

Finding Seeds of Future Discoveries in Current Literature

Tanja Urbančič

University of Nova Gorica, Nova Gorica, Slovenia and
Jozef Stefan Institute, Ljubljana, Slovenia

Ingrid Petrič

University of Nova Gorica, Nova Gorica, Slovenia

Nada Lavrač, Bojan Cestnik, Matjaž Juršič, Borut Sluban

Jozef Stefan Institute, Ljubljana, Slovenia

Marta Macedoni-Lukšič

University Children's Hospital, Ljubljana, Slovenia (previous),
Institute for Autism (now)

Outline

- **Part 1: Introduction**
 - Creativity and knowledge discovery
 - Literature as a source of cross-context knowledge discoveries
 - Open vs. closed discovery
- **Part 2: Open discovery framework**
 - Exploring rarity (RaJoLink)
 - Case 1: RaJoLink in the Migraine domain
 - Case 2: Towards better understanding of Autism
- **Part 3: Closed discovery framework**
 - Speeding-up the process:
exploring outliers, ranking candidate bridging terms, workflows, HCI (CrossBee)
- **Part 4: What next?**

“What could be learned from looking at nothing?”

(F. Slakey)



- **Hubble Space Telescope**

- Carried into orbit by a space shuttle in 1990
- Positioned above the Earth’s atmosphere
- Extremely high-resolution images
- Many Hubble observations led to breakthroughs in astrophysics



- **Hubble Deep Field**

- Image assembled from 342 separate exposures taken over 10 consecutive days (18-28 Dec 1995)
- Focused on a very small region in the constellation of Ursa Major (“an empty black void”)
- The picture revealed over 1,500 galaxies and became a landmark image in the study of the early universe

Creativity and Discoveries

- Boden (The Creative Mind – Myths and Mechanisms, 2004):

Creativity as “the ability to come up with ideas or artifacts that are new, surprising and valuable”

Three types of creativity: combinatorial, exploratory, transformational

- In many discoveries, existing knowledge is creatively combined with new ideas.

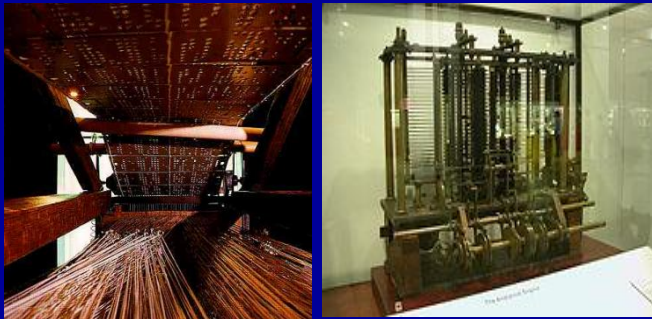
- Koestler (The Act of Creation, 1964):

“Creative act uncovers, selects, re-shuffles, combines, synthesizes already existing facts, ideas, faculties, skills. The more familiar the parts, the more striking the new whole.”

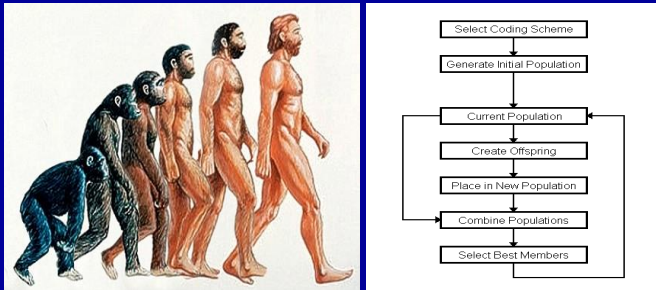
- Ideas often come from other contexts



Examples from the history of computer science



- **From Jacquard's loom (1801) to Babbage's Analytic Engine (1837)**
 - Idea of punched cards representing data, problems etc. (instead of colorful textile patterns)

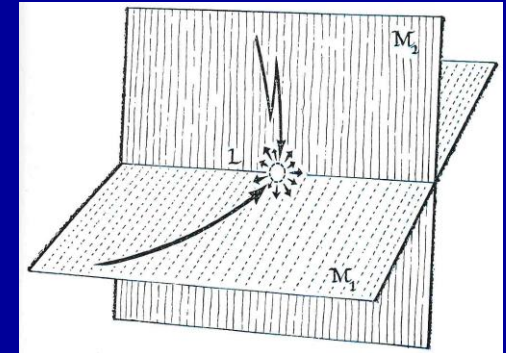


- **From evolution in nature to evolutionary computing (Lawrence J. Fogel, 1964)**
 - Idea of populations of candidate solutions developing through simulated evolution (instead of “survival of the fittest” in nature)

Cross-context link discovery

- **Bisociations as a basis of human creativity**

“... the perceiving of a situation or idea, L , in two self-consistent but habitually incompatible frames of reference, M_1 and M_2 . The event L ... is not merely linked to one associative context but bisociated with two.” (Koestler, 1964)

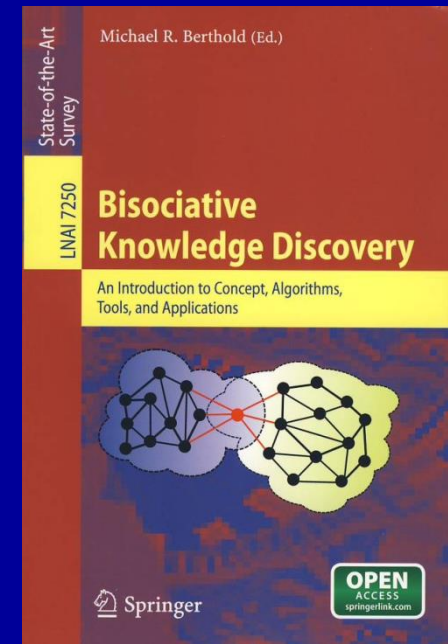


- **Finding links which lead “out-of-the-plane”**

- BISON: Bisociation Networks for Creative Information Discovery, 7FP project (2008-2010)

- **Two concepts are bisociated if and only if:**

- There is **no direct, obvious evidence** linking them
- One has to **cross contexts** to find the link
- This new link provides some **novel insight**

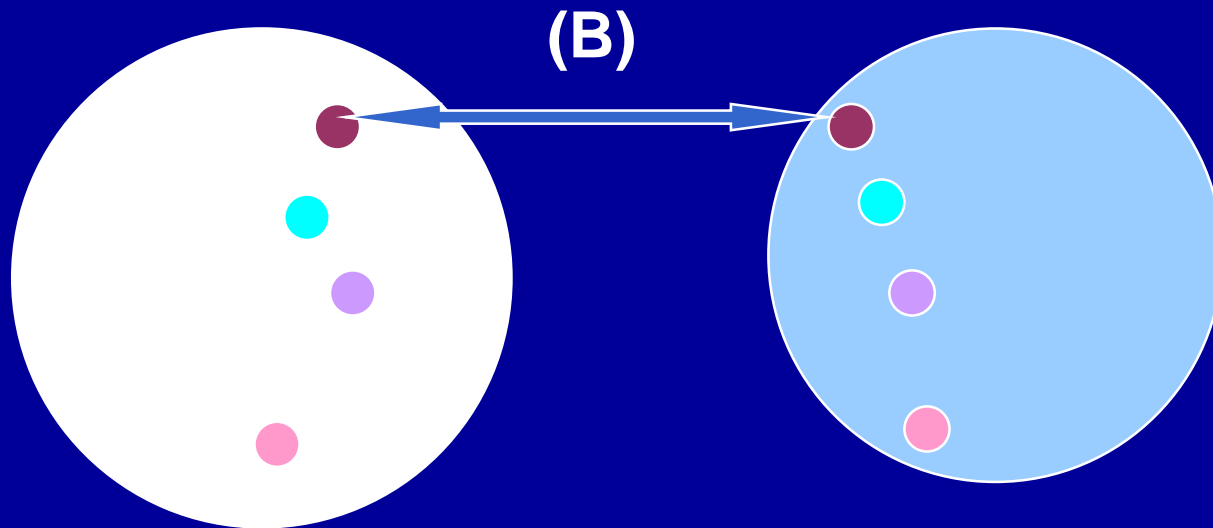


Cross-context link discovery from literature

Swanson, Medical literature as a potential source of new knowledge, 1990

Literature about
magnesium (A)
(38.000 articles)

Literature about
migraine (C)
(4.600 articles)



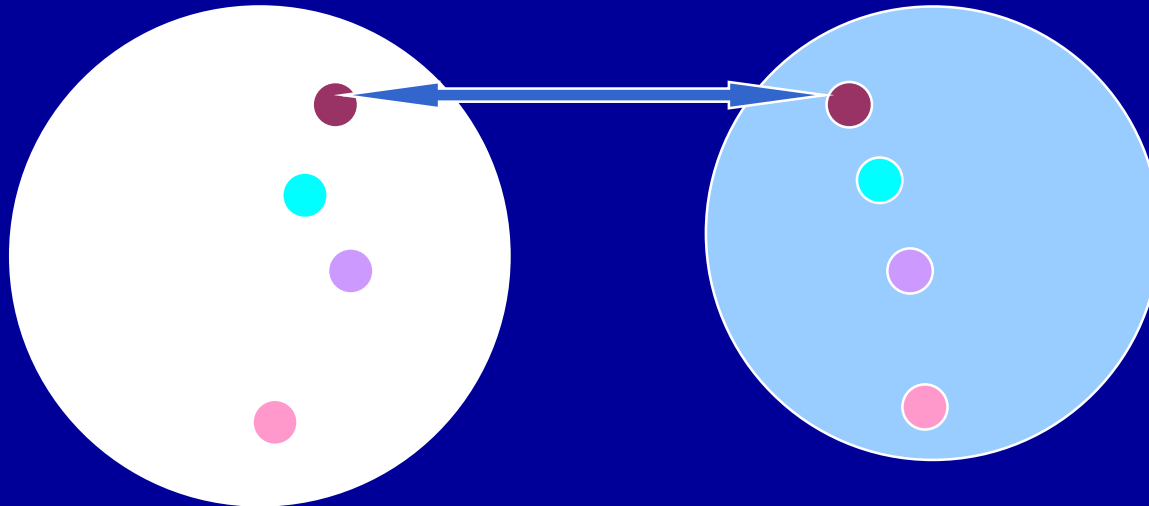
Swanson's ABC model

B: Linking (bridging) terms

Literature about
magnesium (A)

Literature about
migraine (C)

B-terms: calcium channel blockers, stress and A type behavior, ...



Argument 1 (magnesium literature)

- Mg is a natural **calcium channel blocker**.
- **Stress and Type A behavior** can lead to body loss of Mg.
- Magnesium has **anti-inflammatory** properties.
- ...

Argument 2 (migraine literature)

- **Calcium channel blockers** can prevent migraine attacks.
- **Stress and Type A behavior** are associated with migraine.
- Migraine may involve sterile **inflammation** of the cerebral blood vessels.
- ...

Generation of a hypothesis
A may influence C

For a given C, how do we find A?

Swanson:

“Search proceeds via some intermediate literature (B) toward an unknown destination A. ... Success depends entirely on the knowledge and ingenuity of the searcher.”

Can we provide a more systematic support?

Scientific literature as a potential source of new knowledge: What do we have at disposal?

