

# Exploring outlier documents to investigate potential links between gut microbiota and Alzheimer's disease

Bojan Cestnik<sup>1,3</sup>, Elsa Fabbretti<sup>2</sup>,  
Donatella Gubiani<sup>2</sup>, Nada Lavrač<sup>1,2</sup>, Tanja Urbančič<sup>2,1</sup>

<sup>1</sup> Jožef Stefan Institute, Ljubljana, Slovenia

<sup>2</sup> University of Nova Gorica, Nova Gorica, Slovenia

<sup>3</sup> Temida d.o.o., Ljubljana, Slovenia

{bojan.cestnik, nada.lavrac}@ijs.si,  
{donatella.gubiani, elsa.fabbretti, tanja.urbancic}@ung.si

**Abstract.** In this paper we propose an approach to cross domain literature-based discovery using text mining tools OntoGen and CrossBee to search for hidden relations in the scientific papers. The approach is based on exploring outlier documents, aiming to help the expert in uncovering implicit relationships among concepts of two different domains of interest. The novelty is the interplay between OntoGen and CrossBee. Ontogen is used to narrow down the search for bridging terms by identifying the set of outlier documents, making the search in CrossBee more focused, efficient and effective. In our first experiment in “Alzheimer” and “gut microbiota” domain we identified the term “nitric oxide synthase” as a promising novel bridging term, which we will investigate in our further work.

**Keywords:** Literature mining, cross context discovery, outliers, gut microbiota, Alzheimer's disease

## 1 Introduction

Understanding complex phenomena often requires knowledge from different disciplines to be taken into account. Also some disease conditions can't be fully understood unless expertise from different disciplines is put together. Such situations suggest the use of IT tools to integrate data and knowledge from different sectors (human health, biomedicine, microbiology, nutrition, etc.) to provide guidelines for evidence based interdisciplinary biomedical and clinical research with more rapid advancement in the field of health. An example is ageing, where age-related pathologies, such as neurodegenerative diseases, cause big challenges to the society. Our preliminary literature mining investigation in this field suggested links between dietary issues and Alzheimer's disease to be further investigated [1]. Some recent medical articles also indicate exploring connections between the digestive system, gut microbiota and neurodegenerative diseases (like Alzheimer's disease) as a promising area of biomedical research [e.g. 2]. A growing number of recently published research papers in this area is a strong motivation for using text and literature mining methods to research the

hypothesis that, besides causing gut problems, an imbalance of gut microbiota can be associated with memory and cognition dysfunction and brain diseases.

The proposed methodology follows the line of research, known as literature-based discovery (LBD), that have proved to be a powerful tool for overcoming the problem of fragmentation of science and for assisting researchers in their process of cross-discipline knowledge discovery [3]. The field of LBD started to evolve with the early work of Swanson [4] and Smalheiser [5], who developed early approaches to assist the user in detecting interesting cross domain terms with a goal to discover unknown relations between previously unrelated concepts in two different domains (two corpora of medical articles) of interest. Their idea of discovering new hypotheses by connecting fragmented pieces of knowledge from different contexts via bridging terms has proved to be very powerful and has inspired other researchers to develop it further. Several examples are presented in [3].

This view provided also a motivation for developing the main functionality of the CrossBee approach to bridging term (B-term) discovery, implemented through ensemble based term ranking, developed by Juršič et al. [4], where an ensemble heuristic for B-term detection was proposed for B-term detection and ranking, with the aim of finding new connections between two predefined domains, represented by two user define sets of medical articles. In addition, the research conducted by Petrič et al. [5] and Sluban et al. [6] complements this research, as it indicates that bridging terms are more frequent in documents that are outlier documents of their own domain; these outlier documents namely contain a substantially larger amount of bridging/linking terms than the normal, non-outlier documents.

Next generation sequencing techniques has revealed the complexity of human gut microbiome [9], demonstrating that its malfunction is associated to a number of pathologies including brain diseases, inflammaging and immunosenescence in the elderly [10]. Recent findings have disclosed insights on the effects of circulating metabolites derived from food transformation processes occurring at digestive level and mediated by local microbiome activity [11]. These studies have disclosed new functional relations between gut, brain ageing and neurodegenerative diseases [2], supported by neuromodulators synthesized at peripheral level and influencing the central neuronal networks [12]. Despite obvious tissue-specific physiology and characteristics, neurons innervating the gut are likely subjected to similar neurodegenerative processes as brain neurons [13]. Focusing on mediators of common neuro-inflammation mechanisms, could therefore disclose previously not identified molecular terms with multi-target therapeutic potential.

