Author(s): Lahti, Lauri

Title: Supporting diagnostics and decision making in healthcare by modular methods of computational linguistics

Year: 2016

Version: Post print

Please cite the original version:

Lahti, Lauri. 2016. Supporting diagnostics and decision making in healthcare by modular methods of computational linguistics. E-Learn 2016 - World Conference on E-Learning, 14 16 November 2016, Washington, D.C., USA (eds. Ho, C., & Lin, G.), 1513-1519. Association for the Advancement of Computing in Education (AACE). 7. 978-1-939797-25-4 (electronic).

Lahti, Lauri (2016e). Supporting diagnostics and decision making in healthcare by modular methods of computational linguistics. Proc. E-Learn 2016 - World Conference on E-Learning, 14–16 November 2016, Washington, D.C., USA (eds. Ho, C., & Lin, G.), 1513-1519. Association for the Advancement of Computing in Education (AACE), Chesapeake, VA, USA. ISBN 978-1-939797-25-4. https://www.learntechlib.org/p/174196. Open access in Aaltodoc publication archive at http://aaltodoc.aalto.fi.

Supporting diagnostics and decision making in healthcare by modular methods of computational linguistics

Lauri Lahti
Department of Computer Science
Aalto University School of Science, Finland

Abstract: We propose a new framework for development of modular computational methods to support processes of healthcare and health education in diverse settings. Motivated by an evaluation by The National Institute for Health and Welfare in Finland the proposed framework aims to address challenges of analyzing knowledge concerning healthcare services and patient records with computational linguistics. The framework aims to promote implementing personalized care in diagnostics, decision making, patient engagement and self-care. We describe some analysis methods of computational linguistics, natural language processing, statistics, algorithms and data mining. We have built a prototype program enabling representing and modifying health-related knowledge structures for purposes of prevention, diagnosis and care. For 25 most common diagnosis names we have identified dependencies of core symptom concepts in a conceptual co-occurrence network of 57 679 unique conceptual links about healthcare guidelines.

1 Introduction

The processes of healthcare systems are currently facing a significant transformation due to digitalisation that brings new computational practices to support everyday life (Ash et al. 2012). In healthcare domain various monitoring and analysis technologies rely on computational methods to provide results that can help to recognize symptoms and identify preferable ways to provide effective care (Riches et al. 2016; Semigran et al. 2016).

Besides medical professionals who make diagnosis and give treatment there are various other interest groups that can benefit from new computational healthcare solutions, including patients, peer-support groups, rehabilitation workers, nurses, pharmacists, educators, financial leaders and political decision makers (Doumbouya et al. 2015). It has been considered that respecting the patients rights, supporting patient engagement and encouraging the patient's self-care are important ways to increase efficient healing and wellbeing as well as prevention (World Health Organization 2016).

We propose a new framework for development of intuitive modular computational methods to support processes of healthcare and health education in diverse settings. We describe the prospects of digitalisation of healthcare in the context of Finland that has a long tradition of state-funded public educational and healthcare services that have produced good results.

2 Emerging opportunities based on digitalisation of healthcare

One of the primary projects in the current government programme in Finland is the digitalisation of public services (Government programme of Finland 2016). In a recent nationwide evaluation it was suggested that a new national data resource unit should be established to deal with the complex data flows of primary care services in the Finnish healthcare system (Ministry of Social Affairs and Health Finland 2016). In Finland there is currently going on a transition to a new information system for emergency call centers (ERICA) and governmental field operation system (KEJO) (The National Institute for Health and Welfare 2015). The KEJO system includes the national primary care records system (EHK) that is based on the national health archive system (Kanta).

According to the evaluation report made by The National Institute for Health and Welfare (2015) an aim is that a patient can be followed in the patient records system through the whole care chain and support with cumulative information also the processes of primary care and the information system of emergency call centers. The report identifies challenges concerning how the information architecture of primary care records can be defined so that the user interfaces relying on it can operate in a way that is guiding, intelligent, supporting care and harmonizing the care practices.

Based on the previous notions it is important to develop with a multidisciplinary research collaboration new models and methods relying on computational linguistics that can introduce new kind of intelligent functionality to support diagnostics and decision making in the processes of Finnish healthcare system. Since digitalisation of healthcare is a major global trend we expect that the challenges and solutions identified in the context of Finnish healthcare can provide some results that can be usefully considered in other national contexts and can help to create shared knowledge resources that facilitate interoperability of healthcare systems. Motivated by the goals and needs discussed above we propose a new framework for development of intuitive modular computational methods to support processes of healthcare and health education in diverse settings.