Example:

- **Biomedical bibliographical database PubMed**
- **US National Library of Medicine**
- **More than 21M citations**
- **More than 5.600 journals**
- **2.000 – 4.000 references added each working day!**



The screenshot displays the PubMed website interface. At the top, the NCBI logo and 'PubMed' branding are visible, along with the text 'A service of the National Library of Medicine and the National Institutes of Health' and the URL 'www.pubmed.gov'. A search bar contains the query 'autism', with 'Go', 'Clear', and 'Save Search' buttons. Below the search bar, there are tabs for 'Limits', 'Preview/Index', 'History', 'Clipboard', and 'Details'. The 'Display' dropdown is set to 'Summary', and 'Show' is set to '500'. The search results show 'All: 11008' and 'Review: 1632'. The first four results are listed:

- 1: [Fazzi E, Rossi M, Signorini S, Rossi G, Bianchi PE, Lani G.](#) **Leber's congenital amaurosis: is there an autistic component?** *Dev Med Child Neurol.* 2007 Jul;49(7):503-7. PMID: 17593121 [PubMed - in process]
- 2: [Pava B, Fuentes N.](#) **Neurobiology of autism: neuropathology and neuroimaging studies.** *Actas Esp Psiquiatr.* 2007 Jul-Aug;35(4):271-6. PMID: 17592791 [PubMed - in process]
- 3: [Havashi ML, Rao BS, Seo JS, Choi HS, Dolan BM, Choi SY, Chattarji S, Tonegawa S.](#) **Inhibition of p21-activated kinase rescues symptoms of fragile X syndrome in mice.** *Proc Natl Acad Sci U S A.* 2007 Jun 25; [Epub ahead of print] PMID: 17592139 [PubMed - as supplied by publisher]
- 4: [Scheeren AM, Stauder JE.](#) **Broader Autism Phenotype in Parents of Autistic Children: Reality or Myth?** *J Autism Dev Disord.* 2007 Jun 23; [Epub ahead of print] PMID: 17588199 [PubMed - as supplied by publisher]

Do we use such sources of knowledge effectively?

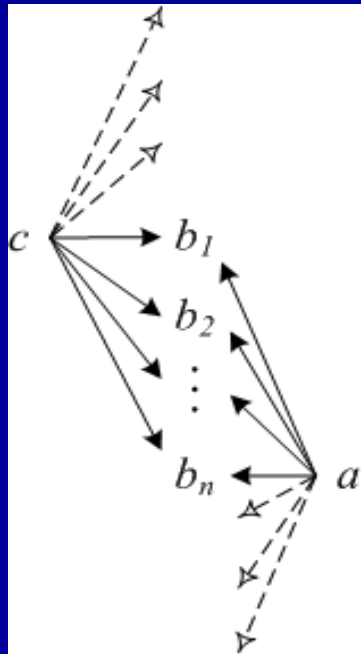
- **Situation:**
 - Speed of knowledge growth, huge amounts of literature available on-line
 - High specialization of researchers
 - Potentially useful connections between “islands” of knowledge may remain hidden
- **Our objective:**
 - To develop methods and SW tools to support researchers in their knowledge discovery

Closed vs. open discovery

(as defined by Weeber et al., 2001)

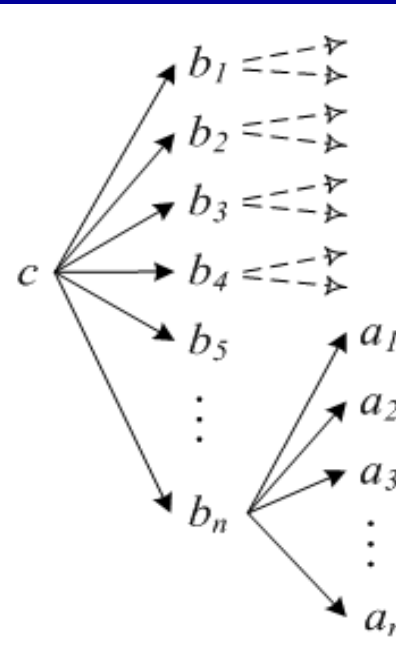
Closed discovery

C and A known in advance



Open discovery

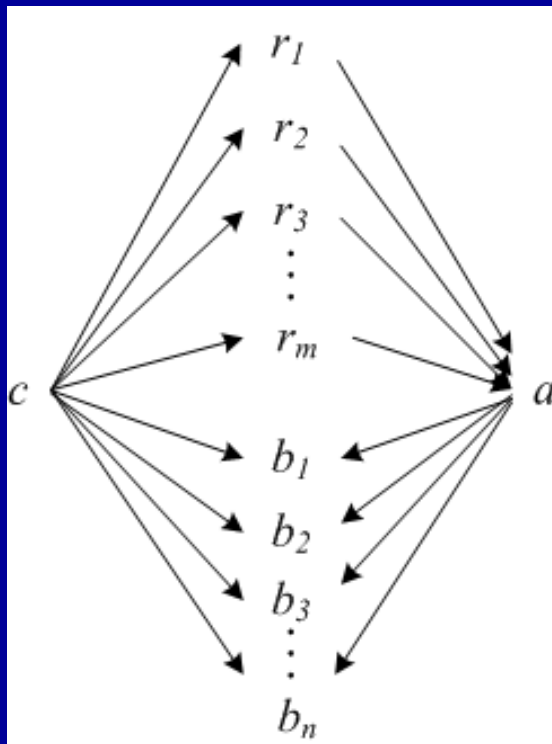
Only C known in advance



Outline

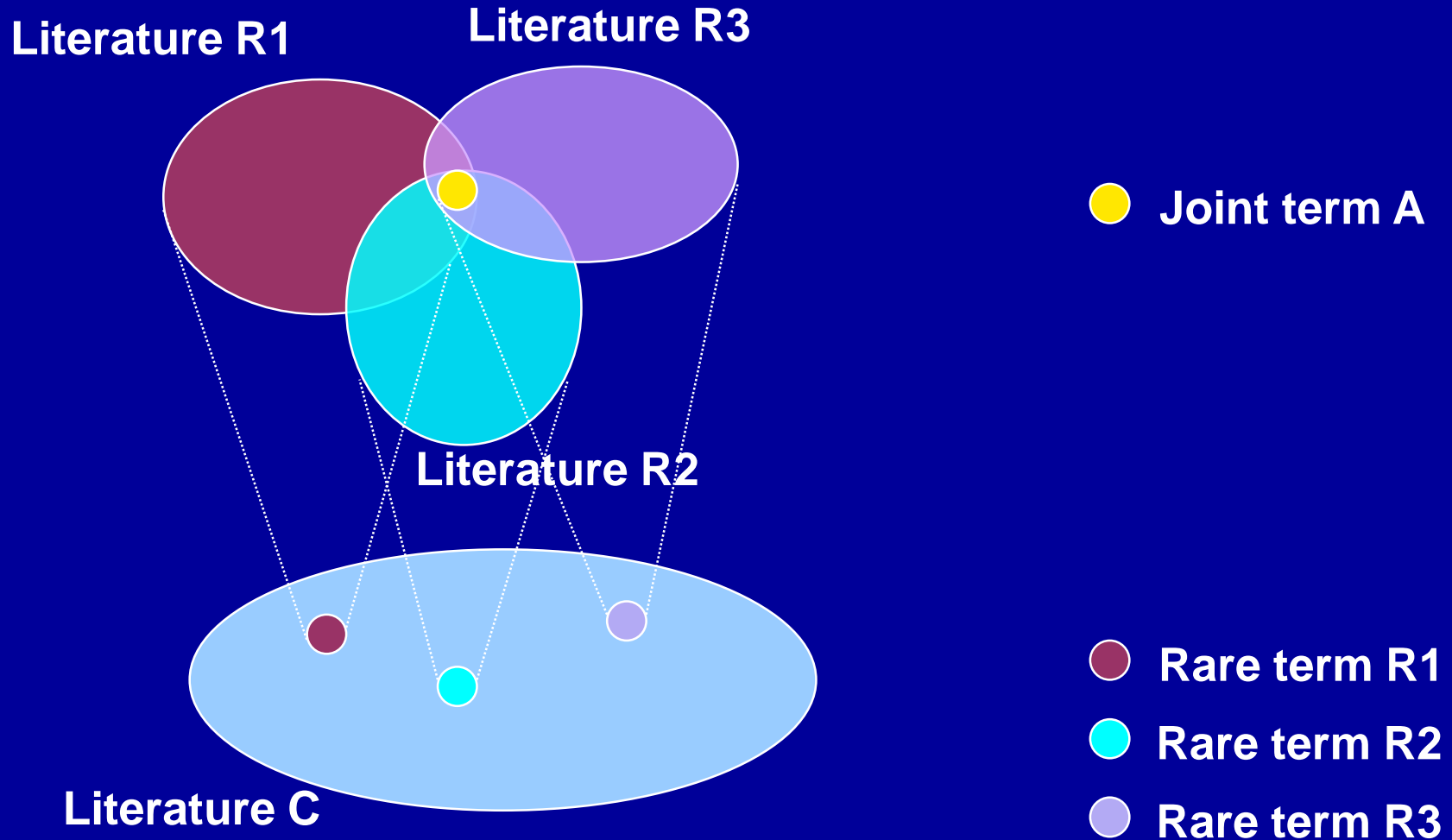
- **Part 1: Introduction**
 - Creativity and knowledge discovery
 - Literature as a source of cross-context knowledge discoveries
 - Open vs. closed discovery
- **Part 2: Open discovery framework**
 - Exploring rarity (RaJoLink)
 - Case 1: RaJoLink in the Migraine domain
 - Story 2: Towards better understanding of Autism
- **Part 3: Closed discovery framework**
 - Speeding-up the process:
exploring outliers, ranking candidate bridging terms, workflows, HCI (CrossBee)
- **Part 4: What next?**

Combined open and closed discovery in RaJoLink (Petrič, Urbančič, Cestnik, Macedoni-Lukšič, 2009)



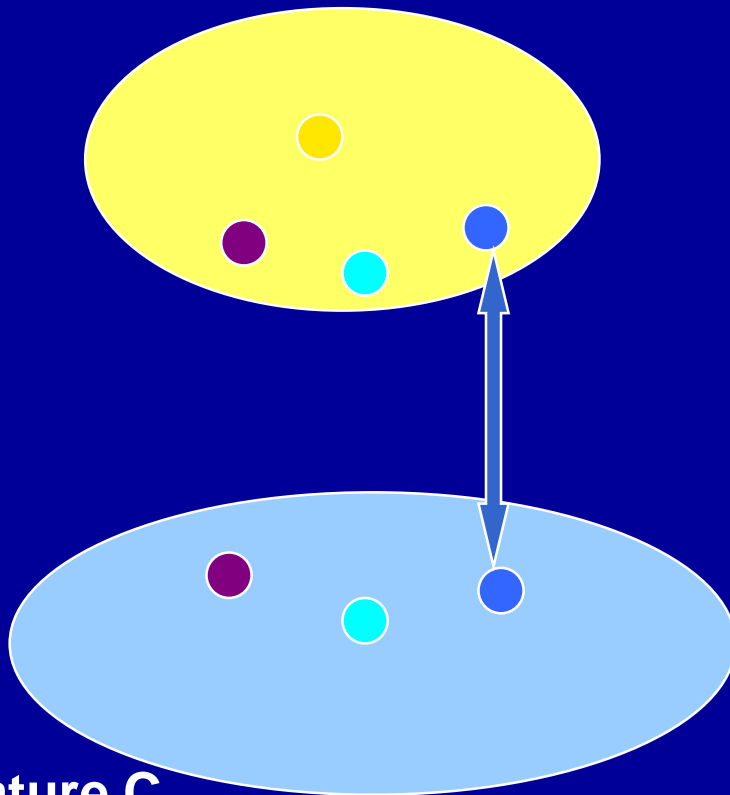
- **Open discovery**
(upper part of the figure):
 - Identifying **Rare** terms r
 - Finding a **Joint** term a
- **Closed discovery**
(lower part of the figure):
 - Searching for **Linking** terms b

RaJoLink method: Idea (open discovery part)



RaJoLink method: Idea (closed discovery part)