In this paper we propose an approach to bisociative cross domain literature-based discovery using text mining tools OntoGen [14] and CrossBee [15] to search for hidden relations in the published scientific papers and apply it to the investigation of potential links between gut microflora and Alzheimer's disease, as two different prototypic fields that are sufficiently specific to identify bridging rare terms related to key mechanistic evidences of the neurodegenerative pathologies. The approach is based on exploring outlier documents in the literature-based discovery process, aiming to help the expert in finding implicit relationships among concepts of two different domains of interest. The underlying assumption is that while the majority of arti-

cles in the given scientific domain describe matters related to a common understanding of the domain or more intensively investigated issues, the exploration of outliers may lead to the detection of scientifically, pharmacologically or clinically relevant bridging concepts among sets of scientific articles from two disjoint domains in a novel, not yet explored way. Additionally, it will be demonstrated that focusing on a potentially interesting subset of outlier papers increases efficiency of the knowledge discovery process by reducing the size of article corpora under investigation.

This paper is organized as follows. Section 2 introduces and describes background technologies used for creative knowledge discovery. Section 3 presents upgraded methodology for narrowing the search space that is based on using outlier documents from two distinct domains. Section 4 illustrates the application of the methodology by presenting results from two domains: “Alzheimer” and “gut microbiota”. Evaluation of the approach is provided in Section 5, discussing and comparing the obtained results with other research, providing a summary and directions for further work.

## 2 Background technologies

In this section describe background technologies that are used in our approach to bisociative cross domain literature-based discovery. First, we give an overview of a closed discovery process used in literature mining. Then, we introduce methods for detection of outlier documents in the set of documents.

### 2.1 Closed discovery with literature mining

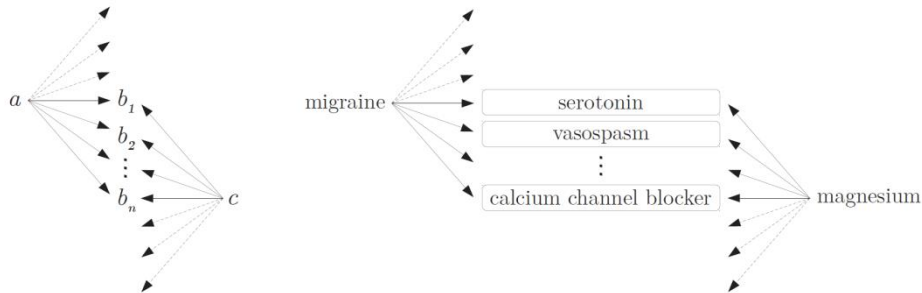
Early work in literature-based discovery (LBD) by Swanson [4] and Smalheiser et al. [5] resulted in an approach to assist the user in literature-based discovery by detecting interesting cross-domain terms with a goal to uncover the possible relations between previously unrelated concepts. The online ARROWSMITH online system, developed by Smalheiser et al. [5], takes as input two sets of titles of scientific papers from disjoint domains (disjoint document corpora)  $A$  and  $C$ , and lists terms that are common to  $A$  and  $C$ ; the resulting bridging terms  $b$  are further investigated by the user for their potential to generate new scientific hypotheses. Their approach, known as the ABC model of knowledge discovery<sup>1</sup>, addresses several settings, including the closed discovery setting, introduced by Weeber et al. [16], where two initially separate domains  $A$  and  $C$  are specified by the user at the beginning of the discovery process, and the goal is to search for bridging concepts (terms)  $b$  in  $B$  in order to support the validation of the hypothesized connection between  $A$  and  $C$ . The closed discovery setting, which is the most frequently addressed LBD setting, is illustrated in Figure 1 (the part on the left).

By studying two separate literatures, the literature on migraine headache and the articles on magnesium, Swanson [17] discovered connections supportive for the hy-

---

<sup>1</sup> In the ABC model, uppercase letter symbols  $A$ ,  $B$  and  $C$  are used to represent concepts (or sets of terms), and lowercase symbols  $a$ ,  $b$  and  $c$  to represent single terms.

pothesis that magnesium deficiency might cause migraine headache. Figure 1 illustrates the closed discovery setting on the Swanson's task of finding the terms linking the migraine and magnesium domains. This well-known example has become the gold standard in the literature mining field and has been used as a benchmark in several studies, including our own work, developed by Juršič et al. [6], Sluban et al. [8] and Petrič et al. [7], which is the basis for the methodology presented in this paper.