3 A framework of computational linguistics to support healthcare

There are many complementing ways to develop new models and methods based on computational linguistics that can promote patient care by supporting diagnostics and decision making relying on patient records in various stages of care process (Szolovits 1995; Arocha et al. 2005; Chmiel et al. 2014; Dasgupta & Chawla 2014; Zhou et al. 2014). We propose a new framework that aims to identify some useful approaches to address these development efforts in respect to health and welfare, information technology and society.

In the domain of health and welfare the framework aims to promote effectiveness of healthcare system, effects-based managerial leadership and patient engagement. In the domain of information technology the framework aims to develop computational (computer-assisted) methods for modeling of health information and natural language in general. In the domain of society the framework aims to support general digitalisation of processes of healthcare and everyday life so that based on this digitalisation it is possible to implement high-quality innovative services in both public and private sectors.

The proposed framework enables defining various hypotheses, including the following two preliminary hypotheses. Firstly, it can be expected that new models and methods relying on computational linguistics can promote systematic analysis about knowledge in patient records thus creating informative reports of trends on populational level and in short time interval that can be used for division of resources, prioritization, scheduling and general operational management. Secondly, it can be expected that new models and methods relying on computational linguistics can promote implementing personalized high-quality care by supporting medical professionals in diagnostics and decision making and patient engagement and self-care. Application of the framework aims to address the tasks identified in the evaluation report made by The National Institute for Health and Welfare (2015).

4 Implementation of the framework

The framework is implemented based on a multistage process that defines preferable analytical methods of computational linguistics (for example natural language processing, statistical, algorithmic and data mining-based methods) and the contents and constraints of the data that is aimed to be analyzed. After getting required research permits the raw data is retrieved from various complementing resources (for example information of patient records in different stages of a care process and supplementing background variables and key measures) and it is combined to form the actual research data that is then analyzed with diverse analysis methods. New models are developed based on the analysis and these models are tested and gained results are reported and harnessed into applied use.

The framework aims to promote public unlimited availability of research results and information with principles of open access and open data although addressing carefully privacy issues and the patient's rights. Besides publishing the results in respected peer-reviewed scientific publications an essential part of the publishing of the results of the framework is carried out by introduction of practical intuitive support tools. Computational models and methods developed with the framework are distributed to applied work of healthcare to support diagnostics and decision making and this distribution can happen in various representations, including concept maps assisting in reasoning, flow charts, algorithm components, statistical formulas, spreadsheet-based form templates and scripts.

The framework aims to develop new models and methods relying on computational linguistics to support processes of healthcare especially related to diagnostics and decision making. Using computational linguistics means exploiting diverse computational methods to interpret and analyze semantics of language. Useful research methods for the framework include developing and applying various natural language processing, statistical, algorithmic and data mining-based methods.

In the framework an important emphasis is given to identify and evaluate preferable analysis methods of computational linguistics to develop new models about the semantics of health-related information in patient records along the care process and to test (i.e. validate) created models. Analysis methods of computational linguistics that can be used for the framework include a broad spectrum of natural language processing, statistical, algorithmic and data mining-based methods that enable to analyze features, dependencies and trends in knowledge expressed with a natural language. Some of the promising analysis methods include using traditional statistical tests, application of big data analytics, exploiting open data collections (for example key measures of healthcare, psychology and semantics) and using text corpuses based on databases of healthcare guidelines and network discussions as well as application of theories concerning diagnostics, decision making and patient engagement.

When making modeling based on computational linguistics for healthcare-related information it is important to consider for the created methods the properties of validity, reliability and responsiveness as well as what are the relationships between the created methods, diagnostic measures, follow-up measures and predictive measures. In the development of methods it is essential to carefully evaluate also the relationships between interpretations of dependency vs. causality, statistical significance vs. clinical significance and the minimal clinically important change.

For the framework when developing models based on computational linguistics it is largely based on analysis of research data which contains information about patient records in different stages of the care process and this is supplemented with background variables and key measures. Table 1 illustrates some methods identified to be useful for analysis of research data in respect to computational linguistics methods when categorized coarsely into three subcategories of natural language processing, statistics as well as algorithms and data mining. Naturally this categorization aims to be only suggestive and furthermore categories and methods have a partial overlap.

Table 1. For the framework some methods for analysis of research data in respect to computational linguistics categorized into natural language processing, statistics as well as algorithms and data mining.

