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

Topic 80 companion sheet

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

This file contains the following supplementary information for Topic 80 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
antiviral treatments	antiviral->drugs (15.3%)	antiviral->drugs->used (1.9%)	antiviral->drugs->used->treatment (0.7%)

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.
value	0.691806 74	0.009335	0.000487	0.115709	0.596658
rank		104	89	85	86

Based on the aforementioned measures, Topic 80 has been classified as a SPECIALIZED topic.

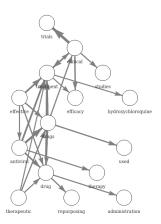


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.



specific administration repurposing efficacy therapies usedrugs currently treatment adverse effects trials effective antiviral drug studies used combination hydroxychloroquine treatments

lopinavir administration hydroxychloroquine effects grantiviral clinical azithromycin treatment or trials studies repurposing or treatment or trials studies repurposing or therapeutic grant therapeutic grant medicine

repurposing
repurposed 50 therapy
efficacy 901 b by therapy
combination 90 b by several effective p
adverse Clinical
antiviral
treatmentuse
therapeutic specific
hydroxychloroquine prevention
prevention

currently
hydroxychloroquine
remdesivir
integration drug antiviral
used ritonavir
stole treatment
stole treatm

administration studies with the rapy of the rape of th

Out-degree Betweenness PageRank

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



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

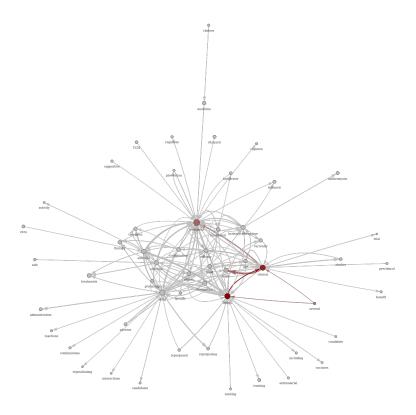


Figure 4: Filtered topic network (by weight). Layout based on Fruchterman-Reingold algorithm. Node size is proportional to topic-specific word probability provided by LDA. Edge width is proportional to topic-specific bigram weight provided by LDA2Net method. Node and edge color represent their betweenness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