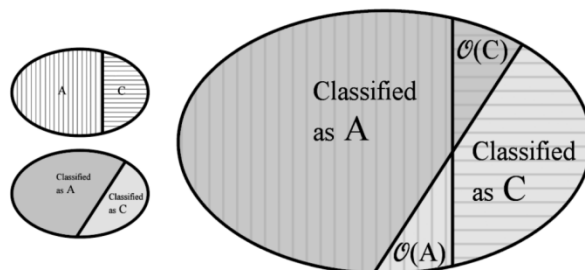


**Fig. 1.** Closed discovery process as defined by Weeber et al. [16]. The schema on the right shows a closed discovery when exploring migraine and magnesium documents, with  $b$ -terms as identified by Swanson [4].

Estimating which of the terms have a high potential for interesting discoveries is a challenging research question. Juršič et al. [6] suggested a solution in which candidate bridging terms are ranked by ensemble voting of heuristics. The methodology was implemented in the CrossBee system, an off-the-shelf solution for finding bridges between two user-defined domains/literatures. Supplementary functionalities and visualizations make CrossBee a user-friendly web application, helping the experts to efficiently investigate cross-domain links.

## 2.2 Outlier document detection

One of the techniques that can be used to detect outlier documents is by using classification algorithms [8]. The technique works as follows. Having documents from two domains of interest we first train a classification model that distinguishes between the documents of these domains. Using the constructed model we classify all the documents. The documents that are misclassified - according to their domain of origin - are declared as outlier documents, since according to the classification model they do not belong to their domain of origin. These domain outliers are actually borderline documents as they were considered by the model to be more similar to the other domain than their originating domain. In other words, if an instance of class  $A$  is classified in the opposite class  $C$ , we consider it to be an outlier of domain  $A$ , and vice versa. We denote the two sets of domain outlier documents with  $O(A)$  and  $O(C)$ , respectively. Figure 2 illustrates this principle.

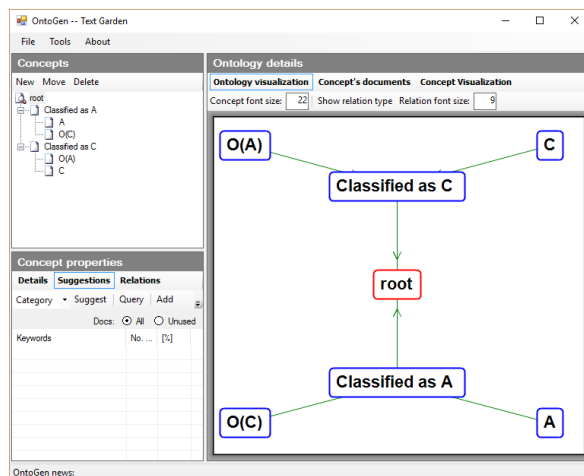


**Fig. 2.** Detecting outliers of a domain pair dataset using document classification.

In [8] we have shown that the majority of bridging terms can be found in outlier documents. We tested this on the gold standard migraine-magnesium domain pair, for which a confirmed list of concept bridging terms (B-terms) is available. The experimental results showed that the sets of detected outlier documents were relatively small - including less than 5% of the entire datasets - and that they contained a great majority of bridging terms, which was significantly higher than in same-sized random subsets. Hence the effort needed for finding cross-domain links is substantially reduced, as it requires exploring a much smaller subset of documents, where a great majority of B-terms are present and more frequent.

A separate approach to outlier document detection is by using a clustering approach. In particular, we used the OntoGen document clustering tool [14] to find outlier documents. In the approach, proposed by Petrič et al. [7], we concentrated on a specific type of outliers - the domain outliers - i.e. the documents that tend to be more similar to the documents of the opposite domain than to those of their own domain. In this approach, we use document clustering to find outlier documents. The approach consists of two steps. In the first step, the OntoGen clustering algorithm [14] is applied to cluster the merged document set  $AUC$ , consisting of documents from both domains  $A$  and  $C$ . The result of unsupervised clustering is a set of two document clusters:  $A'$  (i.e. a set of documents from  $AUC$  classified as  $A$ ), and  $C'$  (i.e. a set of documents from  $AUC$  classified as  $C$ ). Then, in the second step, for each of the clusters a supervised clustering approach is applied taking into account the documents' original domains  $A$  and  $C$ . As a result, a two-level tree hierarchy of clusters is generated, as illustrated in Figure 3.

