è un vocabolario controllato e dinamico di termini correlati semanticamente e genericamente, che copre in modo esaustivo un dominio specifico della conoscenza (Unisist, *Guidelines for the establishment and development of mono-lingual scientific and technical thesauri for information retrieval*, Parigi, Unesco, 1976).

<sup>5</sup> Per la lista tipologica completa delle relazioni rintracciabili in un Thesaurus, si veda la pagina dedicata della Biblioteca Nazionale Centrale di Firenze: <https://thes.bncf.firenze.sbn.it/signle.php#als>.

3) un insieme di criteri e procedure standard per il controllo morfologico dei termini e delle loro relazioni semantiche.

### - **Ontologie**

Le ontologie sono le concettualizzazioni, strutturate in un linguaggio (standardizzato, con una tassonomia rigida) comprensibile dalle macchine, di un dominio d'interesse, utilizzate per rendere espliciti i rapporti tra le risorse nel web. Sono schemi di riferimento, necessari per associare ai dati pubblicati nel web quel grado di struttura (e di semantica) che li rende comprensibili e usabili dalle macchine<sup>6</sup>, utilizzando linguaggi formali come il *Web ontology language* (OWL) sviluppato dal W3C a partire dal 2001. "Le ontologie hanno una struttura concettuale simile a quella dei thesauri, solo che esse possono prevedere un numero di relazioni superiori, generando così una complessa rete di connessioni fra concetti, che può essere visualizzata anche graficamente. Inoltre, loro caratteristica specifica è la capacità di esprimere i concetti in maniera non ambigua e quindi con un alto livello di precisione semantica" (Crupi, 2013).

I tre aggettivi fondamentali che descrivono le caratteristiche dell'idea di ontologia sono: **formale, esplicita e condivisa**.

- **Formale**= le informazioni devono essere descritte formalmente, cioè con regole ben precise che non siano ambigue;
- **Esplicite**= le informazioni devono essere indicate senza darle per sottintese;
- **Condivisa**= vanno incluse solo le informazioni ritenute corrette da coloro che sono interessati ad utilizzarle.

Grazie al ragionamento, che prende il nome di **interferenza**, sarà possibile, per i sistemi automatici, scoprire nuove informazioni ampliando quelle già note: ciò avviene attraverso l'utilizzo dei *Reasoner* (componenti che appunto servono per "ragionare") sia per scoprire nuova conoscenza che per identificare errori nei dati stessi. Lo standard utilizzato per la rappresentazione di questi dati è quello dei linguaggi RDF. Tali linguaggi, tutti ormai standard del W3C sono:

- RDF (Resource Description Framework)

---

<sup>6</sup> Introduzione a RDA : linee guida per rappresentare e scoprire le risorse / Carlo Bianchini, Mauro Guerrini ; prefazione di Barbara B. Tillett ; postfazione di Gordon Dunsire / Mauro Guerrini; Carlo Bianchini. - STAMPA. - (2014).

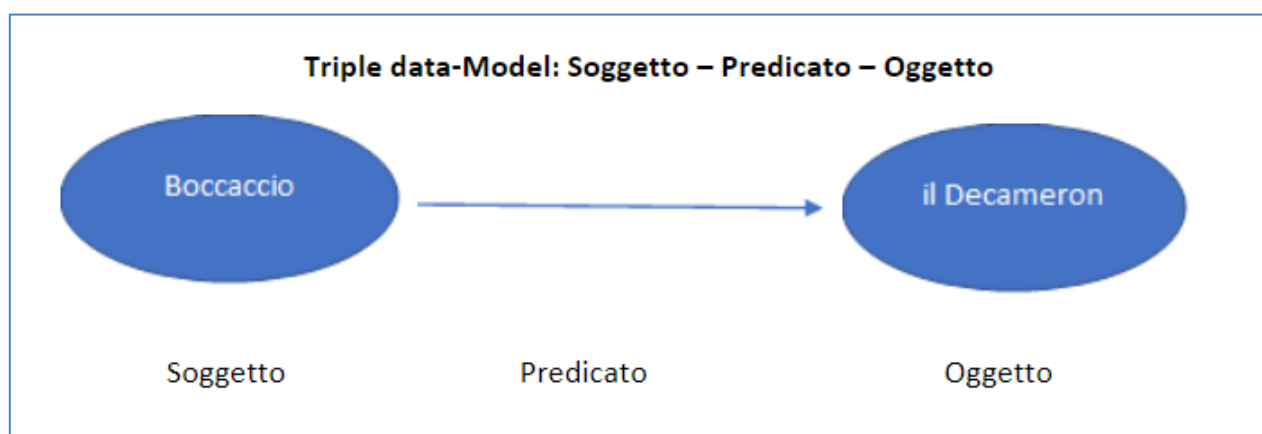
- RDFS o RDF (Resource Description Framework Schema)
- OWL (Web Ontology Language)
- SKOS/SKOS-XL (Simple Knowledge Organization System/SKOS eXtension for Labels<sup>7</sup>)

### **Elementi del linguaggio RDF**

L'elemento fondamentale in RDF è la **tripla**. Questo significa che ogni cosa che verrà rappresentata in essi dovrà essere un insieme di triple. Ogni tripla è formata da tre elementi:

- **Soggetto**: è sempre una risorsa identificata in modo univoco (URI). Ad esempio, può essere una persona, una cosa, un concetto astratto, ecc.
- **Predicato** (detto anche Relazione): consente di creare una relazione tra il soggetto e un oggetto. Un medesimo soggetto può avere n diverse relazioni con altrettanti oggetti.
- **Oggetto**: può essere sia una risorsa identificabile in modo univoco (URI) oppure un dato

Più triple formano un **grafo**, i cui **nodi** sono formati dai soggetti e dagli oggetti, mentre gli **archi**, con verso, sono i predicati (ed il verso va dal soggetto all'oggetto).



