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

Topic 96 companion sheet

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

This file contains the following supplementary information for Topic 96 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
orthopedic traumas	wound->healing (5.2%)	wound->healing->repair (3.7%)	wound->healing->repair->attempt (0.5%)

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	0.521756	0.007171	0.000379	0.103614	0.536496
rank	8	11	74	84	21

Based on the aforementioned measures, Topic 96 has been classified as a SPECIALIZED topic.

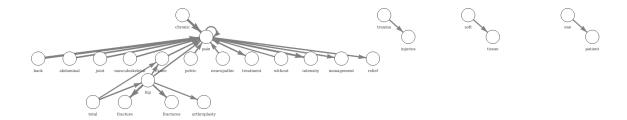


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.

upper primary reconstruction back fractures total headjoint treatment mm injuries position painhip traumatype fracture technique following patient by significant processing tracture to the position patient by the primary reconstruction packets and primary reconstruction primary reconstruction primary reconstruction packets and primary reconstruction primary reconstruction primary reconstruction packets and primary reconstruction primary reco

without interest and primary without interesting and primary without interesting the pack patient total on the pack patient arthroplasty is a packet and primary arthroplasty is a packet and primary in the packet in the primary property in the primary property is a packet in the primary property in the primary propert

complications following outcomespatient hip intensity of the point pall ninjuries of trauma knee treatment used fractures relief arthroplasty outcome management

outcomes following pelvic

Azerona policy pelvic

Azerona pelvic

Azerona policy pelvic

Azerona pelvic

Azerona policy pelvic

Azerona p

complications
head management
study on knee
following intensity
treatment ovia on the study on t

management

pain total

pain time treatment fracture significant intensity fractures outcomes used compared in a compared in a

Out-degree Betweenness PageRank

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



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

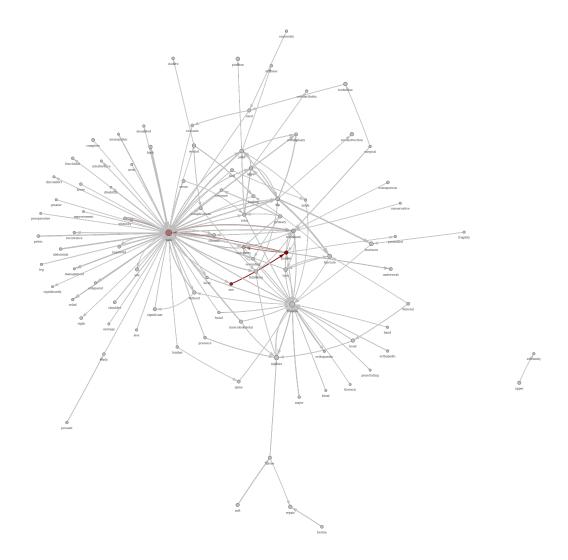


Figure 4: Filtered topic network (by weight). Layout based on Fruchterman-Reingold algorithm. Node size is proportional to topic-specific word probability provided by LDA. Edge width is proportional to topic-specific bigram weight provided by LDA2Net method. Node and edge color represent their betweenness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