From a joint document set  $AUC$ , a similarity graph (Figure 4) between two document sets  $A$  and  $C$  is constructed with OntoGen by ranking and visualizing all the documents from  $AC$  in terms of their similarity to centroid  $a$  of document set  $A$ . The OntoGen tool can then be used to build two document clusters,  $A'$  and  $C'$  (where  $A'UC'=AC$ ) in an unsupervised manner, using OntoGen's 2-means clustering algorithm (see Figure 3). Cluster  $A'$  (labeled *Classified as A* in Figure 3) consists mainly of documents from  $A$ , but may contain also some documents from  $C$ . Similarly, cluster  $C'$  (labeled *Classified as C* in Figure 3) consists mainly of documents from  $C$ , but may contain also some documents from  $A$ .



**Fig. 3.** Target domain documents from literatures *A* and *C*, clustered according to the OntoGen's two step approach, first using unsupervised and then supervised clustering to obtain outlier documents *O(A)* and *O(C)* of literatures *A* and *C*, respectively.

Results obtained by Sluban et al. [8] and by Petrič et al. [7] confirm the hypothesis that most bridging terms appear in outlier documents and that by considering only outlier documents the search space for B-term identification can be largely reduced. In this way, we can substantially reduce the search space for finding B-term candidates.

### 3 The upgraded methodology further narrowing the search space by outlier document detection

The particular methodology applied in this work follows our previous work in outlier detection [7] using a document clustering and exploration tool OntoGen [14]. In this approach, each document from the two literatures is an instance, represented by a set of words using frequency statistics based on the Bag of Words (BoW) text representation [18]. The BoW vector enables to measure content similarity of documents. Content similarity computation is performed with OntoGen, which was designed for interactive data-driven construction of topic ontologies [14]. Content similarity is measured using the standard TF\*IDF (term frequency inverse document frequency) weighting method [19], where high frequency of co-occurring words in documents indicates high document similarity. The similarity between documents is visualized with OntoGen in the document's similarity graph, as illustrated in Figure 4.

The cosine similarity measure, commonly used in information retrieval and text mining to determine the semantic closeness of two documents where document features are represented using the BoW vector space model, is used to position the documents according to their similarity to the representative document (centroid) of a

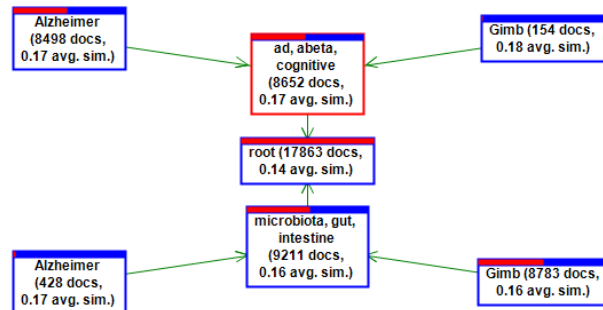
selected domain. Documents positioned based on the cosine similarity measure can be visualized in OntoGen by a similarity graph with cosine similarity values that fall within the  $[0, 1]$  interval. Value 0 means extreme dissimilarity, where two documents share no common words, while value 1 represents the similarity between two semantically identical documents in the BoW representation.

Compared to this early approach to outlier document detection using OntoGen, an upgraded method is proposed in this paper. Instead of a single step outlier detection process, used in Petrič et al. [7], we here use a two-step outlier detection process using OntoGen, illustrated in Figure 5 when applying OntoGen on the actual documents of this application. The method uses domains  $A$  and  $C$ , and builds a joint document set  $AC$  (i.e.  $AUC$ ). For this intention, two individual sets of documents (e.g. titles, abstracts or full texts of scientific articles), one for each term under research (namely, literature  $A$  and literature  $C$ ), are automatically retrieved from bibliographic databases or extracted from other document sources. The documents from the two individual sets are loaded as a single text file (i.e. a joint document set  $AC$ ) where each line represents a document with the first word in the line being its name. We consider all the terms and not just the medical ones. A list of 523 English stop words is then used to filter out meaningless words, and English Porter stemming is applied.

