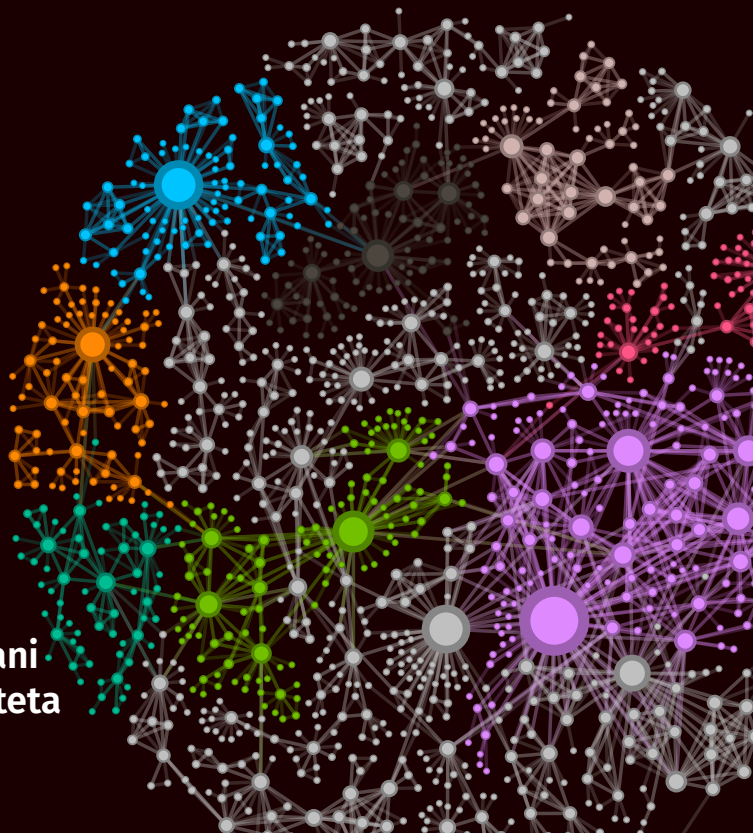


NetSlo '23

VII. srečanje raziskovalcev s področja analize omrežij

Univerza v Ljubljani
Medicinska fakulteta
26. januar 2023



VII. srečanje raziskovalcev s področja analize omrežij

NetSlo '23

Zbornik prispevkov

Ljubljana, 26. januarja 2023

VII. srečanje raziskovalcev s področja analize omrežij (NetSlo '23)

Zbornik prispevkov

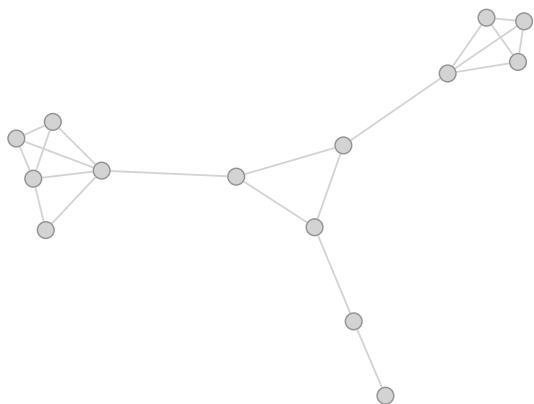
Uredil	Andrej Kastrin
Izdal in založil	Univerza v Ljubljani, Medicinska fakulteta, Inštitut za biostatistiko in medicinsko informatiko Ljubljana, 2023
Elektronski vir	PDF
Način dostopa	https://netslo.mf.uni-lj.si/netslo23zbornik.pdf

Publikacija je nastala v okviru raziskovalnega projekta Napovedovanje sodelovanja med raziskovalci s pomočjo odkrivanja zakonitosti iz literature (J5-2552), ki ga financira Javna agencija za raziskovalno dejavnost Republike Slovenije.

Kataložni zapis o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani

COBISS.SI-ID 164356867

ISBN 978-961-267-258-4 (PDF)



Predgovor

Po dvoletnem premoru bo 26. januarja na Medicinski fakulteti Univerze v Ljubljani potekalo tradicionalno, letos že sedmo, Srečanje raziskovalcev s področja analize omrežij. Analiza omrežij—pogosto označena tudi kot znanost 21. stoletja—je vsebinsko močno razvejano in izrazito transdisciplinarno področje raziskovanja. V preteklem stoletju je analiza omrežij botrovala nekaterim paradigmatskim miselnim preskokom v proučevanju kompleksnih socialnih sistemov. Kasneje je mdr. omogočila razumevanje zgradbe in mehanizmov delovanja svetovnega spleta. Danes analiza omrežij utira nova spoznanja v vedah o živem in je takorekoč pospeševalnik razvoja njihovega védenja.