Some analysis methods for the framework in respect to computational linguistics		
Natural language processing	*Statistics*	*Algorithms and data mining*
parsing of natural language, finite-state transducers, sentiment analysis, frequency distributions, n-grams, co-occurrences and networking properties	parametric and non-parametric tests, covariance and correlation analysis, variance analysis, regression analysis and factor analysis	unsupervised methods (clustering techniques such as hierarchical, k-means and fuzzy clustering and self organizing maps) and supervised methods (comparison methods such as decision trees, k nearest neighbours, linear discriminant analysis, shopping cart analysis, association rules, naive Bayes classification and support vector machines)

5 Experiment

The framework aiming to support healthcare processes relying on computational linguistics can be used for various purposes depending on the current goals and needs. We have made preliminary experiments with healthcare-related data collections to identify and evaluate useful requirements and conditions for practical application of the framework in a healthcare setting. Motivated by previous research results concerning effective adaptive visual decision support methods (Lee 2006; Horsky et al. 2012; Visweswaran et al. 2015) we have built a preliminary prototype program for representing and modifying health-related knowledge structures based on the framework. A modular architecture is an essential design principle suggested for both the framework and its implemented prototype program. Modularity aims to ensure easy configuration and later updating of the functionality of the framework to address diverse needs of analysis and representation of data. Figure 1 illustrates an experimental prototype program relying on a textual input field enabling a person to make a health-related query and two interactive visualisations to represent dependencies concerning the concepts of the query in a conceptual co-occurrence network about healthcare guidelines. Referring to our previous work (Lahti 2016b; Lahti 2016d) we are using a conceptual co-occurrence network that we created based on a set of 93 medical texts containing 85 055 words about healthcare guidelines given by Terveyskirjasto provided by The Finnish Medical Society Duodecim (Käyvän hoidon potilasversiot 2016) containing 57 679 unique conceptual links between 2014 unique nouns having at least 3 occurrences in the set of medical texts. Each pair of nouns occurring in the same sentence in the set of medical texts defines a conceptual link between these nouns in the co-occurrence network. The original medical texts are in Finnish but we report the results in English although some concepts have ambiguation that may emerge in varied translations (Lahti 2016b; Lahti 2016d).

Our preliminary prototype program can be used with a web browser for example with an ordinary computer or a smartphone and it aims to offer interactive support for a user in information retrieval and creative problem solving concerning medical knowledge. When the person provides some health-related concepts as an input text, supplied with predictive typing assistance, the program identifies the shortest paths of consecutive links between these concepts in the co-occurrence network based on nouns occurring in the set of medical texts and illustrates these connecting associative link paths in the form of an adaptive concept map. In Figure 1 based on the input

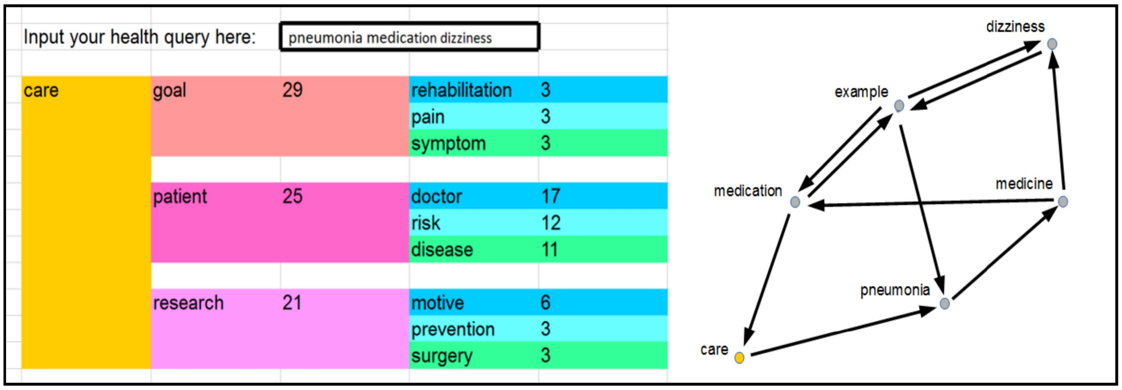

Figure 1. Illustration of a prototype program for representing and modifying health-related knowledge structures based on the framework.

concepts ”pneumonia”, ”medication” and ”dizziness” the program generates a concept map containing six concepts connected with ten directional links.

Besides the concept map there is an interactive visualisation of a decision tree that enables the person to adjust properties of the concept map and thus to modify the perspectives available to the knowledge and extend the exploration in knowledge structures. For example, by selecting certain concepts of the concept map the person can request a possibility to see alternative related concepts and connecting associative link paths generated for the selected concepts. These requested related concepts and link paths emerge as hierarchical items in the updated decision tree and by selecting some of these items then enables further modification of the current knowledge representation. In Figure 1 the decision tree shows the most co-occurring concepts for the selected concept "care" and then for those concepts the most co-occurring concepts, each supplied with the number of occurrences. By adding progressively more concepts as input text enables to gain an increase in the various perspectives and connectivity of concepts shown in the knowledge representation.

