

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

Topic 32 companion sheet

G. Minello

C.R.M.A. Santagiustina

M. Warglien

This file contains the following supplementary information for Topic 32 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
Covid-19 literature	must->considered (6.9%)	one->especially->important (2.5%)	one->especially->important->reasons (1.7%)

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	0.675373	0.008013	0.000121	0.136094	0.540646
rank	69	50	11	90	25

Based on the aforementioned measures, Topic 32 has been classified as a CROSS-CUTTING topic.

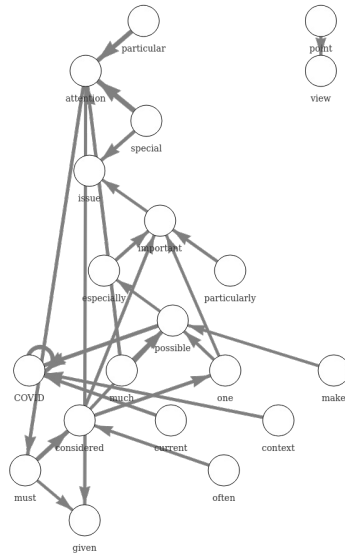


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.

A word cloud of terms related to COVID-19. The word "COVID" is the largest and most central. Other prominent words include "possible", "attention", "important", "given", "difficult", "often", "issue", "since", "special", "present", "context", "question", "various", "conditions", "especially", "considered", "one", "must", "general", "particular", "specific", "certain", "different", and "possibility". The words are arranged in a circular pattern around the central "COVID" word.

A word cloud of COVID-19 related terms. The most prominent words are 'attention', 'COVID', 'possible', 'special', 'important', 'considered', 'one', 'particular', 'given', 'aspects', 'different', 'context', 'various', 'since', 'current', 'especially', 'conditions', 'must', 'issue', 'often', 'certain', 'particularly', 'specific', 'present', and 'various'. The words are arranged in a circular pattern, with 'attention' and 'COVID' being the largest and most central.

A word cloud of COVID-19 related terms. The most prominent words are 'attention', 'possible', 'important', 'issue', 'particular', 'specific', 'question', 'difficult', 'often', 'aspects', 'situations', 'current', 'context', 'reasons', 'must', 'different', 'given', 'considered', 'one', 'since', 'view', 'conditions', 'present', 'especially', and 'different'. The words are arranged in a circular pattern, with 'attention' and 'possible' being the largest and most central.

A word cloud of terms related to the COVID-19 pandemic. The words are arranged in a circular pattern, with 'special' and 'important' being the largest and most central. Other prominent words include 'attention', 'possible', 'particular', 'must', 'often', 'considered', 'especially', 'make', 'important', 'context', 'current', 'various', 'COVID', 'much', 'make', 'considered', 'especially', 'must', 'often', 'particular', 'possible', 'attention', 'specific', 'given', 'different', 'present', 'several', 'partially', 'one', 'conditions', 'difficult', 'certain', and 'various'. The words are in various shades of gray and are of different sizes, indicating their relative frequency or importance in the dataset.

A word cloud of terms related to COVID. The most prominent word is 'COVID' in the center. Other visible words include 'present', 'possible', 'important', 'attention', 'particular', 'another', 'specific', 'related', 'since', 'way', 'different', 'part', 'given', 'must', 'one', 'issue', 'especially', 'casereasons', 'often', 'considered', 'conditions', 'aspects', and 'question'. The words are arranged in a circular pattern around the central 'COVID' word.

COVID

3

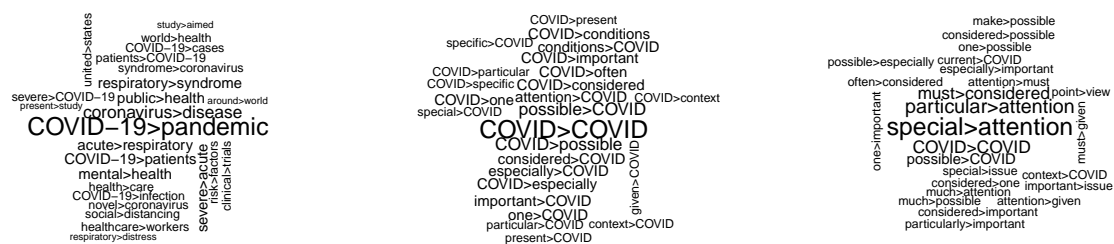


Figure 3: Top 25 bigrams (i.e., edges) by measure.