Literature A



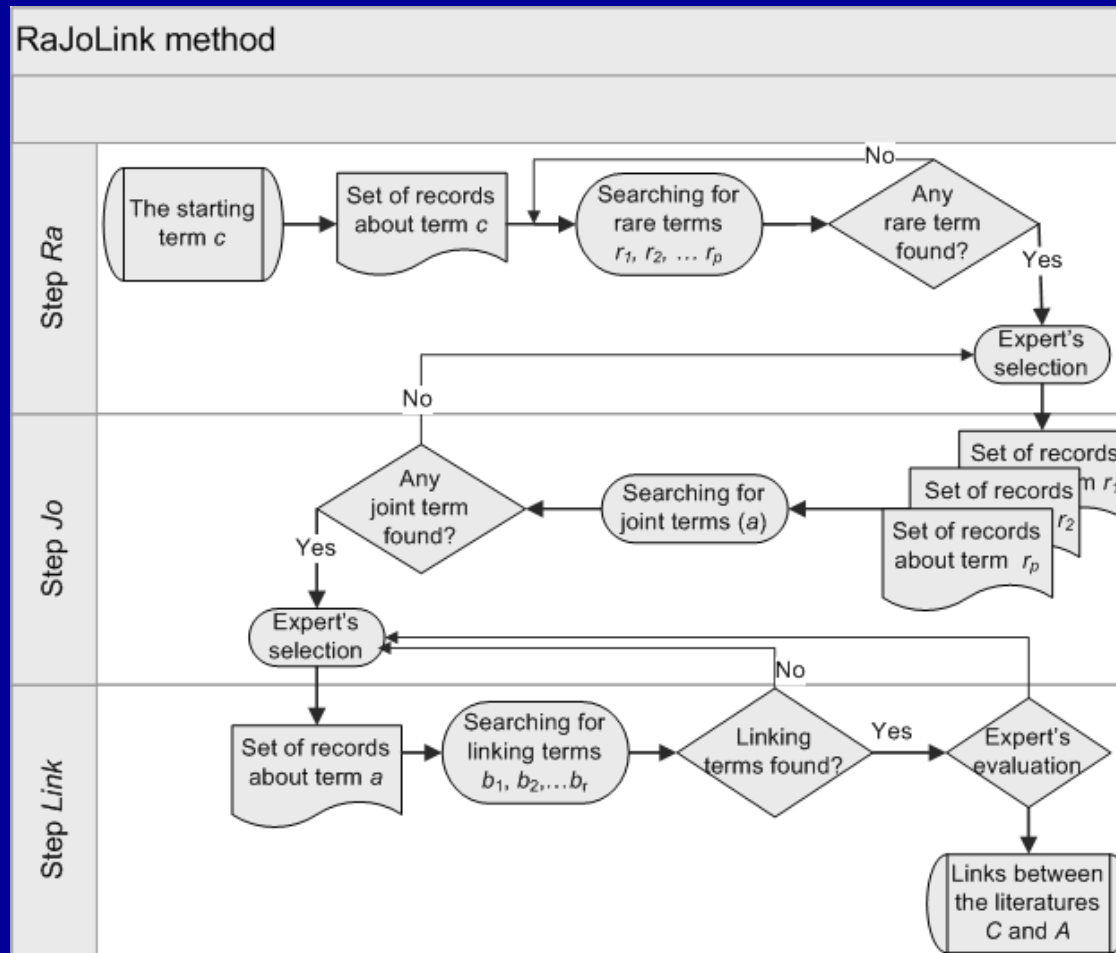
● Joint term A

● Linking term B1

● Linking term B2

● Linking term B3

Procedures of the RaJoLink



Case 1: “Re-investigating” migraine

- **Problem: migraine headaches**
- **The golden standard evaluation**
- **Terms extracted from medical articles published before 1988**
- **Generation of hypotheses related to migraine domain with the support of RaJoLink tool**

Results in migraine experiment

Step Jo

File Help

Previous step Next step

Search for

- VILOXAZINE
- VESICULAR
- VASCULOGENIC
- URACIL
- TYMPANIC
- TUBERCULIN

Retrieve

100 Titles

Go

Input set of records

[Circadian rhythm of plasma norepinephrine
Fatal toxicity of antidepressant drugs in ove
Viloxazine versus imipramine in the treatmen
Administered viloxazine interferes in liquid-ch
Cerebral activating properties of indeloxazin
Viloxazine in depressed women: clinical resp
Effects of indeloxazine hydrochloride (YM-DE
3-[(2-ethoxyphenoxy)methyl]piperidine deri
Effect of some tricyclic and nontricyclic antid
The automated Tail Suspension Test: a comp
Effect of antidepressant and neuroleptic dru
Viloxazine as a betamimetic antidepressant c
Activation and desensitization of presynaptic
[Modeling neurosis in the dog and a study of
Carbamazepine-viloxazine interaction in pati
Effects of antidepressants on histamine H2 r
[Inhibition of noradrenaline deamination by a
Does viloxazine really improve sex drive? A c
Inhibition of monoamine oxidase by viloxazin
Theophylline intoxication following viloxazine
[Comparative evaluation of the therapeutic
Viloxazine hydrochloride in narcolepsy: a pre
Age, therapeutic "milieu" and clinical outcom
Narcolepsy 1985.
Pharmacokinetics of the antidepressant drug
Effects of antidepressants on receptor-activ
The stereoselectivity of serotonin uptake in l

Results

	Term	Frequencies	Sum of Fq	MeSH codes	A & C
<input type="checkbox"/>	27 STABILITY	17:0,0,0,2,...	21	C23:E05:E...	0
<input type="checkbox"/>	40 ESCHERICHIA	15:0,0,0,1,...	30	B03:B04:C...	0
<input type="checkbox"/>	42 COLI	15:0,0,0,3,...	35	B03:B04:C01	0
<input type="checkbox"/>	47 HEPATOCYTE	14:0,0,0,1,...	22	A11:D08:D12	0
<input type="checkbox"/>	64 MUTANT	13:0,2,0,0,...	20	B01:C11:D12	0
<input type="checkbox"/>	69 CYTOCHROME	13:0,0,0,0,...	25	C16:D05:D08	0
<input type="checkbox"/>	83 CLONE	12:0,0,0,0,...	20	A11	0
<input type="checkbox"/>	84 ANTITUMOR	12:0,0,0,0,...	12	D27:E05	0
<input type="checkbox"/>	88 SUSPENSION	11:2,1,0,0,...	13	D03:E01:E05	0
<input type="checkbox"/>	92 RADICAL	11:0,0,1,1,...	18	D01:D03:D...	0
<input type="checkbox"/>	97 MICROSCOPIC	11:0,0,1,0,...	14	C06:E01	0
<input type="checkbox"/>	114 Y	10:0,0,0,0,...	12	A05:A11:B...	0
<input type="checkbox"/>	115 SULFUR	10:0,0,0,0,...	11	B03:D01:D...	0
<input type="checkbox"/>	131 MICROsome	10:0,0,0,0,...	11	A11	0
<input type="checkbox"/>	132 MATRIX	10:0,1,0,0,...	10	A05:A10:A...	0
<input type="checkbox"/>	138 EPIDERMAL	10:0,0,0,0,...	15	A11:C04:C...	0
<input type="checkbox"/>	143 CIS	10:0,0,0,1,...	12	D01:D02	0
<input type="checkbox"/>	145 AQUEOUS	10:0,0,0,1,...	12	A07:A09:D...	0
<input type="checkbox"/>	152 RECOMBINANT	9:0,2,0,1,0,...	14	D06:D12:D...	0
<input type="checkbox"/>	155 NATRIURETIC	9:0,0,0,1,0,...	12	D06:D08:D27	0
<input type="checkbox"/>	160 MAGNESIUM	9:0,0,0,0,0,...	9	C18:D01:D...	0
<input type="checkbox"/>	161 LYMPHOID	9:0,0,0,0,0,...	9	A10:A11:C...	0
<input type="checkbox"/>	166 INTERFERON	9:0,7,0,1,0,...	24	D12:D27	0
<input type="checkbox"/>	169 FRAGMENT	9:0,0,0,0,0,...	12	D08:D12:E...	0
<input type="checkbox"/>	182 VA	8:1,1,0,0,0,...	17	D02:D12	0
<input type="checkbox"/>	183 UNSTABLE	8:0,0,0,0,0,...	17	C14:G14	0
<input type="checkbox"/>	189 SUPEROXIDE	8:0,0,0,0,0,...	12	D01:D02:D08	0
<input type="checkbox"/>	198 PROTON	8:0,1,0,2,0,...	10	D01:D08:D...	0
<input type="checkbox"/>	203 PHOSPHODIESTERASE	8:0,0,0,0,0,...	9	D08:D12:D27	0
<input type="checkbox"/>	209 INTERLEUKIN	8:0,0,0,0,0,...	16	D08:D12	0
<input type="checkbox"/>	211 INDICTRIE	8:0,0,0,0,0...	10	D08:D12	0

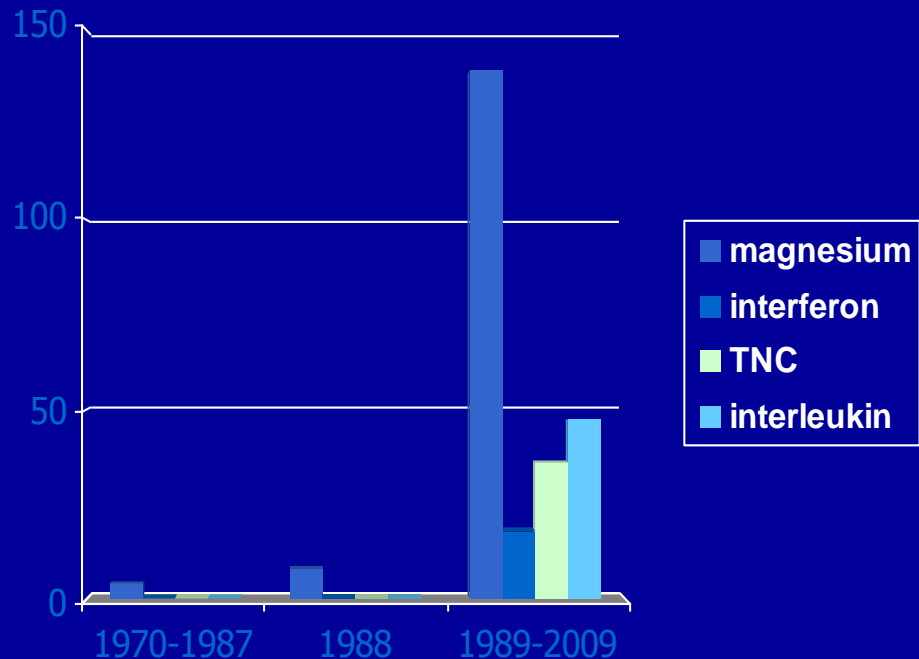
Number of target terms: 3491 All terms: 8455

Results in migraine experiment

- **Magnesium**
“re-discovered”
- **Other 3 important connections identified with RaJoLink:**
 - **interferon**
 - **interleukin**
 - **tumor necrosis factor**

Their relation with migraine appeared in PubMed after 1988!