When requested the prototype program can provide for the person supplementing information describing the details motivating the relationships identified between the given concepts. In the domain of healthcare the prototype program aims to support a person in making diagnosis as well as planning suitable care by letting him analyse personally the interactive visual knowledge representation. Furthermore the visualisation can be used also for representing suggested progressive action steps to be carried out when implementing the care shown in the form of a network and a decision tree updated along the care as it proceeds.

In our previous work (Lahti 2016d) we evaluated some properties of knowledge structures to enable development of modeling patterns of creation, modification and retrieval of medical knowledge. We evaluated some statistical properties of Wikipedia articles in respect to 25 most common diagnosis names for the year 2011 provided by the medical company Practice Fusion based on its electronic health records system having a data set of over 7 million patients largely based on consultations with primary care physicians (Rowley 2011). Now in our current work we decided to carry out further experimental evaluation relying on the same set of 25 most common diagnosis names, shown in Table 2 in the order of decreasing frequency (we did not have an access to the exact frequencies of the diagnosis names). We now evaluated what kind of adaptive concept maps the prototype program generates based on health-related queries that contain concepts typically associated with symptom descriptions for some common medical diagnosis names.

For 25 most common diagnosis names (Rowley 2011) we examined the set of 93 medical texts about healthcare guidelines of Duodecim (Lahti 2016b; Käyvän hoidon potilasversiot 2016) to identify typical symptom descriptions. If we did not identify sufficiently clear typical symptom descriptions in this set of texts we then examined some supplementary guidelines of Duodecim that we retrieved by querying the healthcare guidelines database of Duodecim (Terveyskirjasto 2016) with the diagnosis name. Based on the found symptom descriptions we aimed to select a set of 2-6 concepts to represent the typical symptoms of each diagnosis name. This set of concepts is referred to as core symptom concepts. With the prototype program we generated an adaptive concept map for each diagnosis name based on identifying the shortest paths in the co-occurrence network about healthcare guidelines (Lahti 2016b) between the query concepts that were the core symptom concepts for this diagnosis name.

Our experimental results of health-related queries with the prototype program are illustrated in Table 2 showing the dependencies of core symptom concepts based on symptom descriptions (Lahti 2016b; Käyvän hoidon potilasversiot 2016; Terveyskirjasto 2016) in the conceptual co-occurrence network about healthcare guidelines

Table 2. Dependencies of core symptom concepts based on symptom descriptions (Lahti 2016b; Käyvän hoidon potilasversiot 2016; Terveyskirjasto 2016) in the conceptual co-occurrence network about healthcare guidelines (Lahti 2016) for 25 most common diagnosis names (Rowley 2011).