Prvo srečanje, takrat še pod imenom Mreženje slovenskih networkašev, je potekalo leta 2015 na ljubljanski Fakulteti za računalništvo in informatiko. Kasneje se je srečanje preimenovalo v NetSlo, po zgledu znamenitih NetSci konferenc.

Letošnji dogodek bo potekal v organizaciji Inštituta za biostatistiko in medicinsko informatiko na Medicinski fakulteti Univerze v Ljubljani ter s finančno pomočjo Fakultete za družbene vede in novomeške Fakultete za informacijske študije. Nastopilo bo sedem predavateljev, ki tako ali drugače krojijo smernice razvoja v analizi omrežij. Zaključek bo potekal tradicionalno, z neformalnim druženjem in večerjo pod Rožnikom.

Lepo vabljeni!

Andrej Kastrin

Program

13:30–14:00		Registracija udeležencev
14:00–14:15		Nagovor organizatorja
14:15–15:00	D. Rozman	Sistemska medicina in bolezenska omrežja
15:00–15:20	R. Koželj	Analiza omrežij časovnega razvoja seizmične aktivnosti
15:20–15:45		Odmor
15:45–16:30	K. Rojko	Publicistična raziskovalna uspešnost slovenskih doktorjev znanosti: trendi in razlike
16:30–17:00	M. Plahuta	Analiza slovenskega koščka Twitterja
17:00–17:30	L. Jesenšek	Struktura slovenske politične razprave na Twitterju
17:30–18:00		Odmor
18:00–18:30	C. Bleker	Stress knowledge map: From knowledge graph to Boolean modelling
18:30–19:00	V. Modic	From causal networks to adverse outcome pathways
20:00–		Večerja v Gostilni Čad

Povzetki pripevkov

Vabljeni predavanje

Sistemska medicina in bolezenska omrežja

prof. dr. Damjana Rozman

Univerza v Ljubljani, Medicinska fakulteta

Povzetek

Pojav visoko zmogljivih tehnologij ustvarja velike količine omskih podatkov, kar omogoča preučevanje organizmov na molekulski in celični ravni. Kljub spektakularnemu napredku v pogenomskem obdobju ostaja vrzel med podatki in medicinskim znanjem ter še večja vrzel med novim znanjem v smislu klinične uporabnosti in koristi za bolnike. Še vedno trpimo zaradi večfaktorskih bolezni, kjer ne poznamo vzrokov ali zdravil. Pogosto tudi ne vemo, kako motnje na molekulski ravni vplivajo na bolezenske fenotipe. Fenotipske značilnosti organizma so namreč določene s prepletenimi reakcijami presnovnih in signalnih poti, kar lahko opišemo kot biološka omrežja oziroma sisteme. Za višje razvite organizme, kot je človek, je izjemno težko preučiti vse njihove procese *in vivo* ali *in vitro*. Zato so se razvila številna računska orodja, kot so npr. modeli na ravni genoma, ki nam olajšajo vpogled v delovanje kompleksnih bioloških omrežij.

Razumevanje, kako so posamezni geni in drugi molekulski dejavniki vključeni v patogenezo, je temeljni del posamezniku prilagojene medicine. Pri tem je nujno uporabiti celostne, sistemske pristope, saj geni in njihovi produkti ne delujejo vsak zase, temveč usklajeno. Govorimo lahko o pristopih sistemske medicine oziroma medicine omrežij, ki predstavlja novo trans-disciplinarno področje. Človeško telo obravnava kot sistem, sestavljen iz več delov, s kompleksnimi odnosi na več ravneh, ki je integrirano tudi v okolje. Raziskovanje s pristopi sistemske medicine pomeni razumevanje in združevanje konceptov, ki prihajajo iz diametralno različnih področij, vključno z medicino, biokemijo, biologijo, fiziko, statistiko, modeliranjem in simulacijo ter znanostjo o podatkih, vse več

pa je tudi etičnih in pravnih dilem. Pri delu s človeškimi vzorci ali samo s človeškimi podatki je treba vedno uveljavljati etična pravila in spoštovati zasebnost. V okviru predavanja bo predstavljenih nekaj primerov študija omrežij pri boleznih jeter, ki se lahko razlikujejo tudi med spoloma. Predstavljeni bodo pristopi, ki preko poznavanja bioloških omrežij lahko privedejo do opredelitve novih bolezenskih biooznačevalcev ali možnih tarč za zdravljenje.

