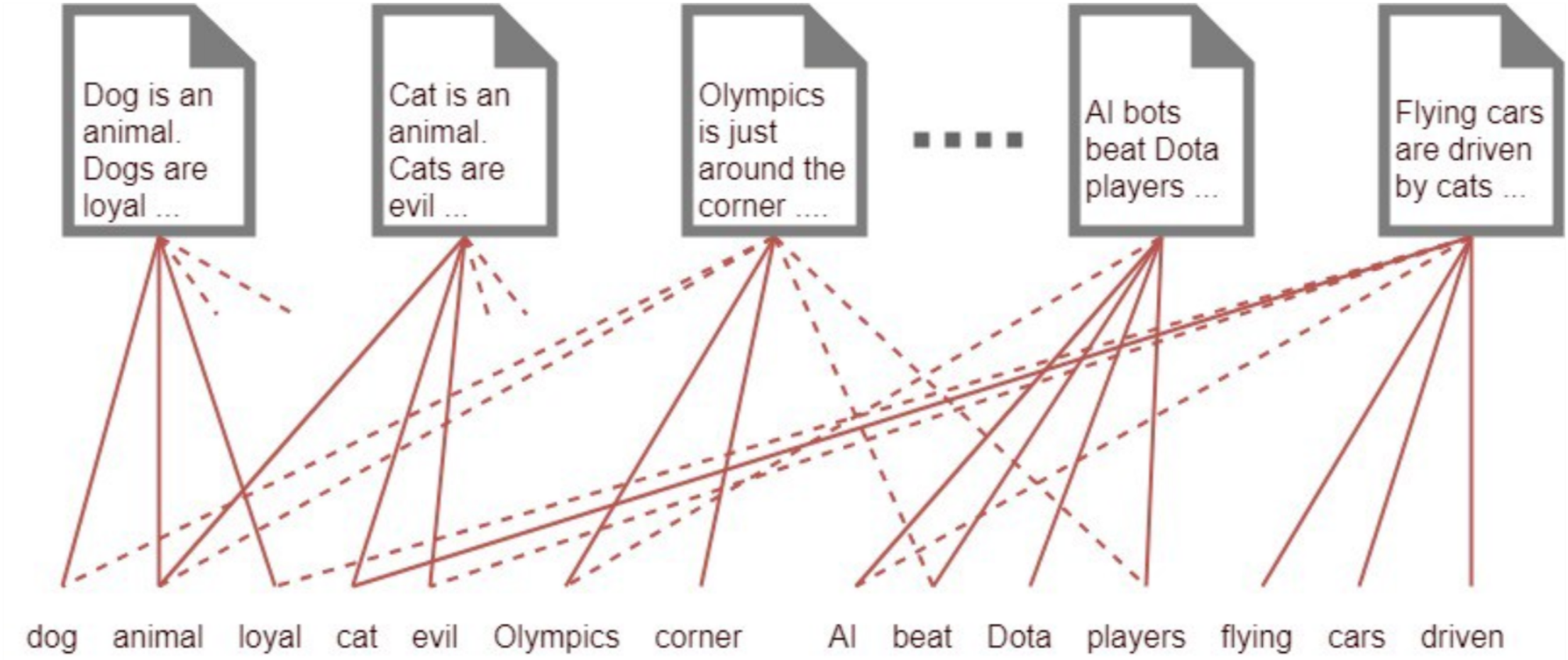
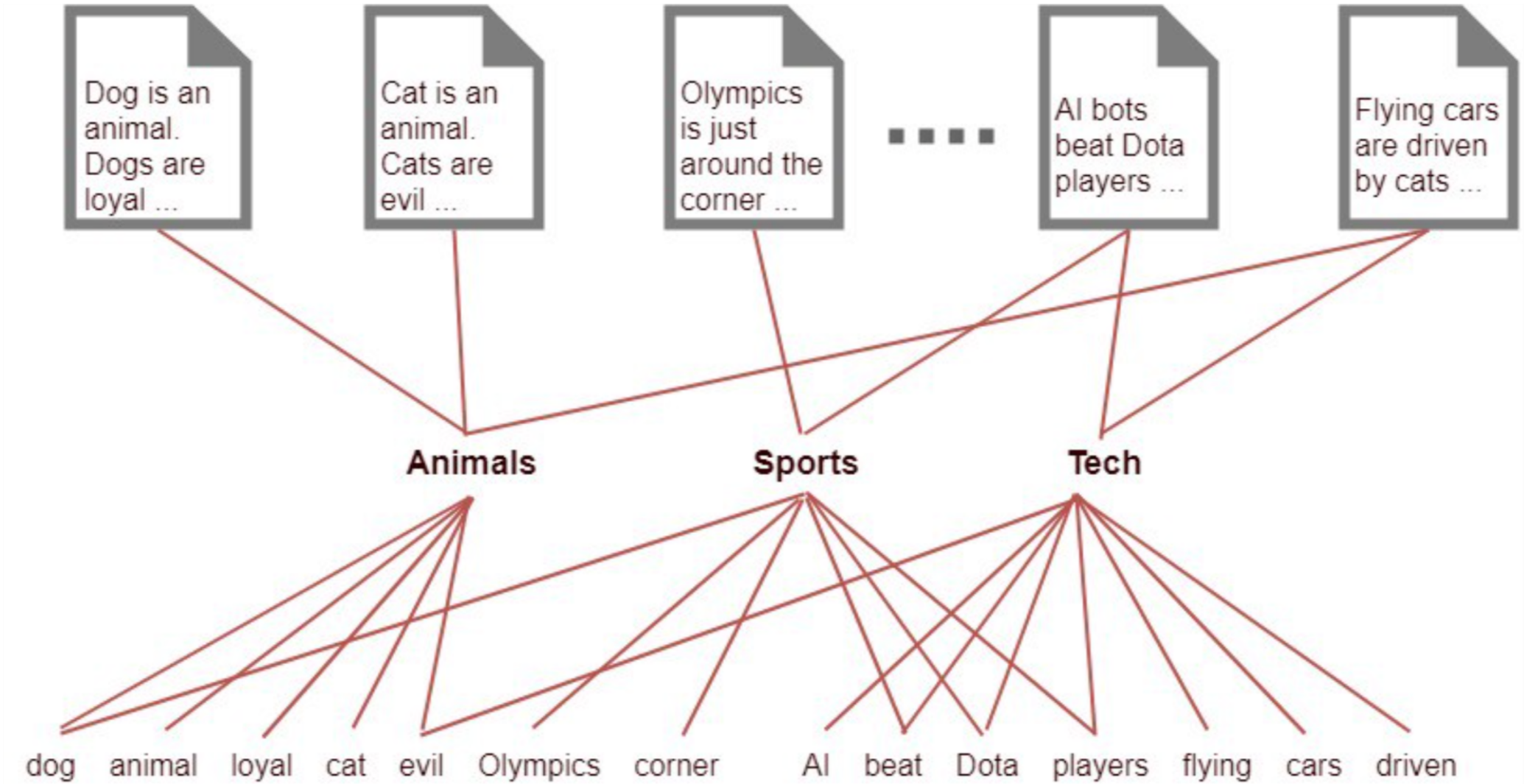


