Language Technologies and Business in the Future

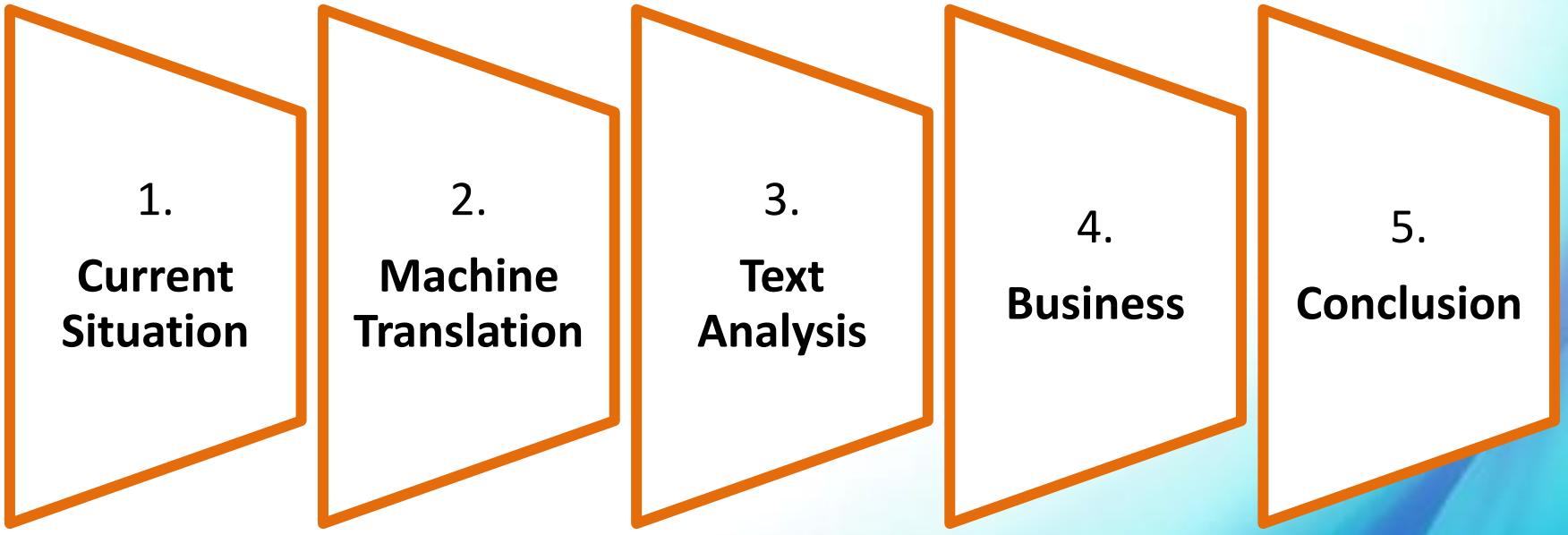
Niko Papula, Multilizer

KITES Symposium, Helsinki 31.10.2013

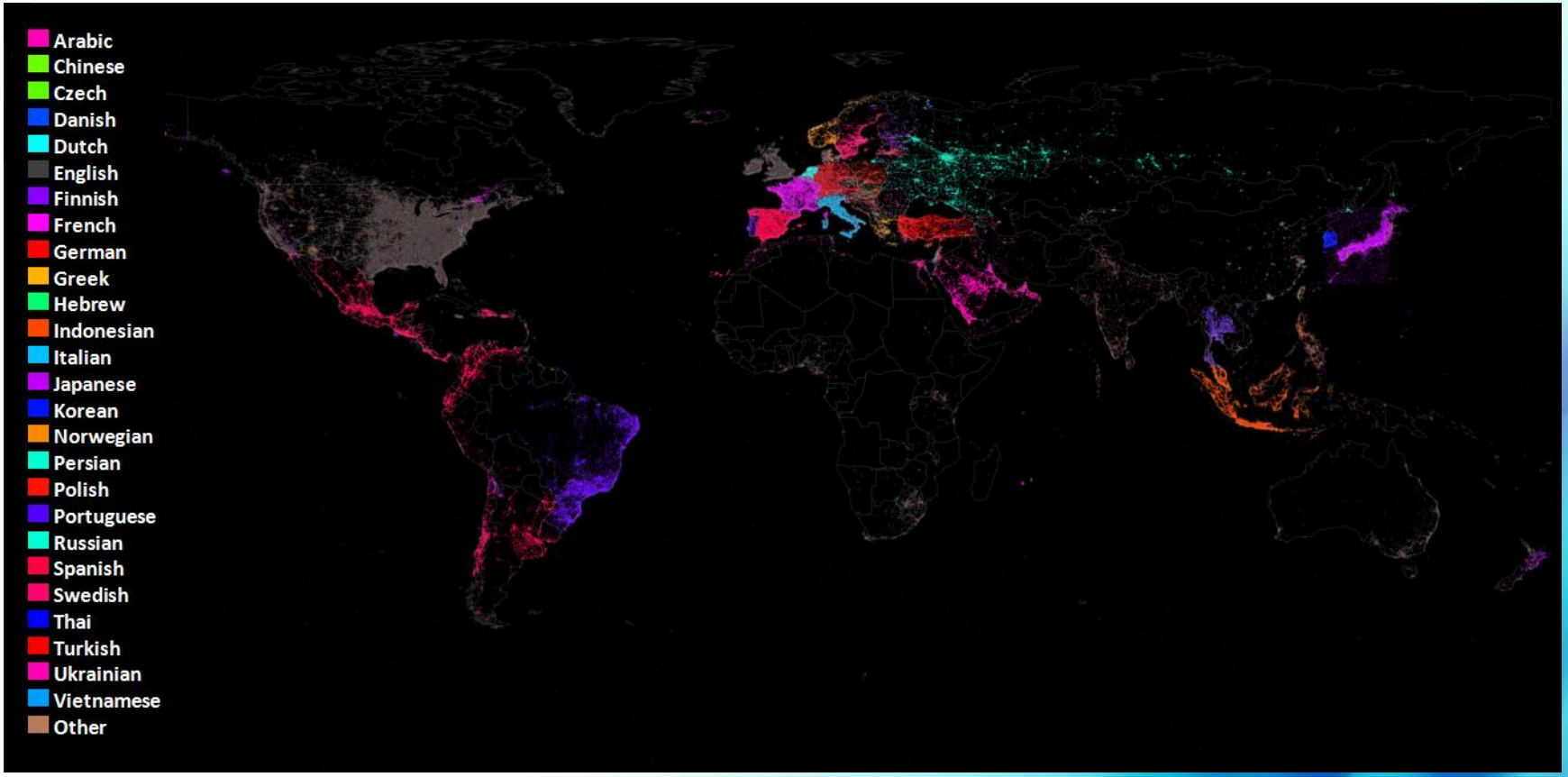
MULTILIZER

- Finnish company providing leading translation & localization technologies since 1995
- Each month more than 120 million words are translated with Multilizer products in more than 150 countries
- Multilizer products:
  - Professional localization tools (software localization in particular)
  - Automatic machine translation qualification (MT-Qualifier)
  - Machine translation for PDF documents

