CoSyne

Multilingual
Content Synchronization
with Wikis

A. Bronner
University of Amsterdam
Netherlands Institute for Sound & Vision

Wikimania 2011, Haifa Israel

Introduction

- CoSyne: Multilingual Content Synchronization with Wikis

- Research & Development Project, Co-Funded by the European Union under the 7th Framework Programme (FP7- ICT)

- March 2010 – February 2013

Consortium

- **4 Research Institutions**

University of Amsterdam (UvA)

Fondazione Bruno Kessler (FBK)

Dublin City University (DCU)

Heidelberg Institute for Theoretical Studies (HITS)

- **3 End User Partners**

Deutsche Welle

Netherlands Institute for Sound & Vision

Wikimedia Netherlands

Objective

The CoSyne project will develop a system for

synchronizing wiki pages in multiple languages

and for

maintaining consistency while pages are edited

in the different languages.

A. Bronner
University of Amsterdam / Netherlands Institute for Sound & Vision
Wikimania 2011, Haifa Israel

Motivation

- **Why** do we want to synchronize wiki pages?

    - Keep content consistent across languages

    - Learn about content in other languages

    - Expand to new languages
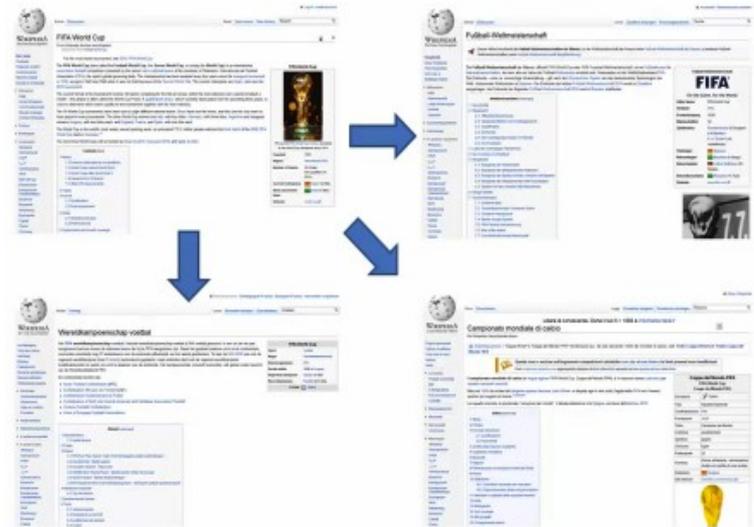
Soccer Example

- 28 May 2011, 19:45

- Wembley stadium, London

- Barcelona vs. Manchester United

- Barcelona wins (3:1)

- Goalkeeper Edwin van der Sar, at the age of 40, plays his last game for Manchester United and retires from football

Soccer Example

- 28 May 2011, 20:35 (already at the half-time ???)

- English user edit: http://**en**.wikipedia.org/wiki/Edwin_van_der_Sar

- "Van der Sar officially retired from football following a 3-1 defeat ..."

Revision as of 18:56, 28 May 2011 (edit)
Ntadams (talk | contribs)
← Previous edit

Revision as of 20:35, 28 May 2011 (edit) (undo)
ShadowTerror10 (talk | contribs)
Next edit →

and [[Fabio Pereira da Silva|Fabio]]. He was included in the [[PFA Team of the Year]], making his third appearance there, others in 2007 and 2008.

and [[Fabio Pereira da Silva|Fabio]]. He was included in the [[PFA Team of the Year]], making his third appearance there, others in 2007 and 2008.

On 22 May 2011, **He** played his final game at Old Trafford in a 4–2 victory over [[Blackpool F.C.|Blackpool]] resulting in Blackpool's relegation from the Premier League. He captained United as a farewell treat. Previous to the last game of the Premier League season he had announced that after the [[2011 UEFA Champions League Final]] against [[FC Barcelona|Barcelona]] he **will** retire.

On 22 May 2011, **he** played his final game at Old Trafford in a 4–2 victory over [[Blackpool F.C.|Blackpool]] resulting in Blackpool's relegation from the Premier League. He captained United as a farewell treat. Previous to the last game of the Premier League season he had announced that after the [[2011 UEFA Champions League Final]] against [[FC Barcelona|Barcelona]] he **would** retire.

Van der Sar officially retired from football following a 3-1 defeat to Barcelona in the 2011 UEFA Champions League Final.

==International career==

==International career==

Soccer Example

- 28 May 2011, 22:47

- Italian user edit: http://**it**.wikipedia.org/wiki/Edwin_van_der_Sar

- "He ~~will play~~ played his last game … "

Soccer Example

- 28 May 2011, 23:10

- German user edit: http://**de**.wikipedia.org/wiki/Edwin_van_der_Sar

- "On May 28 2011 he was defeated in his last professional game … "

[gesichtete Version]
Version vom 28. Mai 2011, 22:54 Uhr (Bearbeiten)
Nicowa (Diskussion | Beiträge)
← Zum vorherigen Versionsunterschied

[ungesichtete Version]
Version vom 28. Mai 2011, 23:10 Uhr (Bearbeiten) (entfernen)
Jan.ruesch (Diskussion | Beiträge)
(→Manchester United)
Zum nächsten Versionsunterschied →

Berbatov strike as United pull away from the pack (4. März 2009)]</ref> Er kündigte an, seine Karriere am Ende der [[Premier League 2010/11|Saison 2010/11]] bei Manchester United zu beenden.<ref>guardian.co.uk: [http://www.guardian.co.uk/football/2011/jan/31/edwin-van-der-sar-united Edwin van der Sar motivated to end Manchester United career on a high (31. Jan. 2011)]</ref>

Berbatov strike as United pull away from the pack (4. März 2009)]</ref> Er kündigte an, seine Karriere am Ende der [[Premier League 2010/11|Saison 2010/11]] bei Manchester United zu beenden.<ref>guardian.co.uk: [http://www.guardian.co.uk/football/2011/jan/31/edwin-van-der-sar-united Edwin van der Sar motivated to end Manchester United career on a high (31. Jan. 2011)]</ref>

+ Am 28.05.2011 unterlag er in seinem letzten profispiel im [[Champions League]]Finale gegen den[[FC Barcelona]] mit 3:1.