In the upgraded methodology, each OntoGen constructed cluster is further divided into two document sub-clusters based on domains  $A$  and  $C$  with the aim to identify outlying documents. For each individual document cluster we proceed as follows: cluster  $A'$  is divided into sub-clusters  $A' \cap A$  and  $A' \cap C$ , while cluster  $C'$  is divided into  $C' \cap A$  and  $C' \cap C$ . In this manner, sub-clusters  $A' \cap C$  (outliers of  $C$ , consisting of documents of domain  $C$  only) and  $C' \cap A$  (outliers of  $A$ , consisting of documents of domain  $A$  only) are the two document sets that consist of outlying documents.

## 4 Results

We demonstrate the utility of outlier detection to finding bisociative links between the literatures from two distinctive domains: gut microbiota and Alzheimer's disease. In the current experiment, we downloaded two input sets of documents from PubMed: 83,322 papers obtained from query "Alzheimer" and 73,960 papers obtained from query "gut microbiota"; the actual query for the later was actually more elaborate in order to cover additional synonyms: "(gut OR intestinal) AND (microbiota OR bacteria)". In the preprocessing phase we eliminated papers older than two years and papers with incomplete title or abstract. As a result, we obtained 8,934 "Alzheimer" papers and 8,937 "gut microbiota" papers. On the joint set of 17,863 papers we constructed a two-level document hierarchy with OntoGen, following the approach to identify outlier documents, proposed in [7]. At the first level, after transforming the documents into a feature vector format, the documents were clustered according to the cosine similarity into two distinct document clusters. At the second level, each of the two clusters was further separated according to the document search origin ("Alzheimer" or "gut microbiota") into two sub-clusters. The constructed hierarchy is shown in Figure 4.



**Fig. 4.** Two-level cluster hierarchy constructed from 17,863 papers with OntoGen. Two first-level clusters are labeled “ad, abeta, cognitive” and “microbiota, gut, intestine”. Four second level sub-clusters separate documents according to their original search keyword.

Based on the constructed cluster hierarchy from Figure 4 we got 582 outlier documents (428 from “Alzheimer” and 154 from “gut microbiota”). We further explored these documents in the on-line Cross-domain Bisociative Exploration Tool CrossBee. By processing the documents from the two separate domains (literatures) of interest, CrossBee extracted 4,723 terms as potential bisociative terms connecting the two domains. The terms in CrossBee were ranked according the estimated bridging term potential [15], tending to push more interesting terms to the top of the list. Even though the list of potential bridging terms is ordered according to the term’s potential, browsing and analyzing the terms from the list still presents a substantial burden for the domain expert and supportive team.

To further reduce the size of the potential bridging term list, the domain expert prepared a special list containing of 289 potentially interesting terms. The use of PubMed search engine with keywords “gut+Alzheimer” generated 42 items (on November 22, 2015). Common terms (such as dementia, ageing, neurodegeneration, inflammation, microbiome, probiotics etc.) and specific molecular factors and pathways (such as the names of transcription factors, trophic factors, proteins misfolded or aggregated, mechanisms of oxidative stress and lipid synthesis, etc.), were manually identified in title, abstracts, and keywords of such publications. 55 terms from this list appeared also among the 4,723 terms extracted by CrossBee. Among those interesting terms, the domain expert identified the term “nitric oxide synthase” [20] as a promising novel bridging term. The combined query “Alzheimer gut nitric oxide” in PubMed revealed that there are no articles that correspond to this query, indicating that this connection was not previously explored in the literature available in PubMed.

Nitric oxide synthases (NOS) are enzymes expressed at different cell levels, like in neurons (nNOS) or endothelial cells (eNOS) where they have important roles for regeneration and neuro- or vaso-active effects in different tissues. NO-mediated innervation is found in the gut nervous system [21, 22] and in the brain, NO controls the function of regions susceptible to neurodegeneration [23, 24, 25].

In Figure 5, the four articles containing the term “nitric oxide synthase” are displayed; three from the “Alzheimer” domain [26, 27, 28] and one from the “gut micro-



biota” domain [29]. By using CrossBee utility to explore bridging terms in more detail (see Figure 5), we identified a possible explanation for the bridging role of “nitric oxide synthase”. In Alzheimer literature the article [27] states: “Furthermore, FA was shown to inhibit the expression and/or activity of cytotoxic enzymes, including inducible nitric oxide synthase, caspases and cyclooxygenase-2.”, while in “gut microbiota” literature the article [29] states: “LSPC and LC combinations ameliorated scopolamine-induced memory impairment by improving total antioxidant capacity (TAOC) level, glutathione peroxidase (GSH-Px) and total superoxide dismutase (T-SOD) activities of brain, serum and colon, suppressing malondialdehyde (MDA) level of brain, serum and colon, and inhibiting brain acetylcholinesterase (AchE), myeloperoxidase, total nitric oxide synthase and neural nitric oxide synthase (nNOS) activities, and nNOS mRNA level.”. Note that for further exploration of the connection between the two domains the role of domain expert is crucial.