Agenda: Language Technologies and Business in the Future



Tweets by Languages

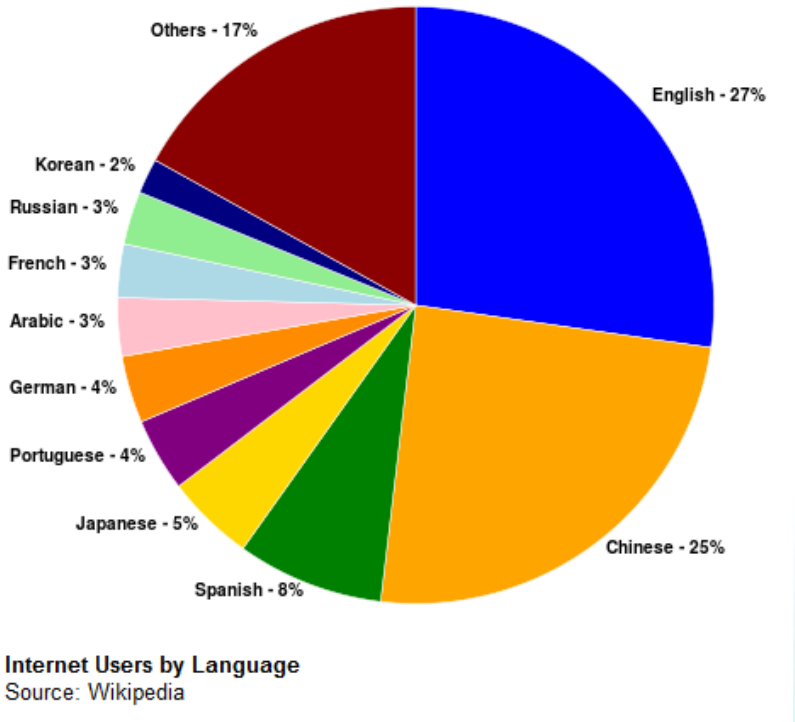


Tweets by languages

Source: http://businesstech.co.za/news/internet/37585/twitter-which-language-dominates-the-twitterverse/

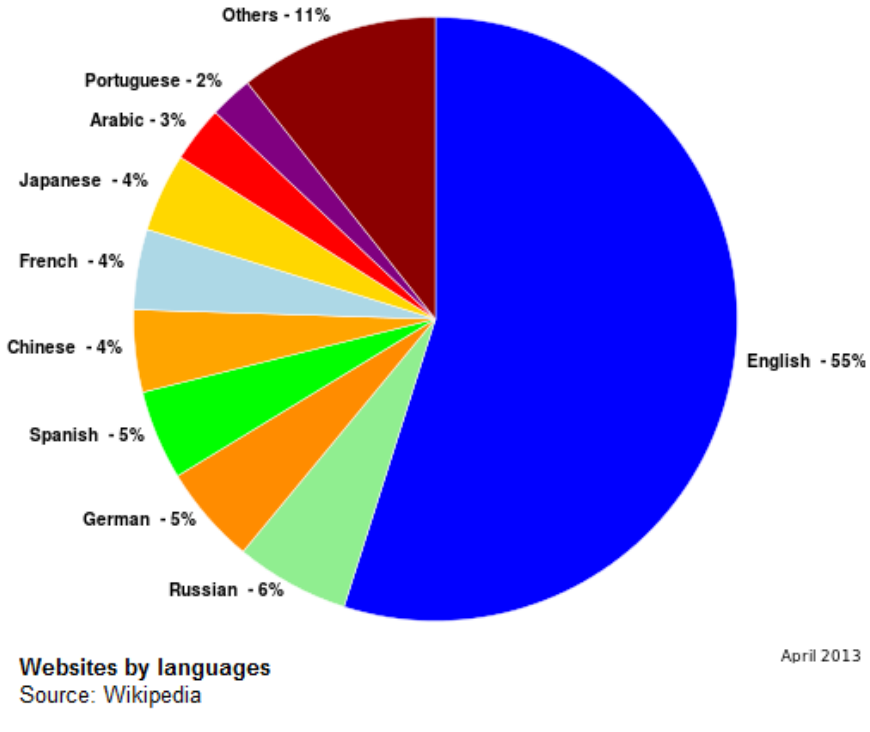
Internet by Languages

Users:



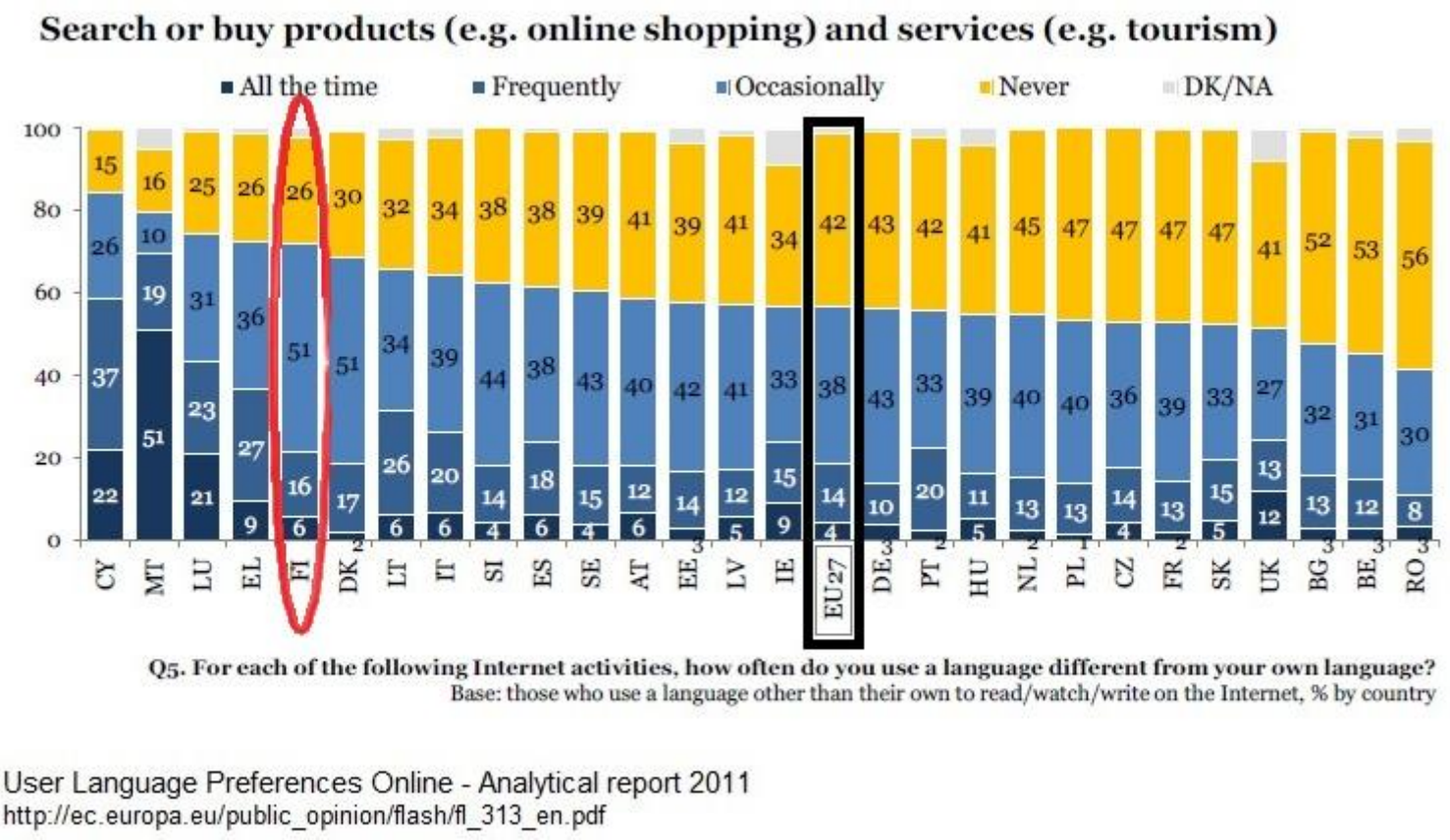
Internet Users by Language
Source: Wikipedia

