**TOPIC MODELLING – LDA**

Pastrand doar adjective si substantive si doar bigrame si trigrame:

remove\_pos = ['ADV', 'PRON', 'PART', 'DET', 'SPACE', 'NUM', 'SYM', 'ADP', 'VERB', 'CCONJ']

stop\_words = ['paper', 'present', 'propose', 'datum', 'people', 'result', 'solution', 'case', 'order', 'base', 'ieee', 'privacy', 'policy', 'new', 'old', 'context']

remove\_entities = ['PERSON', 'NORP', 'FAC', 'ORG', 'GPE', 'LOC', 'LANGUAGE', 'DATE', 'TIME', 'PERCENT', 'MONEY', 'QUANTITY', 'CARDINAL', 'ORDINAL']

lda\_model = LdaMulticore(corpus=corpus, id2word=dictionary, iterations=500, num\_topics=1, workers=4, passes=100)

Author\_id = 562:

0: 0.014\*"cloud computing" + 0.014\*"scheduling algorithm" + 0.010\*"real time" + 0.010\*"resource management" + 0.010\*"large scale" + 0.008\*"satellite image" + 0.006\*"service level" + 0.006\*"task scheduling" + 0.006\*"smart city" + 0.006\*"cloud service"

Author\_id = 534:

0: 0.011\*"large scale" + 0.011\*"mobile device" + 0.008\*"opportunistic network" + 0.005\*"real time" + 0.004\*"cloud computing" + 0.004\*"wide range" + 0.003\*"quality life" + 0.003\*"scale system" + 0.003\*"inconvenience helpful" + 0.003\*"sale account"

Author\_id = 1146:

0: 0.015\*"virtual reality" + 0.008\*"real time" + 0.007\*"virtual environment" + 0.005\*"sensory substitution" + 0.005\*"virtual space" + 0.004\*"sound localization" + 0.004\*"sound source" + 0.004\*"smith chart" + 0.003\*"virtual world" + 0.003\*"augmented reality"

Author\_id = 841:

0: 0.012\*"x ray" + 0.008\*"electron microscopy" + 0.006\*"composite material" + 0.006\*"drug delivery" + 0.005\*"ray diffraction" + 0.004\*"iron oxide" + 0.004\*"sol gel" + 0.003\*"microscopy sem" + 0.003\*"transmission electron" + 0.003\*"thin film"

Author\_id = 872:

0: 0.014\*"neural network" + 0.009\*"real time" + 0.008\*"wireless sensor" + 0.008\*"fractal dimension" + 0.007\*"unmanned aerial" + 0.006\*"sensor network" + 0.006\*"image processing" + 0.006\*"large scale" + 0.005\*"convolutional neural" + 0.005\*"control system"

Author\_id = 1284:

0: 0.011\*"natural language" + 0.008\*"e learning" + 0.006\*"textual complexity" + 0.006\*"chat conversation" + 0.005\*"social network" + 0.004\*"learning process" + 0.003\*"polyphonic model" + 0.003\*"machine learning" + 0.003\*"artificial intelligence" + 0.003\*"web page"

Cu *stop\_words = ['paper', 'present', 'propose', 'datum', 'people', 'result', 'solution', 'case', 'order', 'base', 'ieee', 'privacy', 'policy',*

*'new', 'old', 'context', 'high', 'different', 'new', 'old', 'research', 'type', 'approach', 'important', 'main', 'range',*

*'helpful', 'large', 'difficult', 'available', 'amount', 'useful', 'importance', 'article', 'abstract', 'scale', 'copyright',*

*'real', 'quality', 'inconvenience', 'benefit', 'unavailable', 'term', 'condition', 'interest', 'organization', 'use',*

*'task', 'student', 'professor', 'teacher', 'university']*

lda\_model = LdaMulticore(corpus=corpus, id2word=dictionary, iterations=**500**, num\_topics=1, workers=4, passes=**200**)

Id = 534:  
 0: 0.011\*"**mobile device**" + 0.009\*"**opportunistic network**" + 0.005\*"**cloud computing**" + 0.004\*"information product" + 0.004\*"service updated" + 0.004\*"fault tolerance" + 0.004\*"technology humanity" + 0.004\*"site agreement" + 0.004\*"profit world" + 0.004\*"account management"

Id = 562:

0: 0.017\*"**scheduling algorithm**" + 0.016\*"**cloud computing**" + 0.010\*"**resource management**" + 0.008\*"**satellite image**" + 0.007\*"**service level**" + 0.007\*"**smart city**" + 0.007\*"**cloud service**" + 0.006\*"**time series**" + 0.006\*"**fault tolerant**" + 0.006\*"**genetic algorithm**"

Id = 1146:

0: 0.016\*"**virtual reality**" + 0.007\*"**virtual environment**" + 0.005\*"**sound localization**" + 0.005\*"**sensory substitution**" + 0.005\*"**virtual space**" + 0.005\*"**assistive device**" + 0.004\*"virtual world" + 0.004\*"smith chart" + 0.003\*"fear level" + 0.003\*"sound source"

Id = 841:

0: 0.013\*"**x ray**" + 0.010\*"**electron microscopy**" + 0.007\*"**drug delivery**" + 0.006\*"**composite material**" + 0.005\*"**ray diffraction**" + 0.004\*"iron oxide" + 0.004\*"sol gel" + 0.004\*"microscopy sem" + 0.003\*"thin film" + 0.003\*"mesoporous silica"

Id = 872:

0: 0.016\*"**neural network**" + 0.009\*"**fractal dimension**" + 0.009\*"**wireless sensor**" + 0.007\*"**unmanned aerial**" + 0.006\*"**image processing**" + 0.006\*"**sensor network**" + 0.005\*"**convolutional neural**" + 0.005\*"**artificial intelligence**" + 0.005\*"**energy consumption**" + 0.005\*"**optic disc**"

Id = 1284:

0: 0.012\*"**natural language**" + 0.009\*"**e learning**" + 0.007\*"**textual complexity**" + 0.007\*"**chat conversation**" + 0.006\*"**social network**" + 0.004\*"polyphonic model" + 0.004\*"learning process" + 0.003\*"artificial intelligence" + 0.003\*"machine learning" + 0.003\*"collaborative learning"