

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

## Topic 73 companion sheet

G. Minello

C.R.M.A. Santagiustina

M. Warglien

This file contains the following supplementary information for Topic 73 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
paper info	web->response (47.6%)	web->response->one (16.4%)	web->response->disclosure->including (9%)

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	0.618312	0.006976	0.000135	0.004368	0.585444
rank	48	7	17	56	75

Based on the aforementioned measures, Topic 73 has been classified as a CROSS-CUTTING topic.

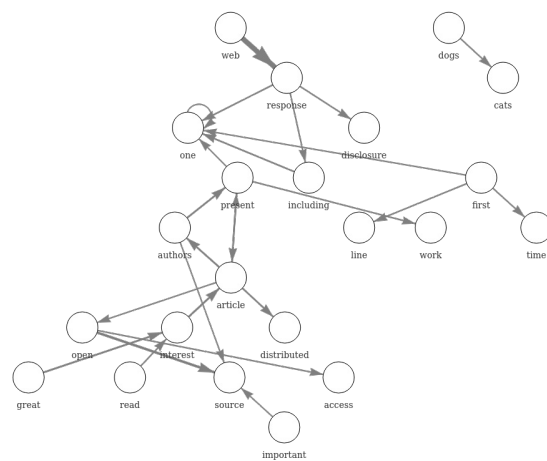


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.

A word cloud of terms related to the research. The most prominent word is 'response'. Other significant words include 'author', 'source', 'present', 'important', 'including', 'cats', 'line', 'one', 'web', 'result', 'article', 'open', 'work', 'relationships', 'part', 'presented', 'table', 'author', 'dogs', 'figure', 'first', 'several', 'interest', 'authors', 'provided', and 'pro'. The words are arranged in a circular pattern, with 'response' at the center.

A word cloud of terms related to the study. The most prominent word is 'response'. Other significant words include 'disclosure', 'presented', 'work', 'recent', 'several', 'present', 'authors', 'interest', 'part', 'describe', 'line', 'report', 'important', 'found', 'open', 'time', 'article', 'cats', 'including', 'access', 'distributed', 'first', 'one', 'source', and 'disclosure'. The words are arranged in a circular pattern around the central word 'response'.

A word cloud of terms related to the research project. The most prominent word is 'present'. Other significant words include 'work', 'distributed', 'open', 'authors', 'web', 'article', 'response', 'part', 'used', 'another one', 'including', 'found', 'several', 'source', 'first', 'interest', 'important', 'line', 'considered', 'describe', 'access', and 'presented'. The words are arranged in a circular pattern around the central word 'present'.

A word cloud of terms related to the article. The most prominent words are 'response', 'one', 'source', 'authors', 'present', 'important', 'additional', 'used', 'part', 'several', 'including', 'recent', 'work', 'time', 'first', 'article', 'reported', 'provided', 'different', 'open', 'found', 'interest', 'new', 'line', 'presented', 'used', 'part', 'several', 'including', 'recent', 'work', 'time', 'first', 'article', 'reported', 'provided', 'different', 'open', 'found', 'interest', 'new', 'line'.

3

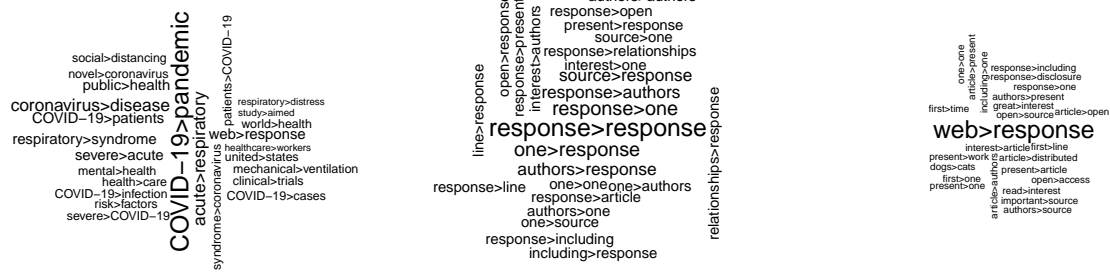


Figure 3: Top 25 bigrams (i.e., edges) by measure.