Websites:



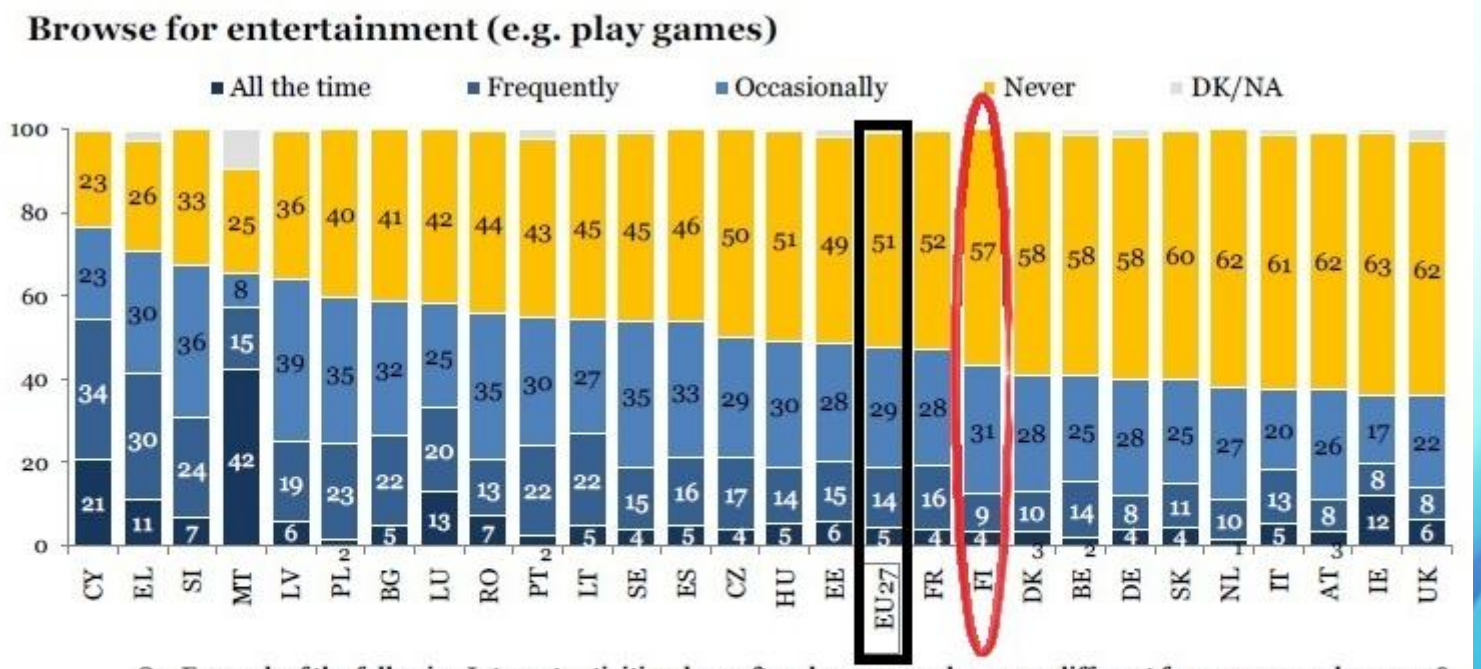
Websites by languages
Source: Wikipedia

Only 18 % of EU27 shop online regularly with other languages



User Language Preferences Online - Analytical report 2011
http://ec.europa.eu/public_opinion/flash/fl_313_en.pdf

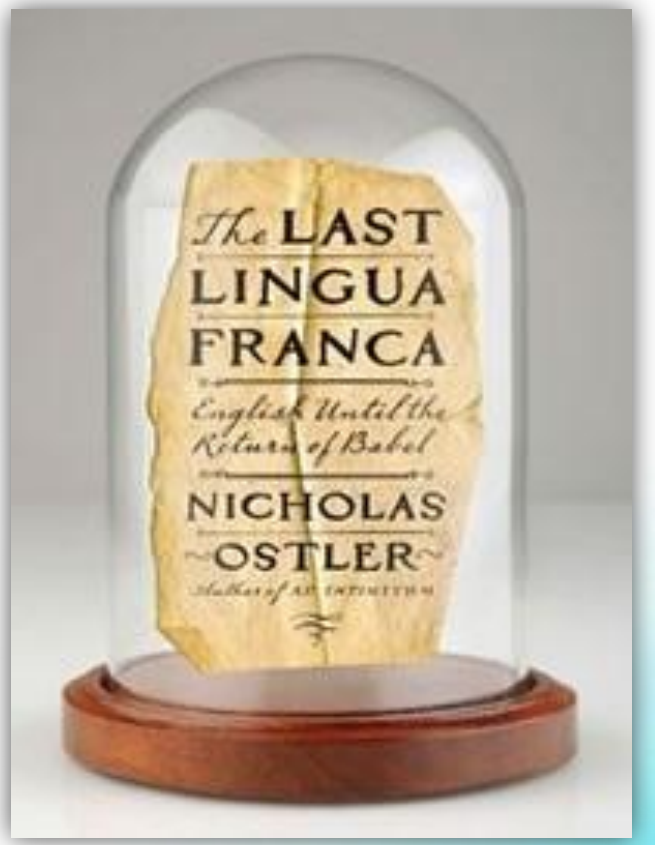
Only 19 % of EU27 browse for entertainment regularly in other languages



