

## Mod. E

### RELAZIONE FINALE DELLE ATTIVITA'

#### Obiettivi raggiunti

Descrivere gli obiettivi del progetto raggiunti.

Compilare la parte sottostante non superando i 500 caratteri per obiettivo

- 1) Creazione di un vocabolario di termini di riferimento utilizzato nell'ambito della compilazione della SDO, allineato alla codifica ICD-10.
- 2) Definizione di regole specifiche per la codifica delle diagnosi, procedure ed interventi nelle SDO.
- 3) Formalizzazione, implementazione e valutazione delle nuove regole di classificazione e codifica.
- 4) Progettazione e implementazione del Web Service e dell'interfaccia web per l'interrogazione e l'accesso al SISCO (versione alfa).

#### Risultati raggiunti, indicatori e fonti di verifica

Obiettivo Specifico	Risultato Raggiunto	Indicatore di valutazione	Fonte di Verifica
1. Validazione e Test del sistema di supporto alla codifica sviluppato per il PSS, nel contesto regionale del <i>Friuli Venezia Giulia</i> .	In relazione all'adozione del Fascicolo Sanitario Elettronico in FVG a fine dicembre 2017 e alla conseguente tardiva implementazione del Profilo Sanitario Sintetico da parte dei MMG della regione, l'attività inizialmente spostata nel secondo anno è tuttavia risultata non effettuabile per indisponibilità di PSS da analizzare.	Non è stato possibile calcolare l'indicatore definito nel progetto (regole implementate sono valide almeno al 70%).	Di conseguenza il documento di analisi della performance del prototipo completo non è stato redatto.
2. Estrazione dei termini da un campione di cartelle cliniche per verifica del linguaggio usato nella descrizione delle diagnosi,	È stata analizzata la lista di 425 condizioni patologiche frequenti in Pronto Soccorso (PS), proposta dal GdL regionale PS FVG. I termini tecnico-	Numeri dei termini estratti: <ul style="list-style-type: none"><li>• 425 condizioni patologiche (1359 parole) nella prima analisi</li><li>• 696 diagnosi (2530 parole) nella seconda analisi</li></ul>	Il database considerato è stato il database SEI con i relativi codici.

procedure ed interventi delle SDO.

scientifici sono stati analizzati, integrati nel vocabolario di cui all'Obiettivo 3 e ne è stata verificata la codificabilità supportata. È stata verificata e corretta la codifica in ICD-9-CM fatta dai medici di PS, a partire dallo scarico del database SEI, analizzando 696 diagnosi.

3. Predisposizione di un vocabolario di termini per facilitare il passaggio a un linguaggio di codifica.

I vocabolari che alimentano le risorse dati di SISCO.web sono stati definiti e digitalizzati. Le principali risorse terminologiche utilizzate sono: l'ontologia dell'Indice Sistemático ICD-9-CM, integrata con l'Indice Alfabetico ICD-9-CM in formato processabile; un vocabolario di termini medici comuni di diagnosi, estratti da cartelle cliniche di MMG; vocabolario iniziale di 696 termini di diagnosi estratti dal database SEI durante la fase 2 (termini estratti dalle cartelle di pronto soccorso); sinonimi e acronimi; termini MeSH mappati a ICD-9-CM;

Rispetto al numero di termini presenti nell'Elenco Sistemático ICD-9-CM e nell'Indice Alfabetico ICD-9-CM, il vocabolario è stato arricchito con n. 1883 termini nuovi.