Number of publications



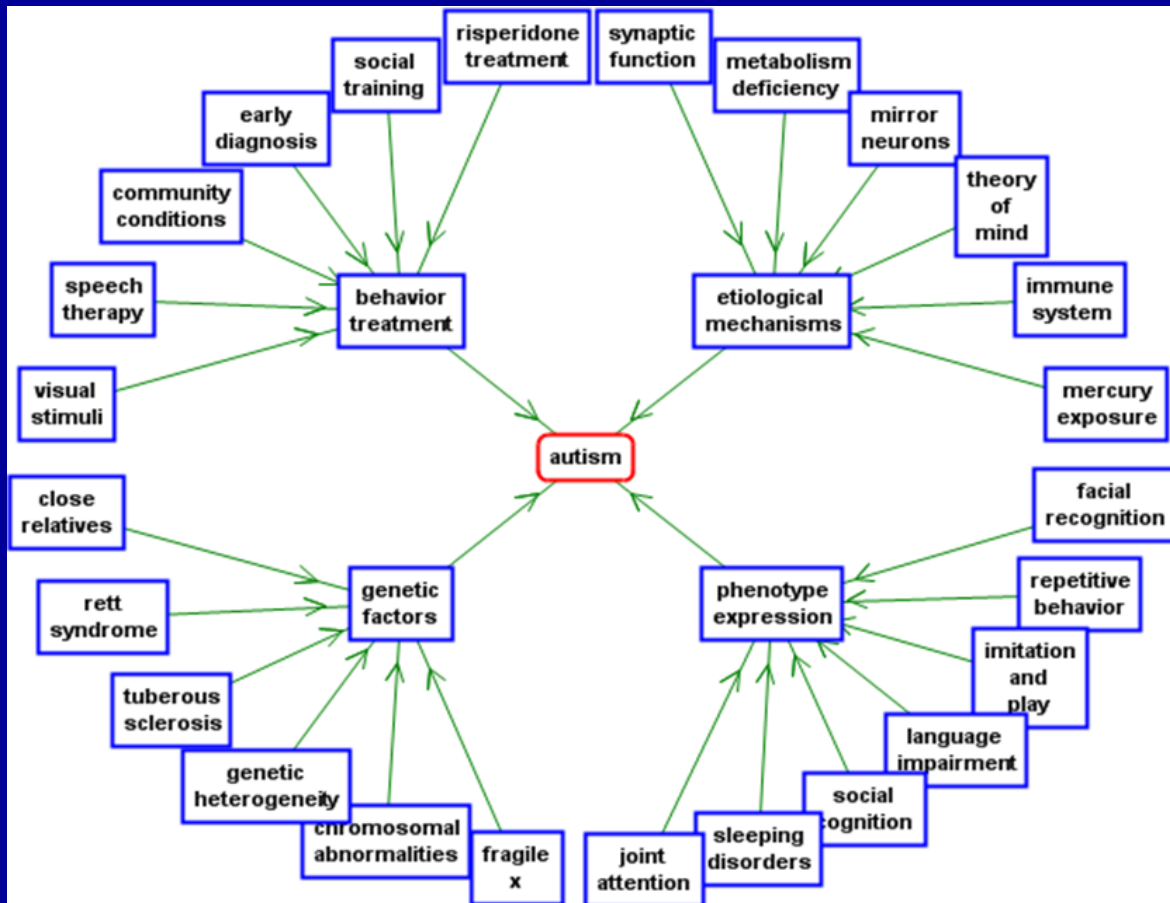
Case 2: Autistic spectrum disorders

- Abnormal development of cognitive, communication and social interaction skills
- Estimated prevalence 5.8 / 1000, increasing
- Causal explanation in 10-15% of cases
- Complex connections of genetic and environmental factors
- Research fragmented (behavioral psychology, genetics, biochemistry, brain anatomy and physiology, ...)
- Early diagnosis and treatment very important



Ontology of autism domain

- Tool: **OntoGen** (Fortuna, Grobelnik, Mladenić, 2006)
- Designed for construction of topic ontologies
- Clustering algorithms used for topic suggestion



RaJoLink in autism domain

- **Data source: PubMed** <http://www.ncbi.nlm.nih.gov/sites/entrez>
 - More than 10.000 documents with autism*
 - 354 articles with entire text in PubMed
 - 214 of them published in the last 10 years
- **RaJoLink experiments**
 - Tool: **RaJoLink** implements RaJoLink method (Petrič et al. 2007, Urbančič et al. 2007, Petrič et al. 2009)
 - Cooperation with medical expert

RaJoLink: Step Ra

RaJoLink File Help

Next step

Search for: autism

Retrieve: 1000 Abstracts

Before: 3 12 2008

Go

Number of all articles: 13187

Input set of records

less use of immature psychological defense r
 A major reason for the slow progress in iden
 in in power to detect disease genes over the
 The classification of autism spectrum disorde

Whether children with autistic spectrum disoi
 Associations between obstetric and parental
 Eye direction detection has been claimed to l
 Although Klinefelter syndrome (47,XXY) has

We examined social anxiety and internalizing
 Autism is a neurodevelopmental disorder tha
 Autism spectrum disorders (autism, Asperger
 The PAX6 gene is a transcription factor expr
 Fragile X Syndrome (FraX) is a broad-spectru
 y learning/memory center. These results der
 Rett Syndrome, an X-linked dominant neuroc
 ltered social behavior and nest building, decr

Dopaminergic neurons of the midbrain are th

Results

		Frequency	MeSH codes	
<input type="checkbox"/>	0	AUTISM	762	C10:F03
<input type="checkbox"/>	1	DISORDER	559	C02:C04:C05:C06:
<input type="checkbox"/>	2	STUDY	498	E01:E05:F02:F04:
<input type="checkbox"/>	3	CHILD	383	C01:C04:C18:E05:
<input type="checkbox"/>	4	USE	376	C10:C21:E05:F01:
<input type="checkbox"/>	5	SPECTRUM	364	C10:E05
<input type="checkbox"/>	6	RESULT	298	V05
<input type="checkbox"/>	7	SUGGEST	262	V05
<input type="checkbox"/>	8	SOCIAL	246	F01:F02:F03:F04:
<input type="checkbox"/>	9	CONTROL	244	D27:E02:E04:E05:
<input type="checkbox"/>	10	HA	232	D08
<input type="checkbox"/>	11	AUTISTIC	230	F03
<input type="checkbox"/>	12	AGE	227	C05:C10:C11:E01:
<input type="checkbox"/>	13	GROUP	223	A08:A11:B01:B03:
<input type="checkbox"/>	14	ASSOCIATE	217	F02:I02
<input type="checkbox"/>	15	FUNCTION	211	C10:C12:C23:D08:
<input type="checkbox"/>	16	HIGH	207	A08:B03:C04:C09:
<input type="checkbox"/>	17	ASD	200	V05
<input type="checkbox"/>	18	INDIVIDUAL	195	F01:N03
<input type="checkbox"/>	19	FINDING	190	C10:E01
<input type="checkbox"/>	20	YEAR	189	H01:I01
<input type="checkbox"/>	21	REPORT	189	E05:H01:N04:V02
<input type="checkbox"/>	22	DEVELOPMENTAL	180	C05:C10:F03:G01:
<input type="checkbox"/>	23	PRESENT	179	V05
<input type="checkbox"/>	24	BEHAVIOR	178	C10:E02:E05:F01:
<input type="checkbox"/>	25	PROVIDE	176	V05
<input type="checkbox"/>	26	EVIDENCE	172	G02

Show terms' frequencies

All Filter

- Anatomy [A]
- Organisms [B]
- Diseases [C]
- Chemicals and Drugs [D]
- Analytical, Diagnostic and Therapeutic Techniques and Equipment [E]
- Psychiatry and Psychology [F]
- Biological Sciences [G]
- Natural Sciences [H]
- Anthropology, Education, Sociology and Social Phenomena [I]
- Technology, Industry, Agriculture [J]
- Humanities [K]
- Information Science [L]
- Named Groups [M]
- Health Care [N]
- Various [V]
- Geographicals [Z]

MeSH: Medical Subject Headings

Show terms' frequencies

1

- Anatomy [A]
- Organisms [B]
- Diseases [C]
- Chemicals and Drugs [D]
- Analytical, Diagnostic and Therapeutic Techniques and Equipment [E]
- Psychiatry and Psychology [F]
- Biological Sciences [G]
- Natural Sciences [H]
- Anthropology, Education, Sociology and Social Phenomena [I]
- Technology, Industry, Agriculture [J]
- Humanities [K]
- Information Science [L]
- Named Groups [M]
- Health Care [N]
- Various [V]
- Geographicals [Z]

- Inorganic Chemicals;D01
- Organic Chemicals;D02
- Heterocyclic Compounds;D03
- Polycyclic Compounds;D04
- Macromolecular Substances;D05
- Hormones, Hormone Substitutes, and Hormone Antagonists;D06
- Enzymes and Coenzymes;D08
- Carbohydrates;D09
- Lipids;D10
- Amino Acids, Peptides, and Proteins;D12
- Nucleic Acids, Nucleotides, and Nucleosides;D13
- Complex Mixtures;D20
- Biological Factors;D23
- Biomedical and Dental Materials;D25
- Pharmaceutical Preparations;D26
- Chemical Actions and Uses;D27

Filtering according to MeSH classification

Show terms' frequencies

1

- Anatomy [A]
- Organisms [B]
- Diseases [C]
- Chemicals and Drugs [D]
- Analytical, Diagnostic and Therapeutic Techniques and Equipment [E]
- Psychiatry and Psychology [F]
- Biological Sciences [G]
- Natural Sciences [H]
- Anthropology, Education, Sociology and Social Phenomena [I]
- Technology, Industry, Agriculture [J]
- Humanities [K]
- Information Science [L]
- Named Groups [M]
- Health Care [N]
- Various [V]
- Geographicals [Z]

- Inorganic Chemicals;D01
- Organic Chemicals;D02
- Heterocyclic Compounds;D03
- Polycyclic Compounds;D04
- Macromolecular Substances;D05
- Hormones, Hormone Substitutes, and Hormone Antagonists;D06
- Enzymes and Coenzymes;D08
- Carbohydrates;D09
- Lipids;D10
- Amino Acids, Peptides, and Proteins;D12
- Nucleic Acids, Nucleotides, and Nucleosides;D13
- Complex Mixtures;D20
- Biological Factors;D23
- Biomedical and Dental Materials;D25
- Pharmaceutical Preparations;D26
- Chemical Actions and Uses;D27

Step Ra - results

Search PubMed for autism

Display: Abstract Show 5 Sort By Send to

Items 1 - 5 of 11761 Page 1 of 2351 Next

1: Behav Modif. 2008 Feb 6 [Epub ahead of print]

Excessive Daytime Sleep: Behavioral Assessment and Intervention in a Child with Autism.

Article Title: Excessive Daytime Sleep: Behavioral Assessment and Intervention in a Child with Autism.

Abstract: Som... excessive daytime sleep but intervention research for this problem has not been conducted. The present study evaluated procedures with a 13 year old boy who has autism and slept for prolonged periods during the day. Classroom staff at a specialized school implemented procedures with the boy according to an ABAB experimental design. Intervention eliminated daytime sleep through a 6-month follow-up assessment. The study

Search word: autism Number of articles: 11761 Select: Abstracts

Before date: 14 2 2008

ID	Freq	Term	MeSH code
20494	1	AMYLOIDOSIS	C10:C16:C18
20506	1	AMPLIFIER	E07
20507	1	AMPHIPHATIC	D12
20508	1	AMPHIBIAN	D12:D20
20510	1	AMOXICILLIN	D02
20514	1	AMNION	A10
20515	1	AMMONIUM	D01:D02:D03:D10:D12
20517	1	AMITROLE	D03
20520	1	AMINOTRANSFERASE	D08
20522	1	AMINOPEPTIDASE	D08
20523	1	AMINOIMIDAZOLECARBOXAMIDE	D08
20524	1	AMINOIMIDAZOLE	D03
20525	1	AMINOGLYCOSIDE	D08
20526	1	AMINOACIDURIA	C12
20534	1	AMENORRHEA	C23
20544	1	AMAZONIA	D01

- 1. Anatomy [A]
- 2. Organisms [B]
- 3. Diseases [C]
- 4. Chemicals and Drugs [D]
- 5. Inorganic Chemicals [D01] +
- 6. Organic Chemicals [D02] +
- 7. Heterocyclic Compounds [D03] +
- 8. Polycyclic Compounds [D04] +
- 9. Macromolecular Substances [D05] +
- 10. Hormones, Hormone Substitutes, and Hormone Antagonists [D06] +
- 11. Enzymes and Coenzymes [D08] +
- 12. Carbohydrates [D09] +
- 13. Lipids [D10] +
- 14. Amino Acids, Peptides, and Proteins [D12] +
- 15. Nucleic Acids, Nucleotides, and Nucleosides [D13] +
- 16. Complex Mixtures [D20] +
- 17. Biological Factors [D23] +
- 18. Chemical and Dental Materials [D25] +
- 19. Pharmaceutical Preparations [D26] +
- 20. Therapeutic Actions and Uses [D27] +

From domain-specific **rare terms** the following were **selected**:

