

# Are Dictionary Definitions of Verbs in Corpora?

Discovering Dead Ends  
in Generating Explanations of Verbs

Marie Stará

December 9, 2022

# Dictionary Definitions of Verbs

How to *define* a verb?

- ▶ hunt: to kill animals for food or for their skin or other parts, or for sport  
—Macmillan Dictionary
- ▶ ulovit: získat lovem (to obtain by hunting)  
—Slovník spisovné češtiny

## Definition to CQL

hunt: to kill animals for food or for their skin or other parts, or for sport

- ▶ kill animals
- ▶ kill (animals) for food
- ▶ kill (animals) for skin
- ▶ kill (animals) for sport

```
lemma = "hunt" []0,7 [lemma = "kill" []0,2 [lemma =  
"animal" ] within <s/> | [lemma = "kill" []0,2 [lemma  
= "animal" ] []0,7 [lemma = "hunt" ] within < s/>
```

# Results

	queries	found	low freq.		queries	found	low freq.
bát (se)	8	4	3	fear	5	3	2
dát	19	16	1	give	7	7	3
dostat	6	4	1	get	5	5	1
existovat	3	3	0	exist	7	5	2
fungovat	5	5	5	function	7	5	1
jednat	11	8	1	act	6	5	3
ležet	4	3	1	lie	4	1	1
namlouvat	6	1	1	insinuate	2	0	-
pojistit	6	1	0	insure	3	2	1
skákat	4	2	1	jump	9	3	2
ulovit	4	3	2	hunt	10	7	3
uvést	10	4	2	initiate	4	1	1
večeřet	2	0	-	dine	1	1	1
vřít	5	4	3	boil	4	2	1
zabít	7	5	3	kill	4	2	1
začít	3	3	1	begin	4	4	2
zapomenout	6	2	2	forget	3	3	2
znemožnit	2	1	1	discredit	2	1	1
žít	4	3	3	live	4	3	0

# So What?

- ▶ the data are there (sort of)
- ▶ usefulness?

# Correction

**Table:** Czech verbs, sorted by percentage of found queries and percentage of low-frequency results.

	% found	% low freq.		% found	% low freq.
existovat	100	0	existovat	100	0
fungovat	100	100	pojistit	16,7	0
začít	100	33,3	dát	84,2	6,3
dát	84,2	6,25	jednat	72,7	12,5
vřít	80	75	dostat	66,7	25
ležet	75	33,3	ležet	75	33,3
ulovit	75	66,7	začít	100	33,3
žít	75	100	skákat	50	50
jednat	72,7	12,5	uvést	40	50
zabít	71,4	60	zabít	71,4	60
dostat	66,7	25	ulovit	75	66,7
bát (se)	50	75	bát (se)	50	75
skákat	50	50	vřít	80	75
znemožnit	50	100	fungovat	100	100
uvést	40	50	namlouvat	16,7	100
zapomenout	33,3	100	zapomenout	33,3	100
namlouvat	16,7	100	znemožnit	50	100
pojistit	16,7	0	žít	75	100
večeřet	0	-	večeřet	0	-