Vocabolario arricchito dei termini di diagnosi e procedure estratti dalle varie fonti integrate nella knowledge base del servizio.

	termini di Malattie rare; neoplasie ICD-9-CM; prefissi e suffissi.		
4. Allineamento dei termini del vocabolario ai termini ICD-10.	A partire dal file word aggiornato di ICD-10, messo a punto dal Centro collaboratore OMS per la Famiglia delle classificazioni internazionali, è stata realizzata la bozza di Indice Alfabetico in formato computabile CSV/Excel/JSON, da utilizzare per supportare la transcodifica in ICD-10.	Il numero di corrispondenze con mapping esatto è 4035.	Il documento strutturato contenente termini guida, parole chiave, codici associati, e codici correlati è parzialmente definito.
5. Definizione di nuove regole di classificazione e codifica delle diagnosi, procedure ed interventi per un sottoinsieme di specialità laddove non sufficienti quelle precedentemente definite per il contesto del PSS.	L'analisi del problema di codifica ha portato all'identificazione di due famiglie di regole da implementare nel sistema: i) regole specifiche dell'aspetto concettuale della codifica delle condizioni di salute; ii) regole per l'estrazione di un set di termini da codificare a partire dal testo, anche parziale, sottoposto dal medico al sistema. Sono stati per ora impostati alcuni algoritmi decisionali conformi a regole OMS, necessari per definire il numero di regole di primo tipo (cfr punto 6).	Sono state definite due famiglie di regole: i) e ii).	Documenti di specifica astratta delle regole di codifica, con descrizione delle regole: a) Albero decisionale per l'individuazione della diagnosi principale; b) Algoritmi di implementazione informatica dell'albero decisionale; c) Clusterizzazione gerarchica delle procedure.
6. Formalizzazione, implementazione e valutazione delle	Le regole di codifica proprie di ICD-9-	Sono state definite le regole di formalizzazione di ICD9-	Ontologia ICD9-CM; codice Java

<p>nuove regole di classificazione e codifica.</p> <p>7. Progettazione e implementazione del Web Service e dell'interfaccia web per l'interrogazione e l'accesso a SISCO.web.</p>	<p>CM sono state in parte formalizzate e implementate in due modi: sotto forma di proprietà ed annotazioni nella versione ontologica dello standard, e come codice Java nel sistema di estrazione dei codici.</p> <p>Le prove effettuate hanno permesso di identificare tecnologie e framework da utilizzare. Sono stati definiti i requisiti utente e formalizzati in UML: Use case e Activity Diagram. Sono stati progettati ed implementati le strutture dati e i componenti software (batch e web) con relativi diagrammi UML: Component e Sequence Diagram. È stata creata, mediante il componente batch, la base di conoscenza di supporto al sistema di codifica. È stato implementato il prototipo web con le interfacce per la codifica delle diagnosi e delle procedure e per la determinazione della condizione</p>	<p>CM in ontologia OWL-DL (classi e proprietà).</p> <p>La formalizzazione di tutta la terminologia ICD9-CM nell'ontologia è stata completata al 70%.</p> <p>La definizione delle regole Drools che formalizzano l'albero decisionale è stata completata ed è stato implementato il motore di esecuzione in codice Java e integrato come modulo via gestore di code RabbitMQ.</p> <p>Esistenza di un prototipo per la codifica delle diagnosi e delle procedure, e per la determinazione della condizione patologica principale.</p>	<p>Documento di analisi dei requisiti, e progettazione del web service.</p> <p>Web Services e Web Application pubblicati e disponibili ad una specifica URL.</p>
---	--	---	--

patologica principale, basato su un motore di regole Drools che realizza un albero decisionale specifico.

