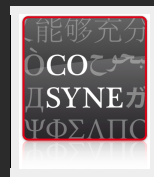


BRIDGING THE GAP OF MULTILINGUAL DIVERSITY



WIKIMEDIA CONFERENTIE NEDERLAND 2012

Amit Bronner

**Nederlands Instituut voor Beeld en
Geluid**

Media Park, Hilversum

COSYNE

content synchronization system for multilingual wikis

identifying overlap between language versions of the same article

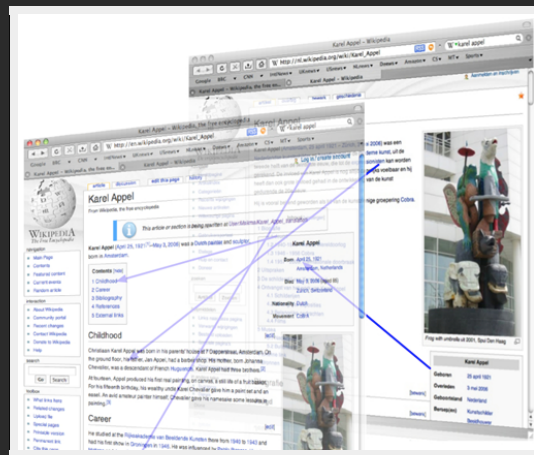
integrating non-overlapping content

going into detail: source, translation, context, comparisons

correct, edit, design and share

keep track and react to changes

machine translation & natural language processing models learn from user corrections



CONTENT OVERLAP

content synchronization system for multilingual wikis

identifying overlap between language versions of the same article

integrating non-overlapping content

going into detail: source, translation, context, comparisons

correct, edit, design and share

keep track and react to changes

machine translation & natural language processing models learn from user corrections



SYNCHRONIZED VIEW

content synchronization system for multilingual wikis

identifying overlap between language versions of the same article

integrating non-overlapping content

going into detail: source, translation, context, comparisons

correct, edit, design and share

keep track and react to changes

machine translation & natural language processing models learn from user corrections

The screenshot shows a web interface for a multilingual wiki. At the top, there's a search bar and navigation links for 'Page', 'Discussion', and 'View history'. The main heading is 'WCN2012'. Below it, there's a paragraph in Dutch: 'Amsterdam is the largest city and the capital of the Netherlands.' followed by several redaction symbols. Below that, there's a paragraph in English: 'Amsterdam has a population of 790,654 within city limits, an urban population of 1,209,419 and a metropolitan population of 2,289,762.' followed by redaction symbols. At the bottom, there's another paragraph in English: 'A new national regulation could greatly affect its famed tolerance for drugs.' followed by redaction symbols. On the left side, there's a sidebar with 'In other languages' and links for 'Nederlands', 'Deutsch', and 'Italiano'. Below that, there's a 'Navigation' section with a 'Main page' link.

The screenshot shows a web interface for a multilingual wiki. At the top, there's a search bar and navigation links for 'Pagina', 'Overleg', and 'enis bekijken'. The main heading is 'WCN2012'. Below it, there's a paragraph in English: 'Amsterdam is de hoofdstad en grootste gemeente van Nederland.' followed by redaction symbols. Below that, there's a paragraph in Dutch: 'De stad, in het Amsterdams Mokum genoemd, ligt in de provincie Noord-Holland, aan de monding van de Amstel en aan het IJ.' followed by redaction symbols. Below that, there's a paragraph in English: 'Op 1 januari 2012 telde de gemeente Amsterdam 790.654 inwoners.' followed by redaction symbols. Below that, there's a paragraph in Dutch: 'Amsterdam heeft een bevolking van 790.654 grenzen, een stad in de stedelijke bevolking van 1.209.419 metropolitane en een bevolking van 2.289.762.' followed by redaction symbols. Below that, there's a paragraph in English: 'Een nieuwe verordening kunnen nationale sterk van invloed zijn op de famed tolerantie voor drugs.' followed by redaction symbols. On the left side, there's a sidebar with 'In andere talen' and links for 'English', 'Deutsch', and 'Italiano'. Below that, there's a 'Navigatie' section with links for 'Hoofdpagina' and 'Gebruikersportaal'.