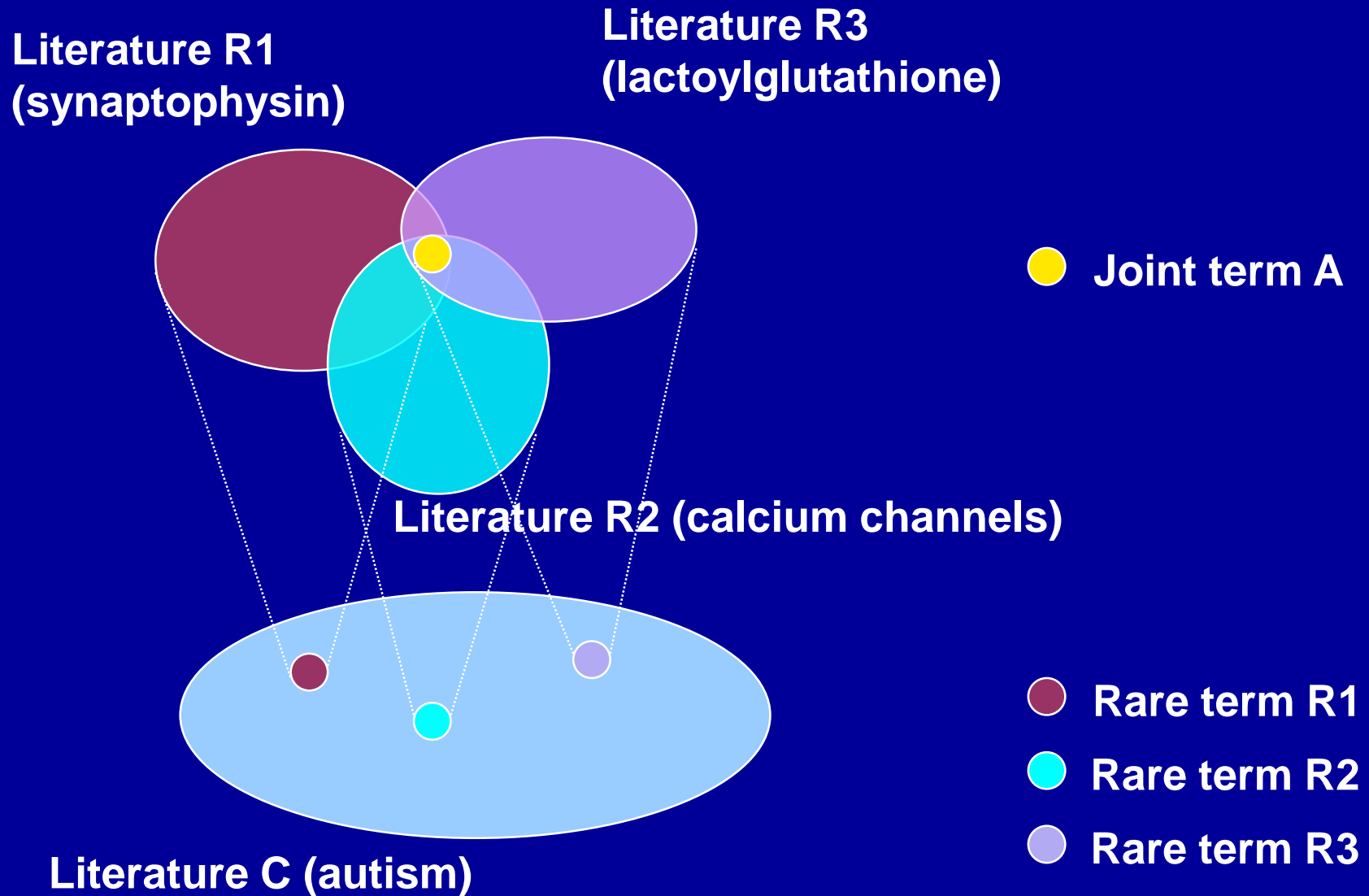
- **lactoylglutathione** (768 articles in PubMed, 1 of them mentioning autism)
- **synaptophysin** (3742 articles in PubMed, 2 of them mentioning autism)
- **calcium channels** (55284 articles in PubMed, 7 of them mentioning autism)

Background knowledge used in selection:

- **increase of polarity of glyoxalase I in autistic brain, glyoxalase system involves lactoylglutathione**
- **altered synaptic function in autism, synaptophysin is a protein localized to synaptic vesicles**
- **abnormal calcium signalling in some autistic children**

Do selected rare terms have something in common?

RaJoLink method: Idea



Step Jo: results in autism experiment

calcium AND channel
chromogranin
cofilin
lactoylglutathione
synaptophysin

Retrieve

2000 Abstracts Go

Input set of records

ex (DGC). Additional results suggest that Prf
Multiple endocrine neoplasms, including an in
We describe two cases of atypical carcinoid
including thymoma have a much better progr
Cancer with endocrine features rarely occur:
Embryonal carcinoma (EC) cells provide a car
the monoclonal antibody A2B5, was express
Endobrevin/VAMP-8 is an R-SNARE localized
Postsynaptic density (PSD)-95, SAP102, anc
e PSD-95 antibody was shown to label exclu
Complement defense 59 (CD59) is a cell surf
cells, which normally underexpress CD59, ar
The vesicular zinc-rich synaptic systems of th
We report two cases of primary large cell ne
of the gallbladder is significant for two reaso
To clarify the neuroendocrine differentiation
r other neuroendocrine markers, including ch
Synaptic vesicle protein 2 (SV2) is a glycoprc
d SV2-immunoreactive cells. The staining pat
The literature on the neuropathology of bipc
prefrontal, and temporal cortices in BD. In th
Prior studies on receptor recycling through le
, the primary effect of early endosomal sorti
We sought to delineate differences between
es.

Results

	Term	Frequencies	Sum of frequencies	MeSH cl
<input type="checkbox"/>	57 ETHYLMALEIMIDE	5:4,1,1,2,11	19	D02:DO
<input type="checkbox"/>	58 ELECTROPHORETIC	5:5,3,1,24,1	34	E05
<input type="checkbox"/>	59 DYE	5:25,4,4,2,12	47	D02:DO
<input type="checkbox"/>	60 DIPLOID	5:1,5,3,1,2	12	D20:G1
<input type="checkbox"/>	61 DEXTRAN	5:2,1,1,1,6	11	D02:DO
<input type="checkbox"/>	64 CONE	5:10,1,27,1,29	68	A08:C1
<input type="checkbox"/>	66 CHLORO	5:5,2,1,3,1	12	D02:DO
<input type="checkbox"/>	68 CATION	5:117,6,2,2,6	133	D12:D2
<input type="checkbox"/>	69 CATHEPSIN	5:2,5,7,1,5	20	D08
<input type="checkbox"/>	70 CARBOXY	5:3,7,4,2,3	19	D08
<input checked="" type="checkbox"/>	71 CALCINEURIN	5:18,6,8,1,4	37	D08
<input type="checkbox"/>	72 C6	5:2,1,2,2,1	8	D08
<input type="checkbox"/>	73 C3	5:3,3,6,9,3	24	D08:D1
<input type="checkbox"/>	75 BRONCHIAL	5:7,26,3,2,18	56	A07:A1
<input type="checkbox"/>	76 BROMIDE	5:4,1,3,3,2	13	D01:DO
<input type="checkbox"/>	77 BISPHOSPHATE	5:9,3,12,2,1	27	D08:D1
<input type="checkbox"/>	78 ATHEROSCLEROSIS	5:14,3,3,1,1	22	C10:C1
<input type="checkbox"/>	80 ALKALINE	5:3,21,4,3,16	47	D01:DO
<input type="checkbox"/>	81 A4	5:6,2,1,2,1	12	D08:D1
<input type="checkbox"/>	82 ZIPPER	4:0,1,1,1,1	4	D12:GO
<input type="checkbox"/>	83 XENOGRAFT	4:0,17,1,1,3	22	E05:E0
<input type="checkbox"/>	85 WEDGE	4:4,4,1,0,3	12	G09:K0
<input type="checkbox"/>	86 VISCOSITY	4:4,0,7,3,1	15	G09:H0
<input type="checkbox"/>	87 VINCRISTINE	4:1,3,0,1,5	10	D03
<input type="checkbox"/>	88 VEGF	4:1,15,5,0,10	31	D08:D1
<input type="checkbox"/>	89 VASODILATOR	4:32,4,2,0,1	39	D01:D2
<input type="checkbox"/>	91 UTERUS	4:2,8,0,2,4	16	A05:CO
<input type="checkbox"/>	92 UROTHELIUM	4:1,3,1,0,2	7	A10
<input type="checkbox"/>	94 UREMIC	4:2,1,0,2,3	8	C12
<input type="checkbox"/>	95 ULTRACENTRIFUGATION	4:0,1,3,1,1	6	E05
<input type="checkbox"/>	07 TIRIIE	4:12,23,9,0,20	64	A05:C1

Number of target terms: 3715 All terms: 16985

Step Link: results in autism experiment

Step Link

File Help

Previous step

calcineurin

Retrieve

5000 Titles Go

Input set of records

Sirolimus and cyclosporin for renal trans
Immunolocalization of calcineurin and FI
Constitutive and acquired resistance to
3,5-Bis(trifluoromethyl)pyrazoles: a nov
Advances in the molecular mechanisms
GABA(A) receptor modulation in rat ce
Phosphatidylinositol 4-phosphate 5-kin
Quantitative immunogold localization of
Clopidogrel-associated thrombotic thror
Connexin43 phosphorylation state and
Hepatitis C after liver transplantation.
Integration of calcineurin and MEF2 sig
Requirement for integration of phorbol
The effect of calcineurin inhibitors and
Competitive inhibition of NMDA recept
A calcineurin-NFATc3-dependent path
[A second target of cyclosporin A and
Cabin1 represses MEF2-dependent Nur
Continuing improvement in cadaver doi
A decade of living donor transplantati
Increased protein synthesis after T cell
Immunophilins may limit calcineurin inh
beta-Adrenergic pathway induces apoc

Results

	Term	Frequencies	Sum of frequencies	MeSH codes
1379	IMBALANCE	2:1,8	9	C18:G13
1331	INFANT	2:4,85	89	C10:C16:C17:C1
1298	INSULIN	2:44,5	49	A03:C10:C18:D0
1251	IRON	2:2,4	6	C15:C18:D01:D0
1234	JUVENILE	2:2,9	11	C01:C04:C05:C0
1174	LIPID	2:9,4	13	C08:C15:C16:C1
1137	MAGNESIUM	2:6,10	16	C18:D01:D02:D0
1069	METABOLIC	2:13,17	30	C05:C10:C16:C1
1068	METABOLISM	2:22,26	48	C16:C18:D12:EC
1047	MILK	2:1,3	4	A12:B06:C18:C2
1039	MITOCHONDRIAL	2:51,22	73	A11:C05:C18:D0
933	NON	2:26,71	97	A08:A13:B01:B0
916	NUTRITION	2:1,6	7	C13:C18:E02:EC
915	NUTRITIONAL	2:2,9	11	C18:E02:E05:GC
913	OBESITY	2:4,2	6	C08:C18:D27
893	ONSET	2:23,54	77	C05:C10:C16:C1
859	OXIDATIVE	2:29,13	42	C18:D12:G04:G1
823	PATHOLOGIC	2:5,1	6	C05:C10:C18:C2
773	PHOSPHORYLATION	2:95,2	97	C18:D12:G06
722	POTASSIUM	2:16,1	17	C18:D01:D02
642	PROTEIN	2:423,85	508	A11:C06:C10:C1
583	REACTIVE	2:10,3	13	C01:C10:C14:C1
523	REPAIR	2:8,6	14	C18:D08:D12:EC
505	RESISTANCE	2:23,5	28	C05:C08:C15:C1
504	RESISTANT	2:18,6	24	B01:C01:C05:C1
499	RESPIRATORY	2:3,5	8	A02:A04:A08:BI
361	SLOW	2:27,7	34	A02:B01:C02:C1
320	STABLE	2:33,1	34	C04:C18:D12:EC
312	STATE	2:17,35	72	C10:C18:C23:D0
288	STRESS	2:92,59	151	A11:C10:C12:C1
269	SUDDEN	2:1,8	7	C09:C14:C18:C2
233	SYNDROME	2:80,876	956	B04:C01:C02:C0
137	TRANSPORT	2:13,4	17	A11:C10:C12:C1
38	VITAMIN	2:2,18	20	C05:C15:C18:D0

Number of target terms: 68 All terms: 2426

Step Link – hypothesis testing

[Expert Opin Pharmacother](#). 2009 Sep;10(13):2127-43.