### Quadro delle attività svolte

Descrivere sinteticamente i contenuti delle attività progettuali svolte, indicando la durata ed i soggetti coinvolti nell'implementazione Compilare per ogni fase progettuale non superando 1000 caratteri per fase					
Fase progettuale	Data prevista di inizio	Data prevista di fine	Attività svolta	Eventuali criticità riscontrate	Soggetti coinvolti nella fase progettuale
I	M1	M6	Non è stato possibile svolgere l'attività di test del sistema di supporto alla codifica del Profilo Sanitario Sintetico dai parte di un campione di MMG del FVG, come indicato nell'Obiettivo 1 della precedente tabella.	a) Tardiva implementazione del Profilo Sanitario Sintetico da parte dei MMG dell'FVG. b) Indisponibilità di PSS da analizzare conformi allo standard HL7 CDA2. (Solo PSS conformi allo standard, infatti, possono essere l'input del servizio di supporto alla codifica realizzato.)	Coinvolgimento di tutte le parti (Area delle classificazioni AAS 2; IIT-CNR Sede di Cosenza, FBK e UNIUD).
II	M7	M9	È stata analizzata la lista di 425 condizioni patologiche frequenti in PS, proposta dal GdL regionale PS FVG. I	L'attività è durata più dei due mesi previsti, anche in relazione all'acquisizione soltanto a metà novembre	Questa attività è stata eseguita da Area delle classificazioni AAS 2.

			termini tecnico-scientifici sono stati analizzati con gli obiettivi di arricchire il vocabolario di cui all'Obiettivo 3 e di verificarne la codificabilità supportata. È stata verificata e corretta la codifica in ICD-9-CM fatta dai medici di PS, a partire dallo scarico del database SEI, analizzando 696 diagnosi.	2017 (M9) di un esperto informatico da parte dell'AAS2.	
III	M10	M12	Definizione e digitalizzazione dei vocabolari che alimentano le risorse dati del sistema di supporto alla compilazione delle SDO. Le fonti principali sono l'Indice Sistemático e l'Indice Alfabetico ICD-9-CM, integrate da: termini medici comuni di diagnosi, estratti da	L'attività ha richiesto più tempo di quanto previsto in proposta a causa della difficoltà nel coinvolgimento dei medici ospedalieri per l'estrazione dei termini medici dalle cartelle cliniche, e la definizione di nuovi vocabolari di supporto (integrativi) dato l'esiguo numero di	Questa attività è stata eseguita da Area delle classificazioni AAS 2, UNIUD, IIT-CNR Sede di Cosenza.

			<p>cartelle cliniche MMG e di Pronto Soccorso FVG; termini MeSH mappati a ICD-9-CM; termini di Malattie rare; neoplasie ICD-9-CM; prefissi e suffissi.</p>	<p>termini di diagnosi estratte dalle cartelle di Pronto Soccorso. L'attività inoltre è iniziata in ritardo a causa dell'impossibilità per i partner di reclutare il personale richiesto dal progetto dato il ritardo di ricezione della documentazione ufficiale di progetto, necessaria ad avviare tutte le procedure di tipo amministrativo.</p>	
IV	M13	M15	<p>Ricostruzione e in forma computabile dell'Indice Alfabetico di ICD-10 a partire dai file PDF ufficiali. La ricostruzione è stata effettuata tramite una prima conversione del PDF in testo, seguita dallo sviluppo ed utilizzo di uno script atto ad</p>	<p>Attività avviata in anticipo, nel corso del primo anno alla luce di alcune peculiarità del documento di partenza. Ricostruzione e difficoltosa per la scarsa omogeneità dell'impaginazione del documento di partenza. Inoltre, il documento di partenza ha</p>	<p>Questa attività è stata eseguita da Area delle classificazioni AAS 2 e UNIUD.</p>

			<p>estrarre termini guida, termini associati, codici ed alcune annotazioni particolari dal testo. Ulteriori controlli manuali si sono resi necessari per ripulire il risultato.</p>	<p>un'organizzazione pensata per l'utente umano, con convenzioni dal significato non sempre omogeneo (es. parentesi tonde).</p>	
V	M7	M9	<p>Analisi e definizione della componente di regole di supporto alla codifica della morbilità, anche con l'esame dei diagrammi di flusso realizzati da CCIOMS. Implementazione all'interno dell'ontologia, tramite Protégé, in buona parte effettuata, implementazione nel codice Java solo parziale. L'analisi delle regole implementate nell'ontologia è stata facilitata dalla realizzazione di un web</p>	<p>L'attività è stata molto impegnativa ed è durata più del previsto. Sono state analizzate le formulazioni diagnostiche dei medici di Pronto Soccorso FVG. È stato chiesto ad ARCS l'accesso a un campione di lettere di dimissioni anonimizzate del Piano nazionale Esiti, correlate a SDO, che è stato analizzato rilevando notevoli difformità riguardanti: format utilizzato; contenuti presenti;</p>	<p>Attività eseguita da FBK e UNIUD, con il supporto di IIT-CNR Sede di Cosenza e Area delle classificazioni AAS 2.</p>