== Nationalmannschaft ==

== Nationalmannschaft ==

Soccer Example

- 29 May 2011, 13:33

- Dutch user edit: http://**nl**.wikipedia.org/wiki/Edwin_van_der_Sar

- "Van der Sar played his last game as goalkeeper …"

Versie van 29 mei 2011 om 13:26 (bewerken)
Letterman (Overleg | bijdragen)

Versie van 29 mei 2011 om 13:33 (bewerken) (ongedaan maken)
Letterman (Overleg | bijdragen)

+ Van der Sar speelde zijn laatste wedstrijd als profkeeper op 28 mei 2011 in [[Londen]], in het [[Wembley]]stadion, waar Manchester United die zaterdagavond de finale om de [[UEFA Champions League]] tegen [[FC Barcelona]] verloor met 1-3. De toen ruim 40-jarige Van der Sar werd daarmee tevens de oudste voetballer uit de geschiedenis die de eindstrijd haalde om het clubkampioenschap van [[Europa]].

+
+

== Loopbaan ==

== Loopbaan ==

Keep Content Consistent (made-up example)
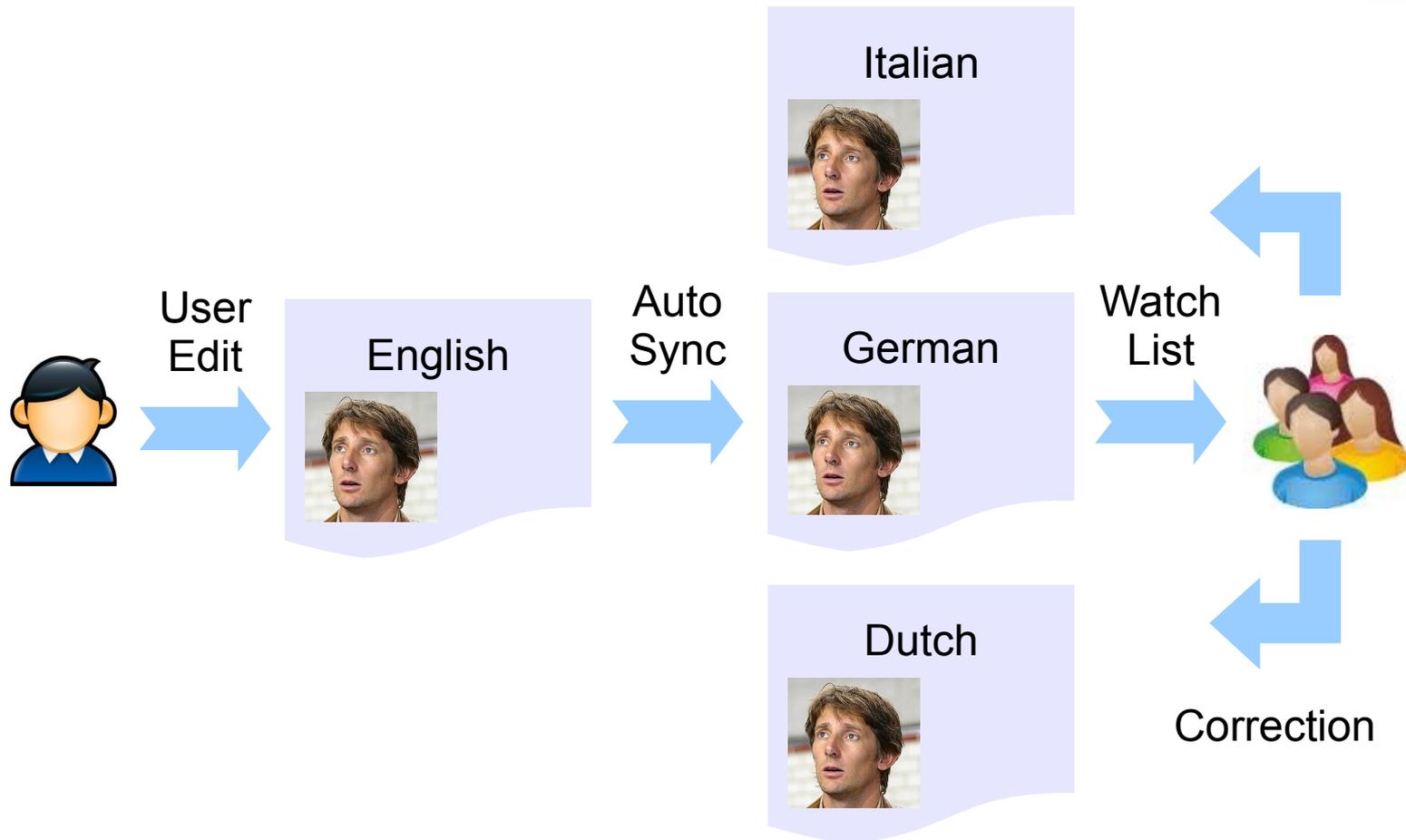
Content from Another Language (made-up example)

→ **Synchronize**

Overlapping Details (Same Info)

Amsterdam

Amsterdam is the largest city and the capital of the Netherlands.

