

CROSSLANG COMPANY PROFILE

- Independent consultants and systems integrators
- Dedicated to Translation Automation Technology
- Founded in 2002
- Based in Gentbrugge, Belgium

Consultancy	Solutions
Language audits	Translation Management Systems
Pilots and Proof of Concepts	Terminology Management
Programme Management	Machine Translation

LANGUAGE TECHNOLOGY MARKET: STATE-OF-THE-ART, TRENDS AND VALUE PROPOSITION

SMART 2016/0103

CONTEXT AND OBJECTIVES

- SMART 2016/0103 is a **service contract** supporting the EC in implementation of Automated Translation building block
- Overall objective:
 - Contribute to the development of CEF Automated Translation as “multilingualism enabler” (beyond simple translation) for:
 - CEF DSIs & Online services linked to CEF DSIs
 - Other relevant public online services

SMART 2016/0103

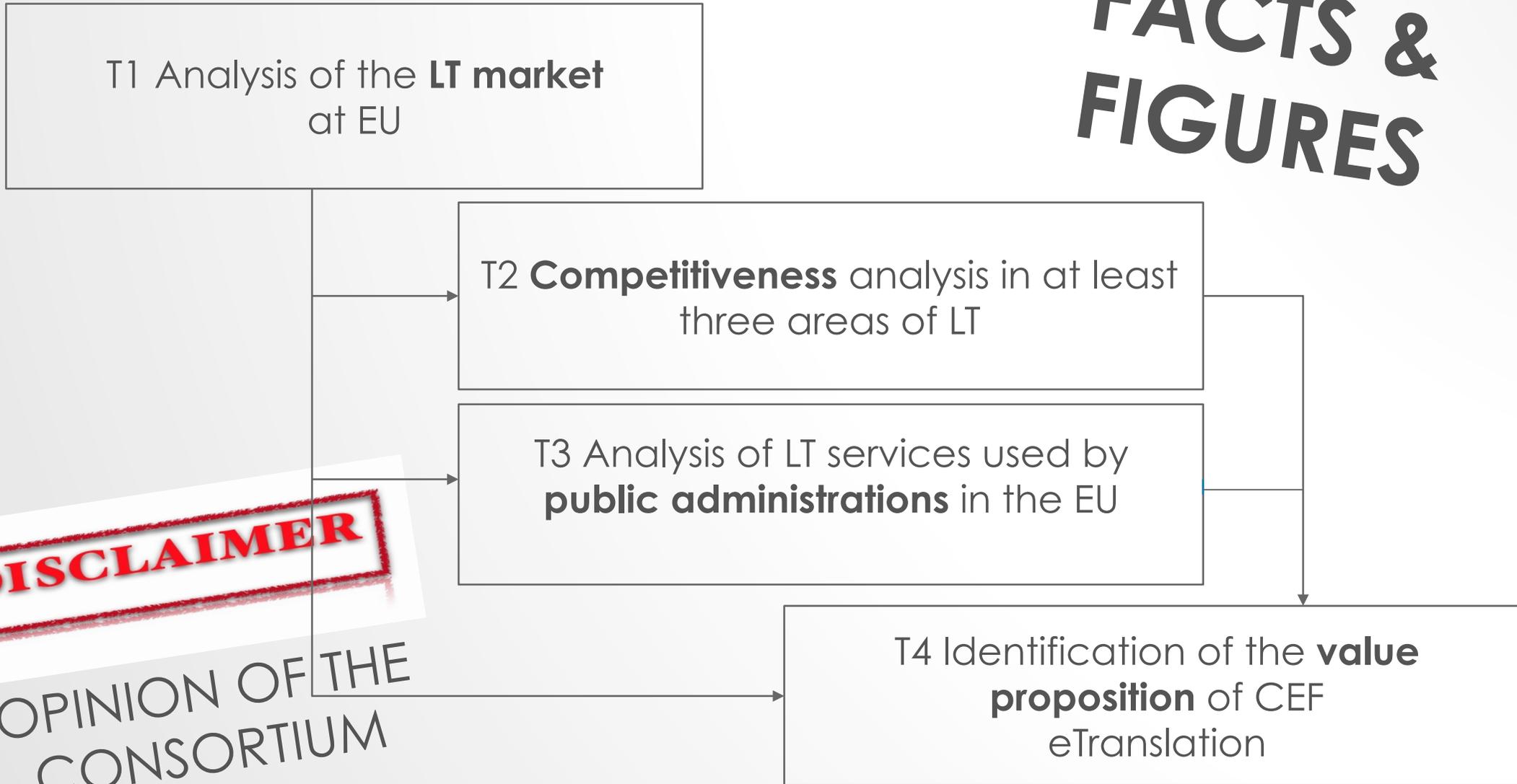
Lot 1 (2018)