Q5. For each of the following Internet activities, how often do you use a language different from your own language?
Base: those who use a language other than their own to read/watch/write on the Internet, % by country

User Language Preferences Online - Analytical report 2011
http://ec.europa.eu/public_opinion/flash/fl_313_en.pdf

Multilingual online world needs MT solutions



Nicholas Ostler: The Last Lingua Franca, English Until the Return of Babel

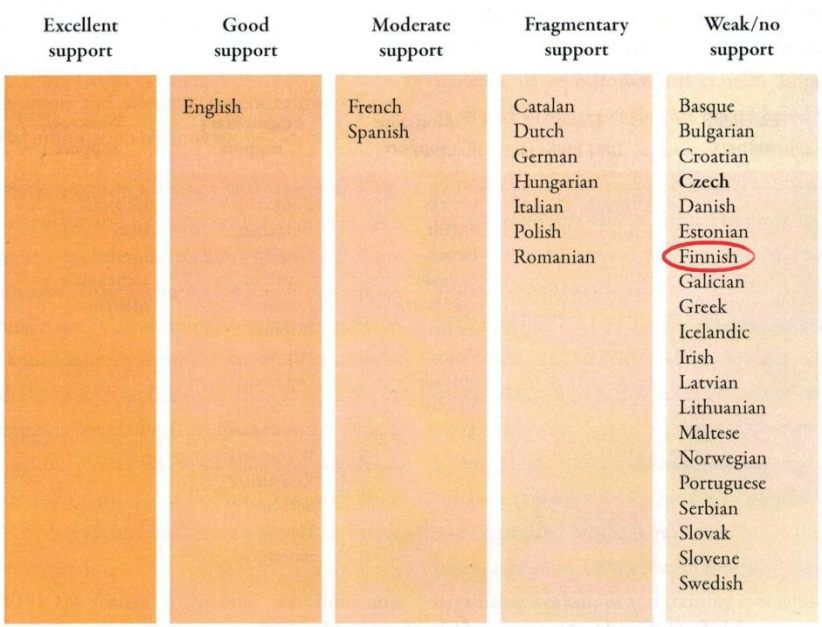
Language Technology in Finland?

Finland is often a leading country what comes to adopting new technologies

→**but** language technologies (LT) for Finnish language are an unfortunate exception.

State of Language Technology Support for 30 European Languages

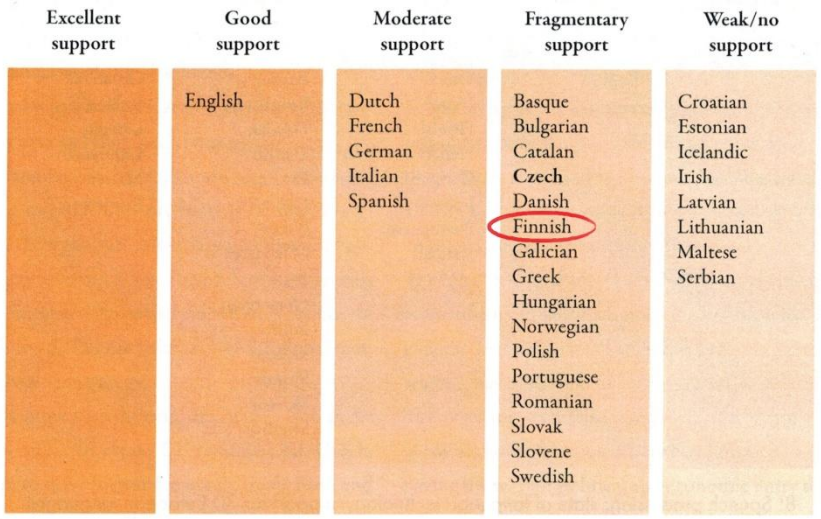
Machine Translation



9: Machine translation: state of language technology support for 30 European languages