De naam van de stad is afgeleid van de ligging bij een in de 13e eeuw aangelegde dam in de Amstel. De plaats kreeg stadsrechten rond het jaar 1300 en groeide later uit tot één van de grootste handelssteden ter wereld in de Gouden Eet Bevolkingsgroei leidde vanaf het eind van de 16e eeuw tot grote stadsuitbreidingen, waaronder de grachtengordel, die op de Werelderfgoedlijst is opgenomen en tot de belangrijkste bezienswaardigheden van de stad behoort. Andere attracties in de h onder meer de musea, zoals het Rijksmuseum, het Stedelijk Museum en het Van Gogh Museum; het Anne Frank Huis; het pr

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes

Article | Discussion   Read | Edit | View history | Se

Amsterdam
From Wikipedia, the free encyclopedia

This article is about the Dutch capital. For other uses, see Amsterdam (disambiguation).

Amsterdam (/ˈæmstərdæm/; Dutch [ɑmstərˈdɑm] (listen)) is the largest city and the capital of the Netherlands. The current position of Amsterdam as capital city of the Kingdom of the Netherlands is governed by the constitution of 24 August 1815 and its successors.[6] Amsterdam has a population of 780,152 within city limits, an urban population of 1,209,419 and a metropolitan population of 2,158,592.[7] The city is in the province of North Holland in the west of the country. It comprises the northern part of the Randstad, one of the larger conurbations in Europe, with a population of approximately 7 million.[8]

Its name is derived from *Amstelredamme*,[9] indicative of the city's origin: a dam in the river Amstel. Settled as a small fishing village in the late 12th century, Amsterdam became one of the most important ports in the world during the Dutch Golden Age, a result of its innovative developments in trade. During that time, the city was the leading center for finance and diamonds.[10] In the 19th and 20th centuries, the city expanded, and many new neighbourhoods and suburbs were formed. The 17th-century canals of Amsterdam (in Dutch: 'Grachtengordel'), located in the heart of Amsterdam, were added to the UNESCO World Heritage List in July 2010.

Amsterdam
— Municipality/city —

More Specific in Dutch

Amsterdam

Amsterdam is the largest city and the capital of the Netherlands.

De naam van de stad is afgeleid van de ligging bij een in de 13e eeuw aangelegde dam in de Amstel. De plaats

The city is in the province of North Holland on the rivers Amstel and IJ.

Article | Discussion

Amsterdam

From Wikipedia, the free encyclopedia

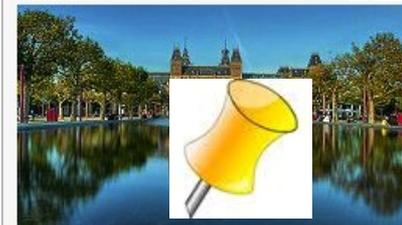*This article is about the Dutch capital. For other uses, see Amsterdam (disambiguation).*

Amsterdam (/ˈæmstərdæm/; Dutch [ˌɑmstərˈdɑm] (listen)) is the largest city and the capital of the Netherlands. The current position of Amsterdam as capital city of the Kingdom of the Netherlands is governed by the constitution of 24 August 1815 and its successors.[6] Amsterdam has a population of 780,152 within city limits, an urban population of 1,209,419 and a metropolitan population of 2,158,592.[7] The city is in the province of North Holland in the west of the country. It comprises the northern part of the Randstad, one of the larger conurbations in Europe, with a population of approximately 7 million.[8]

Its name is derived from *Amstelredamme*,[9] indicative of the city's origin: a dam in the river Amstel. Settled as a small fishing village in the late 12th century, Amsterdam became one of the most important ports in the world during the Dutch Golden Age, a result of its innovative developments in trade. During that time, the city was the leading center for finance and diamonds.[10] In the 19th and 20th centuries, the city expanded, and many new neighbourhoods and suburbs were formed. The 17th-century canals of Amsterdam (in Dutch: 'Grachtengordel'), located in the heart of Amsterdam, were added to the UNESCO World Heritage List in July 2010.

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes

Read | Edit | View history

52°22′23″N 4°53′32″E

Amsterdam
— Municipality/city —

Appears only in Dutch

Amsterdam

Amsterdam is the largest city and the capital of the Netherlands.

De naam van de stad is afgeleid van de ligging bij een in de 13e eeuw aangelegde dam in de Amstel. De plaats

The city is in the province of North Holland on the rivers Amstel and IJ.

Article | Discussion   Read | Edit | View history | Se

Amsterdam

The city is also called 'Mokum' (מקום) in the local dialect.

WIKIPEDIA
The Free Encyclopedia

From W

This article is about the Dutch capital. For other uses, see Amsterdam (disambiguation).