25 most common diagnosis names (Rowley 2011)	Dependencies of core symptom concepts based on symptom descriptions (Lahti 2016b; Käyvän hoidon potilasversiot 2016; Terveyskirjasto 2016) in the conceptual co-occurrence network relying on a set of 93 medical texts about healthcare guidelines provided by The Finnish Medical Society Duodecim (Lahti 2016b)		
Ranking number and name	*Core symptom concepts based on symptom descriptions found in healthcare guidelines of Duodecim (Lahti 2016b; Käyvän hoidon potilasversiot 2016; Terveyskirjasto 2016) for the current diagnosis name*	*Supplementary intermediate concepts emerging in the adaptive concept map based on the shortest paths between the core symptom concepts in the conceptual co-occurrence network*	*Conceptual links in the adaptive concept map based on the shortest paths between the core symptom concepts in the conceptual co-occurrence network (concepts indicated with id codes and the notation A>B denotes a link from concept A to concept B)*
1. Hypertension	salt (id=1566); alcohol (id=76); physical_exercise (id=841); nutriment (id=1362); overweight (id=1990)	amount (id=1024); calcium (id=517); care (id=390); child (id=783); day (id=1320); person (id=348); symptom (id=1108); use (id=761);	1566>76;1566>1024>841;1566>761>783>1362;1566>761>1108>1990;76>1566;76>841;76>783>1362;76>1108>1990;841>76>1566;841>76;841>348>1362;841>348>1990;1362>517>761>1566;1362>390>76;1362>517>841;1362>1990;1990>783>1320>1566;1990>783>76;1990>783>841;1990>1362;
2. Hyperlipidemia	lipid_value (id=1357); smoking (id=1752); blood_pressure (id=1878)	diabetes (id=177); diabetic (id=175); nephropathy (id=1029);	1357>175>1752;1357>1878;1752>1029>1357;1752>177>1878;1878>1357;1878>175>1752;
3. Diabetes	overweight (id=1990); abdominal_obesity (id=568); genome (id=1190); age (id=440); physical_exercise (id=841)	child (id=783); direction (id=1588); estimate (id=120); example (id=256); overgrowth (id=825); patient_version (id=1250); percent (id=1254); person (id=348); risk_factor (id=1392); time (id=20); urinary_incontinence (id=1929);	1990>1752>568;1990>1392>1190;1990>783>440;1990>783>841;568>1752>1990;568>256>825>1190;568>256>440;568>256>841;1190>1392>1990;1190>120>1588>568;1190>440;1190>1254>841;440>1929>1990;440>256>568;440>1190;440>1250>841;841>348>1990;841>1752>568;841>1392>1190;841>20>440;
4. Back pain	work (id=1776); lifting (id=1075); position (id=125)	back (id=1470); back_pain (id=1461); epilepsy (id=235); limb (id=1334); person (id=348); symptom (id=1108);	1776>1461>1470>1075;1776>1334>125;1075>348>235>1776;1075>125;125>1108>1776;125>1075;
5. Anxiety	emotional_life (id=1738); problem (id=1121); behavior (id=759); interaction (id=1947)	child (id=783); guidance (id=1100);	1738>1121;1738>1121>759;1738>1947;1121>1738;1121>759;1121>1947;759>1121>1738;759>1121;759>1100>1947;1947>1738;1947>1121;1947>783>759;
6. Obesity	weight_index (id=1160); waist (id=1951); diet (id=1414); physical_exercise (id=841)	abdominal cavity (id=1863); care (id=390); fat (id=1360); moon (id=725); mother (id=2008); obesity (id=824); weight (id=1161);	1160>2008>1951;1160>390>1414;1160>841;1951>1161>1160;1951>725>1414;1951>725>841;1414>824>1160;1414>1360>1951;1414>841;841>1160;841>1863>1951;841>1414;
7. Allergic rhinitis	pollen (id=1497); animal (id=207)	immunotherapy (id=1493); symptom (id=1108);	1497>1108>1493>207;207>1493>1108>1497;
8. Reflux esophagitis	heartburn (id=1094); overweight (id=1990); smoking (id=1752)	care (id=390); child (id=783); feeling (id=1115); reflux_disease (id=1367); symptom (id=1108);	1094>1367>1108>1990;1094>1115>390>1752;1990>783>1367>1094;1990>1752;1752>1108>1367>1094;1752>1990;
9. Respiratory problems	cough (id=2005); mucus (id=848); dyspnea (id=341); chest_pain (id=1384)	symptom (id=1108);	2005>848;2005>605>341;2005>605>1384;848>2005;848>1108>341;848>1108>1384;341>1108>2005;341>1108>848;341>1384;1384>1108>2005;1384>1108>848;1384>341;
10. Hypothyroidism	fatigue (id=1963); cold (id=737); overweight (id=1990); constipation (id=1793); pulse (id=1599)	child (id=783); connection (id=1974); example (id=256); mouth (id=1585); risk_factor (id=1392); symptom (id=1108);	1963>1585>737;1963>1108>1990;1963>1108>1793;1963>1108>1599;737>1585>1963;737>1392>1990;737>1974>1793;737>256>1599;1990>1108>1963;1990>1392>737;1990>1108>1793;1990>783>1599;1793>1108>1963;1793>1974>737;1793>1108>1990;1793>1108>1599;1599>1108>1963;1599>256>737;1599>1108>1990;1599>1108>1793;