Source: META-NET White Paper Series: The Czech Language in the Digital Age, 2012

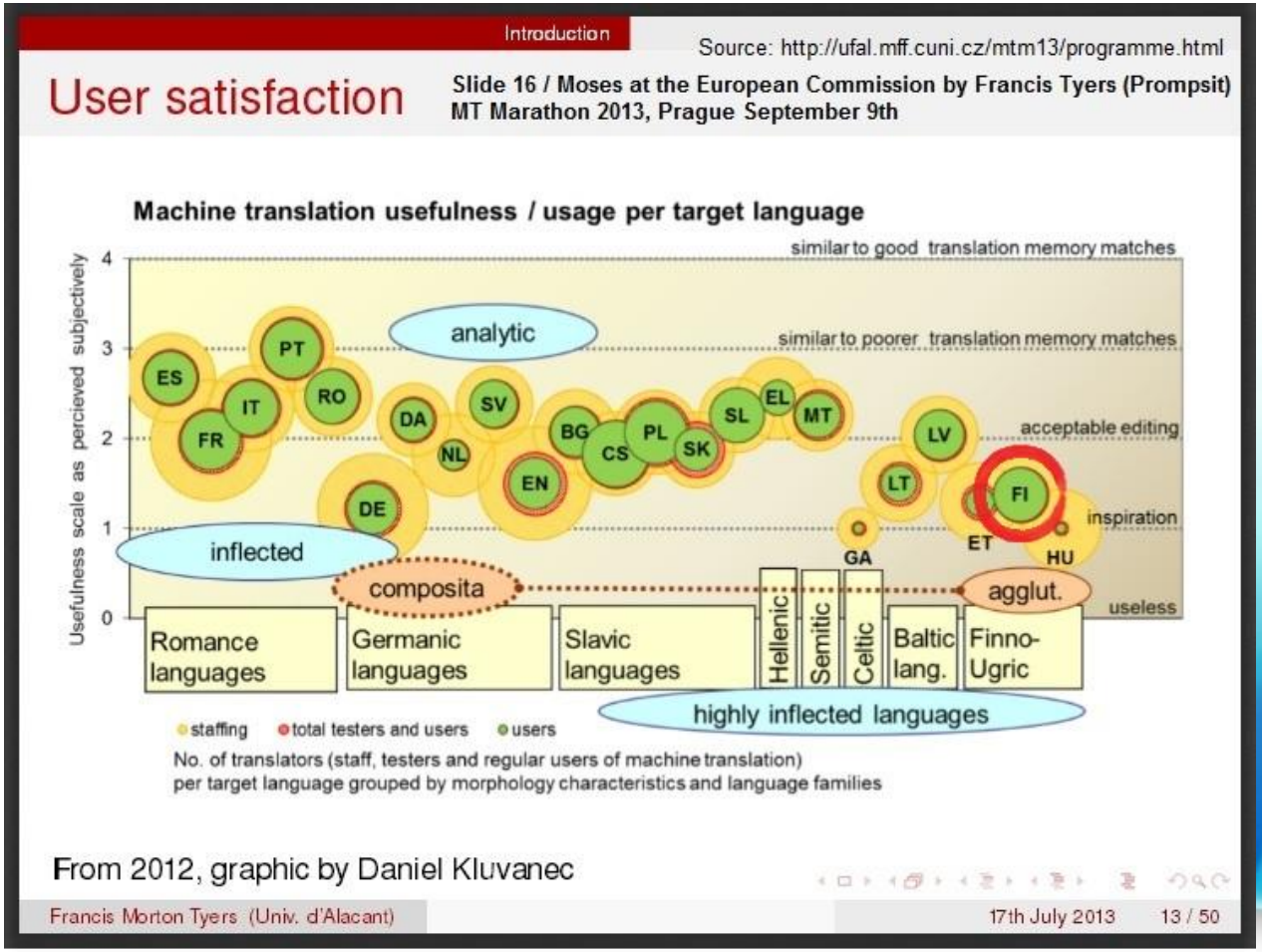
Text Analysis



10: Text analysis: state of language technology support for 30 European languages

Source: META-NET White Paper Series: The Czech Language in the Digital Age, 2012

Machine Translation Usefulness in European Languages



Current State of MT

	Kohdekieli – Target language																					
	EN	BG	DE	CS	DA	EL	ES	ET	FI	FR	HU	IT	LT	LV	MT	NL	PL	PT	RO	SK	SL	SV
EN	–	40.5	46.8	52.6	50.0	41.0	55.2	34.8	38.6	50.1	37.2	50.4	39.6	43.4	39.8	52.3	49.2	55.0	49.0	44.7	50.7	52.0
BG	61.3	–	38.7	39.4	39.6	34.5	46.9	25.5	26.7	42.4	22.0	43.5	29.3	29.1	25.9	44.9	35.1	45.9	36.8	34.1	34.1	39.9
DE	53.6	26.3	–	35.4	43.1	32.8	47.1	26.7	29.5	39.4	27.6	42.7	27.6	30.3	19.8	50.2	30.2	44.1	30.7	29.4	31.4	41.2
CS	58.4	32.0	42.6	–	43.6	34.6	48.9	30.7	30.5	41.6	27.4	44.3	34.5	35.8	26.3	46.5	39.2	45.7	36.5	43.6	41.3	42.9
DA	57.6	28.7	44.1	35.7	–	34.3	47.5	27.8	31.6	41.3	24.2	43.8	29.7	32.9	21.1	48.5	34.3	45.4	33.9	33.0	36.2	47.2
EL	59.5	32.4	43.1	37.7	44.5	–	54.0	26.5	29.0	48.3	23.7	49.6	29.0	32.6	23.8	48.9	34.2	52.5	37.2	33.1	36.3	43.3
ES	60.0	31.1	42.7	37.5	44.4	39.4	–	25.4	28.5	51.3	24.0	51.7	26.8	30.5	24.6	48.8	33.9	57.3	38.1	31.7	33.9	43.7
ET	52.0	24.6	37.3	35.2	37.8	28.2	40.4	–	37.7	33.4	30.9	37.0	35.0	36.9	20.5	41.3	32.0	37.8	28.0	30.6	32.9	37.3
FI	49.3	23.2	36.0	32.0	37.9	27.2	39.7	34.9	–	29.5	27.2	36.6	30.5	32.5	19.4	40.6	28.8	37.5	26.5	27.3	28.2	37.6
FR	64.0	34.5	45.1	39.5	47.4	42.8	60.9	26.7	30.0	–	25.5	56.1	28.3	31.9	25.3	51.6	35.7	61.0	43.8	33.1	35.6	45.8
HU	48.0	24.7	34.3	30.0	33.0	25.5	34.1	29.6	29.4	30.7	–	33.5	29.6	31.9	18.1	36.1	29.8	34.2	25.7	25.6	28.2	30.5
IT	61.0	32.1	44.3	38.9	45.8	40.6	26.9	25.0	29.7	52.7	24.2	–	29.4	32.6	24.6	50.5	35.2	56.5	39.3	32.5	34.7	44.3
LT	51.8	27.6	33.9	37.0	36.8	26.5	21.1	34.2	32.0	34.4	28.5	36.8	–	40.1	22.2	38.1	31.6	31.6	29.3	31.8	35.3	35.3
LV	54.0	29.1	35.0	37.8	38.5	29.7	8.0	34.2	32.4	35.6	29.3	38.9	38.4	–	23.3	41.5	34.4	39.6	31.0	33.3	37.1	38.0
MT	72.1	32.2	37.2	37.9	38.9	33.7	48.7	26.9	25.8	42.4	22.4	43.7	30.2	33.2	–	44.0	37.1	45.9	38.9	35.8	40.0	41.6
NL	56.9	29.3	46.9	37.0	45.4	35.3	49.7	27.5	29.8	43.4	25.3	44.5	28.6	31.7	22.0	–	32.0	47.7	33.0	30.1	34.6	43.6
PL	60.8	31.5	40.2	44.2	42.1	34.2	46.2	29.2	29.0	40.0	24.5	43.2	33.2	35.6	27.9	44.8	–	44.1	38.2	38.2	39.8	42.1
PT	60.7	31.4	42.9	38.4	42.8	40.2	60.7	26.4	29.2	53.2	23.8	52.8	28.0	31.5	24.8	49.3	34.5	–	39.4	32.1	34.4	43.9
RO	60.8	33.1	38.5	37.8	40.3	35.6	50.4	24.6	26.2	46.5	25.0	44.8	28.4	29.9	28.7	43.0	35.8	48.5	–	31.5	35.1	39.4
SK	60.8	32.6	39.4	48.1	41.0	33.3	46.2	29.8	28.4	39.4	27.4	41.8	33.8	36.7	28.5	44.4	39.0	43.3	35.3	–	42.6	41.8
SL	61.0	33.1	37.9	43.5	42.6	34.0	47.0	31.1	28.8	38.2	25.7	42.3	34.6	37.3	30.0	45.9	38.2	44.1	35.8	38.9	–	42.7
SV	58.5	26.9	41.0	35.6	46.6	33.3	46.6	27.4	30.9	38.9	22.7	42.0	28.2	31.0	23.7	45.6	32.2	44.2	32.7	31.3	33.5	–

7: Konekäännös 22 EU-kielen välillä – Machine translation between 22 EU-languages [32]

Source: META-NET White Paper Series 2012 The Finnish language in the Digital Age

MT will be part of translators' work

- MT already improves translation efficiency and speed:
  - Combination of MT and post-editing is faster than pure human translation. Standard text: English to German **30% faster**, English to Spanish **25% faster** (source: SDL)
- MT combined with other LTs will revolutionize translation work:
  - MT evaluation technologies will help in MT post-editing pricing
  - Qualified MT will make post-editing faster
  - Customized MT with controlled language will produce best quality
  - Speech recognition with MT will make interpretation more efficient
  - MT will complement TM