The screenshot displays the CrossBee interface for a B-Term analysis of "nitric oxide synthase". The main content area shows a list of documents with their titles and domains. The first document is titled "Anti-inflammatory effects of glaucocalicin B in microglia cells. Over-activated microglia is involved in various kinds of neurodegenerative process including Parkinson, Alzheimer and HD dementia" with domain "Alzheimer". The second document is "Lactobacillus casei 01 facilitates the ameliorative effects of probiotics extracted from lily seed on learning and memory impairment in scopolamine-induced amnesia mice" with domain "GIMB". The interface also includes a search bar, navigation controls, and a sidebar with options like "Start", "Downloads", "Term View", "Document View", and "B-Term".

Fig. 5. Analysis of term “nitric oxide synthase” in CrossBee: three papers from Alzheimer literature [26, 27, 28] and one from Gut microbiota literature [29].

## 5 Discussion and conclusion

The interplay between the two tools OntoGen and CrossBee is the novelty of the proposed approach. Apart from constructing the document clustering from the input set of documents and finding the keywords describing the two document classes, Ontogen is used to narrow down the search for bridging terms in CrossBee by identifying the set of outlier documents. In such way, the search in CrossBee can be more focused, efficient and effective. This opens the door for many new applications in which the size of the search space is a severe limiting factor.

The role of the expert from the biomedical field was crucial in the document exploration process. Besides providing a list of potentially interesting terms for further

exploration, interdisciplinary collaboration allowed for efficient investigation of bridging terms suggested by CrossBee. Among them, in our first experiment we identified and decided to focus on the term “nitric oxide synthase” [20] as a promising novel bridging term, interesting in terms of mechanistic suggestions and drug-discovery new potential [30, 31, 32]. From the set of articles that were used as input to CrossBee the expert carefully reviewed the four articles containing this term; three from the “Alzheimer” domain [26, 27, 28] and one from the “gut microbiota” domain [29]. In addition, the combined query “Alzheimer gut nitric oxide” in PubMed revealed that there are no articles that correspond to this query, which is especially interesting, indicating that this connection was not previously explored in the literature.

Interpretation of pointing ‘nitric oxide synthase’ (i.e. NOS) with current literature knowledge, can be interpreted in view of microbiome contribution in NOS-mediated inflammation at gut level [33, 29, 13]. Although not yet demonstrated, it is likely that such effects dysregulate the brain-gut communication. In pathological conditions, such as during inflammation or in the presence of environmental stressors or ageing, abnormal iNOS function results in oxidative effects and neurodegeneration. In particular, in the gut, iNOS induces intestinal barrier damage [34], and in the brain causes nitrosylation of proteins and cell death with neurological consequences like dementia, Alzheimer or Parkinson diseases [35, 36]. Even though NO is locally produced, its effects can diffuse systemically, thus influencing the progression of the disease [26, 27, 28].

Due to the rapid grow of scientific literature IT represents a useful support to solve complex interdisciplinary questions. The identification of cross-context links leading to new insights and discoveries is not an easy task due to the huge search space of possibilities. In this paper we suggest a methodology that combines two possible approaches to overcome this problem. First it identifies the parts of the search space with increased probability of finding good candidate terms/concepts thus restricting the huge amounts of existing literature to more manageable sources to be explored, and then it estimates the potential of candidate links for new discoveries, enabling the user to concentrate on the most promising ones.

The potentialities of personalized diet and microbiome-based therapeutics are of great interest. Bioactive nutrients possibly modulate individual microbiome responses limiting inflammation and free radical synthesis, including NO [37]. Although further studies are necessary to clarify the consequences of pathological NO signaling in different tissues, the study of NO/iNOS-targeted therapeutic strategies might have a clinical benefit [38]. Our study suggests that a such validated IT process identify new common molecular targets that can be used as multi-purpose drugs, able to block multiple processes such as oxidative and inflammatory effects with high neuro-protection relevance for peripheral and central neurons.