Modeling documents just with words. You can see that we can't really infer any useful information due to the large amount of connections



Words are modeled by a set of topics and documents are modeled by a set of topics. The relationships are clearer than the first example because there's a fewer connections than the first example.

	Word1	word2	word3	word4	.....
Topic1	0.01	0.23	0.19	0.03	
Topic2	0.21	0.07	0.48	0.02	
Topic3	0.53	0.01	0.17	0.04	

Each topic contains a score for all the words in the corpus.

A layman's example

Suppose you have various photographs(*documents*) with captions(*words*). You want to display them in a gallery so you decide to categorize the photographs on various themes(*topics*) based on which you will create different sections in your gallery.

Photo by Soragrit Wongsa on Unsplash

	Word1	word2	word3	word4	.....
Topic1	0.01	0.23	0.19	0.03	
Topic2	0.21	0.07	0.48	0.02	
Topic3	0.53	0.01	0.17	0.04	

Each topic contains a score for all the words in the corpus.

Figure 1. The intuitions behind latent Dirichlet allocation. We assume that some number of "topics," which are distributions over words, exist for the whole collection (far left). Each document is assumed to be generated as follows. First choose a distribution over the topics (the histogram at right); then, for each word, choose a topic assignment (the colored coins) and choose the word from the corresponding topic. The topics and topic assignments in this figure are illustrative—they are not fit from real data. See Figure 2 for topics fit from data.