Autism: an emerging 'neuroimmune disorder' in search of therapy.

[Theoharides TC](#), [Kempuraj D](#), [Redwood L](#).

RESULTS/CONCLUSION: Increased oxidative stress and immune dysregulation are present in ASDs. Mast-cell activation may contribute to gut-blood-brain barrier disruption and brain inflammation. No effective treatments have emerged. Well-designed clinical trials with nonpsychotropic drugs were few and ASD characteristics varied considerably, making conclusions difficult. Psychotropic drugs are often used for stereotypic and aggressive behaviors. Unique combinations with antioxidant and anti-inflammatory flavonoids hold promise. New potential translational research areas and possible treatments are suggested.

[Neurobiol Dis](#). 2007 May;26(2):342-52. Epub 2007 Jan 25.

Beta-amyloid causes downregulation of calcineurin in neurons through induction of oxidative stress.

[Celsi F](#), [Svedberg M](#), [Unger C](#), [Cotman CW](#), [Carri MT](#), [Ottersen OP](#), [Nordberg A](#), [Torp R](#).

Calcineurin is an abundant cytosolic protein that is implicated in the modulation of glutamate release. Here we show that the expression level of this enzyme is reduced in primary neuronal cultures treated with beta-amyloid. Parallel experiments in ETNA cell lines expressing SOD1 suggested that the effect of beta-amyloid on calcineurin expression is mediated by oxidative stress. The relevance of the in vitro experiments was

Argument 1 (calcineurin literature)

- Erin et al. (2003) observed that calcineurin occurred as a complex with **Bcl-2** in various regions of rat and mouse brain.
- Chen et al. (2003) reported about the decrease in protein ubiquitination in synaptosomes and in nonneural cells that may play role in the regulation of **synaptic** function by a calcineurin antagonist FK506.
- Winder and Sweatt (2001) described the critical role of protein phosphatase 1, and calcineurin in the activity-dependent alterations of **synaptic plasticity** that depends on calcineurin.

Argument 2 (autism literature)

- Fatemi et al. (2001) reported a reduction of **Bcl-2** (a regulatory protein for control of programmed brain cell death) levels in autistic cerebellum.
- Belmonte et al. (2004) reviewed neuropathological studies of cerebral cortex in autism indicating abnormal **synaptic** and columnar structure and neuronal migration defects.
- Huber et al. (2002) showed evidences about an important role of **fragile X** protein, an identified cause of autism, in regulating activity-dependent **synaptic plasticity** in the brain.

Do identified pairs of documents point towards useful hypotheses?

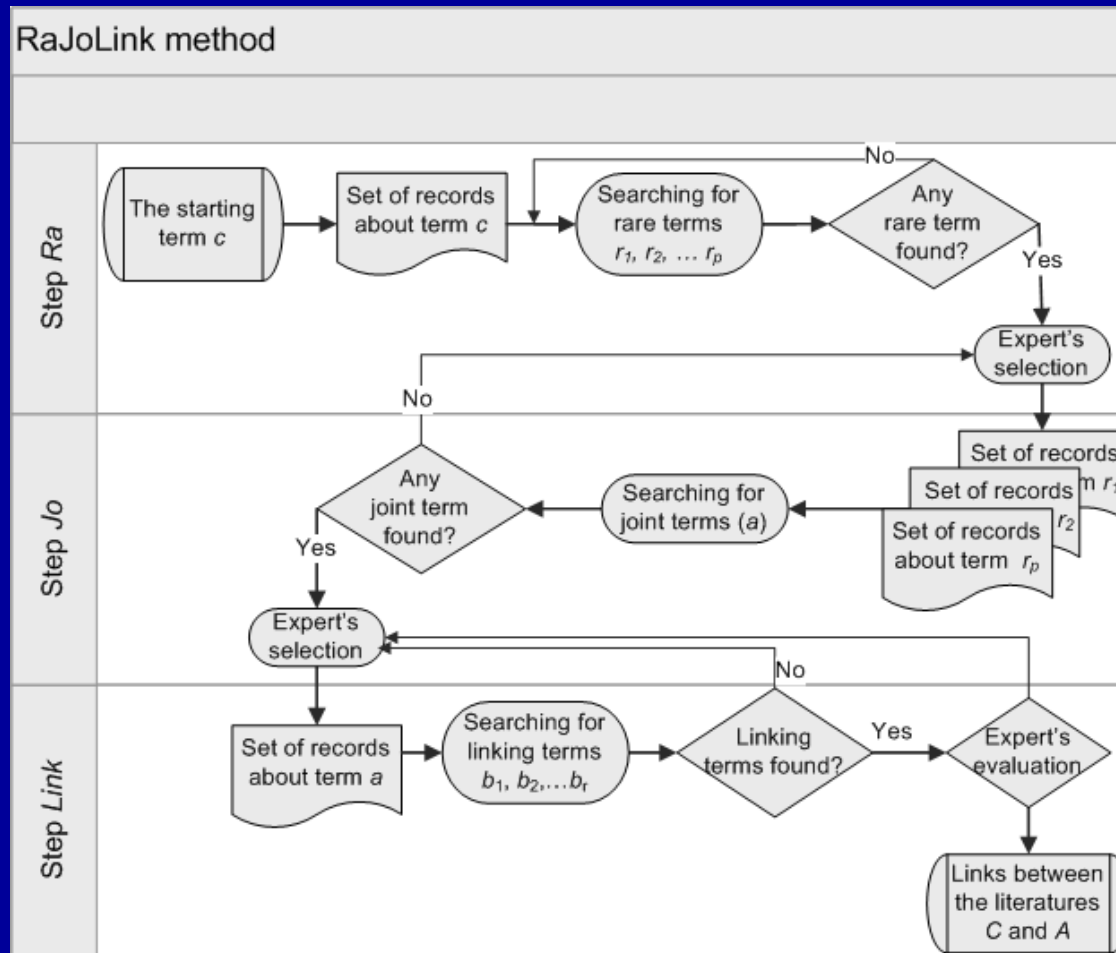
Expert's evaluation

Recent studies indicate that **calcineurin** participates in intracellular signaling pathways, which regulate **synaptic plasticity** and neuronal activities. An impaired **synaptic plasticity** is thought to be also a consequence of the lack of FMR1 protein in fragile X syndrome which is one of the identified cases of **autism**.

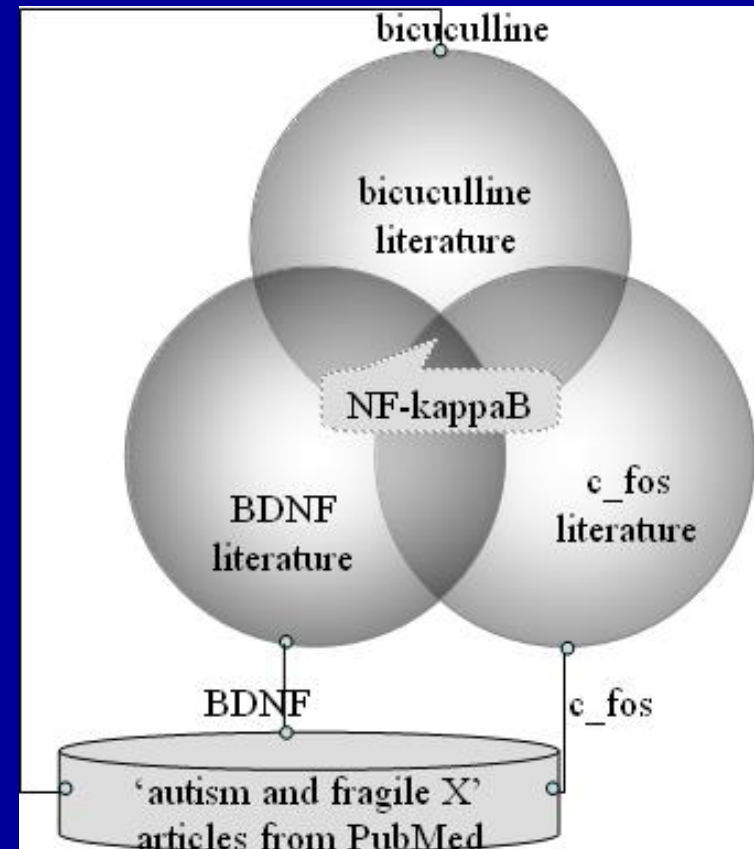
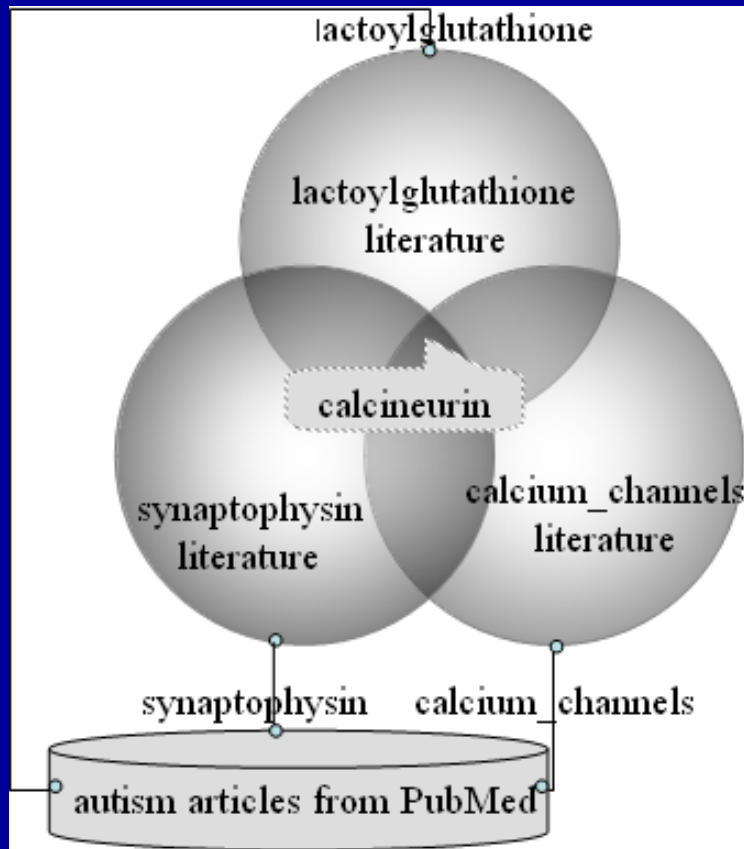
Expert's guidance

“Continue with fragil X, this would really be interesting...”

Procedures of the RaJoLink



Results obtained on autism domain, and on autism+fragile_X domain



Autism literature

Araghi-Niknam and Fatemi [2] showed reduction of *Bcl-2*, an important marker of apoptosis, in frontal, parietal and cerebellar cortices of autistic individuals.

NF-kappaB literature

Mattson [12] reported in his review that activation of NF-kappaB in neurons can promote their survival by inducing the expression of genes encoding antiapoptotic proteins such as *Bcl-2* and the antioxidant enzyme Mn-superoxide dismutase.

Vargas et al. [21] reported altered *cytokine* expression profiles in brain tissues and cerebrospinal fluid of patients with autism.

Ahn and Aggarwal [1] reported that on activation NF-kappaB regulates the expression of almost 400 different genes, which include enzymes, *cytokines* (such as TNF, IL-1, IL-6, IL-8, and chemokines), adhesion molecules, cell cycle regulatory molecules, viral proteins, and angiogenic factors.

