Document or database? Searching for the perfect storage paradigm for lexical data

Michal Měchura
Masaryk University, Brno, Czech Republic

Euralex, Mannheim, July 2022

www.lexiconista.com

Edit Configure Download Upload

someone@example.com ▼

⚡ New ID 2 ▶ 💾 Save ⟵ Cancel ⊘ Delete

Reload 3

able
arm
ask

```xml
<entry>
    <headword>ask</headword>
    <partOfSpeech>v</partOfSpeech>
    <sense>
        <definition>To look for an answer to a question.</definition>
        <example>I need to <h>ask</h> you a question.</example>
        <example>I don't know, <h>ask</h> your mother.</example>
    </sense>
    <sense>
        <definition>To talk to someone to see if they will do something.</definition>
        <example>She <h>asked</h> me to help her.</example>
        <example>We will have to <h>ask</h> for more money.</example>
    </sense>
</entry>
```
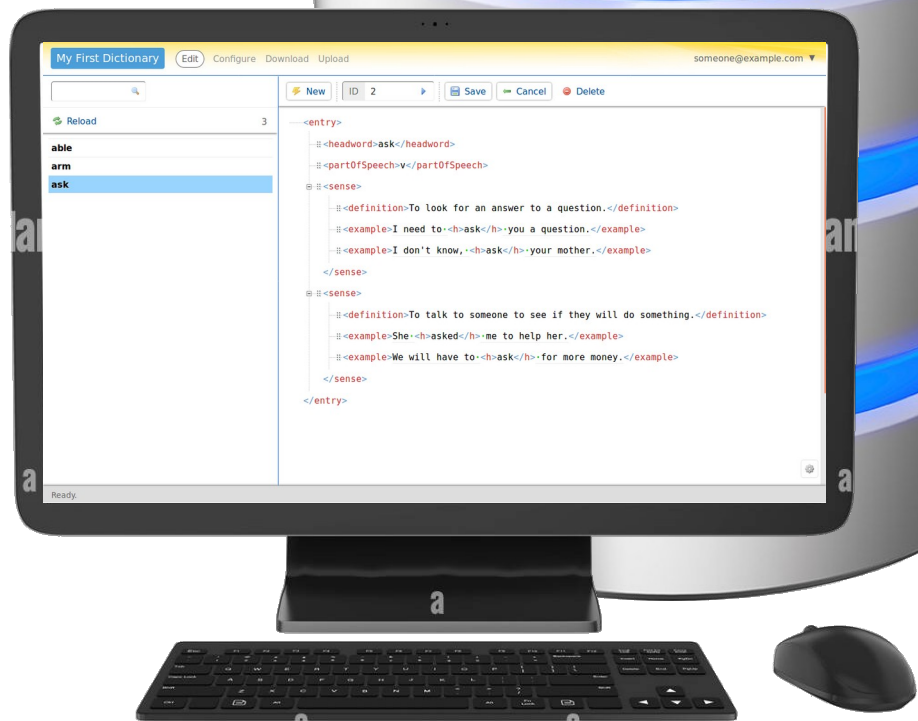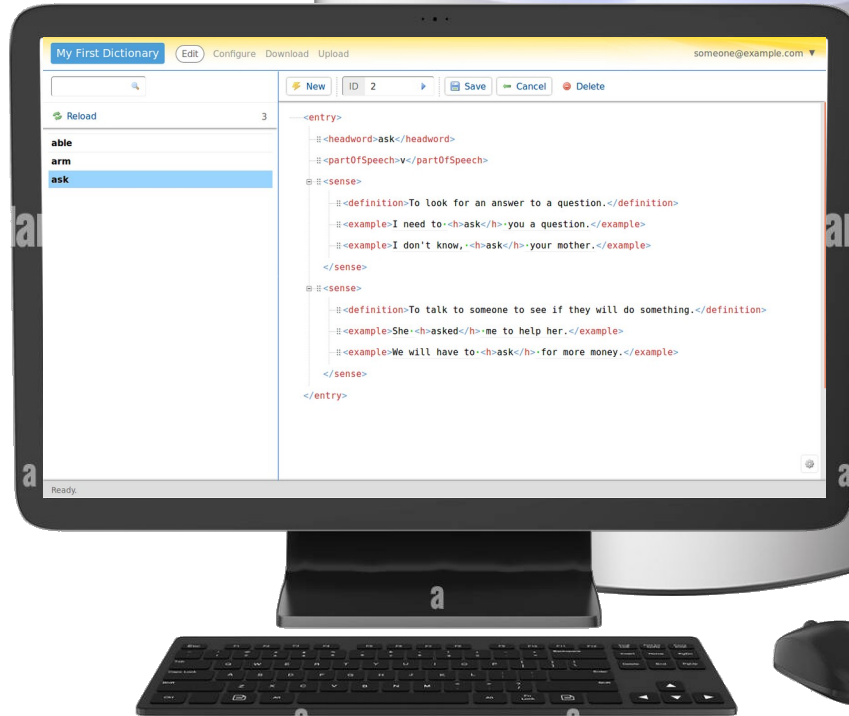
Ready.