*Figura 2. RDF*

<sup>7</sup> In realtà questo linguaggio è utilizzato per rappresentare i thesauri e non le ontologie, ma è comunque basato sulle strutture degli altri, essendo a sua volta un'estensione di RDF.

### 3. ONTOLOGIA DI SOPHIA: VERSO UN'ONTOLOGIA SUPPORTATA DELL'INTELLIGENZA ARTIFICIALE

Sophia è un progetto *Cross-domain* tra gli ambiti del mercato del lavoro e quello dell'istruzione e formazione. Si è quindi reso necessario un lavoro di integrazione di diversi sistemi afferenti a entrambi i domini. Il lavoro di analisi è consistito nell'individuazione di tali classificazioni, tassonomie e sistemi utili al progetto per individuare e gestire le informazioni relative allo *Skill mismatch*.

L'ontologia di Sophia integra diverse classificazioni, termini ed entità affini alle attività e ai contenuti del progetto e le relative tabelle di corrispondenza. Serve per due ragioni fondamentali:

- per il fine di ordinare il database delle fonti dati avendo a disposizione risorse per un raccordo tra le varie classificazioni;
- per avere una fonte dati utile al training dello strumento cognitivo utilizzato.

#### 3.1. Modello cognitivo Watson

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 del progetto. I servizi utilizzati sono i seguenti:

- 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;
- Watson Knowledge Studio (WKS) crea 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): tramite uno o più modelli cognitivi costruiti con WKS, analizza il testo ed estrae in tempo reale 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 ecc.) e fornirne il lemma;
- Watson Language Translator traduce da una lingua all'altra il testo che gli viene sottomesso. Nel caso del progetto Sophia, traduce in italiano i testi delle offerte di lavoro o delle offerte formative scritti in altre lingue.

### 3.2. Architettura della soluzione

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, richiama un modulo su IBM Cloud (Orchestrator) per l'analisi del tipo di richiesta e, in base alla tipologia dell'utente, visualizza 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 mercato del lavoro.

### 3.3. IBM Watson Knowledge Studio

Al fine di estrarre le informazioni rilevanti dai dati non strutturati (sia inclusi nell'ontologia, che nella base dati su offerta formativa e professionale), Watson Knowledge Studio (WKS) costruisce un modello cognitivo Machine learning in lingua italiana.

WKS è un'applicazione basata sul cloud che consente a sviluppatori ed esperti del dominio di insegnare a Watson il linguaggio del loro settore. Gli utenti formano Watson mediante esempi, annotando dei documenti campione con le informazioni pertinenti al loro dominio.

Watson Knowledge Studio è poi in grado di formare un modello di machine learning che può analizzare documenti nuovi, non visti in precedenza, consentendo così un'adozione in scala della conoscenza degli esperti del dominio.

Per la comprensione del contenuto dei testi non strutturati viene sviluppato un modello cognitivo che utilizza un approccio statistico per trovare entità e relazioni di interesse nei testi.

Un'**entità** corrisponde alla categorizzazione di una cosa reale, nel caso in esame si tratta di entità per la classificazione di informazioni incluse nel testo, come il tipo di corso di laurea effettuato o le Skill apprese.

Una **relazione**, invece, definisce un binario, la relazione ordinata tra due entità. Perché sia presente una citazione di relazione, il testo deve definire esplicitamente la relazione e associare le citazioni delle due entità tra loro all'interno di una frase. Le relazioni vengono utilizzate per specificare meglio la natura di una determinata entità.



In generale, il processo per creare un modello di machine learning con WKS è il seguente:

- si definiscono i tipi di entità e di relazione per l'estrazione delle informazioni di interesse;
- si annotano documenti di esempio, etichettando le parole che rappresentano i tipi di entità, per identificare i tipi di relazione e per definire le coreferenze, che identificano le citazioni differenti che fanno riferimento alla stessa cosa, ovvero, la stessa entità;
- WKS prepara un modello sulla base dei documenti annotati, che può venire utilizzato per trovare le entità, le relazioni e le coreferenze nei nuovi e mai visti prima documenti.

È possibile analizzare le prestazioni del modello tramite le seguenti metriche (minimo valore: 0 – massimo valore: 1), fornite da WKS:

- **F1**, può essere interpretato come una media ponderata dei valori di precision e recall. Un punteggio basso indica che il modello genera annotazioni errate e non riesce a trovare le annotazioni che dovrebbero venire trovate;
- **Precision**, frazione dell'output del modello di Machine learning accurata. La *Precision* viene determinata dal numero di annotazioni etichettate correttamente diviso il numero totale di annotazioni aggiunte dal modello. Un punteggio di *Precision* di 1.0 per il tipo di entità A significa che tutte le citazioni che sono state etichettate come tipo di entità A appartengono veramente a tale classificazione.