Ming et al. [13] reported about the increased urinary excretion of an *oxidative stress* biomarker - 8-iso-PGF2alpha in autism.

Zou and Crews [24] reported about an increase in NF-kappaB DNA binding following *oxidative stress* neurotoxicity.

Expert's evaluation

It is thought that autism could result from an **interaction between genetic and environmental factors** with an **oxidative stress and immunological disorders as potential mechanisms linking the two** [3], [13]. Both of the mechanisms are related to NF-kappaB as the result of our analysis. The activation of the transcriptional factor NF-kappaB was shown to prevent neuronal apoptosis in various cell cultures and in vivo models [12]. Oxidative stress and elevation of intracellular calcium levels are particularly important inducers of NF-kappaB activation. In addition, various other genes are responsive to the activation of the NF-kappaB, including those for cytokines. In this way the NF-kappaB can be involved in the complex linkage between the immune system and autism [3], [21]. So, **according to our analysis one possible point of convergence between “oxidative stress” and “immunological disorder” paradigm in autism is NF-kappaB.**

Results - Towards better understanding of autism

- Urbančič T., Petrič I., Cestnik B., Macedoni-Lukšič M. Literature mining: towards better understanding of autism. In: Bellazzi, R. et al., editors. AIME 2007. Proceedings of the 11th Conference on Artificial Intelligence in Medicine in Europe; Amsterdam, The Netherlands. 217-226 (2007).
“... according to our analysis one possible point of convergence between oxidative stress and immunological disorder paradigm in autism is NF-kappaB.”
- Naik U.S., Gangadharan, Abbagani K., Nagalla B., Dasari N., Manna S. K. A Study of Nuclear Transcription Factor –Kappa B in Childhood Autism. *PLoS ONE* 6(5): e19488 (2011).
“We have noticed significant increase in NF-kB DNA binding activity in peripheral blood samples of children with autism.”
“This finding has immense value in understanding many of the known biochemical changes reported in autism.”

A Study of Nuclear Transcription Factor-Kappa B in Childhood Autism

Usha S. Naik¹, Charitha Gangadharan², Kanakalatha Abbagini¹, Balakrishna Nagalla³, Niranjan Dasari¹, Sunil K. Manna^{2*}

¹ Department of Psychiatry, Osmania Medical College, Hyderabad, India, ² Laboratory of Immunology, Centre for DNA Fingerprinting and Diagnostics, Nampally, Hyderabad, India, ³ National Institute of Nutrition, Hyderabad, India

Abstract

Background: Several children with autism show regression in language and social development while maintaining normal motor milestones. A clear period of normal development followed by regression and subsequent improvement with treatment, suggests a multifactorial etiology. The role of inflammation in autism is now a major area of study. Viral and bacterial infections, hypoxia, or medication could affect both foetus and infant. These stressors could upregulate transcription factors like nuclear factor kappa B (NF- κ B), a master switch for many genes including some implicated in autism like tumor necrosis factor (TNF). On this hypothesis, it was proposed to determine NF- κ B in children with autism.

Methods: Peripheral blood samples of 67 children with autism and 29 control children were evaluated for NF- κ B using electrophoretic mobility shift assay (EMSA). A phosphor imaging technique was used to quantify values. The fold increase over the control sample was calculated and statistical analysis was carried out using SPSS 15.

Results: We have noted significant increase in NF- κ B DNA binding activity in peripheral blood samples of children with autism. When the fold increase of NF- κ B in cases ($n=67$) was compared with that of controls ($n=29$), there was a significant difference (3.14 vs. 1.40, respectively; $p<0.02$).

Conclusion: This finding has immense value in understanding many of the known biochemical changes reported in autism. As NF- κ B is a response to stressors of several kinds and a master switch for many genes, autism may then arise at least in part from an NF- κ B pathway gone awry.

Citation: Naik US, Gangadharan C, Abbagini K, Nagalla B, Dasari N, et al. (2011) A Study of Nuclear Transcription Factor-Kappa B in Childhood Autism. PLoS ONE 6(5): e19488. doi:10.1371/journal.pone.0019488

Editor: Monica Uddin, University of Michigan, United States of America

Received: September 29, 2010; **Accepted:** April 8, 2011; **Published:** May 9, 2011

Copyright: © 2011 Naik et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The study was funded by a Central Government Grant from the Department of Biotechnology, Ministry of Science and Technology, Government of India. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: manna@cfdi.org.in

Open discovery - Conclusions

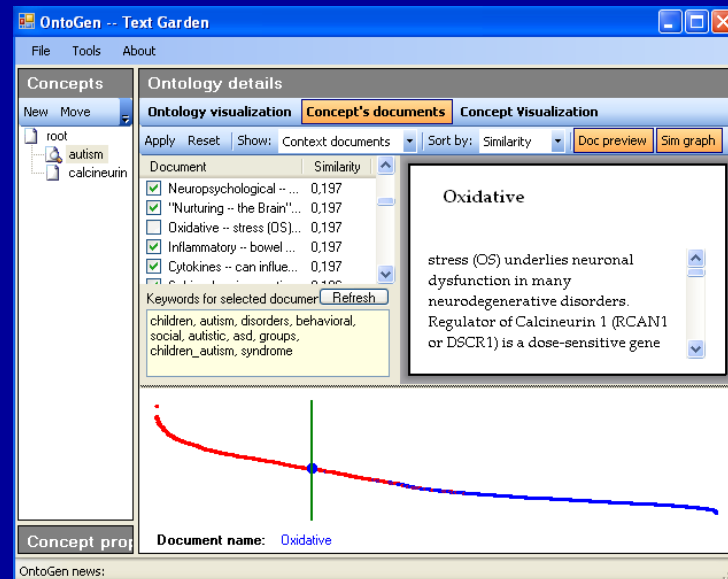
- It is worthwhile to explore **rare terms for generation of hypotheses** (RaJoLink method).
- **Expert's involvement** is crucial for speeding up the process (selections) and for evaluations of candidate hypotheses, but it's well supported by the method.
- Expert evaluation confirmed the **relevance of discovered relations in the autism domain** (published in medical journals).
- **RaJoLink: Method for finding seeds of future discoveries in nowadays literature**

Outline

- **Part 1: Introduction**
 - Creativity and knowledge discovery
 - Literature as a source of cross-context knowledge discoveries
 - Open vs. closed discovery
- **Part 2: Open discovery framework**
 - Exploring rarity (RaJoLink)
 - Case 1: RaJoLink in the Migraine
 - Case 2: With RaJoLink Towards better understanding of Autism
- **Part 3: Closed discovery framework**
 - Speeding-up the process:
exploring outliers, ranking candidate bridging terms, workflows, HCI (CrossBee)
- **Part 4: What next?**

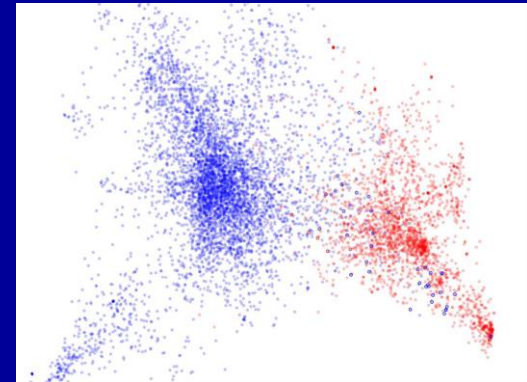
Closed discovery: improving search for bridging terms

- **Exploring outliers** (Petrič, Cestnik, Lavrač, Urbančič, 2012)
 - Similarity graph of all documents from domains A and C, obtained with Ontogen (Fortuna et al., 2006)
 - Outliers from literature C are positioned among the documents from literature A.
 - Outliers can be successfully used for detecting bridging terms.

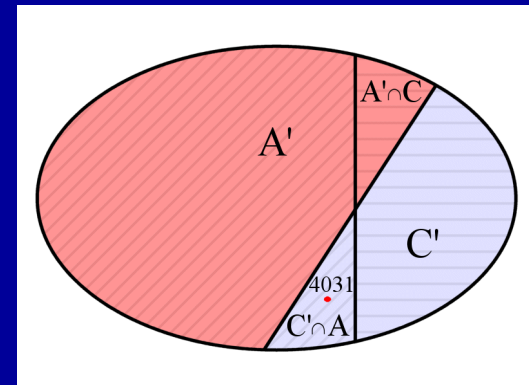


Closed Discovery – Role of Outlier documents

- The majority of bridging concepts can be found in outlier documents.
(Sluban, Juršič, Cestnik, Lavrač, 2012)



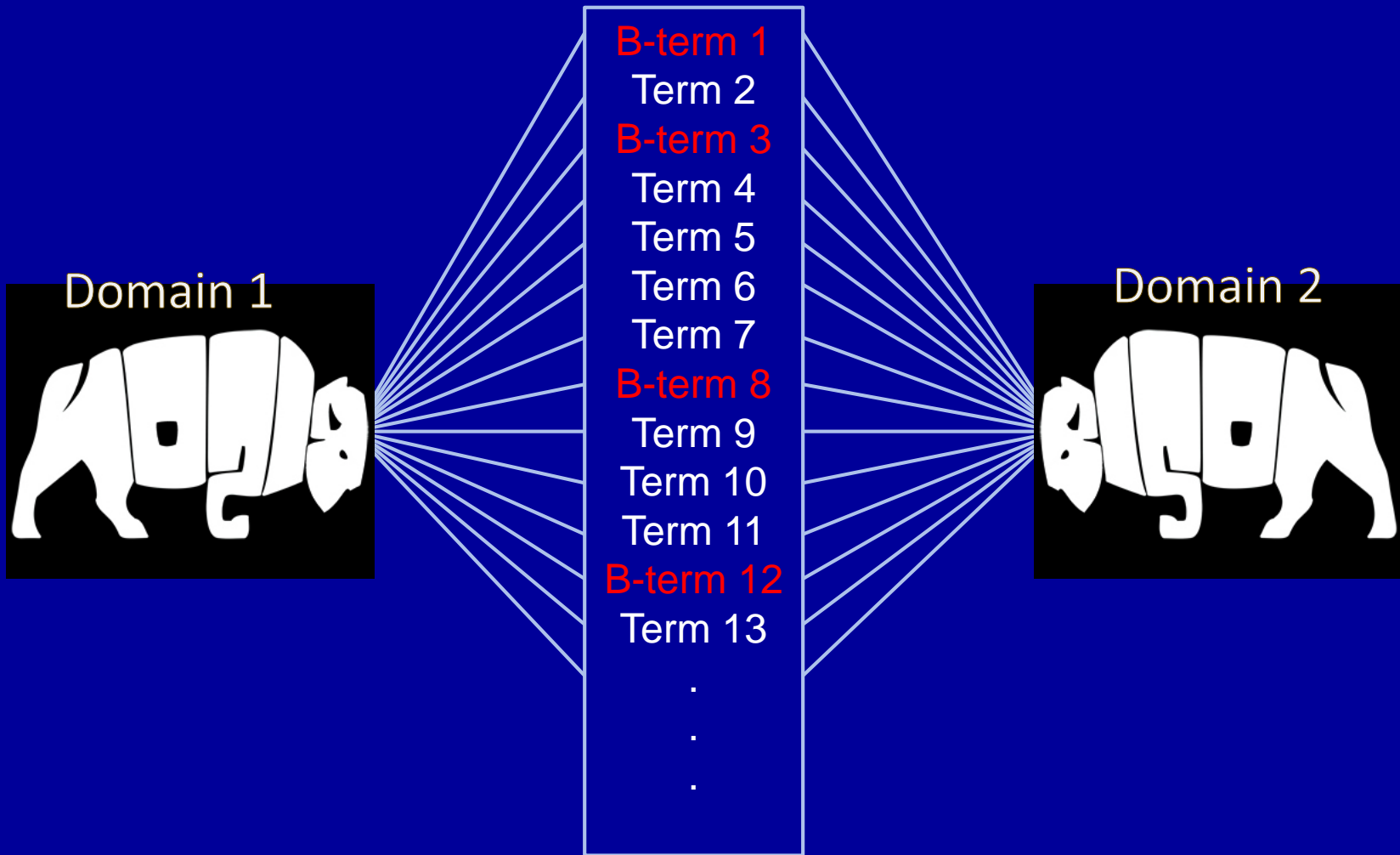
- Focusing on outliers from literatures *A* and *C* resulted in substantial search reduction.
(Petrič, Cestnik, Lavrač, Urbančič, 2012)