My First Dictionary   Edit   Configure   Download   Upload   someone@example.com ▼

⚡ New   ID 2 ▶   💾 Save   ↩ Cancel   ⛔ Delete

🔍

♻ Reload   3

able
arm
ask

```
<entry>
    <headword>ask</headword>
    <partOfSpeech>v</partOfSpeech>
    <sense>
        <definition>To look for an answer to a question.</definition>
        <example>I need to·<h>ask</h>·you a question.</example>
        <example>I don't know,·<h>ask</h>·your mother.</example>
    </sense>
    <sense>
        <definition>To talk to someone to see if they will do something.</definition>
        <example>She·<h>asked</h>·me to help her.</example>
        <example>We will have to·<h>ask</h>·for more money.</example>
    </sense>
</entry>
```
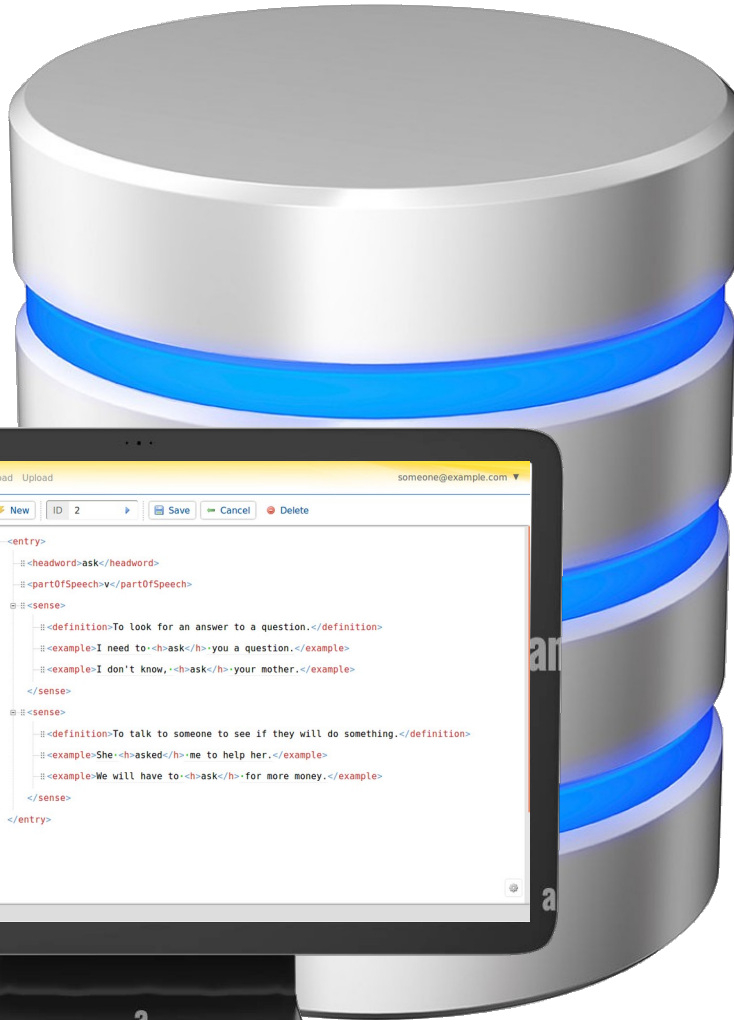
Ready.

back-end

front-end

My First Dictionary | Edit Configure Download Upload | someone@example.com ▼

⚡ New | ID 2 ▶ | 💾 Save | ↩ Cancel | ⊘ Delete

🔍

♻ Reload   3

able
arm
ask

```
<entry>
  <headword>ask</headword>
  <partOfSpeech>v</partOfSpeech>
  <sense>
    <definition>To look for an answer to a question.</definition>
    <example>I need to·<h>ask</h>·you a question.</example>
    <example>I don't know,·<h>ask</h>·your mother.</example>
  </sense>
  <sense>
    <definition>To talk to someone to see if they will do something.</definition>
    <example>She·<h>asked</h>·me to help her.</example>
    <example>We will have to·<h>ask</h>·for more money.</example>
  </sense>
</entry>
```

Ready.