Un punteggio di precisione basso indica che il modello di Machine learning ha generato annotazioni non corrette;

- **Recall**, quante citazioni che dovrebbero essere state annotate da un'etichetta selezionata sono state effettivamente annotate con tale etichetta. La *Recall* viene determinata dal numero di annotazioni etichettate correttamente diviso il numero di annotazioni che dovrebbero essere state create. Un punteggio di *Recall* di 1.0 significa che tutte le citazioni che dovrebbero essere state etichettate come tipo di entità A lo sono state correttamente.

Un punteggio di *Recall* basso indica che il modello non è riuscito a creare le annotazioni che dovevano venire create.

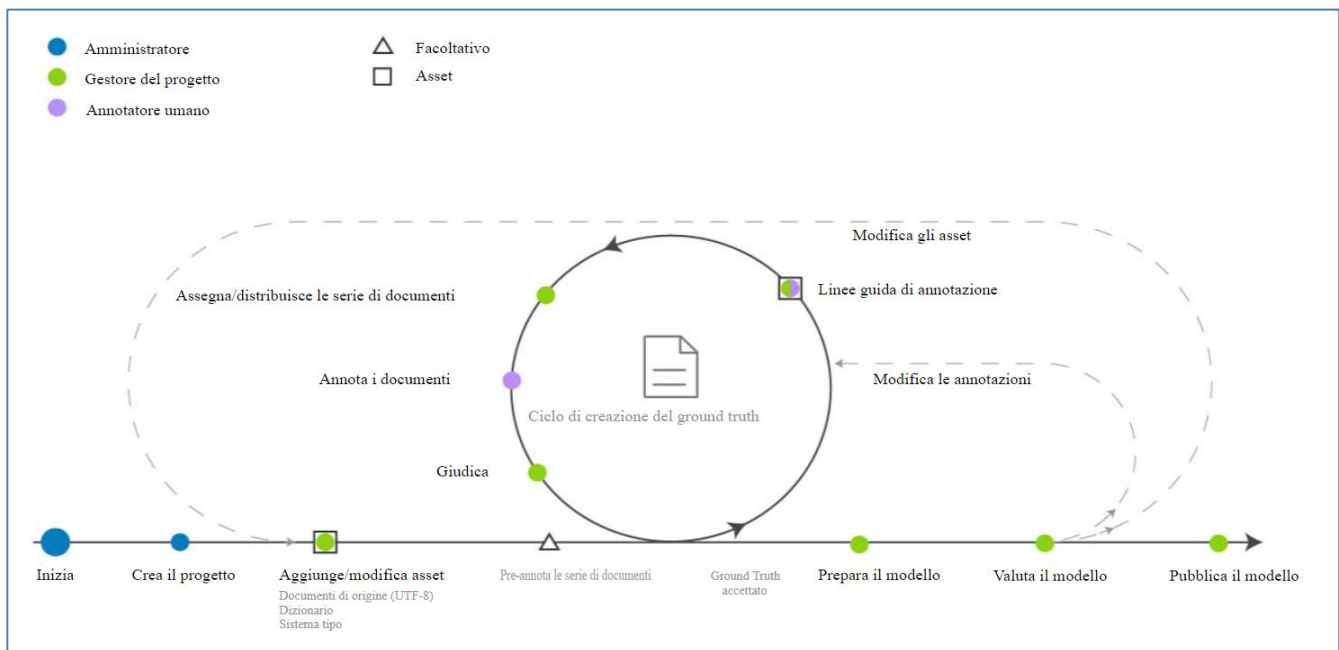


Figura 3. Il flusso di lavoro per lo sviluppo di un modello di machine learning con WKS. Fonte: IBM

### 3.4. IBM Watson Natural Language Understanding

WKS, abbinato a Watson Natural Language Understanding (NLU), permette di elaborare il linguaggio naturale per l'analisi avanzata del testo e la sua classificazione, estrarre *Insight* personalizzati utilizzando il modello Machine learning di WKS e comprendere quali informazioni rilevanti sottoposte al modello.

All'arrivo, ad esempio, di un CV dell'utente utente, l'orchestratore estrae dal documento il testo e lo invia a NLU tramite chiamata API (per più informazioni sulle API utilizzate per l'integrazione Watson NLU ed Engine, si rimanda alla documentazione ufficiale: <https://cloud.ibm.com/apidocs/natural-language-understanding>).

In generale, il modello Machine learning permette di interpretare il linguaggio naturale tramite entità, ovvero categorie di concetti, e relazioni, collegamenti ordinati tra due entità.

#### 4. PRINCIPALI CLASSIFICAZIONI, TASSONOMIE E SISTEMI CONFLUITI NELL'ONTOLOGIA INIZIALE DI SOPHIA

Il progetto Sophia rappresenta un LMI *cross-domain* e come tale ha bisogno di un'ontologia trasversale ai domini presi in considerazione. È possibile definire la metodologia sperimentale del progetto come segue:

*"Dati due o più silos di documenti afferenti domini diversi e accomunati da un'ontologia trasversale è possibile rilevare il grado di affinità di un documento di un dominio con tutti i documenti degli altri domini ed esprimere tale grado di affinità mediante concetti dell'ontologia adottata."*

Nel caso specifico di Sophia i domini presi in esame sono il mercato del lavoro da una parte e l'istruzione e formazione dall'altra. Ne consegue che:

*"Dato un silos di documenti del dominio mercato del lavoro (es. web job vacancies), un altro silos di documenti del dominio istruzione e formazione (es. schede SUA-CdS) accomunati da un'ontologia contenente le tassonomie CP, ATECO, ISCED, ESCO ecc., è possibile rilevare le corrispondenze di un documento appartenente ad un dominio con tutti i documenti appartenenti l'altro dominio, esprimere tali corrispondenze in gradi di vicinanza e classificarli mediante concetti dell'ontologia adottata".*

Lo sviluppo futuro di Sophia prevede una copertura e distribuzione transnazionale. È stato quindi necessario includere nell'ontologia iniziale un ampio numero delle fonti internazionali ed europee, le loro declinazioni e curvature nazionali e regionali, progetti e strumenti interni INAPP, tutti utili a creare una rete neurale artificiale complessa – l'ontologia del progetto.