Analiza omrežij časovnega razvoja seizmične aktivnosti

Romi Koželj

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko

Povzetek

Potres je naravni pojav, ki nastane kot posledica notranje Zemljine dinamike globoko pod površjem in ga z današnjim znanjem še ni mogoče napovedati. V raziskavi poizkušamo z uporabo omrežne znanosti pridobiti nova znanja o značilnostih in razvoju seizmične aktivnosti skozi čas. Implementiramo in med seboj primerjamo različne modele omrežij, ki temeljijo na interakciji med potresi v času in kraju ter na predpostavki o podobnostih potresne aktivnosti na izbranih geografskih območjih. Iz omrežij, ki jih konstruiramo v več manjših zaporednih časovnih oknih, izračunamo nabor značilk ter prikažemo njihovo spreminjanje skozi čas. Na koncu z uporabo modela ARIMA za napovedovanje časovnih vrst preverimo, ali je iz dobljenih vzorcev moč sklepati o značilnostih seizmičnega dogajanja v prihodnosti.

Analiza dobljenih omrežij ter generiranih časovnih vrst pokaže, da preko večine obravnavanih omrežij dobimo zanesljiv in predvidljiv odziv vrednosti v časovnih vrstah. Iz rezultatov napovedovanja vrednosti časovnih vrst je razvidno, da oblike časovnih vrst, predvsem v intervalih, v katerih se zgodi močnejši potres, z uporabljenim modelom ARIMA ne moremo dobro napovedati.

Publicistična raziskovalna uspešnost slovenskih doktorjev znanosti: trendi in razlike

izr. prof. dr. Katarina Rojko

Fakulteta za informacijske študije v Novem mestu

Povzetek

Prvi del predstavitve bo vključeval izsledke raziskave publicistične raziskovalne uspešnosti slovenskih doktorjev znanosti s poudarkom na vpliv Bolonjske reforme, drugi pa izsledke raziskave te uspešnosti s poudarkom na posamezne znanstvene discipline. Vzorec zajema vse, ki so od leta 2007 oz. 1991 doktorirali, a za vsakega le obdobje 2 leti pred in 2 oz. 7 let po letu doktoriranja. Njihova publicistična uspešnost je analizirana upoštevajoč produktivnost, sodelovanje in neodvisnost od mentorja. Rezultati prvega dela raziskave kažejo, da razlike med pred-bolonjskimi in bolonjskimi doktoranti niso vidne pri polovici indikatorjev, je pa vseeno opazen negativen trend.

Drugi del raziskave, ki je med indikatorji iskal tudi korelacije, vključil podatke o citatih, zajel daljše časovno obdobje (28 generacij ter prvih 10 let kariere) in raziskovalce razdelil v 234 disciplin, pa je razkril v večini pričakovane razlike med disciplinami po vseh kazalnikih. Identificirani so bili tudi večinoma podobni trendi— zmanjševanje produktivnosti ter neodvisnosti, a vedno več sodelovanja in citatov.

Analiza slovenskega koščka Twitterja

Marko Plahuta

Virostatiq

Povzetek

Raziskovanje družabnega omrežja Twitter je zanimivo tako z vidika besedilnega rudarjenja po vsebinah objav kot tudi z vidika analize strukturnih lastnosti omrežja samega in proučevanja njegovega razvoja v času. V prispevku avtor obravnava podatkovne tehnologije, ki jih uporablja za zbiranje, shranjevanje, analizo in vizualizacijo omrežja Twitter. Omrežje trenutno sestavlja okrog 28 000 uporabniških imen in več kot 90 milijonov tвитov. V nadaljevanju avtor predstavi metapodatkovno shemo, ki jo uporablja za gradnjo omrežja in opiše algoritem za ugotavljanje spremembe uporabniškega imena uporabnika. Posebej obravnava dileme, na katere mora podatkovni analitik odgovoriti ob zajemu omrežja, npr. kolikšna je vrednost retvita v primerjavi z neposrednim odgovorom ali pa kako vrednotimo pripisan ključnik v odnosu do dodanega URL naslova. V tem okviru predstavi tudi shemo, ki jo uporablja za uteževanje povezav v svoji zbirki. Drugi del predstavitve avtor začne s kratkim uvodom v dve tehniki za vlaganje besed (word2vec) oz. vozlišč omrežja (node2vec) v vektorski prostor. Na izbranih primerih iz slovenske tвитosfere prednosti uporabe vektorskih vložitev tudi nazorno ilustrira; predstavi primere gručenja po uporabniških imenih, domenah, ter politični in tematski sorodnosti tвитov.

Struktura slovenske politične razprave na Twitterju

Luka Jesenšek

Univerza v Ljubljani, Fakulteta za družbene vede

Povzetek

Z metodami za analizo socialnih omrežij preučujem strukturo slovenske politične razprave na spletnem družbenem omrežju Twitter pred volitvami v Evropski parlament 2019.