## Literature

1. Gubiani D, Petrič I, Fabbretti E, Urbančič T: Mining scientific literature about ageing to support better understanding and treatment of degenerative diseases. In: Mladenić, D.,

- Grobelnik, M. (eds). Data mining and data warehouses (SiKDD 2015): Proceedings of the 18th International Multiconference Information Society - IS 2015, October 5th, 2015, Ljubljana, Slovenia: volume E. Ljubljana: Institut Jožef Stefan, (2015)
2. Ghaisas S, Maher J, Kanthasamy A: Gut microbiome in health and disease: Linking the microbiome-gut-brain axis and environmental factors in the pathogenesis of systemic and neurodegenerative diseases. *Pharmacol Ther* pii: S0163-7258(15)00225-9 (2015)
  3. Bruza P, Weeber M (eds.): *Literature-Based Discovery*. Springer Science & Business Media (2008)
  4. Swanson DR: Medical literature as a potential source of new knowledge. *Bulletin of the Medical Library Association* 78(1), 29 (1990)
  5. Smalheiser N, Swanson DR, et al.: Using ARROWSMITH: a computer-assisted approach to formulating and assessing scientific hypotheses. *Computer Methods and Programs in Biomedicine* 57(3), 149–154 (1998)
  6. Juršič M, Cestnik B, Urbančič T, Lavrač N: Bisociative literature mining by ensemble heuristics. In: *Bisociative Knowledge Discovery*, pp. 338–358. Springer (2012)
  7. Petrič I, Cestnik B, Lavrač N, Urbančič T: Outlier Detection in Cross-Context Link Discovery for Creative Literature Mining. *The Computer Journal.*; 55(1): 47-61, (2012)
  8. Sluban B, Juršič M, Cestnik B, Lavrač N: Exploring the power of outliers for crossdomain literature mining. In: *Bisociative Knowledge Discovery*. pp. 325–337. Springer (2012)
  9. Ji B, Nielsen J: From next-generation sequencing to systematic modeling of the gut microbiome. *Front Genet* 6:219 (2015)
  10. Peterson CT, Sharma V, Elmén L, Peterson SN: Immune homeostasis, dysbiosis and therapeutic modulation of the gut microbiota. *Clin Exp Immunol* 179(3):363-377 (2015)
  11. Tiso M, Schechter AN: Nitrate reduction to nitrite, nitric oxide and ammonia by gut bacteria under physiological conditions. *PLoS One* 10(3):e0119712. (2015)
  12. Rhee SH, Pothoulakis C, Mayer EA: Principles and clinical implications of the brain-gut-enteric microbiota axis. *Nat Rev Gastroenterol Hepatol* 6(5):306-314 (2009).
  13. Derkinderen P, Rouaud T, Lebouvier T, Bruley des Varannes S, Neunlist M, De Giorgio R: Parkinson disease: the enteric nervous system spills its guts. *Neurology* 77(19):1761-1767 (2011)
  14. Fortuna B, Grobelnik M, Mladenčić D: Semi-automatic data-driven ontology construction system. In: Bohanec M, Gams M, Rajkovič V, Urbančič T, Bernik M, Mladenčić D, Grobelnik M, Heričko M, Kordeš U, Markič O, Musek J, Osredkar M J, Kononenko I, and Novak Škarja B (eds) *Proceedings of the 9th International multi-conference Information Society*. Ljubljana, Slovenia. pp 223-226, (2006)
  15. Juršič M, Cestnik B, Urbančič T, Lavrač N: Cross-domain literature mining: Finding bridging concepts with CrossBee. In: *Proceedings of the 3rd International Conference on Computational Creativity*. Pp 33–40; University College Dublin, Dublin, Ireland, (2012)
  16. Weeber M, Klein H, de Jong-van den Berg L, Vos R, et al.: Using concepts in literature-based discovery: Simulating Swanson’s Raynaud–fish oil and migraine–magnesium discoveries. *Journal of the American Society for Information Science and Technology* 52(7), 548–557 (2001)
  17. Swanson DR: Migraine and magnesium: Eleven neglected connections. *Perspectives in Biology and Medicine* 78(1), 526–557 (1988)
  18. Sebastiani F: Machine learning in automated text categorization. *ACM Comput. Surv.* 34(1), pp. 1–47, (2002)
  19. Salton G, Buckley C: Term Weighting Approaches in Automatic Text Retrieval. *Information Processing and Management* 24(5), pp. 513–523, (1988)