	Istruzione e formazione	Professioni e occupazione
ATECO 2007	20	90
<b>Atlante Professioni</b>	20	80
Classi di Laurea	90	10
CP2011	30	70
EQF	10	90
ESCO	30	70
ISCED	95	20
ISCO-08	15	95
NACE	5	95
<b>Repertorio Nazionale delle Qualificazioni</b>	80	20
<b>Sistema informativo sulle Professioni</b>	20	80

N.B: Il grafico rappresenta la copertura delle varie fonti usate per la creazione dell'ontologia rispetto ai due domini: 1. Istruzione e formazione, 2. Mercato del lavoro. I valori riportati nel grafico sono indicativi e hanno un carattere illustrativo per la rappresentazione della metodologia adottata. I valori realisaranno desunti dall'occorrenza dei termini in relazione alle fonti prese in considerazione.

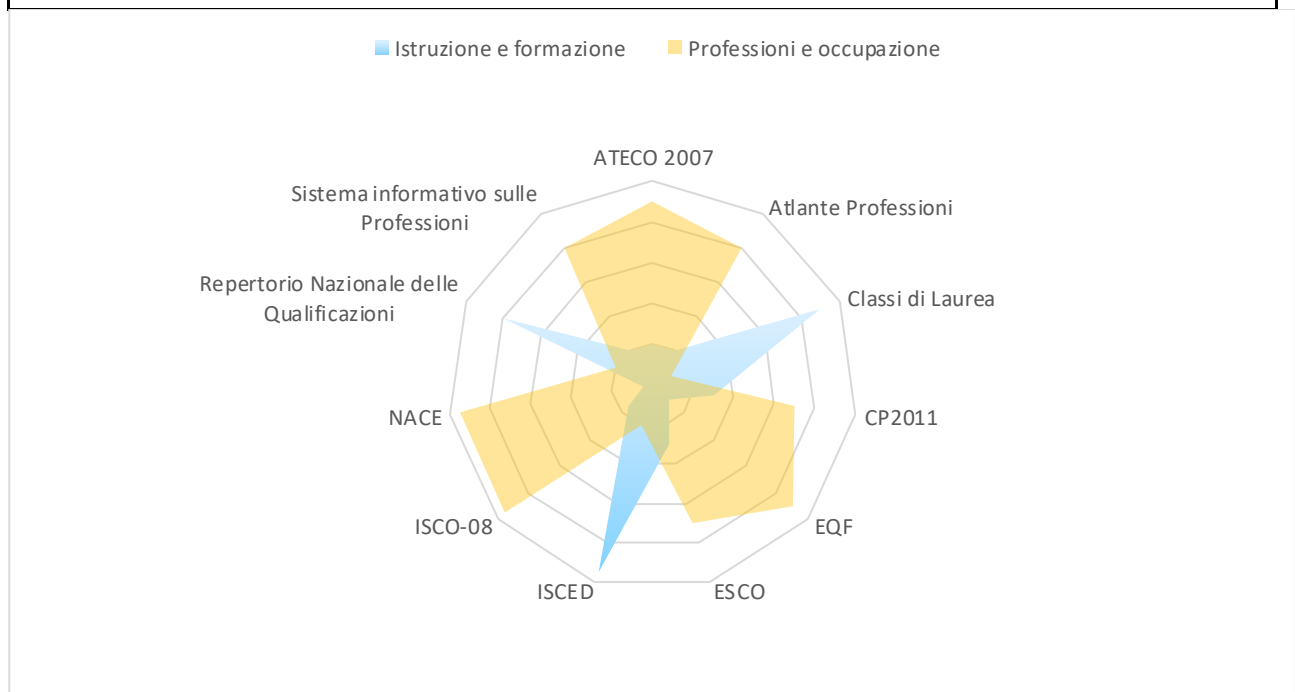


Figura 4. Ontologia del progetto: le fonti

#### 4.1. ISCO-08

La classificazione ISCO-08, fornita dall'International Labour Organization (ILO), è strutturata in un documento che alla prima colonna presenta il codice ISCO 2008, alla seconda il codice della precedente classificazione e all'ultima il nome della professione in inglese (come riportato nell'immagine che segue).