Amsterdam (/ˈæmstərdæm/; Dutch [ɑmstərˈdɑm] (◄listen)) is the largest city and the capital of the Netherlands. The current position of Amsterdam as capital city of the Kingdom of the Netherlands is governed by the constitution of 24 August 1815 and its successors.[6] Amsterdam has a population of 780,152 within city limits, an urban population of 1,209,419 and a metropolitan population of 2,158,592.[7] The city is in the province of North Holland in the west of the country. It comprises the northern part of the Randstad, one of the larger conurbations in Europe, with a population of approximately 7 million.[8]
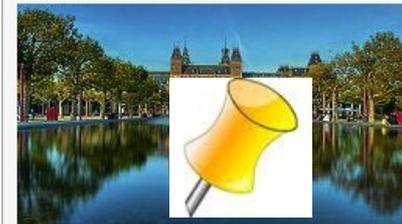
Its name is derived from *Amstelredamme*,[9] indicative of the city's origin: a dam in the river Amstel. Settled as a small fishing village in the late 12th century, Amsterdam became one of the most important ports in the world during the Dutch Golden Age, a result of its innovative developments in trade. During that time, the city was the leading center for finance and diamonds.[10] In the 19th and 20th centuries, the city expanded, and many new neighbourhoods and suburbs were formed. The 17th-century canals of Amsterdam (in Dutch: 'Grachtengordel'), located in the heart of Amsterdam, were added to the UNESCO World Heritage List in July 2010.

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia

Interaction
Help
About Wikipedia
Community portal
Recent changes

Amsterdam
— Municipality/city —

Appears only in English

Amsterdam

Zie Amsterdam (doorverwijspagina) voor andere betekenissen van Amsterdam.

Amsterdam is the largest city and the capital of the Netherlands.

De naam van de stad is afgeleid van de ligging bij een in de 13e eeuw aangelegde dam in de Amstel. De plaats

The city is in the province of North Holland on the rivers Amstel and IJ.

Article | Discussion | Read | Edit | View history | Se

Am

From W

The city is also called 'Mokum' (מקום) in the local dialect.

52°22′23″N 4°53′32″E

WIKIPEDIA
The Free Encyclopedia

This article is about the Dutch capital. For other uses, see Amsterdam (disambiguation).

Main p
Conten
Feature
Curren

Amsterdam's coffee shops will not be closing their doors to tourists any time soon.

Random article
Donate to Wikipedia

The city is in the province of North Holland in the west of the country. It comprises the northern part of the Randstad, one of the larger conurbations in Europe, with a population of approximately 7 million.[8]

Interaction
Help
About Wikipedia
Community portal
Recent changes

Its name is derived from *Amstelredamme*,[9] indicative of the city's origin: a dam in the river Amstel. Settled as a small fishing village in the late 12th century, Amsterdam became one of the most important ports in the world during the Dutch Golden Age, a result of its innovative developments in trade. During that time, the city was the leading center for finance and diamonds.[10] In the 19th and 20th centuries, the city expanded, and many new neighbourhoods and suburbs were formed. The 17th-century canals of Amsterdam (in Dutch: 'Grachtengordel'), located in the heart of Amsterdam, were added to the UNESCO World Heritage List in July 2010.

Contradiction

Amsterdam

Amsterdam is the largest city and the capital of the Netherlands.

The city is in the province of North Holland on the rivers Amstel and IJ.

The city is also called 'Mokum' (מקום) in the local dialect.

Amsterdam's coffee shops will not be closing their doors to tourists any time soon.

The city is in the province of North Holland in the west of the country. It comprises the northern part of the Randstad, one of the larger conurbations in Europe, with a population of approximately 7 million.[8]

Amsterdam has a population of 780,152 within city limits
On April 1st 2011 the municipality counted 783.363 residents.

Take Action? (Not Necessarily...)

and the capital of the Netherlands.

Holland on the rivers Amstel and IJ.

kum' (מקום) in the local dialect.

osing their doors to tourists any time soon.

