

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

Topic 31 companion sheet

G. Minello

C.R.M.A. Santagiustina

M. Warglien

This file contains the following supplementary information for Topic 31 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
digital technology	digital->version (6.1%)	digital->scanning->calorimetry (3.2%)	digital->scanning->calorimetry->technology (3.2%)

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	0.601469	0.008887	0.000443	0.000002	0.552114
rank	37	89	83	41	29

Based on the aforementioned measures, Topic 31 has been classified as a SPECIALIZED topic.

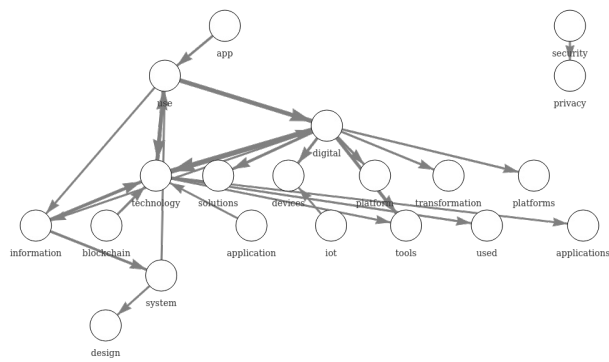


Figure 1: Network of top 25 bigrams (i.e., edges) by weight.

[illegible]

3



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