Nell'esempio che segue:

- la prima colonna riporta il codice della classificazione aggiornata al 2008;
  - la seconda il codice della vecchia classificazione 1988, il cui fine si esaurisce nella mera conversione con i codici più aggiornati;
  - la terza colonna, invece, fa riferimento al nome della professione indicata dai codici visti prima.
- La professione "Abess" (alla prima riga) avrà quindi associati il codice 2636, secondo la classificazione ISCO-08, e il codice 2460, secondo la classificazione ISCO-88:

ISCO-08	ISCO-88	English title
2636	2460	Abess
2636	2460	Abbot
1345	1229	Academic, university: head of department or faculty
2310	2310	Academic, university: lecturer
3433	3439	Accessioner, library
2652	2453	Accompanist
2411	2411	Accountant
2411	2411	Accountant, certified
2411	2411	Accountant, chartered
2411	2411	Accountant, management
2411	2411	Accountant, tax
3117	3117	Acidiser, oil and gas well
2659	3474	Acrobat
2655	2455	Actor
2120	2121	Actuary
2230	3241	Acupuncturist
3315	3417	Adjuster, claims
7311	7311	Adjuster, precision instrument

Figura 5. ISCO-08



### 4.3. 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; L'oggetto della classificazione, la professione, è definito come un insieme di attività lavorative concretamente svolte da un individuo, che richiamano conoscenze, competenze, identità e statuti propri.

La logica utilizzata per aggregare professioni diverse all'interno di un raggruppamento si basa sul concetto di competenza, visto nella sua duplice dimensione del livello e del campo delle competenze richieste per l'esercizio della professione. Il livello di competenza è definito in funzione della complessità, dell'estensione dei compiti svolti, del livello di responsabilità e di autonomia decisionale che caratterizza la professione; il campo di competenza coglie, invece, le differenze nei domini settoriali, negli ambiti disciplinari delle conoscenze applicate, nelle attrezzature utilizzate, nei materiali lavorati, nel tipo di bene prodotto o servizio erogato nell'ambito della professione. Il criterio della competenza delinea un sistema classificatorio articolato su 5 livelli di aggregazione: **grandi gruppi professionali, gruppi professionali, classi professionali, categorie e infine le unità professionali.**

Codice professione	Nome professione	descr_5
cod_5	nome_5	descr_5
1.1.1.1.0	Membri di organismi di governo e di assemblee nazionali con potestà legislativa e regolamentare	Le professioni classificate in questa unità definiscono, discutono e approvano leggi, regolamenti e statuti validi sull'intero territorio nazionale, definiscono e implementano le politiche di governo individuando e allocando le risorse necessarie alla loro realizzazione.
1.1.1.2.0	Membri di organismi di governo e di assemblee regionali e di Province autonome con potestà legislativa e regolamentare	Le professioni classificate in questa unità applicano sul territorio leggi nazionali e, negli ambiti previsti dall'organizzazione dello Stato, definiscono, discutono e approvano leggi, regolamenti e statuti validi sul territorio regionale e delle Province autonome, definiscono e implementano le politiche di governo locale individuando e allocando le risorse necessarie alla loro realizzazione.

Figura 7. CP2011

#### 4.4. 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. Offre informazioni in merito a:

- rappresentazioni dei contenuti professionali
- aspetti evolutivi e di fabbisogno dei contenuti professionali nel breve e nel medio termine
- previsioni di occupazione a medio termine a livello nazionale e regionale
- previsioni sugli andamenti dell'economia e dell'occupazione settoriale nei prossimi anni.

Il progetto fa parte del Sistema informativo integrato sulle professioni che coinvolge diversi soggetti pubblici e non che, per loro finalità e a vario titolo, producono ed erogano nel nostro Paese informazioni inerenti le professioni, il lavoro, l'occupazione e la formazione.

Collegate al set informativo su professioni, fabbisogni professionali e previsioni di occupazione ci sono i dati forniti da **Istat**, **Unioncamere**, **Inail** e **Regioni**.

Il progetto, inoltre, estende la classificazione CP2011 fornendo un ulteriore livello di approfondimento – “voce professionale”:

Codice Grande Gruppo Prof.le	Descrizione Grande Gruppo Prof.le	Codice Gruppo Prof.le	Descrizione Gruppo Prof.le	Codice Classe Prof.le	Descrizione Classe Prof.le	Codice Categoria Prof.le	Descrizione Categoria Prof.le	Codice Unità Prof.le	Descrizione Unità Prof.le	Codice Voce Prof.le	Descrizione Voce Prof.le
1	LEGISLATORI, IMPRENDITORI E ALTA DIRIGENZA	1.1	Membri dei corpi legislativi e di governo, dirigenti ed equiparati dell'amministrazione pubblica, nella magistratura, nei servizi di sanità, istruzione e ricerca e nelle organizzazioni di interesse nazionale e sovranazionale	1.1.1	Membri di organismi di governo e di assemblee con potestà legislativa e regolamentare	1.1.1.1	Membri di organismi di governo e di assemblee con potestà legislativa e regolamentare a livello nazionale	1.1.1.1.0	Membri di organismi di governo e di assemblee nazionali con potestà legislativa e regolamentare	1.1.1.1.0.1	membro del consiglio dei ministri
										1.1.1.1.0.2	membro del senato
										1.1.1.1.0.3	membro della camera dei deputati
										1.1.1.2.0.1	membro del consiglio regionale
										1.1.1.2.0.2	membro della giunta di Provincia autonoma
										1.1.1.2.0.3	membro della giunta regionale
1.1.1.2.0.4	membro di consiglio di Provincia autonoma										
						1.1.1.2	Membri di organismi di governo e di assemblee con potestà legislativa e regolamentare a livello regionale e di Province autonome	1.1.1.2.0	Membri di organismi di governo e di assemblee regionali e di Province autonome con potestà legislativa e regolamentare	1.1.1.2.0.5	presidente della giunta regionale
										1.1.1.2.0.6	presidente di Provincia autonoma

Figura 8. SIP: voci professionali della CP2011



## 4.5. 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. L'Atlante è ulteriormente articolato in due sottosezioni: Atlante e Qualificazioni e Atlante e Professioni.