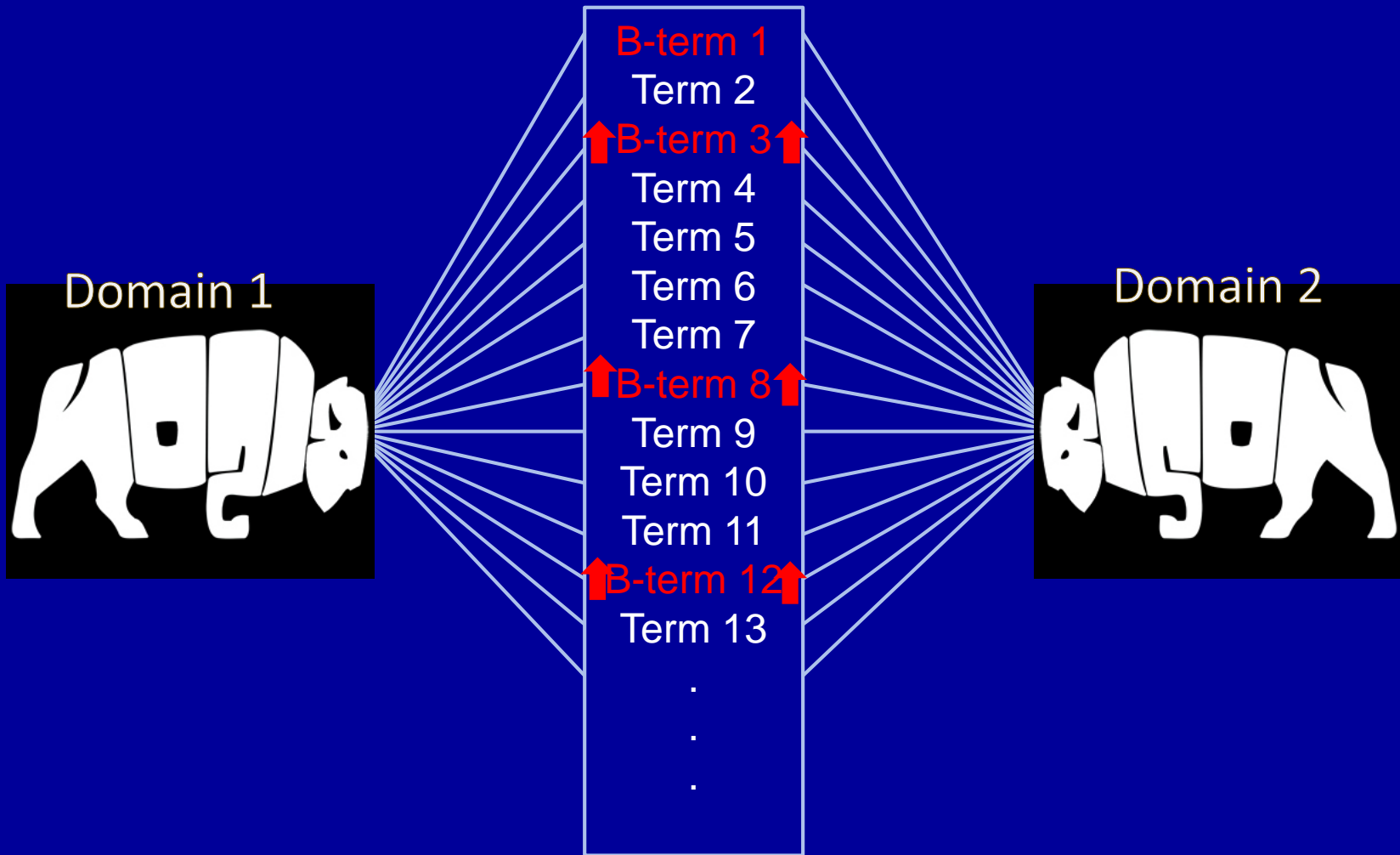
Closed discovery – improving search for bridging terms

- **Web application CrossBee (Juršič 2013):**
 - Ranking of candidate b-terms according to their bisociation potential
 - Side-by-side document view, customized for expert inspection of potential cross-domain links
 - Standard document navigation – filtering, keyword and similarity search
 - Color-based domain separation scheme
 - User interface customization (emphasizing features on/off, ...)
 - Hierarchical semi-automated document clustering (*Top Circle*)

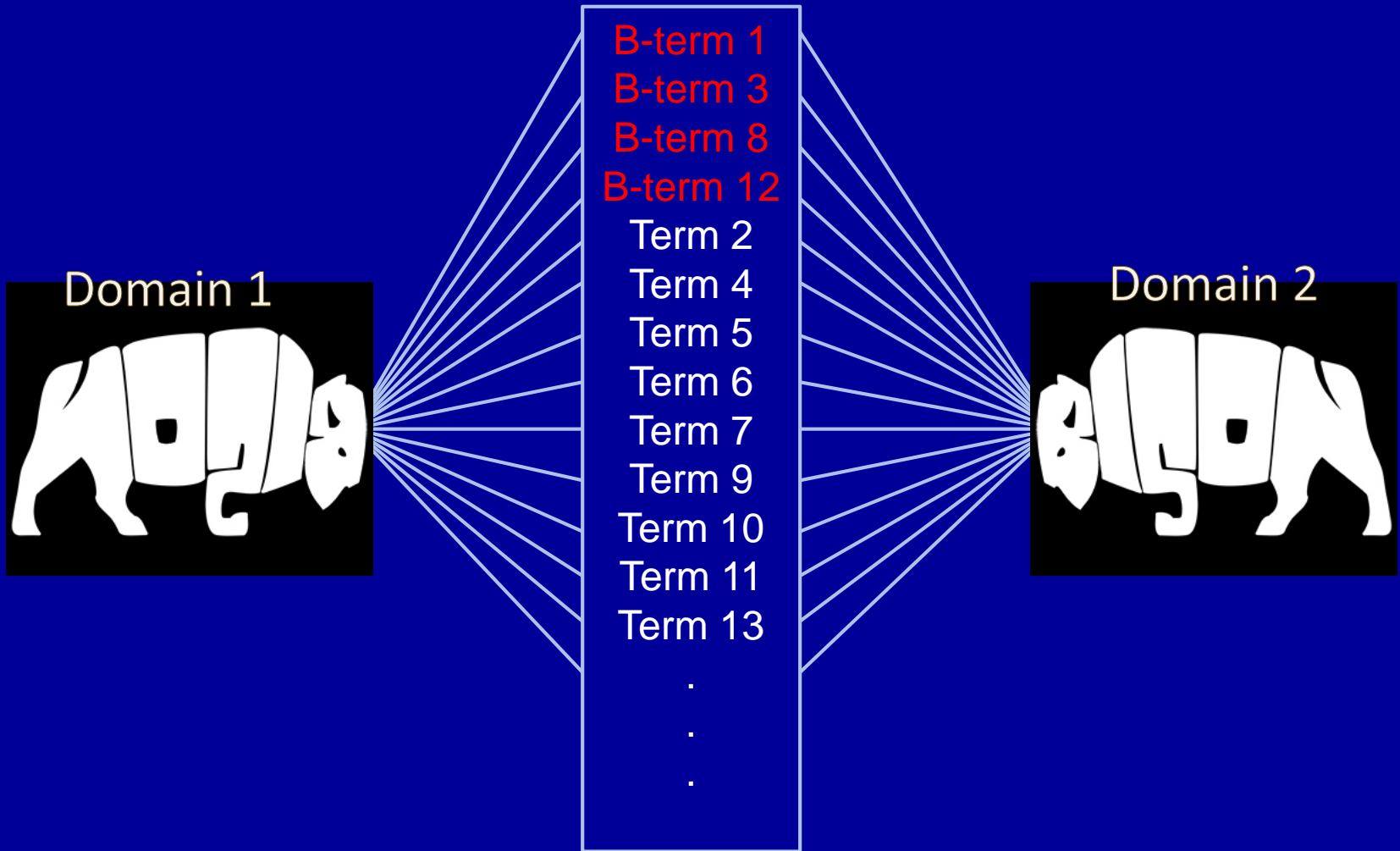
B-TERM DISCOVERY



B-TERM DISCOVERY



B-TERM DISCOVERY

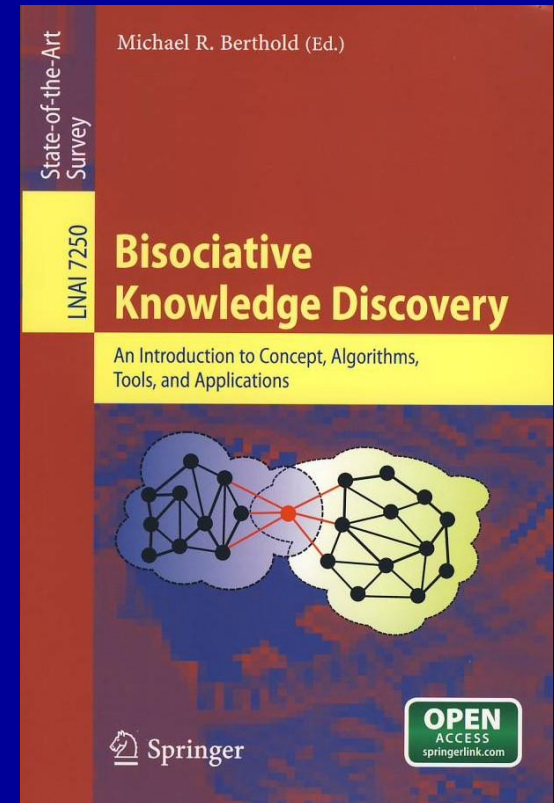


Syde-by-side document inspection in CrossBee

The screenshot displays the CrossBee web application interface. At the top, the logo 'CROSSBEE CROSS CONTEXT DISSOCIATION EXPLORER' is on the left, and 'Supported by BISON' with logos for the 7th Framework Programme and the European Union is on the right. A navigation bar contains buttons for 'Start', 'Downloads', 'Term View', 'Document View', and 'BTerms'. The main content area is titled 'B-Term Identify (Term "paroxysmal" Analysis)'. It features a search box, a main menu with options like 'Start', 'Downloads', and 'Term View', and an item basket. The central part of the page shows two columns of document results. The left column, highlighted in pink, is for Document #2270, 'Paroxysmal and other features of the electroencephalogram in migraine', with a domain of 'MIG'. The right column, highlighted in green, is for Document #3456, '[A case of paroxysmal tachycardia o...]', with a domain of 'MAG'. Each document entry includes a list of important terms ordered by importance and alphabetically. At the bottom, a footer contains the text: 'The research was supported by the European Commission under the 7th Framework Programme FP7 ICT 2007 C FET Open project BISON 211898.', 'CrossBee: Application version: 3.0, built on: 17.1.2012', 'In synch with the results published in the Bison book.', and 'Copyright © 2010 Jozef Stefan Institute. Style designed by Free CSS Templates. SiteMap.'

Conclusions

- **New discoveries** based on cross-domain connections uncovered from literature
- **Literature mining tools**
- **Open discovery**
(suggesting which domains to connect)
- **Closed discovery**
(supportive evidence through bridging terms)
- **Applications of literature mining so far mostly in biomedical domains**
(e.g. genes and diseases, mechanisms underlying different disorders, etc.)
- Potential for **other domains**



What next?

- **Open discovery and CrossBee – upgrades and improvements**
- **Continued research in computational creativity for knowledge discovery**
- **ConCreTe: Concept Creation Technology**
- **Case 3: Call for ideas**