on of 780,152 within city limits
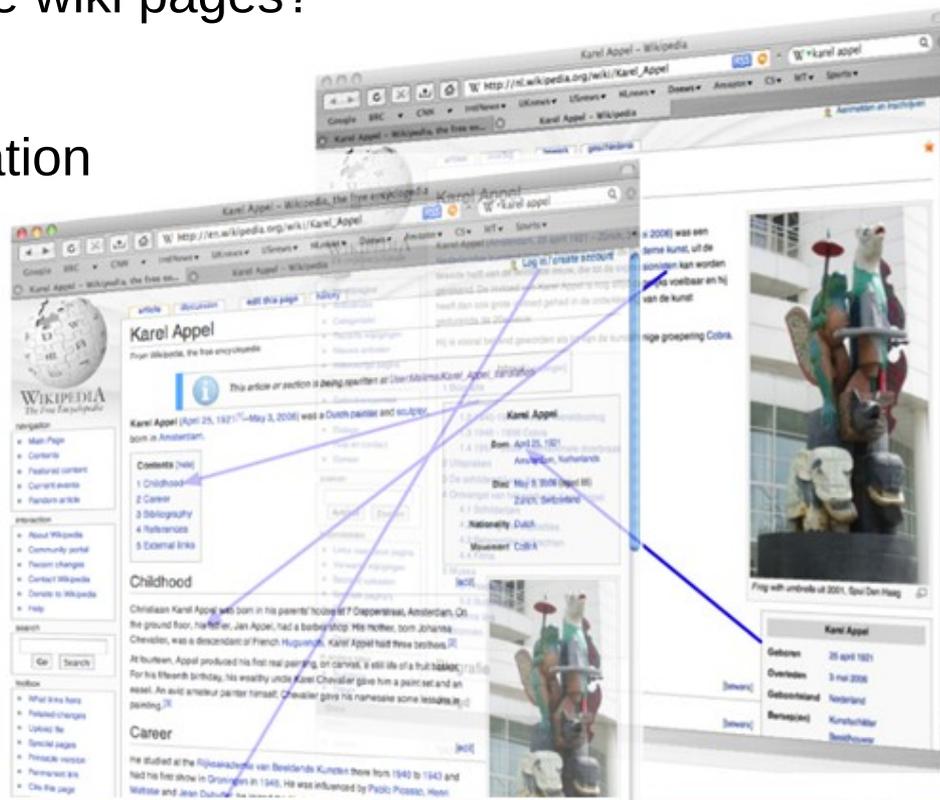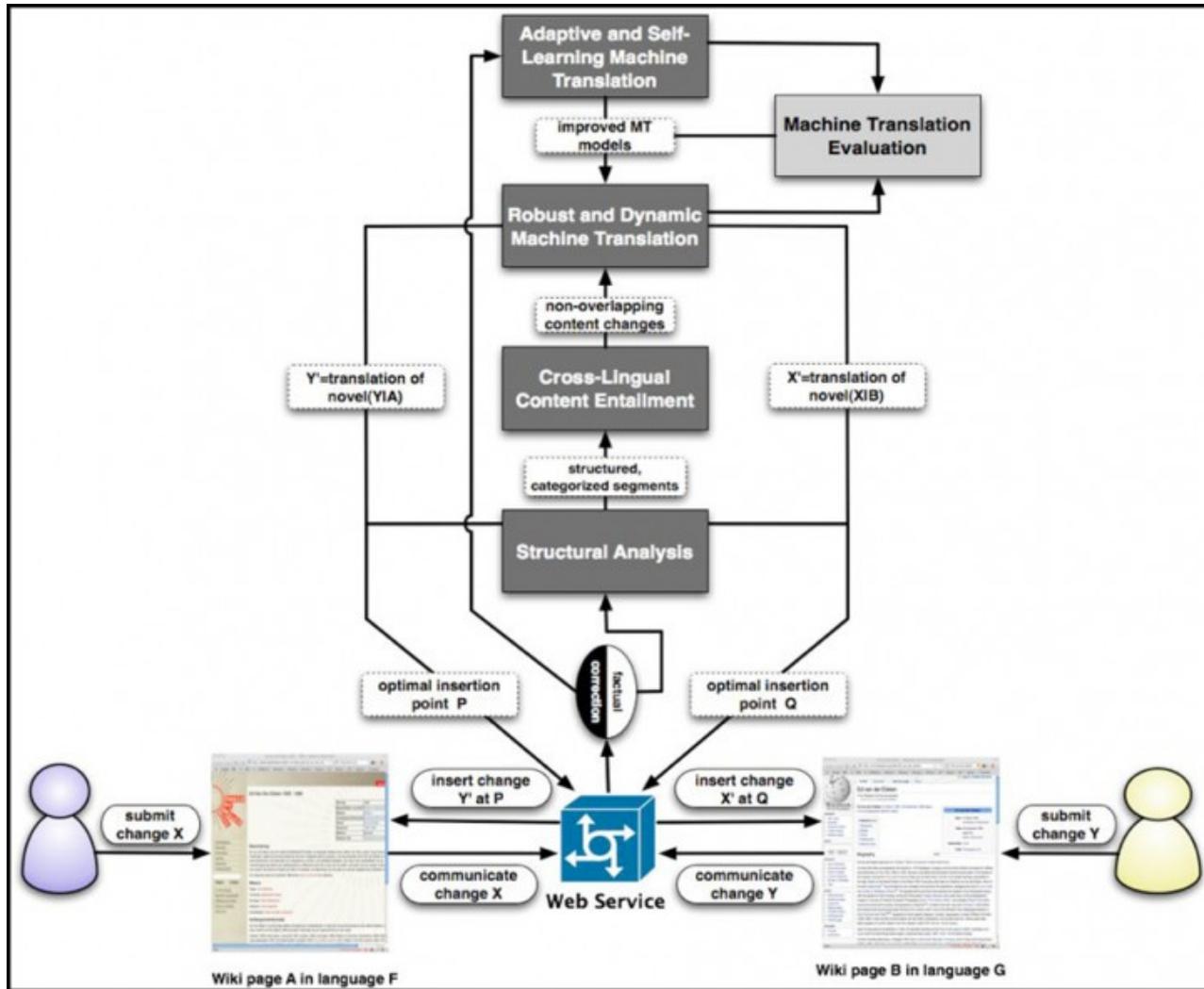ality counted 783.363 residents.

Article | Discussion

Editing Amsterdam

From Wikipedia, the free encyclopedia

aracters ▸ Help ▸ Cite

→ Append

→ Add

→ Replace

le|icon'|'|æ|m|s|t|ər|d|æ|m}}; Dutch {{IPA
]]. The current position of Amsterdam as c
erlands|**constitution**]] of 24 August 1815
ngrijk der Nederlanden (1815) '''(Dutch)''
/wiki/Grondwet_voor_het_Koningrijk_der_Nederlanden_(1815) | accessdate =2
city limits, an urban population of 1,209,419 and a metropolitan populatio
/touristinformation/aboutamsterdam/factsandfigures |title=Fact
The city is in the [[Provinces of the Netherlands|**province**]]
part of the [[**Randstad**]], one of the larger [[**conurbation**]]s
on.<ref name="R2040">{{cite web||http://www.rijksoverheid.nl/o
randstad-2040-facts-figures-wat-komt-er-op-de-randstad-af.html
(Netherlands)|VROM]]|title=Randstad2040; Facts & Figures (p.26

