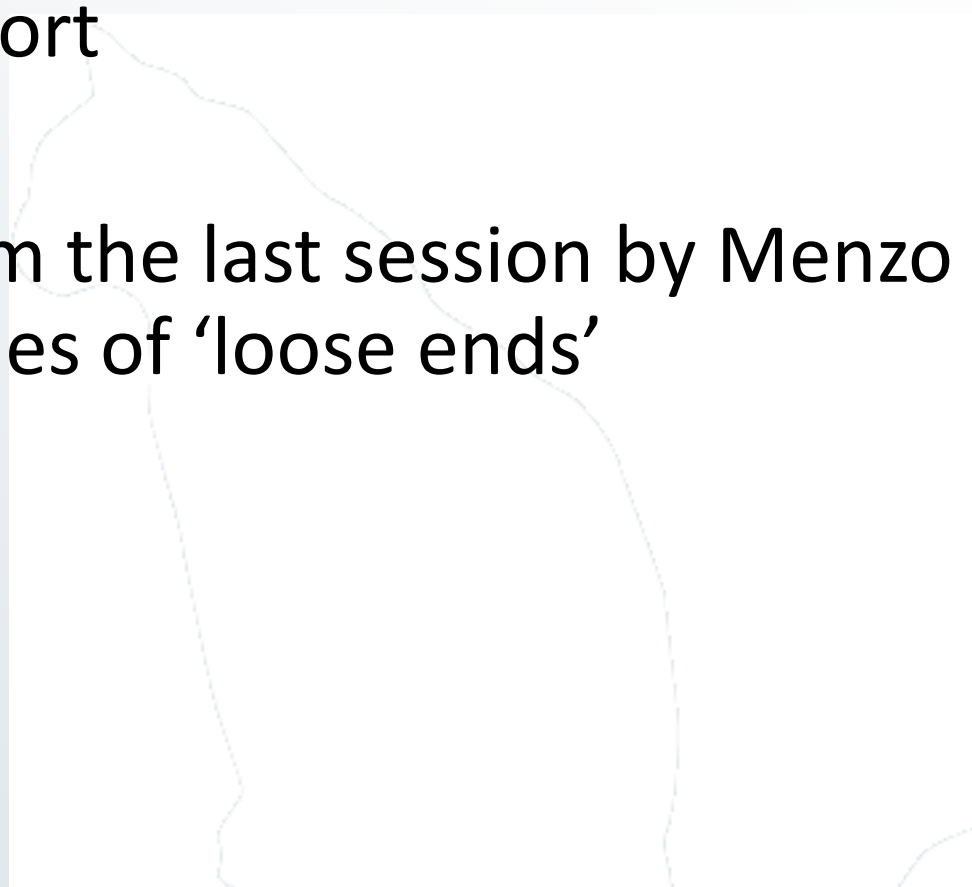