You have two options
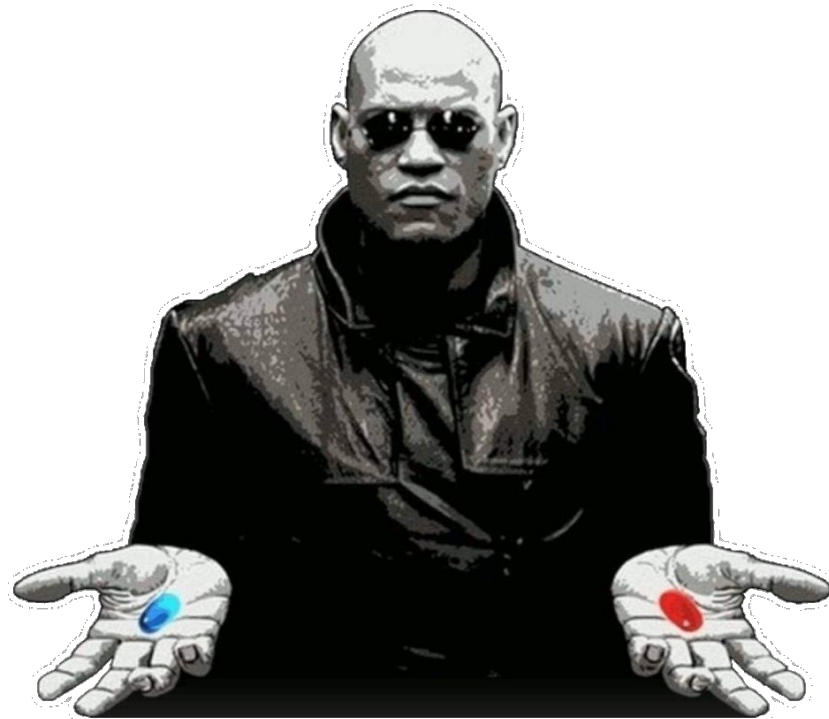
'Document'…   …or 'database'?

The dictionary entry as a 'document'

```xml
<entry>
    <headword>absolutely</headword>
    <pos>adv</pos>
    <sense>
        <gloss>completely</gloss>
        <translation>go hiomlán</translation>
        <translation>go huile agus go hiomlán</translation>
        <exampleContainer>
            <example>I absolutely agree</example>
            <exampleTranslation>aontaím go huile agus go hiomlán</exampleTranslation>
        </exampleContainer>
    </sense>
    <sense>
        <gloss>very</gloss>
        <translation>amach is amach</translation>
        <translation>ar fad</translation>
        <exampleContainer>
```
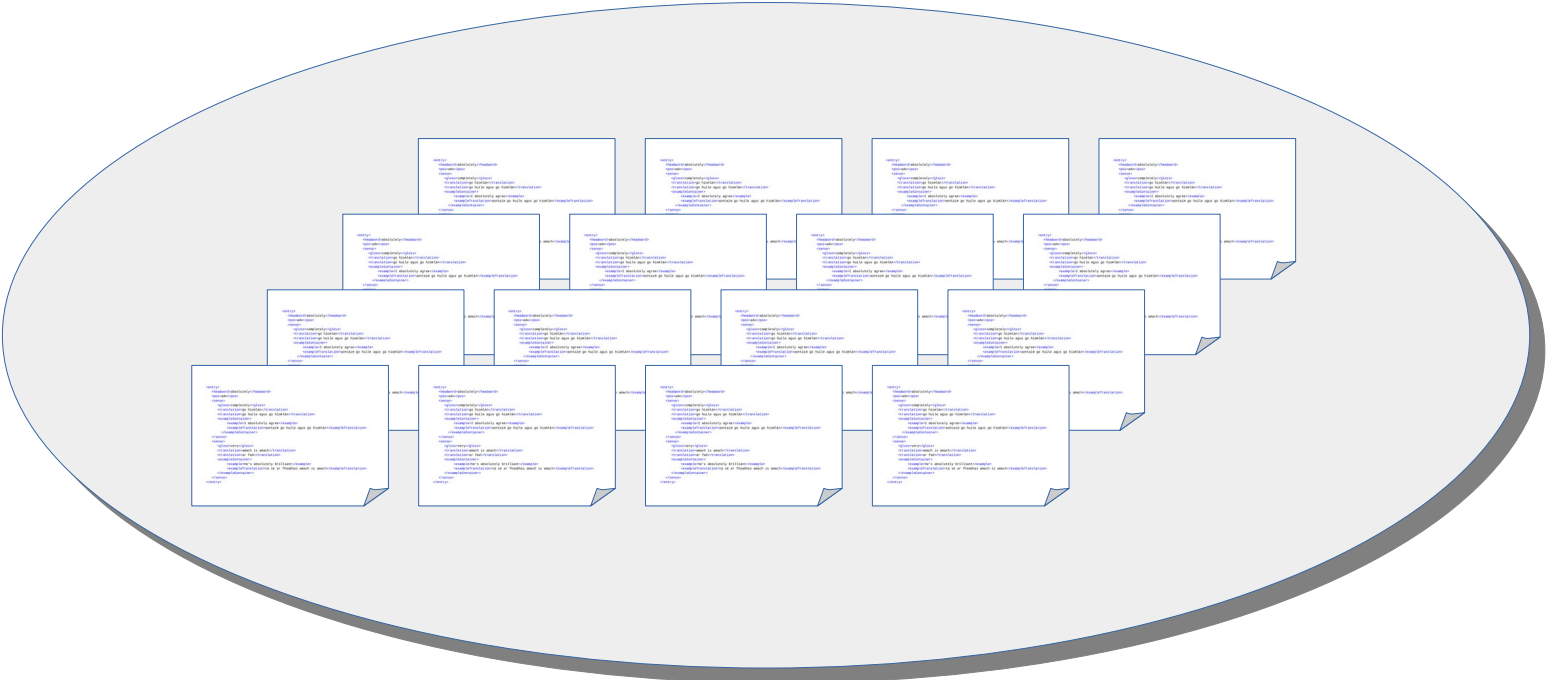
The dictionary as a collection of 'documents'

