

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

Topic 101 companion sheet

G. Minello

C.R.M.A. Santagiustina

M. Warglien

This file contains the following supplementary information for Topic 101 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.

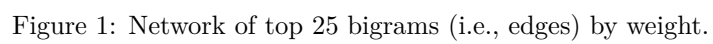
Human label	2-gram label	3-gram label	4-gram label
health care personnel and pandemic	healthcare->workers (58.3%)	healthcare->workers->work (21.4%)	healthcare->workers->work->providers (3.9%)

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	0.706331	0.008158	0.000310	0.000001	0.572800
rank	81	57	62	39	56

Based on the aforementioned measures, Topic 101 has been classified as a CROSS-CUTTING topic.



A word cloud of healthcare-related terms. The most prominent word is 'healthcare' in the center. Other large words include 'work', 'nurses', 'doctors', 'hospitals', 'physicians', 'care', 'workers', 'support', 'providers', 'occupational', 'employees', 'pandemic', 'personnel', 'among', 'job', 'nursing', 'home', 'protect', 'burnout', 'work', 'health', 'care', 'hospital', 'hcws', 'worker', 'professional', 'nursing', 'occupational', 'employees', 'pandemic', 'personnel', 'among', 'job', 'nursing', 'home', 'protect', 'burnout', 'work', 'health', 'care', 'hospital', 'hcws', 'worker', 'professional'.

A word cloud of COVID-19 related terms. The most prominent word is 'healthcare' in the center. Other large words include 'hospitals', 'nursing', 'physicians', 'support', 'burnout', 'work', 'job', 'doctors', 'exposure', 'occupational', 'pandemic', 'workers', 'home care', 'hospital', 'equipment', 'employees', 'protection', 'protect', 'health', 'amplify', 'hcs', 'nurses', and 'protection'. The words are arranged in a circular pattern around the central 'healthcare' word.

A word cloud visualization showing various terms associated with healthcare workers. The most prominent words are "healthcare", "workers", and "work". Other visible words include "nurses", "burnout", "providers", "conducted", "hospital", "occupational", "conditions", "employees", "health", "provided", "strain", "hows", "professional", "personnel", "home", "support", "settings", "pandemic", "among", "worker", "care", "nursing", and "wound". The words are arranged in a circular pattern around a central point, with varying font sizes indicating their relative frequency or importance.

A word cloud of terms related to the COVID-19 pandemic. The words are arranged in a circular pattern, with 'healthcare workers' and 'work' being the most prominent. Other visible words include 'experienced', 'involved', 'employees', 'burnout', 'job', 'care', 'among', 'high', 'occupational', 'worked', 'hospitals', 'pandemic', 'nurses', 'hospital', 'hours', 'hows', 'support', 'home', 'exposure', 'caring', and 'especially'.

3



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