### ADA.01.01.01 (ex ADA.1.236.772) - Progettazione dell'impianto di coltivazione in pieno campo e in serra

**ATTIVITÀ**

Valutazione dell'idoneità ambientale per la coltivazione (es. caratteristiche pedo-climatiche, morfologia del terreno, disponibilità di acqua irrigua, ecc.)

Valutazione dei fabbisogni infra-strutturali (es. macchine agricole, impianti di essiccazione, magazzini stoccaggio, ecc.)

Analisi dei costi e valutazione della convenienza economica del processo produttivo nel mercato di riferimento

**RISULTATI ATTESI**

RA1: Elaborare lo studio di fattibilità dell'impianto di coltivazione, valutando la sostenibilità economica dell'investimento e identificando i vincoli di natura ambientale e le esigenze tecnico-strutturali

**Codici ISTAT CP2011 associati all'ADA**

Codice	Titolo
2.3.1.1.5	Botanici
2.3.1.3.0	Agronomi e forestali

**Codici ISTAT ATECO associati alla sequenza di processo**

Codice Ateco	Titolo Ateco
01.11.30	Coltivazione di legumi da granella
01.11.20	Coltivazione di semi oleosi

Figura 9. Atlante Lavoro

L'Atlante e Qualificazioni contiene il Repertorio Nazionale dei titoli di istruzione e formazione e delle Qualificazioni professionali. Il Repertorio è costituito da tutti i repertori dei titoli di istruzione e formazione, e delle Qualificazioni professionali rilasciati in Italia da un Ente titolare o rilasciati in esito ad un contratto di Apprendistato.

### Tecnico superiore responsabile delle produzioni e delle trasformazioni agrarie, agro-alimentari e agro-industriali

- 📍 SETTORE 01. [Agricoltura, silvicoltura e pesca](#)
- 📍 REPERTORIO - Istruzione Tecnica Superiore - ITS
- 📍 EQF - 5
- 📍 Visualizza scheda descrittiva

AdA associate alla Qualificazione 📄

📌 ADA.02.01.05 (ex ADA.2.136.410) - Gestione della qualità dei processi e prodotti alimentari

Competenze
🎓

**Titolo:** Gestire i processi di produzione e trasformazione nell'ambito di specializzazioni e peculiarità del 'Made in Italy'

**Titolo:** Applicare sistemi di controllo su materiali, processi e prodotti per il miglioramento della qualità

📌 Codici ISTAT CP2011 associati

Codice	Titolo
3.2.2.3.2	Tecnici dei prodotti alimentari
3.1.5.4.2	Tecnici della produzione alimentare

📌 Codici ISTAT ATECO associati

Codice Ateco	Titolo Ateco
10.11.00	Produzione di carne non di volatili e di prodotti della macellazione (attività dei mattatoi)
10.12.00	Produzione di carne di volatili e prodotti della loro macellazione (attività dei mattatoi)

Figura 10. Atlante e Qualificazioni

L'Atlante e Professioni è la sezione che descrive il mondo delle professioni e si articola in Repertorio delle professioni dell'Apprendistato, Professioni regolamentate, Qualificazioni regionali abilitanti e Professioni non organizzate in ordini e collegi.

## 4.6 ATECO

La classificazione ATECO, declinazione ISTAT del NACE europeo, che a sua volta deriva dalla classificazione ISIC dell'ONU, è aggiornata al 2021. La classificazione si articola in sei livelli, ogni livello è caratterizzato da codice, titolo e descrizione:

```
<gruppo>
  <codice>3</codice>
  <titolo>LAVORAZIONE E CONSERVAZIONE DI FRUTTA E ORTAGGI</titolo>
  <descrizione></descrizione>
  <classe>
    <codice>1</codice>
    <titolo>Lavorazione e conservazione delle patate</titolo>
    <descrizione></descrizione>
    <categoria>
      <codice>0</codice>
      <titolo>Lavorazione e conservazione delle patate</titolo>
      <descrizione></descrizione>
      <sottocategoria>
        <codice>0</codice>
        <titolo>Lavorazione e conservazione delle patate</titolo>
        <descrizione>- produzione di patate surgelate preparate
- produzione di purè di patate disidratato
- produzione di snack a base di patate
- produzione di patatine fritte
- produzione di farina e fecola di patate
- sbucciatura industriale delle patate</descrizione>
```

Figura 11. ATECO

#### 4.7. NACE

La classificazione NACE dell'Eurostat è disponibile sia in italiano che in inglese. È utile sottolineare che in Italia, l'ISTAT traduce i codici NACE, successivamente alla Rev. 2 del 2007, nei codici ATECO.

La classificazione si divide in quattro livelli, come osservabile nell'immagine di seguito riportata:

A	Agriculture, forestry and fishing
A1	Crop and animal production, hunting and related service activities
A11	Growing of non-perennial crops
A111	Growing of cereals (except rice), leguminous crops and oil seeds
A112	Growing of rice
A113	Growing of vegetables and melons, roots and tubers
A114	Growing of sugar cane
A115	Growing of tobacco
A116	Growing of fibre crops
A119	Growing of other non-perennial crops
A12	Growing of perennial crops
A121	Growing of grapes
A122	Growing of tropical and subtropical fruits
A123	Growing of citrus fruits
A124	Growing of pome fruits and stone fruits
A125	Growing of other tree and bush fruits and nuts
A126	Growing of oleaginous fruits
A127	Growing of beverage crops
A128	Growing of spices, aromatic, drug and pharmaceutical crops
A129	Growing of other perennial crops
A13	Plant propagation
A130	Plant propagation

*Figura 12. NACE*

#### 4.8. ISCED e le classi di laurea

Per le classi ISCED, fornite da UNESCO, vi è a disposizione una tabella di corrispondenza con associati ad ogni corso di formazione in Italia sia il codice della classe (che è il tipo di classificazione che si utilizza in Italia) che il codice ISCED associato.

