

L'ontology-Based Data Management per l'eHealth: studio sui malati di cancro alla tiroide

Maurizio Lenzerini², Marina Morelli^{1,2}, Antonella Poggi^{1,2}, Riccardo Rosati²

¹ Dipartimento di Lettere e Culture Moderne

² Dipartimento di Ingegneria Informatica, Automatica e Gestionale "Antonio Ruberti"
Sapienza Università di Roma

{lenzerini,morelli,poggi,rosati}@diag.uniroma1.it

Abstract

Questo progetto mira ad applicare modelli, metodi e tecnologie per l'Ontology-Based Data Management (OBDM) per raccogliere, integrare e analizzare i dati sulla qualità della vita, correlata alla salute dei pazienti affetti dal cancro della tiroide. I nostri obiettivi specifici sono: (i) la costruzione di un'ontologia sulla qualità della vita correlata alla salute dei pazienti affetti da cancro; (ii) l'applicazione di tale ontologia nello sviluppo di un'applicazione per la raccolta, l'integrazione e l'analisi di informazioni sulla qualità della vita dei pazienti con carcinoma della tiroide metastatico. L'adozione dei metodi e delle tecniche di OBDM consente soluzioni flessibili ed estendibili, che integrano risorse e standard di dati e metadati esistenti (vocabolari, ontologie) nel campo della qualità della vita correlata alla salute di malati di cancro. Inoltre, prevediamo di estendere l'attuale tecnologia OBDM per tenere conto delle questioni relative alla privacy e alla protezione dei dati del suddetto dominio applicativo.

Nel campo della gestione della conoscenza per l'assistenza medica, una pletora di strumenti semantici, quali vocabolari, classificazioni e ontologie, sono comunemente usati per la codifica delle descrizioni testuali (che coinvolgono, per esempio, diagnosi, procedure, effetti collaterali), al fine di facilitare l'integrazione delle cartelle cliniche elettroniche nei sistemi di supporto alle decisioni. Inoltre le fonti originali dei dati sono molteplici e varie (note del medico, risultati di laboratorio, risultati radiologici, questionari compilati dal paziente). Pertanto, il successo e l'efficacia di un'applicazione per la gestione dei dati in campo medico sono legati a tre requisiti cruciali: (i) la flessibilità e l'estendibilità della soluzione per la gestione dei dati; (ii) la capacità di integrare facilmente e correttamente le risorse semantiche esistenti nel campo medico; (iii) la capacità di affrontare le questioni relative alla privacy e alla protezione dei dati. Riteniamo che nessuno degli approcci proposti finora sia in grado di affrontare tutte le questioni di cui sopra in un quadro coerente.

Il recente paradigma dell'Ontology-Based Data Management (OBDM) [Poggi *et al.*, 2008] fornisce principi e tecniche per affrontare questa sfida. Un sistema di OBDM è costituito

da un'ontologia, un insieme di fonti di dati e un mapping tra l'ontologia e le fonti. L'ontologia è una descrizione formale e astratta del dominio di interesse, espressa in termini di concetti rilevanti (e dei loro attributi), relazioni tra concetti e asserzioni logiche che descrivono formalmente la conoscenza del dominio. Le fonti di dati sono i repository in cui sono archiviati i dati di interesse. Il mapping è una specifica formale della corrispondenza tra i dati contenuti nelle fonti di dati e gli elementi (concetti, attributi e relazioni) dell'ontologia.

I metodi, gli algoritmi e la tecnologia dell'OBDM, nella loro versione attuale, rispondono già a due dei tre requisiti principali sopra descritti. In particolare, il fatto che un sistema di OBDM sia basato su una specifica dichiarativa consente di modificare ed estendere facilmente l'insieme delle fonti di dati, nonché l'ontologia, che funge da interfaccia tra il livello dei dati e gli utenti del sistema [Xiao *et al.*, 2018]. Inoltre, l'approccio OBDM consente di collegare facilmente dati esterni e risorse di meta-dati al sistema di OBDM attraverso la specifica di mapping dichiarativi e/o tecniche di allineamento ontologico [Harrow *et al.*, 2017]. D'altra parte, la ricerca sul supporto alla privacy e sui problemi di protezione dei dati in un sistema di OBDM è ancora in una fase iniziale e fornisce solo contributi teorici iniziali per la gestione dei problemi di privacy nell'ambito del paradigma dell'OBDM [Grau *et al.*, 2015].

