LDA2Net: Digging under the surface of COVID-19 topics in literature

Topic 82 companion sheet

G. Minello C.R.M.A. Santagiustina M. Warglien

This file contains the following supplementary information for Topic 82 of the manuscript "*LDA2Net*: Digging under the surface of COVID-19 topics in scientific literature":

- Human label and automatic n-gram label proposals (Table 1)
- Summary measures (Table 2)
- Network of top 25 bigrams (Figure 1)
- Wordclouds of top 25 words by node relevance measure (Figure 2)
- Wordclouds of top 25 bigrams by edge relevance measure (Figure 3)
- Filtered (0.99 percentile) topic network (Figure 4)

Table 1: Human and automatic label proposals. Automatic label candidate for largest word community of the topic. In parenthesis: absolute frequency of the walk out of a sample of size 1000.

Human label	2-gram label	3-gram label	4-gram label
mental health	psychological->stress (18.9%)	stress->disorder->symptoms (7.5%)	psychological->stress->disorder->symptoms (5.7%)

Here follows the set of topic-specific measures that have been used to classify the topic and to analyse its structural properties (see manuscript for details):

Table 2: Summary measures

JSD	Mean propensity	Variance propensity	Modularity	Barrat Clustering Coeff.
0.648831	0.010014	0.000669	0.156571	0.598007
55	114	104	97	90

Based on the aforementioned measures, Topic 82 has been classified as a SPECIALIZED topic.

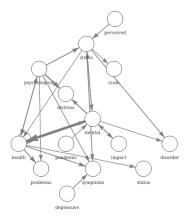


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.



prevalence

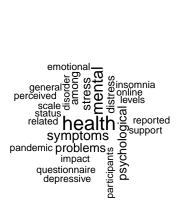
status tress levels

support by psychological
pandemic health by greatly
general among mental symptoms
symptoms
perceived depressive beneficial
wellbeing psychiatric

consequences
questionnaire
outcomesimpact
scalepsychological
levels problems
status distress crisis
physical health
supports
possible disorder
possible disorder
among
depressive



interventions pandemic psychological levels online stress scale mental reported health spokes among stress spokes sample interventions pandemic psychological levels on the stress scale mental reported among stress spokes spokes stress spokes among spokes spoke



Out-degree Betweenness PageRank

Figure 2: Top 25 unigrams (i.e., nodes) by measure.

general-population social-support physical-activity COVID-19-patients health-problems public-health cross-section-problems public-health cross-section-section-petitive-care covers and population of the covers of

symptoms>stress
symptoms>mental
symptoms-health stress>stress
stress>symptoms health>stress>stress
mental>symptoms stress>health
mental>psychological
psychological>stress mental>stress
mental>distress health>mental
mental>mental
health>distress mental>health>health>symptoms
stress>mental
health>lealth>health>symptoms
stress>mental
health>health>listress-mental
distress>mental
health>psychological>mental
health>psychological>metal
health>psychological>filtresshealth
health>psychologicals
stress>psychological

psychological-symptoms psychological-symptoms psychological-impact depressive-symptoms and depressive-symptoms are seen and seen and seen are seen and seen and seen are seen as trees-symptoms psychological-series psycho

Figure 3: Top 25 bigrams (i.e., edges) by measure.

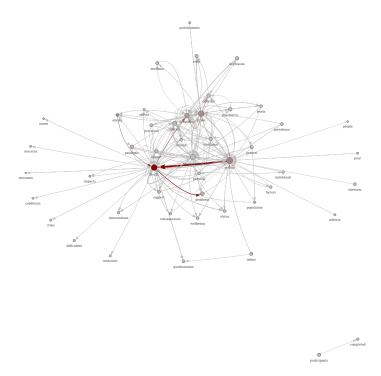


Figure 4: Filtered topic network (by weight). Layout based on Fruchterman-Reingold algorithm. Node size is proportional to topic-specific word probability provided by LDA. Edge width is proportional to topic-specific bigram weight provided by LDA2Net method. Node and edge color represent their betweenness centrality. Isolated nodes have been removed after filtration.