Classificazione delle Classi di Laurea secondo la FOET_2013					
Numero	Denominazione	Ordinamento	ISCED_F_1dgt	ISCED_F_2dgt	ISCED_F_3dgt
-	Corsi del vecchio ordinamento	pre-509	-	-	-
01	Biotecnologie	D.M. 509/99	05	051	0512
01/S	Antropologia culturale ed etnologia	D.M. 509/99	03	031	0314
10	Ingegneria industriale	D.M. 509/99	07	071	0719
10/S	Conservazione dei beni architettonici e ambientali	D.M. 509/99	07	078	0731
100/S	Tecniche e metodi per la società dell'informazione	D.M. 509/99	06	068	0688
101/S	Teoria della comunicazione	D.M. 509/99	03	032	0321
102/S	Teoria e tecniche della normazione e dell'informazione giuridica	D.M. 509/99	04	042	0421
103/S	Teorie e metodi del disegno industriale	D.M. 509/99	02	021	0212
104/S	Traduzione letteraria e in traduzione tecnico-scientifica	D.M. 509/99	02	023	0231
11	Lingue e culture moderne	D.M. 509/99	02	023	0231
11/S	Conservazione dei beni scientifici e della civiltà industriale	D.M. 509/99	05	053	0588

#### **0314 Sociology and cultural studies**

**Sociology and cultural studies** is the study of human beings and the way they behave in groups and in relation to society. The study of ethnology and social anthropology are included here, likewise the study of human and social geography.

Programmes and qualifications with the following main content are classified here:

- Criminology
- Cultural geography
- Cultural studies
- Demography/population studies

Figura 13. ISCED

## 5. CONCLUSIONI

---

La velocità di cambiamento del mercato del lavoro non è l'unico fattore determinante dello *skill mismatch*. Anche la frammentazione di sistemi territoriali e la bassa intensità di dialogo tra il mondo del lavoro e quello dell'istruzione e formazione contribuiscono al fenomeno impattando in maniera altrettanto forte.

Per dare un impulso al cambiamento in grado di ridurre tale divario è necessario cambiare i paradigmi attualmente in vigore e sfruttare appieno tutti i presupposti oggi disponibili (dati, classificazioni, tecnologie ecc.) e arrivare a un modello *Cross-domain*, aperto, sinergico e interoperabile condiviso da tutti gli attori in gioco.

Nel corso degli ultimi decenni siamo stati testimoni di uno sviluppo straordinario delle tecnologie informatiche. La crescita esponenziale della potenza di calcolo dei computer di ultima generazione, la disponibilità della banda larga, VoIP, 4G, l'avvento dello smartphone e delle app, la nascita di *Cloud Computing*, *Linked Open Data*, *Big Data* e *Internet of Things* hanno stravolto il paradigma della società dell'informazione. Da un uso prevalente delle tecnologie e attrezzature informatiche per la videoscrittura, fogli di calcolo e stampa dei documenti la società digitale si è evoluta verso il networking e verso l'uso quotidiano delle piattaforme che implementano l'intelligenza artificiale (AI)<sup>8</sup>, le reti neurali artificiali (ANNs)<sup>9</sup>, l'apprendimento automatico o *Machine Learning* (ML)<sup>10</sup>, *Deep Learning* (DL)<sup>11</sup>, l'elaborazione del linguaggio naturale o *Natural Language Processing* (NLP)<sup>12</sup> e *Labour Market Intelligence* (LMI)<sup>13</sup> impossibili da immaginare pochi anni fa.

Se qualche anno addietro i dati e le tecnologie a nostra disposizione non permettevano la costruzione del sistema proposto nei punti precedenti, lo scenario di oggi è radicalmente cambiato. Pubblicando i dati in formato aperto (Open Data), usando le API, implementando le ontologie comuni, usando le tecnologie di ultima generazione di NLP, ML e AI – il risultato è raggiungibile, un sistema pubblico di Real-time LMI incentrato sul cittadino e Multistakeholder oggi è possibile.

---

<sup>8</sup> Abilità di un computer di ragionare e svolgere funzioni simulando il comportamento umano.

<sup>9</sup> Modello computazionale composto di "neuroni" artificiali, ispirato a una rete neurale biologica.

<sup>10</sup> Sottoinsieme dell'intelligenza artificiale che ha come scopo la creazione di algoritmi in grado di apprendere o migliorare le proprie performance con l'esperienza.

<sup>11</sup> Sottoinsieme dell'apprendimento automatico basato sull'apprendimento dai dati derivati dalle reti neurali artificiali.

<sup>12</sup> Implementazione dell'intelligenza artificiale nel campo della processazione automatica delle informazioni scritte o parlate in una lingua naturale.

<sup>13</sup> Refers to the use and design of AI algorithms and frameworks to analyze Labour Market Data for supporting decision making (Boselli et al. ed. orig.).