D'altro canto, mentre in campo medico si riscontra un interesse crescente per la qualità della vita (QoL) dei malati di cancro, con un numero crescente di articoli e persino riviste specializzate, è stato riportato che i medici sottovalutano la quantità di sintomi fisici associati ai trattamenti per il cancro della tiroide [James *et al.*, 2018] e che i dati sulla QoL non sono inclusi negli studi controllati randomizzati degli effetti dei trattamenti oncologici. Pertanto, mentre l'efficacia e la sicurezza di terapie specifiche per il trattamento del cancro della tiroide possono essere facilmente ricavate da studi randomizzati controllati, i dati sulla qualità della vita, la compliance e i costi dei pazienti sono carenti o non riflettono la pratica della vita reale.

Il progetto mira ad applicare le metodologie e le tecniche dell'OBDM al dominio dei pazienti affetti da cancro della tiroide. Gli obiettivi specifici del progetto sono:

1. costruire un'ontologia sulla qualità della vita correlata alla salute dei pazienti affetti da cancro alla tiroide;

2. integrare l'ontologia con vocabolari e classificazioni mediche standard, quali ad esempio: la classificazione internazionale delle patologie ICD-10¹; il CTCAE² per la valutazione degli effetti avversi dei trattamenti da parte del medico e il PRO-CTCAE³ che è la versione corrispondente per il paziente; il RECIST 1.1. per la valutazione della risposta radiologica [Eisenhauer *et al.*, 2009];
3. estendere la tecnologia OBDM esistente al supporto delle politiche di protezione dei dati che sono rilevanti nel suddetto dominio di interesse;
4. costruire un sistema di OBDM, basato sull'ontologia sviluppata al punto 1, che integra ed estende i dataset esistenti (es. Il database ITCO dei pazienti affetti da cancro alla tiroide) concentrandosi sulla qualità della vita del paziente unita all'efficacia e alla sicurezza dell'approccio terapeutico.

Il sistema di OBDM faciliterà la trasmissione, il riutilizzo, la condivisione, l'integrazione e l'associazione di dati da più fonti; consentirà aggregazioni statistiche per fornire informazioni sull'efficacia, la sicurezza, la compliance dei pazienti e la qualità della vita (e le loro relazioni) che seguono strategie terapeutiche specifiche; e garantirà allo stesso tempo il supporto alle politiche di protezione dei dati.

Riferimenti bibliografici

- [Eisenhauer *et al.*, 2009] Elizabeth A Eisenhauer, Patrick Therasse, Jan Bogaerts, Lawrence H Schwartz, D Sargent, Robert Ford, Janet Dancey, S Arbuck, Steve Gwyther, Margaret Mooney, et al. New response evaluation criteria in solid tumours: revised recist guideline (version 1.1). *European journal of cancer*, 45(2):228–247, 2009.
- [Grau *et al.*, 2015] Bernardo Cuenca Grau, Evgeny Kharlamov, Egor V Kostylev, e Dmitriy Zheleznyakov. Controlled query evaluation for datalog and owl 2 profile ontologies. In *Twenty-Fourth International Joint Conference on Artificial Intelligence*, 2015.
- [Harrow *et al.*, 2017] Ian Harrow, Ernesto Jiménez-Ruiz, Andrea Splendiani, Martin Romacker, Peter Woollard, Scott Markel, Yasmin Alam-Faruque, Martin Koch, James Malone, e Arild Waaler. Matching disease and phenotype ontologies in the ontology alignment evaluation initiative. *Journal of biomedical semantics*, 8(1):55, 2017.
- [James *et al.*, 2018] Benjamin C James, Briseis Aschebrook-Kilfoy, Michael G White, Megan K Applewhite, Sharon P Kaplan, Peter Angelos, Edwin L Kaplan, e Raymon H Grogan. Quality of life in thyroid cancer - assessment of physician perceptions. *Journal of Surgical Research*, 226:94–99, 2018.
- [Poggi *et al.*, 2008] Antonella Poggi, Domenico Lembo, Diego Calvanese, Giuseppe De Giacomo, Maurizio Len-

zerini, e Riccardo Rosati. Linking data to ontologies. *J. on Data Semantics*, X:133–173, 2008.

[Xiao *et al.*, 2018] Guohui Xiao, Diego Calvanese, Roman Kontchakov, Domenico Lembo, Antonella Poggi, Riccardo Rosati, e Michael Zakharyashev. Ontology-based data access: A survey. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI)*, pages 5511–5519, 2018.

¹<https://icd.who.int/browse10/2016/en#/II>

²https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm

³<https://healthcaredelivery.cancer.gov/pro-ctcae/instrument.html>