20. Katusic ZS, Austin SA: Endothelial nitric oxide: protector of a healthy mind. *European Heart Journal*, 35(14): 888-894, (2014)
21. Rivera LR, Poole DP, Thacker M, Furness JB: The involvement of nitric oxide synthase neurons in enteric neuropathies. *Neurogastroenterol Motil* 23(11):980-988 (2011)
22. Phillips RJ, Powley TL: Innervation of the gastrointestinal tract: patterns of aging. *Auton Neurosci* 36(1-2):1-19 (2007)
23. Toda N, Okamura T: Cerebral blood flow regulation by nitric oxide in Alzheimer's disease. *J Alzheimers Dis.* 32(3):569-78 (2012)
24. Blomeley CP, Cains S, Bracci E: Dual Nitroergic/Cholinergic Control of Short-Term Plasticity of Corticostriatal Inputs to Striatal Projection Neurons. *Front Cell Neurosci* 9:453 (2015)
25. Steinert JR, Chernova T, Forsythe ID: Nitric oxide signaling in brain function, dysfunction, and dementia. *Neuroscientist* 16(4):435-452 (2010)
26. Gan P, Zhang L, Chen Y, Zhang Y, Zhang F, Zhou X, Zhang X, Gao B, Zhen X, Zhang J, Zheng LT: Anti-inflammatory effects of glaucocalyxin B in microglia cells. *J Pharmacol Sci.* 2015 May; 128(1): 35-46, (2015)
27. Mancuso C, Santangelo R: Ferulic acid: pharmacological and toxicological aspects. *Food Chem Toxicol.* 2014 Mar; 65, (2014)
28. Rannikko EH, Weber SS, Kahle PJ: Exogenous  $\alpha$ -synuclein induces toll-like receptor 4 dependent inflammatory responses in astrocytes. *BMC Neurosci.* 2015 Sep 7;16:57. doi: 10.1186/s12868-015-0192-0, (2015)
29. Xiao J, Li S, Sui Y, Wu Q, Li X, Xie B, Zhang M, Sun Z: Lactobacillus casei-01 facilitates the ameliorative effects of proanthocyanidins extracted from lotus seedpod on learning and memory impairment in scopolamine-induced amnesia mice. *PLoS One.* 9(11), (2014)
30. Pierucci M, Galati S, Valentino M, Di Matteo V, Benigno A, Pitruzzella A, Muscat R, Di Giovanni G: Nitric oxide modulation of the basal ganglia circuitry: therapeutic implication for Parkinson's disease and other motor disorders. *CNS Neurol Disord Drug Targets* 10(7):777-791 (2011)
31. Dijkstra G, van Goor H, Jansen PL, Moshage H: Targeting nitric oxide in the gastrointestinal tract. *Curr Opin Investig Drugs* 5(5):529-36 (2004)
32. Zhihui Q: Modulating nitric oxide signaling in the CNS for Alzheimer's disease therapy. *Future Med Chem* 5(12):1451-1468 (2013)
33. Baruch K, Kertser A, Porat Z, Schwartz M: Cerebral nitric oxide represses choroid plexus NF B-dependent gateway activity for leukocyte trafficking. *The EMBO Journal.* 2015; 34(13): 1816-1828, (2015)
34. Grishin A, Bowling J, Bell B, Wang J, Ford HR: Roles of nitric oxide and intestinal microbiota in the pathogenesis of necrotizing enterocolitis. *J Pediatr Surg* 51(1):13-7 (2016)
35. Hess DT, Matsumoto A, Kim SO, Marshall HE, Stamler JS: Protein S-nitrosylation: purview and parameters. *Nat Rev Mol Cell Biol* 6(2):150-166 (2005)
36. Horn TF, Wolf G, Duffy S, Weiss S, Keilhoff G, MacVicar BA: Nitric oxide promotes intracellular calcium release from mitochondria in striatal neurons. *FASEB J* 16(12):1611-1622 (2002)
37. Jeong JJ, Woo JY, Ahn YT, Shim JH, Huh CS, Im SH, Han MJ, Kim DH: The probiotic mixture IRT5 ameliorates age-dependent colitis in rats. *Int Immunopharmacol.* 26(2):416-22. (2015)
38. Broom L, Marinova-Mutafchieva L, Sadeghian M, Davis JB, Medhurst AD, Dexter DT: Neuroprotection by the selective iNOS inhibitor GW274150 in a model of Parkinson disease. *Free Radic Biol Med* 50(5):633-640 (2011).