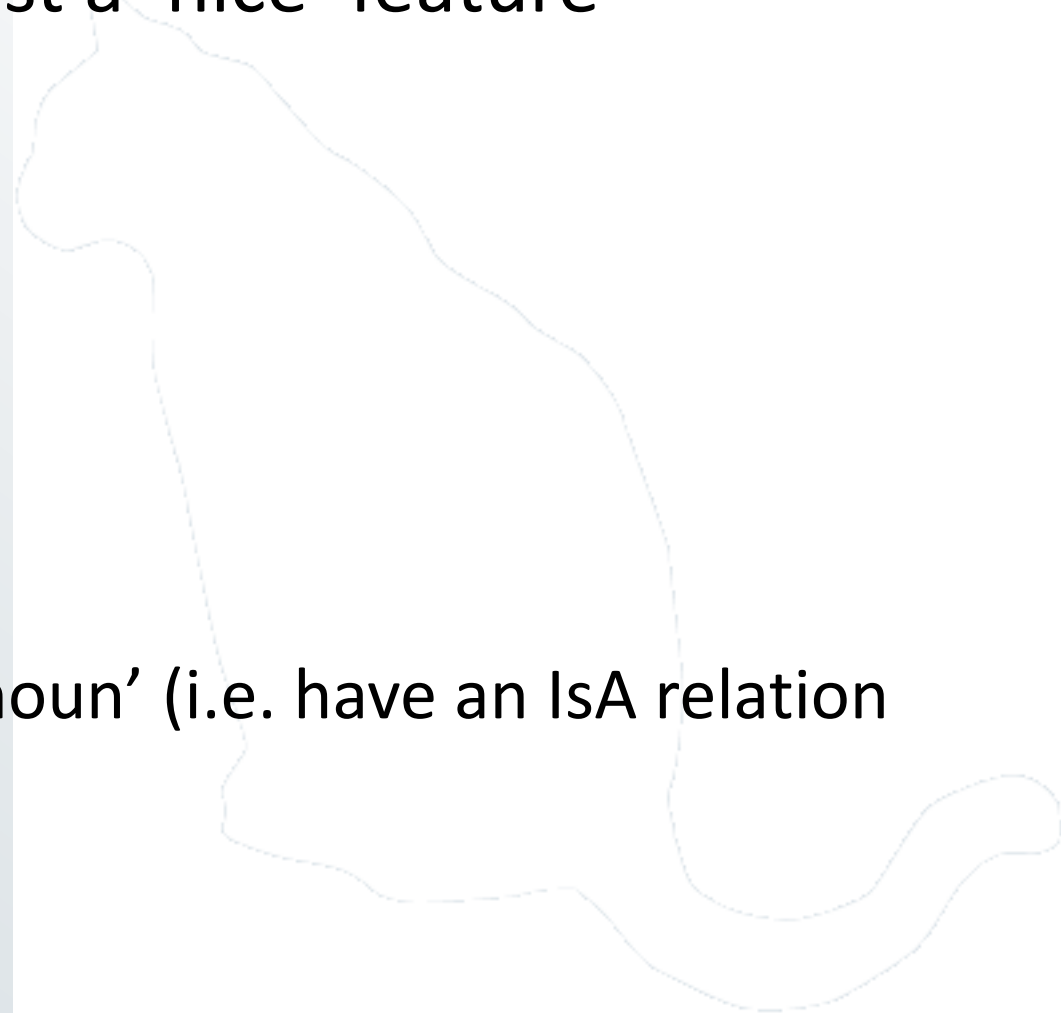# ISOcat: known issues