Opisujem omrežje 2867 uporabnikov in 40 670 objav (tvitov), središčne uporabnike, pretok informacij prek posredovanih objav (retvitov), skupnosti uporabnikov in širjenje medijskih vsebin na platformi. Med najaktivnejšimi uporabniki prevladujejo širši javnosti manj znani posamezniki, med najvplivnejšimi pa politične stranke, politiki in znane osebnosti. Polarizacija močno zaznamuje pretok informacij v omrežju: analiza retvitov, vzorci komuniciranja in analiza deljenih vsebin kažejo na obstoj jasno določenih, mnenjsko homogenih skupnosti. Uporabniki Twitterja pogosteje komunicirajo z osebami, s katerimi si delijo podobna prepričanja, pogosteje delijo vsebine, ki prihajajo iz njihovih skupnosti ter pogosteje vključujejo povezave do vsebin, ki se skladajo z njihovimi pogledi. Medijske vsebine se pretežno razširjajo znotraj skupnosti uporabnikov in ne preko celotnega omrežja, znotraj različnih skupnosti pa se razširjajo medijske vsebine iz različnih virov.

Politiki imajo velik vpliv na širjenje vsebin znotraj svojih skupnosti, a niso med središčnimi uporabniki, ki bi spodbujali komunikacijo med različnimi deli omrežja.

Stress knowledge map: From knowledge graph to Boolean modelling

Carissa Bleker, Ph.D.

National Institute of Biology / Nacionalni inštitut za biologijo

Abstract / Povzetek

With pressure on global food security set to increase due to a growing human population and the increasingly apparent effect of climate change on agriculture, our understanding of the complexity of plant responses to biotic and abiotic stressors is becoming ever more important. Knowledge on molecular processes occurring within the plant cell is currently scattered across various sources, and thus not easily accessible for analysis or modelling. Stress Knowledge Map (SKM) is an attempt at integrating this dispersed information into a freely available resource. The types of entities (nodes) within SKM include genes and gene products, complexes, and metabolites, connected to one another via molecular interactions. Implemented in a Neo4j database, and translated to a number of standard formats, it supports the interactive exploration of its contents, and represents a basis for various modelling approaches (<https://skm.nib.si/>).

While the topology of SKM is well described, many detailed kinetics are lacking, making classical dynamical modelling impractical. Instead, Boolean networks and Boolean modelling allows for the analysis of complex systems without the need for detailed mechanistic knowledge. In a Boolean network, nodes are assigned one of two states: “active” or “inactive”, defined from a set of regulator nodes. Realistically, the nodes have a degree of activity, and are not merely “active” or “inactive”. To model this, the functions in Boolean space can be translated to qualitative ordinary differential equations (ODEs) with the use of various transforms. We developed a Python package BoolDoG (<https://nib-si.github.io/BoolDoG/>) that can read in regulatory networks, as well as

Boolean networks; run Boolean simulations and steady state analysis; transform the Boolean network to continuous ODEs; and run continuous simulations. The package is designed to be interoperable and easily extendable.

From causal networks to adverse outcome pathways

Vid Modic

National Institute of Biology / Nacionalni inštitut za biologijo

Abstract / Povzetek

A decade ago, adverse outcome pathways (AOPs) and networks of AOPs have been put forward as a tool for organizing toxicological knowledge across different levels of biological organization, from the initial interaction of chemicals with the biological system to the individual and population level effects relevant for environmental risk assessment. There currently only around 500 AOPs in the AOP-Wiki (<https://aopwiki.org/>), which don't even closely cover all the different ways in which chemicals can adversely affect biological systems.

Today we will present a semiautomatic approach for development of AOPs from causal biological networks, which are available on the Causal Biological Network database (<https://www.causalbionet.com/>). The approach, demonstrated on a developmental neurotoxicity networks, includes additional curation of the causal toxicological networks, removal of disconnected part of the networks and feedback loops, and separations into simple paths-candidate AOPs. In the second part we will show how to perform a weight of evidence analysis of an example pathway, which is necessary before such pathways are used in regulation of chemical safety.

Donatorji

Univerza v Ljubljani
Medicinska fakulteta



Univerza v Ljubljani
Fakulteta *za družbene vede*



Fakulteta za
informacijske študije
Faculty of information studies

Organizacija in izvedba	Andrej Kastrin • Univerza v Ljubljani, Medicinska fakulteta, Inštitut za biostatistiko in medicinsko informatiko
Programski odbor	Andrej Kastrin • Univerza v Ljubljani Gorazd Drevenšek • Univerza v Ljubljani Nataša Kejžar • Univerza v Ljubljani Petra Kralj Novak • Institut "Jožef Stefan" Zoran Levnajić • Fakulteta za informacijske študije



<https://netslo.mf.uni-lj.si>