11. Visual refractive errors	eyesight (id=1090); eye (id=1507); attention (id=397); weakness (id=340)	air (id=444); difficulty (id=1815); factor (id=1676); length (id=1219); patient (id=1251); understanding (id=2001); year (id=1948);	1090>1507;1090>1251>397;1090>444>340;1507>1090;1507>1676>397;1507>1948>340;397>1251>1090;397>1219>1507;397>2001>340;340>444>1090;340>1815>1507;340>1815>397;
12. General medical exam	risk (id=1393); lifestyle (id=213); ability (id=735)	disease (id=1667); illness (id=1446); research (id=1760); symptom (id=1108);	1393>1667>213;1393>1760>735;213>1108>1393;213>1446>735;735>1446>1393;735>1446>213;
13. Osteoarthritis	joint_pain (id=1061); physical_exercise (id=841); rest (id=807); rigidity (id=505)	connection (id=1974); joint (id=1071); mouth (id=1585); performance (id=1707); time (id=20); trace (id=497);	1061>841;1061>497>807;1061>1707>505;841>1061;841>20>807;841>1071>505;807>497>1061;807>1974>841;807>1585>505;505>1707>1061;505>1707>841;505>605>807;
14. Fibromyalgia/myositis, neuritis	fitness (id=695); wellbeing (id=411); pain (id=605); soreness (id=119)	cancer (id=1618); care (id=390); cheek (id=1246); health_ministry (id=1689); help (id=114); patient (id=1251); quality_of_life (id=210); specialized_doctor (id=242); symptom (id=1108); tooth (id=320);	695>210>411;695>320>605;695>1108>119;411>390>695;411>390>605;411>1689>1251>119;605>1618>695;605>114>411;605>1246>119;119>1251>695;119>1251>242>411;119>1251>605;
15. Malaise and fatigue	vomiting (id=1111); diarrhea (id=1390); spasm (id=675); fever (id=728)	mouth (id=1585); patient (id=1251); symptom (id=1108);	1111>1390;1111>675;1111>1251>728;1390>1111;1390>1108>675;1390>1251>728;675>1111;675>1585>1390;675>1108>728;728>1251>1111;728>1251>1390;728>1108>675;
16. Pain in joint	pain (id=605); ache (id=1624); rigidity (id=505); oedema (id=1755); soreness (id=119); mobility (id=835)	care (id=390); cheek (id=1246); child (id=783); foot (id=480); joint (id=1071); patient (id=1251); symptom (id=1108); vision (id=1093);	605>1624;605>505;605>1755;605>1246>119;605>835;1624>605;1624>605>505;1624>783>1755;1624>1108>119;1624>1093>835;505>605;505>605>1624;505>605>1755;505>1071>119;505>605>835;1755>605;1755>783>1624;1755>605>505;1755>119;1755>390>835;119>1251>605;119>1108>1624;119>1071>505;119>1755;119>1071>835;835>605;835>841>1624;835>1071>505;835>480>1755;835>1071>119;
17. Acute laryngopharyngitis	pharynx (id=1048); cough (id=2005); rhinitis (id=1077)	care (id=390); liquid (id=1033);	1048>390>2005;1048>390>1077;2005>1033>1048;2005>1077;1077>390>1048;1077>2005;
18. Acute maxillary sinusitis	rhinitis (id=1077); cough (id=2005); sore_throat (id=714); fever (id=728)	care (id=390); cheek (id=1246); child (id=783); symptom (id=1108);	1077>2005;1077>1108>714;1077>783>728;2005>1077;2005>1108>714;2005>783>728;714>1108>1077;714>1108>2005;714>1108>728;728>390>1077;728>1246>2005;728>1108>714;
19. Major depressive disorder	mood (id=962); fatigue (id=1963); concentration (id=567)	care (id=390); human (id=424); illness (id=1446); performance (id=1707); rehabilitation (id=704); symptom (id=1108);	962>1446>1963;962>704>1121>567;1963>390>962;1963>390>1707>567;567>1947>424>962;567>1738>1108>1963;
20. Acute bronchitis	cough (id=2005); mucus (id=848); dyspnea (id=341)	symptom (id=1108);	2005>848;2005>605>341;848>2005;848>1108>341;341>1108>2005;341>1108>848;
21. Asthma	cough (id=2005); rhinorrhea (id=849); dyspnea (id=341); strain (id=1348)	blood (id=1891); child (id=783); emotion (id=1739); symptom (id=1108); week (id=1905);	2005>849;2005>605>341;2005>783>1348;849>2005;849>1739>605>341;849>1739>1891>1348;341>1108>2005;341>1108>1624>849;341>1348;1348>1905>2005;1348>1905>2005>849;1348>341;
22. Depressive disorders	mood_disorder (id=961); change (id=1014); depression (id=943)	care (id=390); example (id=256); pregnancy (id=1355);	961>256>1014;961>943;1014>390>961;1014>1355>943;943>961;943>841>1014;
23. Nail fungus	nail (id=740); pain (id=605); dandruff (id=364)	care (id=390); cheek (id=1246); need (id=1655); removal (id=1234); use (id=761);	740>1246>605;740>761>1234>364;605>1246>740;605>390>364;364>1234>1655>740;364>1234>605;
24. Coronary atherosclerosis	smoking (id=1752); cholesterol (id=645); blood_pressure (id=1878)	care (id=390); diabetes (id=177); diabetic (id=175);	1752>841>645;1752>177>1878;645>390>1752;645>1878;1878>175>1752;1878>645;
25. Urinary tract infection	need_to_urinate (id=1926); stinging (id=610); fever (id=728)	adverse_effect (id=309); patient (id=1251); symptom (id=1108); time (id=20); woman (id=1025);	1926>1025>1108>610;1926>1025>1251>728;610>309>1025>1926;610>20>728;728>1251>1025>1926;728>20>610;