Its name is derived from ''Amstelredamme'',<ref name="Britannica Eleven">[
</ref> indicative of the city's origin: a dam in the river [[**Amstel**]]. Set
ortant ports in the world during the [
was the leading center for finance an
=exc Cambridge.org], Capitals of Capit
/ref> In the 19th and 20th centuries,
ods and suburbs were formed.
The 17th-century [[**canals of Amsterdam**]] (in Dutch: 'Grachtengordel') lo

Challenges

- **How** do we synchronize wiki pages?
- Research domains:
  - User Edits Classification
  - Structural Analysis
  - Content Entailment
  - Machine Translation

Source text(s) → User edits → Structural analysis → Content entailment → Machine translation → Target text

A. Bronner
University of Amsterdam / Netherlands Institute for Sound & Vision
Wikimania 2011, Haifa Israel

Synchronization Process

A. Bronner
University of Amsterdam / Netherlands Institute for Sound & Vision
Wikimania 2011, Haifa Israel

User Edits Classification

◆ Decides whether the meaning has changed:
  ◆ Factual change → Synchronize
  ◆ Fluency change → Improve (self learning)

A. Bronner
University of Amsterdam / Netherlands Institute for Sound & Vision
Wikimania 2011, Haifa Israel
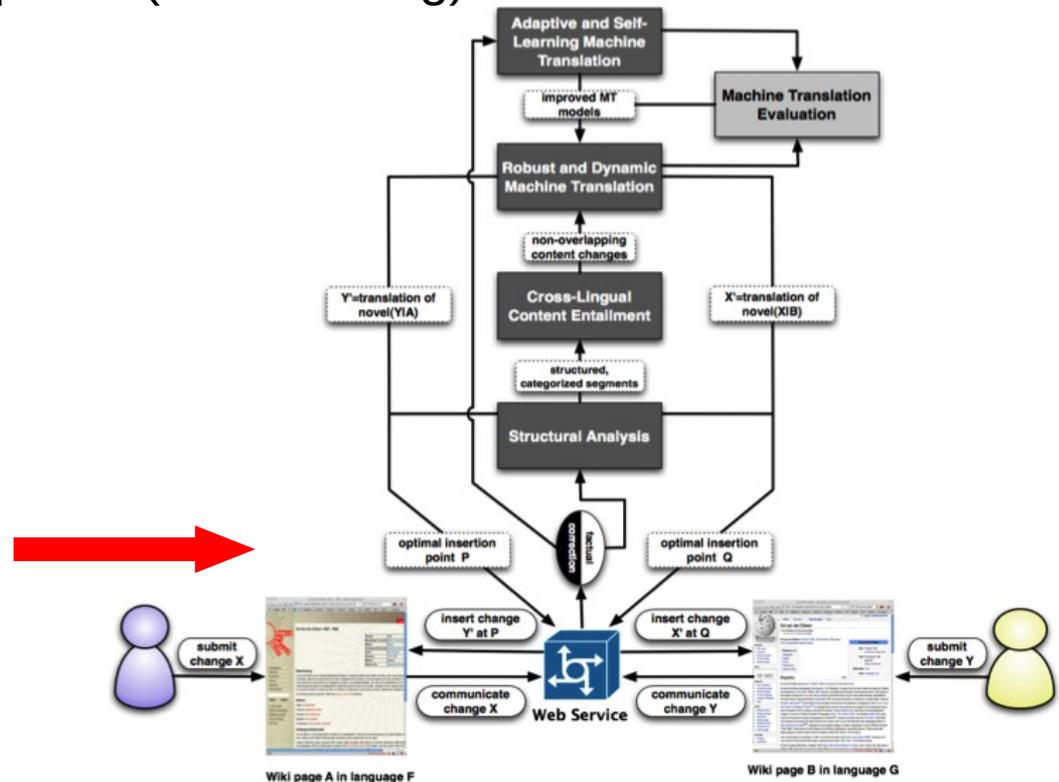
Structural Analysis

- Decides which segments to compare
- Instructs how to reconstruct the wikitext
- Decides where to insert the modifications

A. Bronner
University of Amsterdam / Netherlands Institute for Sound & Vision
Wikimania 2011, Haifa Israel
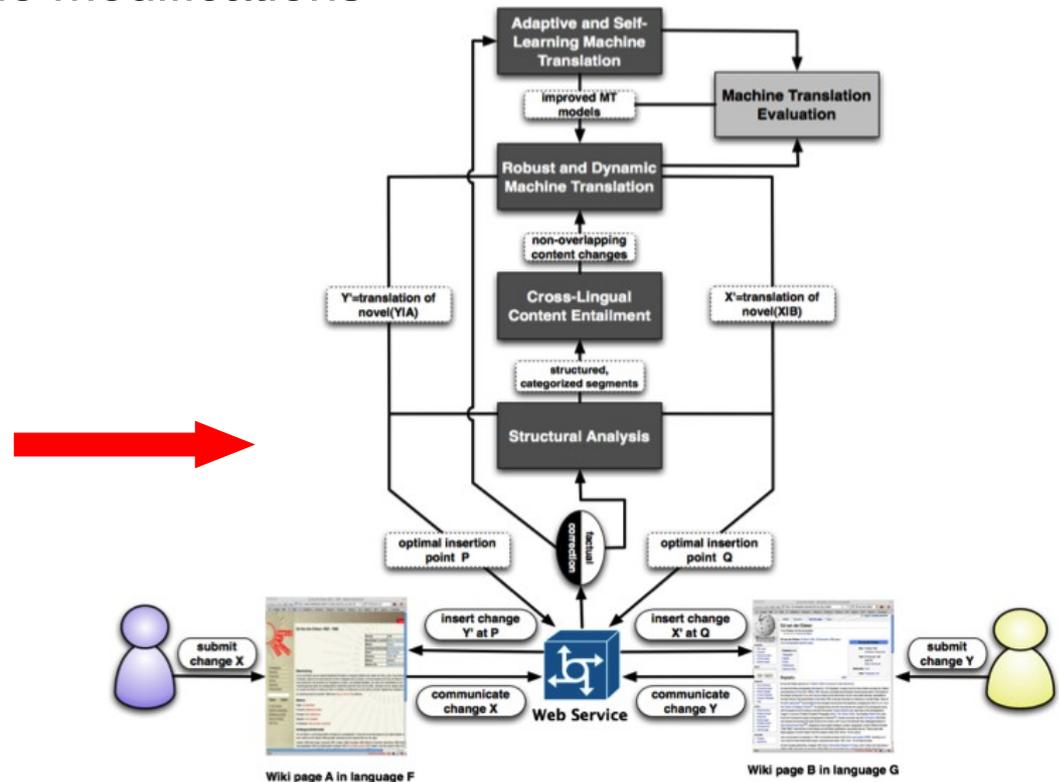
Cross-lingual Content Entailment