EXPLORE

content synchronization system for multilingual wikis

identifying overlap between language versions of the same article

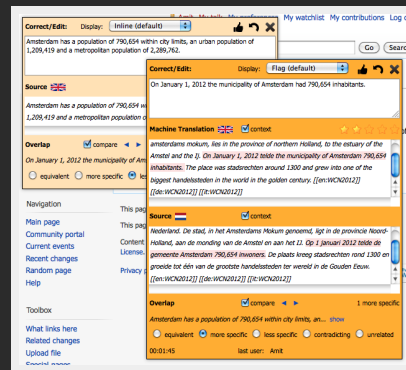
integrating non-overlapping content

going into detail: source, translation, context, comparisons

correct, edit, design and share

keep track and react to changes

machine translation & natural language processing models learn from user corrections



CREATE

content synchronization system for multilingual wikis

identifying overlap between language versions of the same article

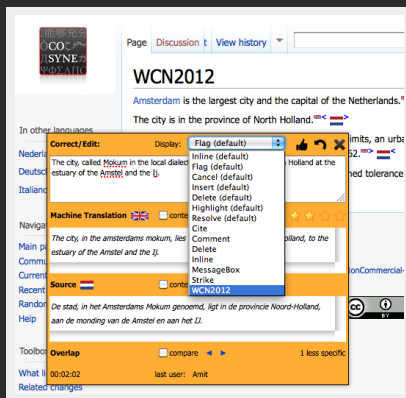
integrating non-overlapping content

going into detail: source, translation, context, comparisons

correct, edit, design and share

keep track and react to changes

machine translation & natural language processing models learn from user corrections



WATCH USER EDITS

content synchronization system for multilingual wikis

identifying overlap between language versions of the same article

integrating non-overlapping content

going into detail: source, translation, context, comparisons

correct, edit, design and share

keep track and react to changes

machine translation & natural language processing models learn from user corrections

The screenshot shows a wiki page for 'WCN2012' with a comparison of two versions. The left version is from November 8, 2012, and the right version is from November 17, 2012. The page is in Dutch. The comparison highlights changes in the text, such as the addition of a plus sign and the word 'grootste' in the second sentence of the first paragraph. The page includes a sidebar with navigation options and a search bar.

The screenshot shows a wiki page for 'WCN2012' with a comparison of two versions. The left version is from November 8, 2012, and the right version is from November 17, 2012. The page is in English. The comparison highlights changes in the text, such as the addition of a plus sign and the word 'grootste' in the second sentence of the first paragraph. The page includes a sidebar with navigation options and a search bar. A 'Machine Translation' section is visible, showing the text in Dutch and its English translation.

LEARNING TECHNOLOGY

content synchronization system for multilingual wikis

identifying overlap between language versions of the same article

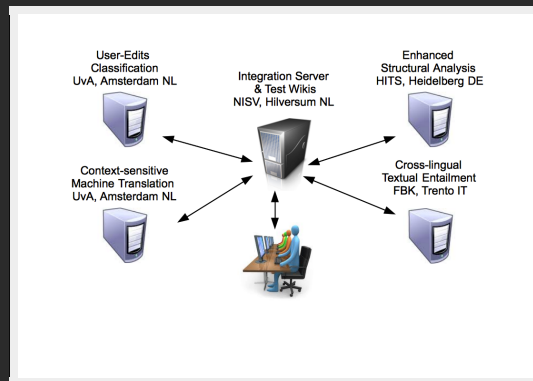
integrating non-overlapping content

going into detail: source, translation, context, comparisons

correct, edit, design and share

keep track and react to changes

translation & language processing models learn
from user corrections

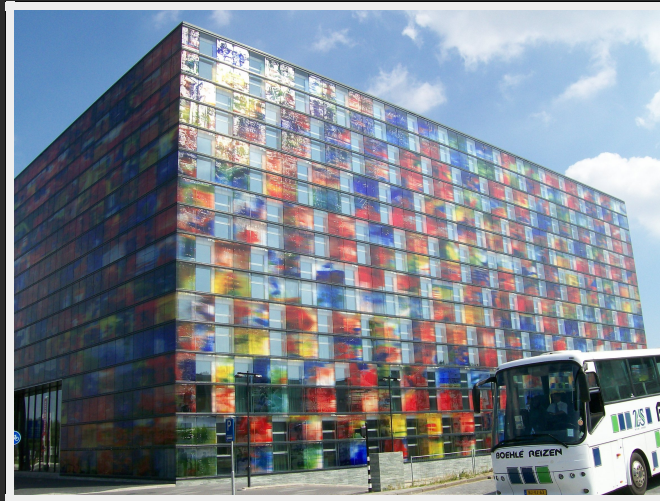


BEELD EN GELUID

Dutch audiovisual archive (one of the largest in Europe)

R&D department

Beeld en Geluid wiki



Bron: http://upload.wikimedia.org/wikipedia/commons/e/e8/Beeld_en_Geluid_zijkant.jpg

THE END