			service sperimentale, basato su OWL API, che simula le ricerche.	possibilità di estrarre informazioni utili all'arricchimento dei vocabolari. Non sono state utilizzate tecniche di machine learning in quanto non previste dal Progetto.	
VI	M10	M15	Le regole sono state formalizzate e implementate e in due modi: sotto forma di proprietà ed annotazioni nella versione ontologica di ICD-9-CM, e come codice Java nel sistema di estrazione dei codici.	L'attività è stata svolta in stretto raccordo con la fase precedente.	Questa attività è stata svolta congiuntamente da FBK, UNIUD e Area delle classificazioni AAS 2.
VII	M15	M20	Analisi dei requisiti e definizione dei casi d'uso del servizio SISCO.web: a) processi e attività d'interazione tra utente e sistema durante la compilazione della diagnosi principale nella SDO; b) progettazione dell'architettura del Web	Analisi anticipata a M11, poiché il supporto SISCO.web può essere pilotato dalla LDO, per quanto attiene all'identificazione e codifica della diagnosi principale, con conseguente	Attività svolta congiuntamente da IIT-CNR Sede di Cosenza, FBK, UNIUD e Area delle classificazioni AAS 2.

Service e delle strutture dati del servizio; c) struttura dell'interfaccia a web; d) test preliminari; e) implementazione di un prototipo web per la codifica delle diagnosi e delle procedure, e la successiva determinazione e della condizione patologica principale delle LDO; f) in progress, numerosi debug con aggiustamento progressivo di interfaccia e algoritmi di ricerca; g) link <http://146.48.65.37:8080/SiscoWeb/>.

### **Eventuale diffusione dei Risultati, trasferimento delle conoscenze già svolti**

Descrivere come la ricerca è stata già divulgata e con quali mezzi  
Compilare la parte sottostante non superando i 1000 caratteri

- Formazione ECM: a) Corso AAS2\_17141 La codifica delle condizioni di salute in Pronto Soccorso e nei reparti ospedalieri, mag-dic 2017; b) arruolamento di 2 Medici di PS per approfondimenti.
- Presentazione del Progetto al Wonca International Classification Committee Meeting - Lviv 25-29 ago 2018.
- Presentazione di parte del Progetto al WHO-FIC Annual Meeting 2018 - Seoul 22-27 ott 2018, con il poster Morbidity coding in ICD-11 (and ICHI): a decision tree to identify the main condition, pubblicato

negli atti.

- Partecipazione alla 9th International Digital Public Health Conference - Marseille, 23-26 nov 2019, tenutosi insieme alla 12<sup>a</sup> Conferenza Europea di Public Health, con disseminazione dei risultati parziali di Progetto.
- Presentazione del Progetto al Workshop icts4health nell'ambito del IEEE Symposium on Computers and Communications - Barcelona 30 giu-3 lug 2019 con l'intervento Towards the Development of a Web Support System for Improving Accuracy in Coding Discharge Diagnosis, pubblicato negli atti.
- Invito a pubblicare il lavoro sullo Special Issue on Multidisciplinary Innovation in Engineering Science & Technology della rivista ASTESJ.

**Firma e timbro  
del Proponente**

---



**Dott.ssa Lucilla Frattura**  
Coordinatore  
Centro Collaborativo Italiano dell'OMS  
per la Famiglia delle Classificazioni Internazionali