

# From Categorical Words to Numbers

- Map each word to a vector of 0s and 1s (“1-hot repr.”):

$$\text{cat} \mapsto (0, 0, \dots, 0, 1, 0, \dots, 0)$$

- Sentence is then a matrix:

		the	cat	is	on	the	mat
Vocabulary size: 1.3M English 2.2M Czech	↑ a	0	0	0	0	0	0
	about	0	0	0	0	0	0
	...	...	...	...	...	...	...
	cat	0	<b>1</b>	0	0	0	0
	...	...	...	...	...	...	...
	is	0	0	<b>1</b>	0	0	0
	...	...	...	...	...	...	...
	on	0	0	0	<b>1</b>	0	0
	...	...	...	...	...	...	...
	the	<b>1</b>	0	0	0	<b>1</b>	0
	↓ zebra	0	0	0	0	0	0