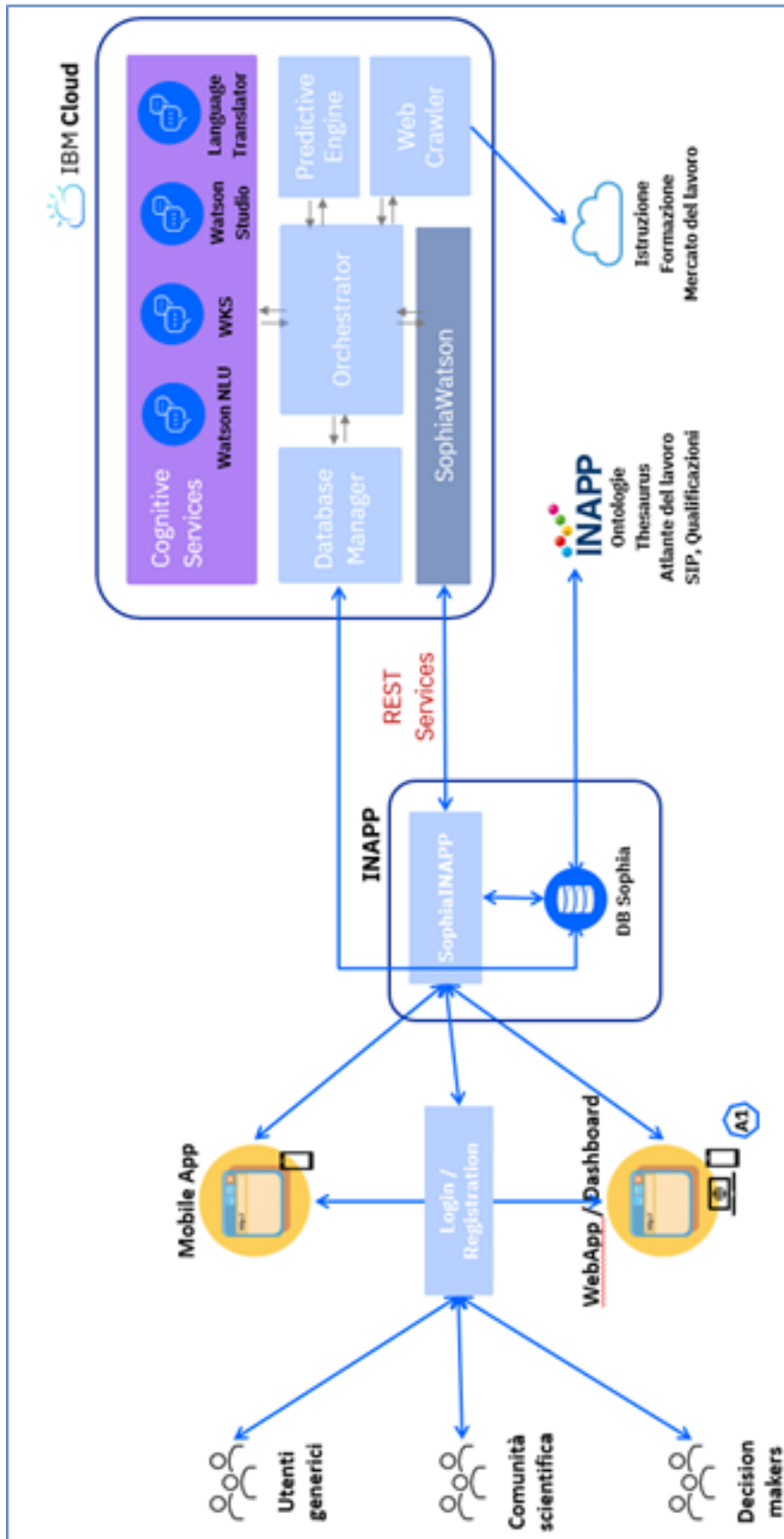


Figura 14. Architettura di SOPHIA

## BIBLIOGRAFIA

Boston Consulting Group (2020), *Fixing the Global Skill Mismatch* <https://on.bcg.com/3aFlzPz>

Boselli R., Cesarini M., Mercorio F., Mezzanzanica M. (2017), *Using Machine Learning for Labour Market Intelligence*, paperID391, Milano, Dip. of Statistics and Quantitative Methods, Univ. of Milano-Bicocca, CRISP Research Centre, <https://bit.ly/3cPP05p>

Cedefop (2018), *Real-time labour market information on skill requirements: setting up the EU system for online vacancy analysis. Global report* <https://bit.ly/3rww67O>

Chiriatti M. (2021), *Incoscienza artificiale: come fanno le machine a prevedere per noi*, Roma, Luiss University Press

Kaur J., Gill N. S. (2020), *Artificial Intelligence and Deep Learning for Decision Makers*, New Delhi, BPB Publications

*La biblioteca piattaforma della conoscenza: collaborative, inclusiva, reticolare. Convegno delle Stelline 2021*, Milano, Editrice bibliografica

Marmo R. (2020), *Algoritmi per l'intelligenza artificiale: progettazione dell'algoritmo Dati e Machine Learning Neural Network-Deep Learning*, Milano, Hoepli

Mezzanzanica M., Mercorio F. (2019), *Big data for Labour Market Intelligence: an introductory guide*, Torino, European Training Foundation <https://bit.ly/3pZq1Ar>