Even MT alone brings savings and efficiency

- **60-70% of respondents prefer imperfect machine translation** to the alternative of no translation at all. (Source: Symantec)
- Machine translation can help you **serve more international customers** while dramatically reducing the cost of call center operations by averting a minimum of 25% of calls. (Source: Cisco)
- Machine translation is capable of **resolving nearly as many customer support problems** as a significantly more costly human translation. (Source: Microsoft, Intel)

MT Example:

TripAdvisor

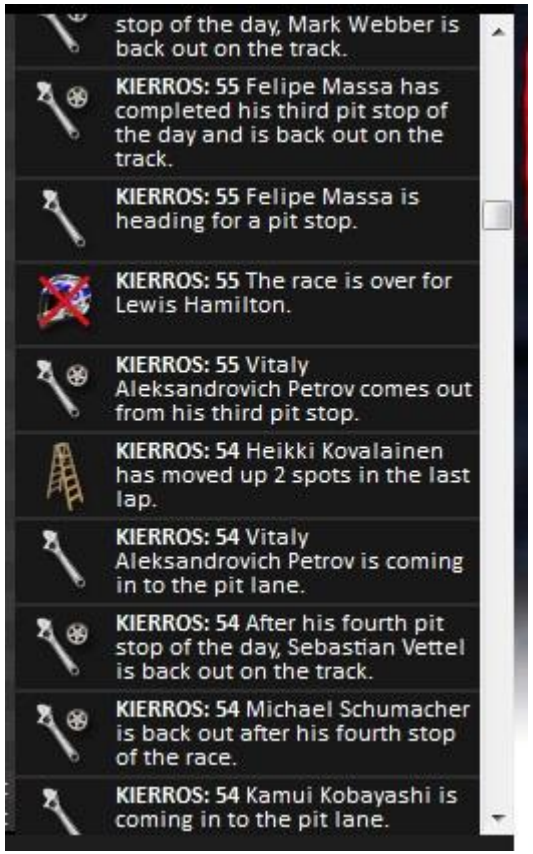
Customer reviews

Although MT is far from perfect its volumes are enormous

- Microsoft Office users translate over **1.3 million characters each hour** using the built-in translation features.
  - http://microsoft-news.com/bing-translator-for-word-powerpoint-2013-launching-today/
- Google Translate serves **200 million people daily**, and breaks down language barriers a **billion times a day**.
  - http://news.cnet.com/8301-1023_3-57585143-93/google-translate-now-serves-200-million-people-daily/

- *"Language technology is the missing piece of the puzzle that makes up the digital single market,"* said Jochen Hummel, Chairman of LT-Innovate
  - *http://www.lt-innovate.eu/page/lt-innovate-summit-2012-summary*
- *"With more than two dozen languages spoken across the EU's 28 countries, communication remains a key impediment to the European economic success."*
  - http://associationsnow.com/2013/10/language-trade-group-teams-with-eu-on-translation-tech/

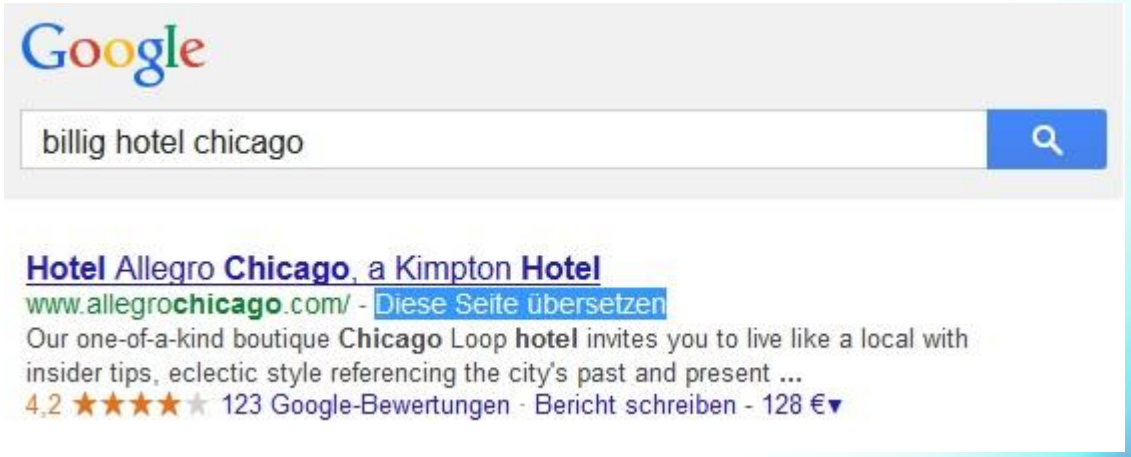
MT + Text Synthesis

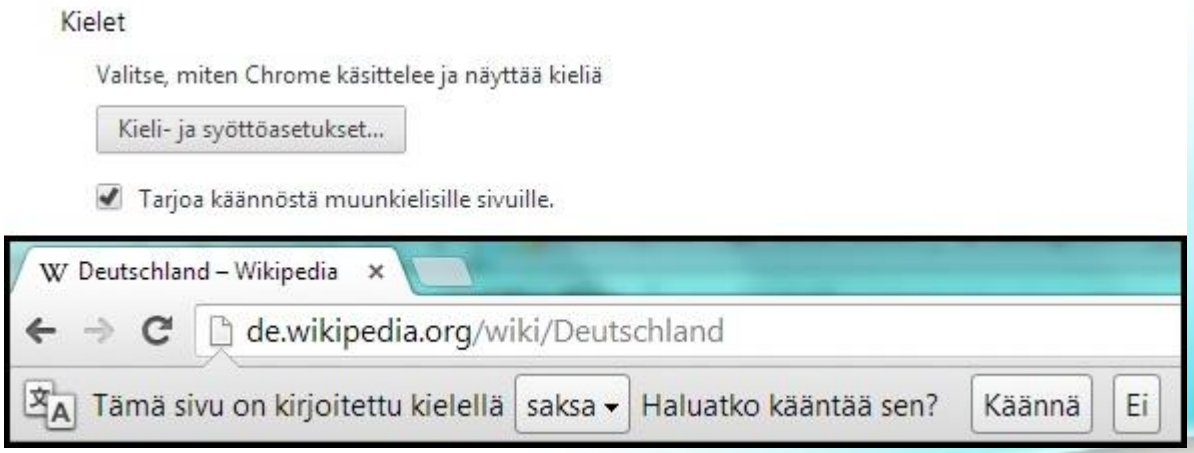


Example: MTV3 F1 Tulospalvelu

- Text synthesis produces controlled language.
  - e.g. weather reports, sport news and descriptions
- Controlled language translates very well with machine translation.