- Determines whether text segments are equivalent
  - Equivalent → No change
  - More specific → Replace
  - Unrelated → Add
  - Contradiction → Ask the user
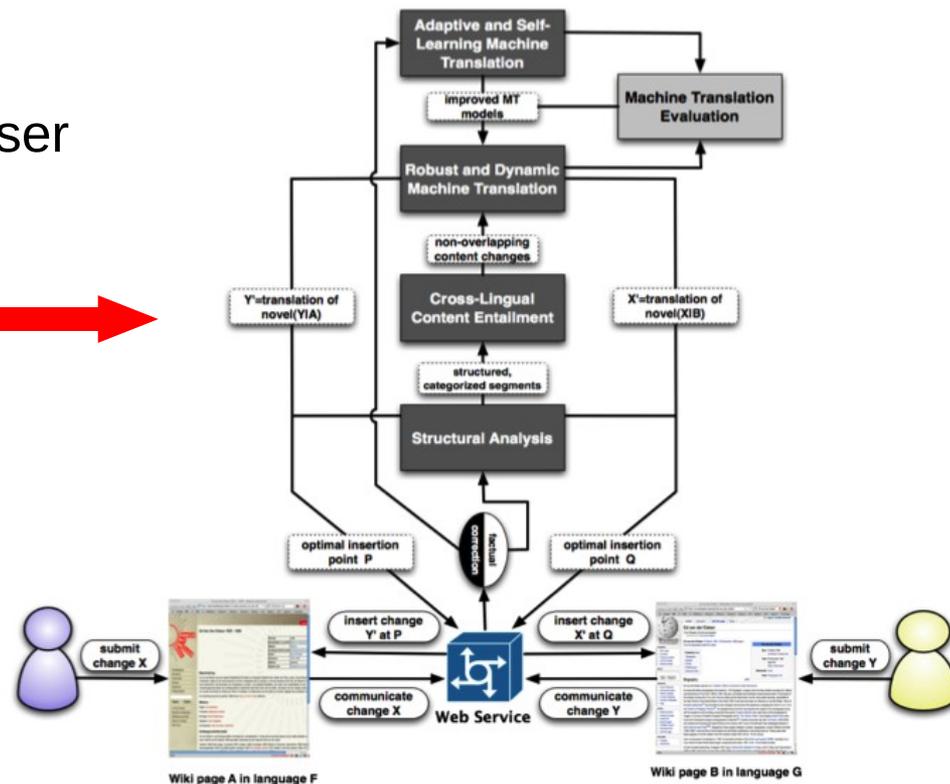
Machine Translation

- Translates segments that need to be added or replaced
- Collects "fluency" corrections and learns improved models

A. Bronner
University of Amsterdam / Netherlands Institute for Sound & Vision
Wikimania 2011, Haifa Israel
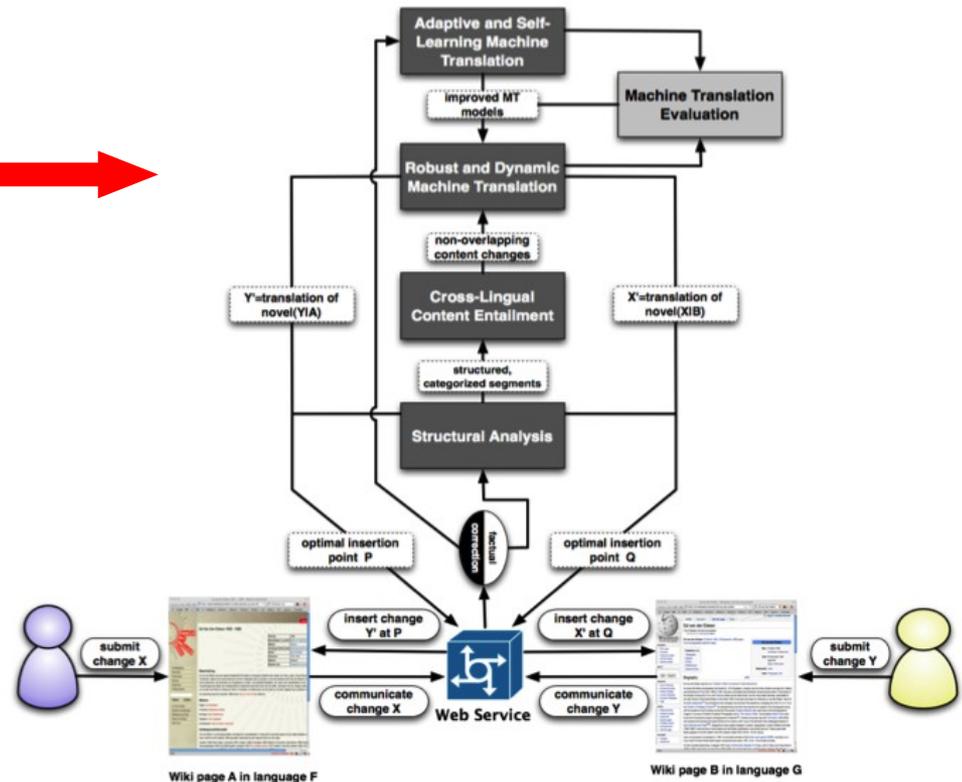
Status of Development

