



**WIKIMEDIA**

# Wikis Are Cool !

<http://ultimategerardm.blogspot.com/>  
GerardM

People who wiki are part of a community  
People who are part of a community  
make a difference



# Conclusion

- Our aim is to provide free, NPOV information in all languages
- We are successful where it is easy and are working on it where it is harder
- When asked: “In 10 years what would Wikipedia be like” Jimbo said: “What will the world be like in 10 years” 😊



# Introduction

- Wikis enable people to cooperate on the information that is out there
- Wikis are changing our world
- Mediawiki is one implementation of a wiki application
- Wiki is Hawaiian for fast
- Wiki is what this talk is about



# The way of the Wiki

- What you know as Wikipedia is only one out of many projects
- Wikipedia is an encyclopedia
- We want Wikipedia's in every language
- All Wikipedia need to adhere to the NPOV (our magic word)



# Wikipedia is an encyclopaedia

- Wikipedia is a free and NPOV encyclopaedia
- Its content doubles in size every four to six months both in bandwidth and in storage
- When we get more capacity we find it is quickly lapped up by more users



# Wikipedia is an encyclopaedia

- We want to have a Wikipedia in all languages
- We are considered a resource that has up to date information
- We do have problems.. So does the “competition”
- Eh, we do not compete



# Wikipedia is an encyclopaedia

- We are weak in many subjects
- We suffer from POV pushers
- We suffer from trolls and vandals
  
- We are not “academic”.. We are anarchistic.. We are a meritocracy.. Hell, it seems to work



# Wiktionary is a dictionary

- Originally it was to rid Wikipedia of lexicological (English) content
- Wiktionaries in many languages have been started
- They are “one to many” dictionaries as well as everything else





# Wiktionary is a dictionary

- The current wiktionaries suffer from a lack of cooperation among the projects
- There has been no development specifically for Wiktionary until recently
- The current Mediawiki software does not cater for the needs of a dictionary project



# Ultimate Wiktionary

- As 80% of lexicological data is metadata, we should integrate all wiktionaries
- Adding a translation creates a “placeholder” for that other language
- With an XML import and export facility, we open our content up



# Ultimate Wiktionary

- Placeholders will make it the first resource for Papiamentu on the Internet
- We actively look for glossaries, thesauri to include in UW
- Like the telephone, a single one is useless, two make a novelty, many make a network where every additional node adds more value than its own



# Relevancy

- UW does not (yet) compare to any of the traditional dictionaries
- With the ability to add to GEMET, people learn the vocabulary to do business with the EC
- There are many small resources, by combining them they gain relevance too



# Wikidata

- Incorporating relational data in Mediawiki.
- It allows us to add “structured data” to an unstructured methodology
- It will enable Ultimate Wiktionary and other projects
- It is an experiment but many of us believe in its potential



# Wikidata

- Wikidata is the first major addition to Mediawiki that is not specifically build for Wikipedia
- Wikidata is developed for money, this is a first within our community
- It is paid by a grant, that was given to the Wiktionary community



# Commons

- A repository of digital material
- It has over 100.000 files since October
- Contains some 6000+ sound files for pronunciations
- All material is under a free license
- All material has to be useful to a Wikimedia project



# Commons

- It was a major innovation as we now are able to share images 😊
- It is in need of metadata for search and localisation
- It will double in size every 6 months
- The software needs to scale





# Organisation

- Foundation - Jimbo, Anthere, Angela
  - National chapters
  - The projects
  - Language versions
- 
- We have people and communities



# Communication

- Talk pages for every article
- The village pump
- Mailing lists
- IRC (several specialized channels)
- Skype (telephony over the internet)



# Problems

- The NPOV is one inhibitor of problems
- Loads of discussion possible
- As we grow, trolls become a bigger problem
- The communities do not agree on anything to start with, getting consensus is hard work



# Money

- We are dependent on donations
- We need your brains AND your money
- We are getting more corporate sponsors, but also we get more partners



# Kennisnet

- Servers are being installed as I speak
- Throws bread on the water
- Share in the kids of the Netherlands a **big** share of our audience
- Ultimate Wiktionary, Commons localization, Queen Beatrix project
- ..the first Wikimedia Foundation partner



# Speculation time

- LISA & its community
- OmegaT & its community
- Open Office & its community
- Lucene & its community
- The importance of the Wikimedia Foundation



# Speculation time - LISA

- TMX – Translation Memory eXchange
- Wikimedia has the ultimate localization projects .. All knowledge in all languages 😊
- Future number of their publication
- Big organizations like HP, IBM, Sun, World Bank are its members ..



# Speculation time - OmegaT

- It is a CAT – TMX level 1
- It is written in Java, integrates with OO
- External glossaries
- .. Integration in UW
- We have a contact in Sabine, but we have not really contacted them yet ..





# Speculation time – OO org

- Integrates with OmegaT
- Has a spellchecker in all its languages
- What if we collaborate on dictionaries..
- What if we collaborate on localizations ..



# Speculation time - Lucene

- Is a tool that the UW already uses
- Has a language savvy community
- Is a great crowd to find likeminded people in.. For our content and for what you can do with that stuff



# Speculation time

- I am here, have many business cards
- What if some of them prove exciting
- What if we can get the content that is out there ?
- What if we can be here next year and tell you about our walking dream and our dreams..