MT will be transparent

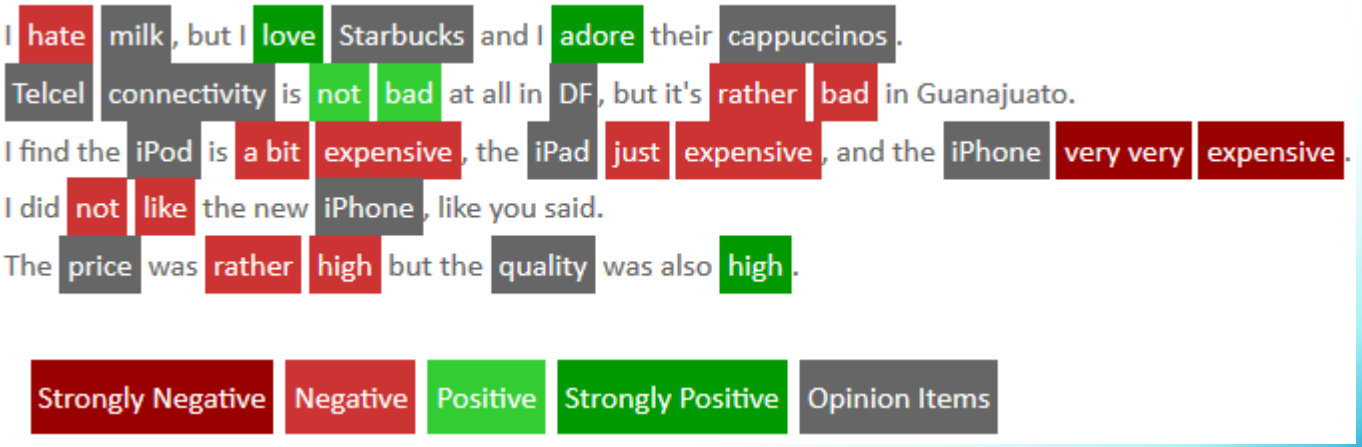




Text Analysis is the key to Data Mining and understanding Big Data

- *"The ability to manage and process the tsunami of data across the world's languages is one of the biggest challenges in the new ICT ecosystem, and one for which LT is a critical enabling technology."* says LT-Innovate
  - http://www.lt-innovate.eu/resources/document/infographic-lt2013-report-status-and-potential-european-language-technology-marke

A simple example of Text Analysis

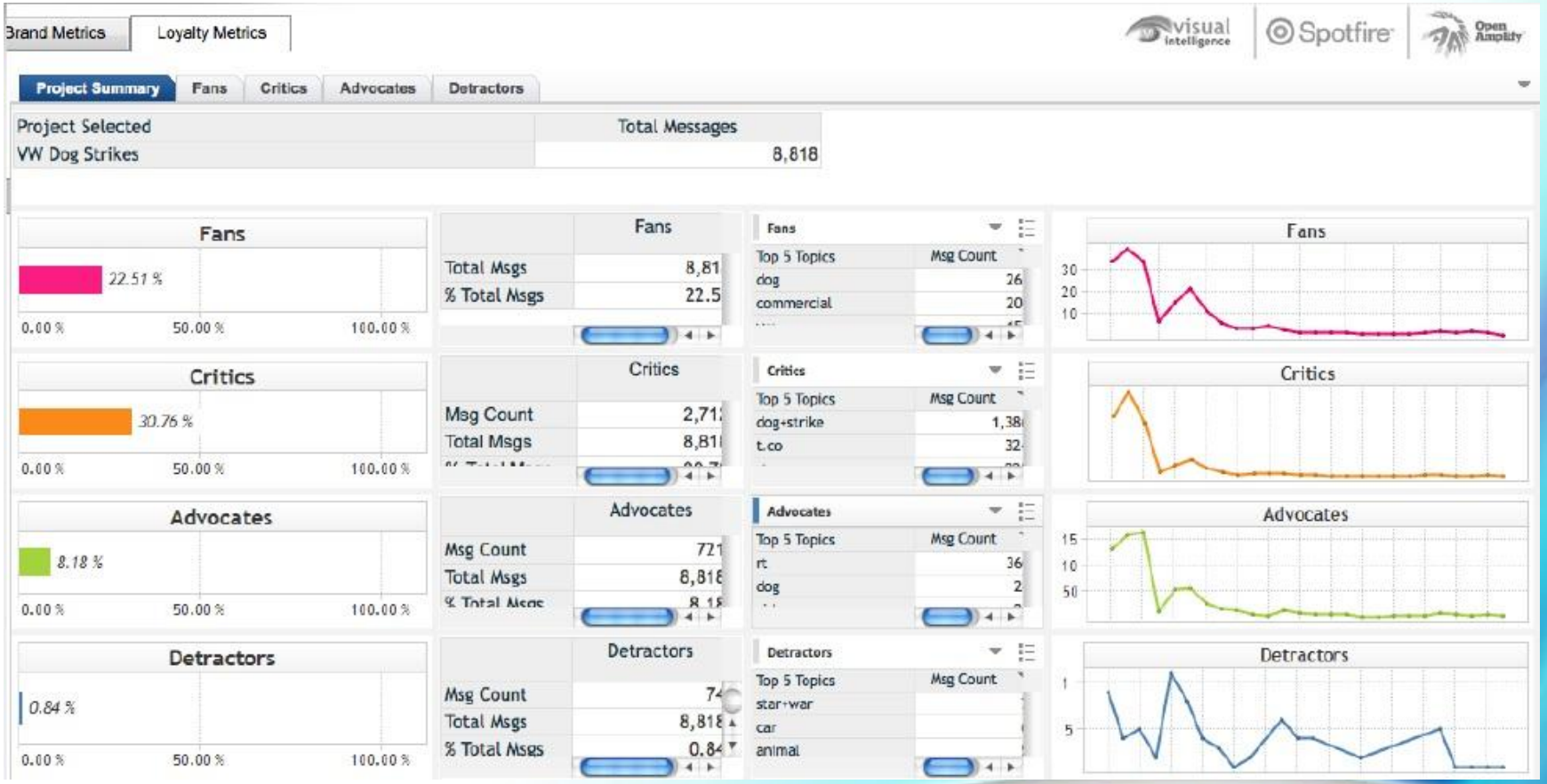


Another example of Text Analysis

”I love my BMW! It’s such a comfortable car so I’ll certainly check out the new 5-series next week”

- Automatic text analysis of the above sentence results in:
  - it’s about cars
  - strong like of BMW
  - positive towards comfort
  - intent with certainty, to assess the ‘new 5 series’
  - time scale, next week

Text Analysis helps to monitor data, understand trends and find leads



Text analysis will eventually enable **computers to understand normal text** which will have very far-reaching consequences!