(Lahti 2016b) for 25 most common diagnosis names (Rowley 2011). For example for the diagnosis name ”allergic rhinitis” we identified two core symptom concepts that are “pollen” and “animal”. The shortest paths of consecutive links in the co-occurrence network about healthcare guidelines between these core symptom concepts are pollen→symptom→immunotherapy→animal and animal→immunotherapy→symptom→pollen and these concepts and conceptual links together form the adaptive concept map. Thus the supplementary intermediate concepts emerging in the adaptive concept map are immunotherapy and symptom. Due to space constraints this article can show only a part of the results, more results are available as open data in a separate publication (Lahti 2016f).

6 Discussion and future work

A fundamental aim of the adaptive visualisation of medical knowledge provided by the prototype program is to support a person to manage with complex interrelated pieces of information, to evaluate them from diverse perspectives and to make a synthesis based on them relying on both intuitive and logical reasoning. Thus the prototype program can be considered to have features of a decision support system and a recommender system. It needs to be emphasized that besides co-occurrence networks also various other computational methods can be expected to be used in the prototype program due to the modularity of operational principles of the proposed framework. A general aim of our proposal is to encourage efforts to build collaboratively adaptive knowledge processing tools that can be freely used and further developed by anyone in a way that fruitfully supports creating ecosystems benefiting from organized modularity and interoperability of tools thus embracing the philosophy of open access and open knowledge.

There is a need for the future research to carefully identify, formulate and validate how various complementing modular analysis and representation methods can be incorporated into the framework to support healthcare processes (see some related analysis for example in Hoffman et al. 2013). To enable the framework to produce new models and methods that can effectively support healthcare and health education in diverse contexts there is a need to establish and strengthen collaboration in evidence-based research among various complementing organizations in the healthcare domain. In the Finnish context there are many respected health information providers such as The National Institute for Health and Welfare, Ministry of Social Affairs and Health Finland, The Finnish Medical Society Duodecim and the Finnish Federation for Social Affairs and Health (SOSTE). In the international context some interesting health information providers include for example non-profit organizations Cochrane (http://www.cochrane.org) that promotes open evidence-based medicine and Guidelines International Network (http://www.g-i-n.net) that promotes formation of clinical practice guidelines as well as the quality criteria collection of Health On Net foundation (http://www.hon.ch).

References

Arocha, J., Wang, D., & Patel, V. (2005). Identifying reasoning strategies in medical decision making: A methodological guide. Journal of Biomedical Informatics. Volume 38, Issue 2, April 2005, Pages 154–171. http://dx.doi.org/10.1016/j.jbi.2005.02.001

Ash, J., Sittig, D., Guappone, K., Dykstra, R., Richardson, J., Wright, A., Carpenter, J., McMullen, C., Shapiro, M., Bunce, A., & Middleton B. (2012). Recommended practices for computerized clinical decision support and knowledge management in community settings: a qualitative study. BMC Medical Informatics and Decision Making, 12(6).

Chmiel, A., Klimek, P., & Thurner, S. (2014). Spreading of diseases through comorbidity networks across life and gender. New Journal of Physics 16, 115013.

Dasgupta, D., & Chawla, N. (2014). Disease and medication networks: an insight into disease-drug interactions. 2nd International Conference on Big Data and Analytics in Healthcare, Singapore 2014. http://www3.nd.edu/~nchawla/papers/bdah_2.pdf

Doumbouya, M., Kamsu-Foguem, B., Kenfack, H., & Foguem, C. (2015). Combining conceptual graphs and argumentation for aiding in the teleexpertise. Computers in Biology and Medicine, Elsevier, 2015, vol. 63, pp. 157-168.

