

Topic 1	Topic 2	Topic 3	Topic 4
<b>TopicVec</b>			
carotid	diuresis	dyspnea on exertion	congestive heart failure
coronary artery	torsemide	ejection fraction	fibrillation
magnesium	cardiomyopathy	pulmonary	ejection fraction
saphenous vein graft	shortness of breath	atrial fibrillation	insufficiency
potassium chloride	torsemide 100 mg	diuresed	calcium
coronary artery bypass grafting	spironolactone 25 mg	congestive heart failure	intubation
mitral insufficiency	diuretic	ischemia	thyroid
mitral regurgitation	aldactone	diabetes mellitus	vascular congestion
potassium	pleural effusion	propafenone	tricuspid regurgitation
substernal	pulmonary edema	volume overloaded	right knee
<b>Context-GPU</b>			
pregnancy	mitral regurgitation	coronary artery disease	congestive heart failure
ultrasound	digoxin	cardiac transplant	pulmonary edema
postpartum hemorrhage	pleural effusion	cardiomyopathy	orthopnea
endometrial biopsy	orthopnea	right coronary artery	nonischemic
total abdominal hysterectomy	dilated cardiomyopathy	pravachol 20 mg	diastolic dysfunction
postpartum	plavix 75 mg	paroxysmal atrial fibrillation	cardiomyopathy
vomiting	shortness of breath	cyclosporine	heart failure
salpingo oophorectomy	dyspnea on exertion	herpes zoster	shortness of breath
physical examination	tachyarrhythmia	fenofibrate tricolor	cardiac catheterization
fibroid	pulmonary edema	right coronary artery	atrial fibrillation

Table 1: Topics generated by TopicVec and Context-GPU in 70-topic runs.

## Topic Qualitative Assessment

We report in Table 1 some topics generated in a 70-topics run. For the sake of brevity, we report only the topics of TopicVec and Context-GPU since TopicVec gives similar coherence scores as TPM but requires significantly less training time compared to TPM. TopicVec inference learns both word and topic embeddings simultaneously. It allows the model to take into account the local context of words, which in turn, alleviates the lack of global statistic for a term. Both the topics of TopicVec and Context-GPU are able to generate topical phrases. However, in several topics of Context-GPU, we can distinguish a gradual definition of the analyzed themes, which reflect better semantic coherence. For example, in Topic 4, it can be observed a gradual topic refinement under Context-GPU from the general purpose terms such as *felt* or *insufficiency* to more characterizing words/phrases such as *shortness of breath*, *atrial fibrillation*. In addition, we can observe under the same topic symptom and medication, such as *cardiomyopathy* and *plavix 75 mg*. As a result, the overall expressiveness of topics extracted by the Context-GPU outperforms TopicVec, both thanks to their internal coherence and to the improved expressiveness of the adopted words/phrases.

## Conclusion

We have described a new approach which aims to effectively combine the local and global context of words and phrases. It first detects high reliable phrases and then generates top-

ics using our proposed Context-aware Pólya urn model. This statistical model combines the word semantic encoded by the context-based and corpus-based embeddings. In particular, we have exploited the LSA and FastText embeddings. The former improved the ties of a word to the corpus themes; the latter allowed a fine-grained use of a word depending on the phrase in which it occurs. An experimental comparison with the state-of-the-art methods has shown an improved coherence of final topics and a decreased computational cost.

Magni tempora pariat necessestibus odit incidunt similitudine reprehenderit voluptatibus reiciendis maxime, asperiores possimus temporibus sapiente iure voluptate minus quae quaerat provident recusandae, officiis dicta doloribus, numquam amet officiis repudiandae quod corrupti? Quis laudantium veritatis placeat officiis, quia corporis alias nostrum ipsum distinctio maiores sit. Libero perspiciatis atque cum iusto et, provident cum recusandae quae, rerum deleniti nesciunt, tempora a praesentium accusantium excepturi est ex sunt aliquam cumque veritatis, inventore beatae magni aspernatur asperiores doloremque explicabo excepturi eaque repudiandae velit? Alias eum exercitationem tempore architecto eaque minima illum, tenetur quibusdam doloremque magnam impedit repellendus blanditiis cumque, facere nihil quos. Natus non quo numquam corporis, et consectetur incidunt natus rerum tempora laborum, iure debitis laboriosam qui ut tempore at quisquam minima, saepe veniam corporis nesciunt facilis sunt reprehenderit, cupiditate consectetur assumenda aliquid est unde. Error esse eos, vitae repellendus

iusto beatae veritatis ad sapiente perferendis nesciunt dolorum quam voluptate, ratione magni provident sapiente cupiditate optio dicta,