3 global trends that create huge need for LT based solutions

- **Unified Communication:** Connectivity across devices and platforms will offer business and consumer users seamless communications.
- **Unified Information Access:** Content sharing in any language and across languages.
- **Unified User Experience:** Access, use and understanding of information from large volumes of data.

Source: LT2013: Status and Potential of the European Language Technology Markets by LT-Innovate

LT is an opportunity to language services industry

- Language technologies are bringing efficiency to situations which can be improved.
- LT is not a threat to highest quality services.
- Language technologies will stretch the language service industry to new areas.
- Innovative companies can **step out of the box and find new business possibilities**.

Worldwide LT Software & Services Market (€B)

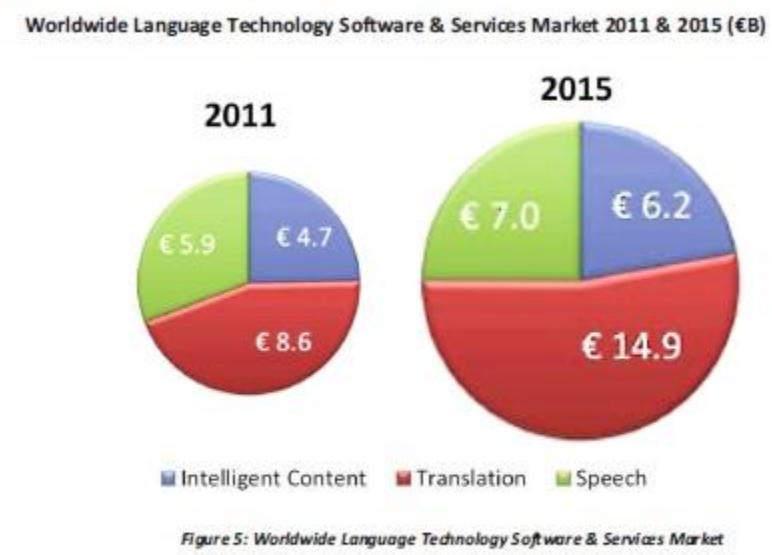


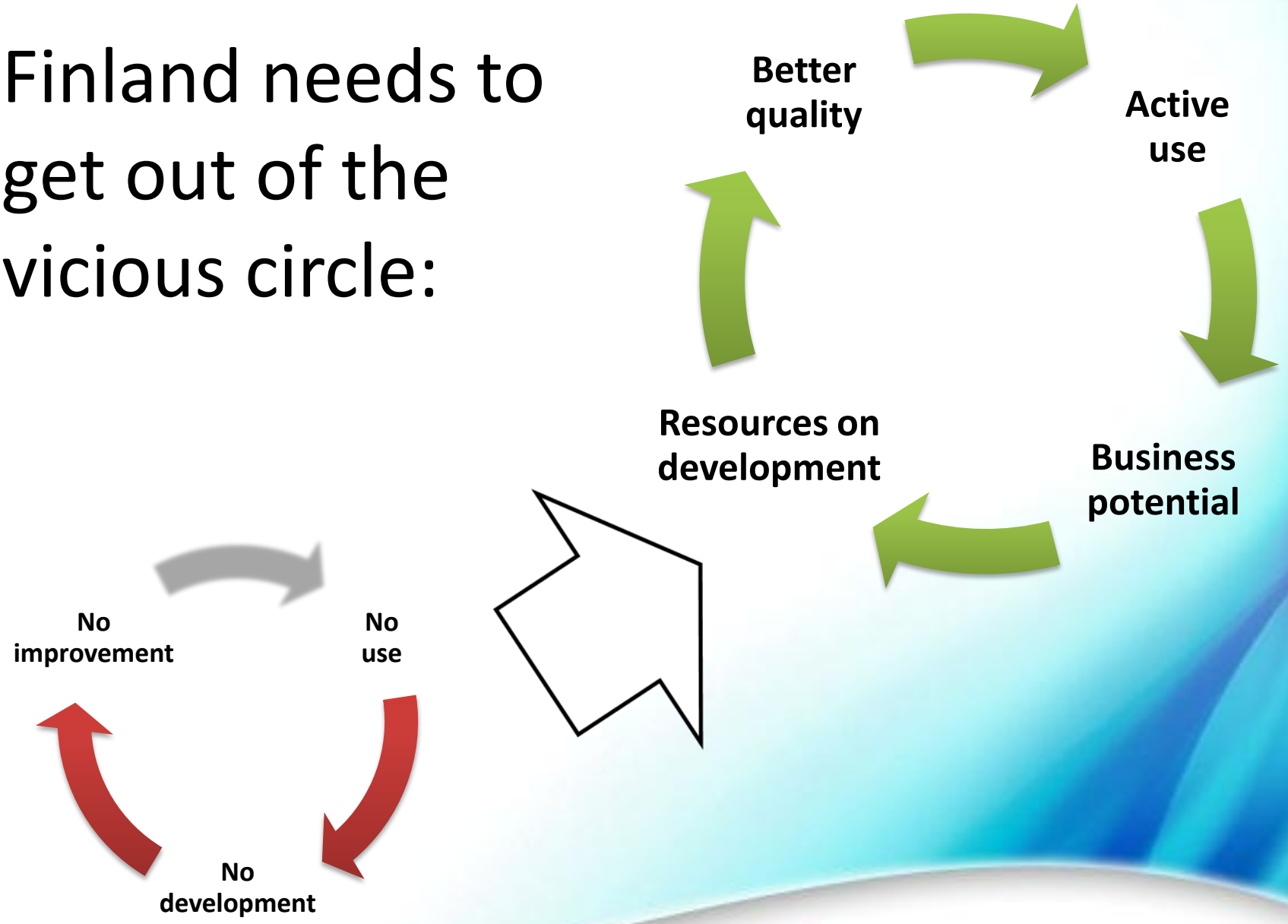
Figure 5: Worldwide Language Technology Software & Services Market

Source: LT-Innovate: LT2013: Status and Potential of the European Language Technology Markets, page 16.

Natural Language Processing (NLP) Market Worth **$9,858.4 Million** by 2018

- Natural Language Processing (NLP) Market includes IVR, OCR, Pattern Recognition, Auto Coding, Text Analysis, Speech Analysis, Machine Translation, Information Extraction, Question Answer, Report Generation.
  - Worldwide Market Forecast and Analysis (2013 - 2018) by MarketsandMarkets
  - http://finance.yahoo.com/news/natural-language-processing-nlp-market-133000650.html

Finland needs to get out of the vicious circle:



Summary

- Global competition favors quick adopters who can offer solutions to all types of language processing needs.
- The economical focus is expanding to areas where people don't communicate in English. **Global companies and individuals will find solutions – with or without us.**

How do you see **the future**?

Contact information:

Niko Papula, Multilizer

Email: niko.papula@multilizer.com

Mobile: +358 50 586 9145