Government programme of Finland (2016). Digitalisation, experiments and dissolving norms. Finnish Government. Published in Finnish. http://valtioneuvosto.fi/hallitusohjelman-toteutus/digitalisaatio

Hoffman, A., Volk, R., Saarimaki, A., Stirling, C., Li, L., Härter, M., Kamath, G., & Llewellyn-Thomas, H. (2013). Delivering patient decision aids on the Internet: definitions, theories, current evidence, and emerging research areas. BMC Medical Informatics and Decision Making, 13(2):S13. doi: 10.1186/1472-6947-13-S2-S13 http://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/1472-6947-13-S2-S13#CR44

Horsky, J., Schiff, G., Johnston, D., Mercincavage, L., Bell, D., & Middleton, B. (2012). Interface design principles for usable decision support: A targeted review of best practices for clinical prescribing interventions. Journal of Biomedical Informatics. Volume 45, Issue 6, December 2012, Pages 1202–1216. http://dx.doi.org/10.1016/j.jbi.2012.09.002

Käyvän hoidon potilasversiot (2016). Patient versions of healthcare guidelines given by Terveyskirjasto provided by The Finnish Medical Society Duodecim. ”Käypä hoito, potilasversiot”. Retrieved in January 2016 from http://www.terveyskirjasto.fi/terveyskirjasto/tk.koti?p_osio=109&p_teos=khp

Lahti, Lauri (2016b). Semantic modeling of healthcare guidelines to support health literacy and patient engagement. Proc. Global

Learn 2016: Global Conference on Learning and Technology, 28-29 April 2016, Limerick, Ireland. Association for the Advancement of Computing in Education (AACE), Chesapeake, VA, USA. https://www.learntechlib.org/p/172737; http://urn.fi/URN:NBN:fi:aalto-201603291477
Lahti, Lauri (2016d). Supporting online health queries by modeling patterns of creation, modification and retrieval of medical knowledge. Proc. EdMedia 2016 - World Conference on Educational Media and Technology, 28–30 June 2016, Vancouver, B.C., Canada (ed. Veletsianos, G.), 705–712. Association for the Advancement of Computing in Education (AACE), Chesapeake, VA, USA. ISBN 978-1-939797-24-7. http://www.learntechlib.org/p/173023; http://urn.fi/URN:NBN:fi:aalto-201606202682
Lahti, Lauri (2016f). Supplement to Lauri Lahti's conference article ”Supporting diagnostics and decision making in healthcare by modular methods of computational linguistics”, to appear online at http://aaltodoc.aalto.fi.
Lee, S. (2006). Visualization of clinical practice guidelines and patient care process. Doctoral dissertation. The George Washington University. Washington, DC, USA. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.473.5825&rep=rep1&type=pdf
Ministry of Social Affairs and Health Finland (2016). Nationwide report of primary care service operation. Intermediary report 2. Published in Finnish. ISBN 978-952-00-3822-9. http://urn.fi/URN:ISBN:978-952-00-3822-9
Riches, N., Panagioti, M., Alam, R., Cheraghi-Sohi, S., Campbell, S., Esmail, A., & Bower, P. (2016). The Effectiveness of Electronic Differential Diagnoses (DDX) Generators: A Systematic Review and Meta-Analysis. PLoS ONE 11(3): e0148991. doi:10.1371/journal.pone.0148991
Rowley, R. (2011). The 25 most common diagnoses. Robert Rowley, Chief Medical Officer, Practice Fusion EMR. Practice Fusion Blog, posted on 9 February 2011. http://www.practicefusion.com/blog/25-most-common-diagnoses/
Semigran, H., Levine, D., Nundy, S., & Mehrotra, A. (2016). Comparison of Physician and Computer Diagnostic Accuracy. JAMA Internal Medicine. Published online October 10, 2016. doi:10.1001/jamainternmed.2016.6001
Szolovits, P. (1995). Uncertainty and Decisions in Medical Informatics. Methods of Information in Medicine, 34:111–21.
Terveyskirjasto (2016) Healthcare guidelines provided by The Finnish Medical Society Duodecim. Health library healthcare guidelines database. http://www.terveyskirjasto.fi/terveyskirjasto/tk.koti
The National Institute for Health and Welfare (2015). Preliminary report of nationwide data management of primary care services - Towards national information applicability. Published in Finnish. ISBN 978-952-302-498-4. http://urn.fi/URN:ISBN:978-952-302-498-4
Visweswaran, S., Ferreira, A., Ribeiro, G., Oliveira, A., & Cooper, G. (2015). Personalized Modeling for Prediction with Decision-Path Models. PLoS ONE 10(6): e0131022. doi:10.1371/journal.pone.0131022
World Health Organization (2016). From innovation to implementation – eHealth in the WHO European Region. WHO Regional Office for Europe. http://www.euro.who.int/en/publications/abstracts/from-innovation-to-implementation-ehealth-in-the-who-european-region-2016
Zhou, X., Menche, J., Barabasi, A., & Sharma, A. (2014). Human symptoms-disease network. Nature Communications, 5, 4212. http://www.nature.com/articles/ncomms5212