The dictionary as a 'database'

entries

entryID	headword	pos
1	finger	noun
2	nail	noun

senses

entryID	senseID	definition
1	1	the end of your hand......
2	2	hard end of your finger...
2	3	pointed metal object...

examples

senseID	text
2	Cut your nails!
2	Don't bite your nails.
3	Hit the nail with a hammer.

The dictionary as a 'database'

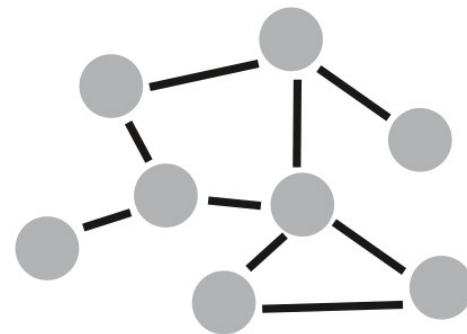|	The 'document' paradigm	The 'database' paradigm
Building an editing UI	easy	difficult
Changing the entry schema	easy	difficult
Cross-references	difficult	easy
Subentries	difficult	easy

❶ ⟶ **❷** ⟶ **❸**

Formatted text

Tree structure
XML, JSON...

Graph structure
SQL, RDF...

'document'

'database'

*"**Digital Natives** are used to receiving information really fast. They like to parallel process and multi-task. They prefer hypertext. **Digital Immigrants** go slowly, step-by-step, one thing at a time."*

— Marc Prensky (2001) 'Digital Natives, Digital Immigrants'

Digital Natives, Digital Immigrants

By Marc Prensky

From *On the Horizon* (MCB University Press, Vol. 9 No. 5, October 2001)
© 2001 Marc Prensky

It is amazing to me how in all the hoopla and debate these days about the decline of education in the US we ignore the most fundamental of its causes. *Our students have changed radically. Today's students are no longer the people our educational system was designed to teach*.

Today's students have not just changed *incrementally* from those of the past, nor simply changed their slang, clothes, body adornments, or styles, as has happened between generations previously. A really big *discontinuity* has taken place. One might even call it a "singularity" – an event which changes things so fundamentally that there is absolutely no going back. This so-called "singularity" is the arrival and rapid dissemination of digital technology in the last decades of the 20th century.

Today's students – K through college – represent the first generations to grow up with this new technology. They have spent their entire lives surrounded by and using computers, videogames, digital music players, video cams, cell phones, and all the other

Thank you.

michmech@lexiconista.com