Known issues

- ISOcat: ongoing effort

- As will be clear from the last session by Menzo there are still a series of 'loose ends'

  - RELcat
  - Searching
  - Mapping
  - Definitions

- Linking DC is not just a 'nice' feature

  – Proper noun
  – Common noun
  – Mass noun
  – Count noun

  are all instances of 'noun' (i.e. have an IsA relation with it)

RelCat

- **Essential** for several Dutch tagsets

N(soort, ….) comes with 2 DCs:

1. Noun

2. Common

How to relate this with one of the DCs for 'common noun', even in case we would find the definition perfect?

Good news: in progress!

Searching

- <u>How to detect</u> which DCs are **Standardized**?
- Or have a Dutch language section?

- How to search using the keys? And what about language of keywords?

- How to detect which DCs 'belong together'
  (unless one mentions the tag set in the definition)

Searching

- How to search for alternative names (Data Element Names): transitief/vergankelijk; prepositie/ voorzetsel

- And the results: when not using 'exact' match and a specific field, MANY results come up, apparently unordered.

- While using exact + field may make you miss relevant entries.

Consequences of mapping

- Suppose, you map with a specific DC, and some essential changes are made to that DC
  - You may no longer want to map, but how do you know?

- Suppose there are several relevant DCs, you select one and just that one doesn't get standardized
  - You have to redo your work (but you first are to be aware that …)

Ill-defined DCs

- Profile: morphosyntax
  - Definition: semantic

- 'concept' in definition not defined in ISOcat , or

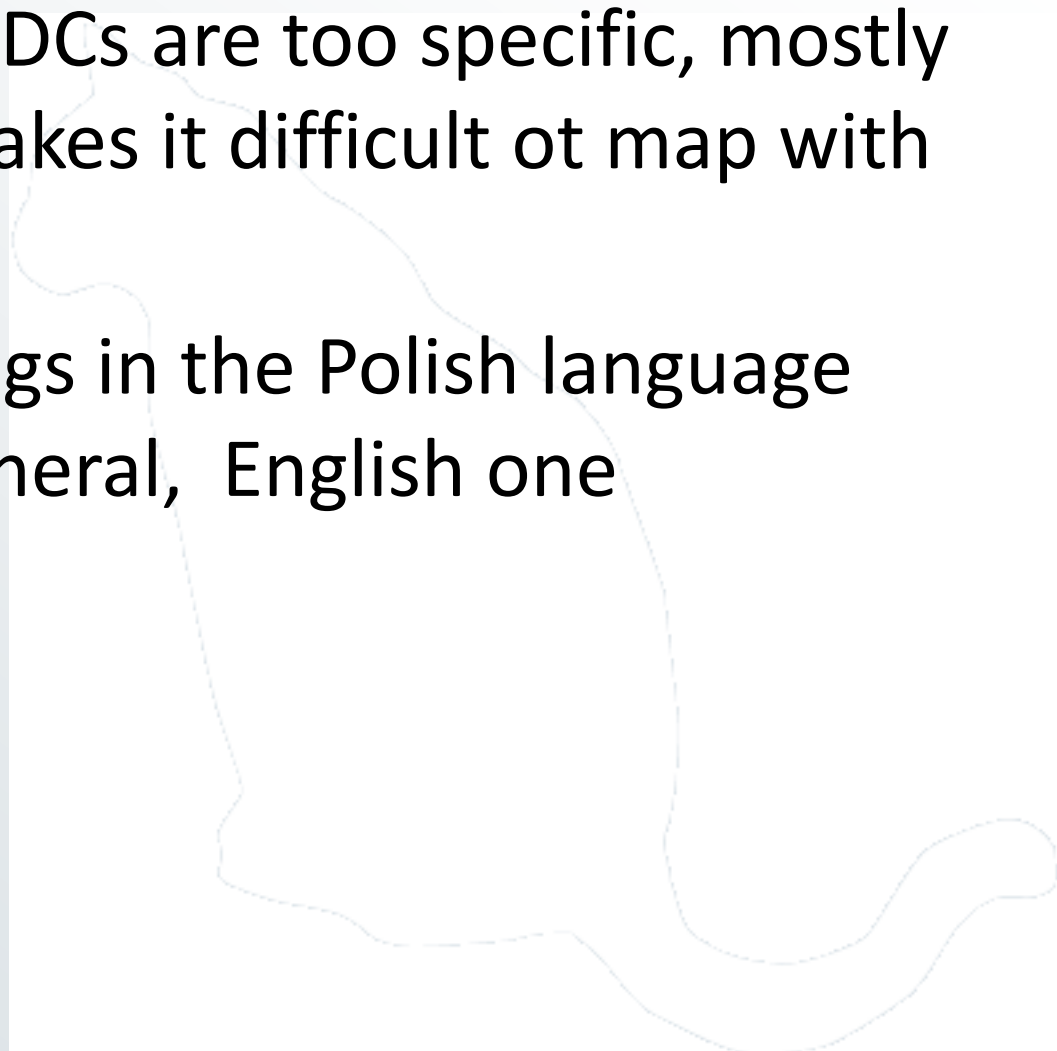- That concept comes with several DCs (which one was meant?)

Too many DCs

- There are too many 'almost the same' DCs, even within the same profile

Too vague DCs

- There are many DCs with rather 'empty' definitions
  - Proper noun: a noun or adjective denoting a single object
  - Common noun: a noun or adjective denoting a class of objects

Too language-specific DCs

- Quite a number of DCs are too specific, mostly Polish ones, this makes it difficult ot map with them

- i.e. stuff that belongs in the Polish language section is in the general, English one

Therefore, while for some technical issues solutions will come up

YOU should also be very careful yourself,

especially wrt the 'soundness' of the DCs, in particular

as far as definitions, profile, and translation are concerned!

Only in that case ISOcat can become a success story!

Thanks !