- 3 prototypes planned
    - 1st prototype released
    - Deployed on end user's (private) wiki

- Target Languages:
    - English, German, Italian, Dutch (core)
    - Bulgarian, Turkish

- Machine Translation Quality
    - Evaluation work
    - Ongoing improvement
    - Open source & open data

Use Cases 1 & 2

◆ The wiki of the Netherlands Institute for Sound & Vision

◆ Deutsche Welle's Kalenderblatt (Today in History)
  ◆ Demo

1ˢᵗ Prototype Demo:
Kalenderblatt on MediaWiki

1.1.1618: Bartolomé Esteban Murillo (†3.4.1682)

Bartolomé Esteban Murillo war ein spanischer Maler des Barock.

Bartolomé Esteban Murillo wurde in Sevilla geboren.

CoSyne
- Step by Step
- Sync > English
- Sync Report

Suche

[Seite] [Suchen]

Navigation
- Hauptseite
- Gemeinschafts-Portal
- Aktuelle Ereignisse
- Letzte Änderungen
- Zufällige Seite
- Hilfe

Werkzeuge
- Links auf diese Seite
- Änderungen an verlinkten Seiten
- Datei hochladen
- Spezialseiten
- Druckversion

The user (DW) edits German content for Kalenderblatt
There is no English page
New entry: Bartolomé Esteban Murillo, a Spanish painter
Born 1.1.1618, Died 3.4.1682

Welcome to CoSyne

Content Synchronization with Wikis

This is a prototype version. It is still under development.

You will be redirected to the modified page in a few seconds.

Executing analysis .. in_process

The user triggers CoSyne

Seite | Diskussion | Bearbeiten | Versionen/Autoren | Löschen | Verschieben | Schützen | nicht mehr beobachten

1.1.1618: Bartolomé Esteban Murillo (†3.4.1682)

Bartolomé Esteban Murillo ⧉ war ein spanischer Maler des Barock.

Bartolomé Esteban Murillo wurde in Sevilla geboren.

In anderen Sprachen
- English

CoSyne
- Step by Step
- Sync > English
- Sync Report

Suche

[Seite] [Suchen]

Navigation
- Hauptseite
- Gemeinschafts-Portal
- Aktuelle Ereignisse
- Letzte Änderungen
- Zufällige Seite
- Hilfe

Werkzeuge
- Links auf diese Seite
- Änderungen an verlinkten Seiten

CoSyne creates a new English page and adds
language links between the German & English versions

CoSyne provides German translation to English and reconstruction of Wikitext formats (hyperlinks, headings, bold/italic, images, templates, ...)

CoSyne adds a warning message about machine generated content and asks users to review & correct the content

1.1.1618: Bartolomé Esteban Murillo (†3.4.1682)

Bartolomé Esteban Murillo 🔗 war ein spanischer Maler des Barock.

Bartolomé Esteban Murillo wurde in Sevilla als jüngster Sohn in einer Familie von 14 Kindern geboren.

Der gläubige Maler schuf vor allem Porträts und Gemälde mit stark religiösen Motiven.

In anderen Sprachen
- English

CoSyne
- Step by Step
- Sync > English
- Sync Report

Suche

Seite   Suchen

Navigation
- Hauptseite
- Gemeinschafts-Portal
- Aktuelle Ereignisse
- Letzte Änderungen
- Zufällige Seite
- Hilfe

Werkzeuge
- Links auf diese Seite
- Änderungen an verlinkten Seiten

The German version is now linked to an English version
The user provides additional information to the German version
adding a new sentence & changing an existing sentence

Welcome to CoSyne

Content Synchronization with Wikis

This is a prototype version. It is still under development.

You will be redirected to the modified page in a few seconds.

Executing analysis .. in_process

The user triggers CoSyne

CoSyne creates a new version of the English page

**CoSyne does not change the 1ˢᵗ sentence,
it replaces the 2ⁿᵈ by a translation of the modified German
& it adds a translation of the new German sentence**

Use Case 3

- Wikimedia Netherlands
  - Options:
    - MediaWiki manuals
    - Copyright lobby texts (?)
    - Specific domains (technical, soccer, …)
  - Languages:
    - EN, DE, IT, NL
  - CoSyne users mailing list

Want to Get Involved?

Visit our website:

cosyne.eu

A. Bronner
University of Amsterdam / Netherlands Institute for Sound & Vision
Wikimania 2011, Haifa Israel