Study on service portfolio development and business case for CEF Automated Translation (**“Market Study”**)

Lot 2 (2018-2020)

Implementation of the ‘service desk’ part of the CEF Automated Translation platform and a program support action (**“Service Desk”**)

FACTS & FIGURES



DISCLAIMER

OPINION OF THE CONSORTIUM

TASK 1. SUPPLY SIDE ANALYSIS

- **Objectives**

- Sizing the market and identifying the main trends
- Estimating the number of LT related companies, including market shares and revenues
- Identifying existing offering of LT tools/services

- **Approach**

- Desk research and IDC's Software Tracker
- Through online questionnaires & telephone interviews

TASK 1. SUPPLY SIDE ANALYSIS

Updated list of LT vendors per country and per type of technology

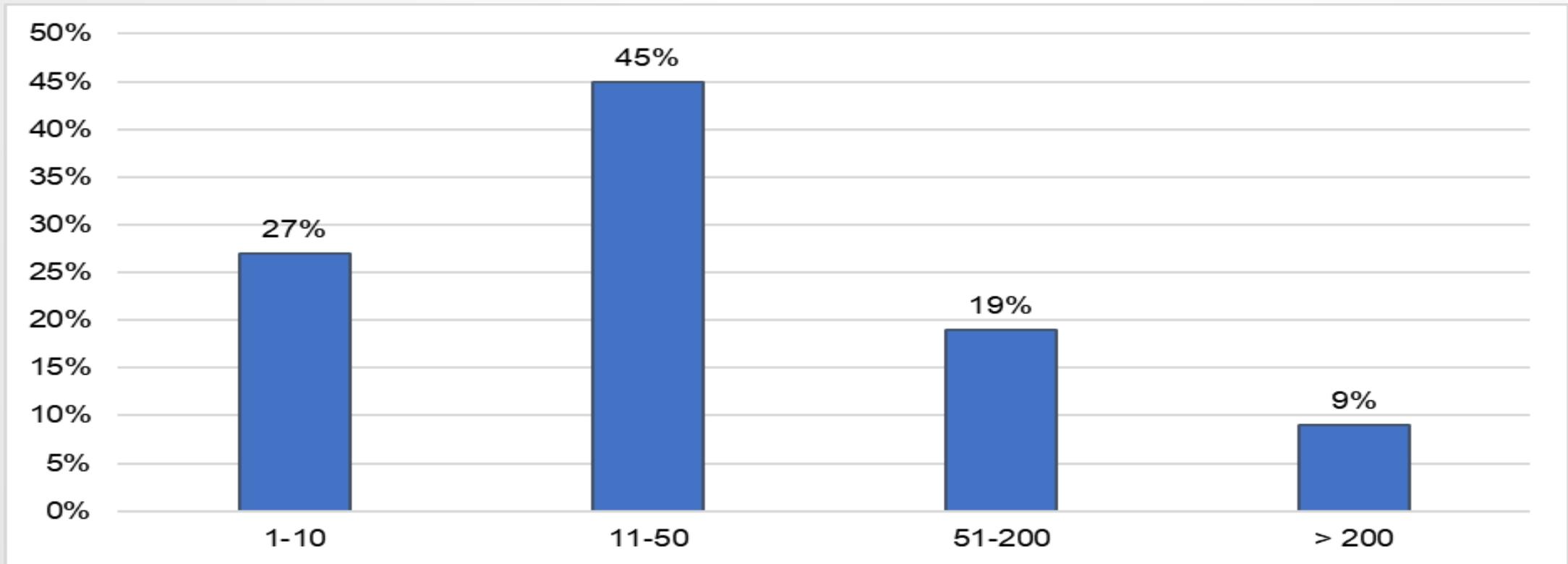
- **Translation Technology:**
MT, TM, TMS
- **Speech Technology:**
ASR, Synthesis
- **Search Technology:**
Multi-lingual & Semantic search
- **Natural Language Understanding:**
Chatbots, Anonymization, Virtual assistants
- **Analytics:**
Text analytics, Keyword extraction, Sentiment analysis

LANGUAGE TECHNOLOGY FORECAST 2018-2020

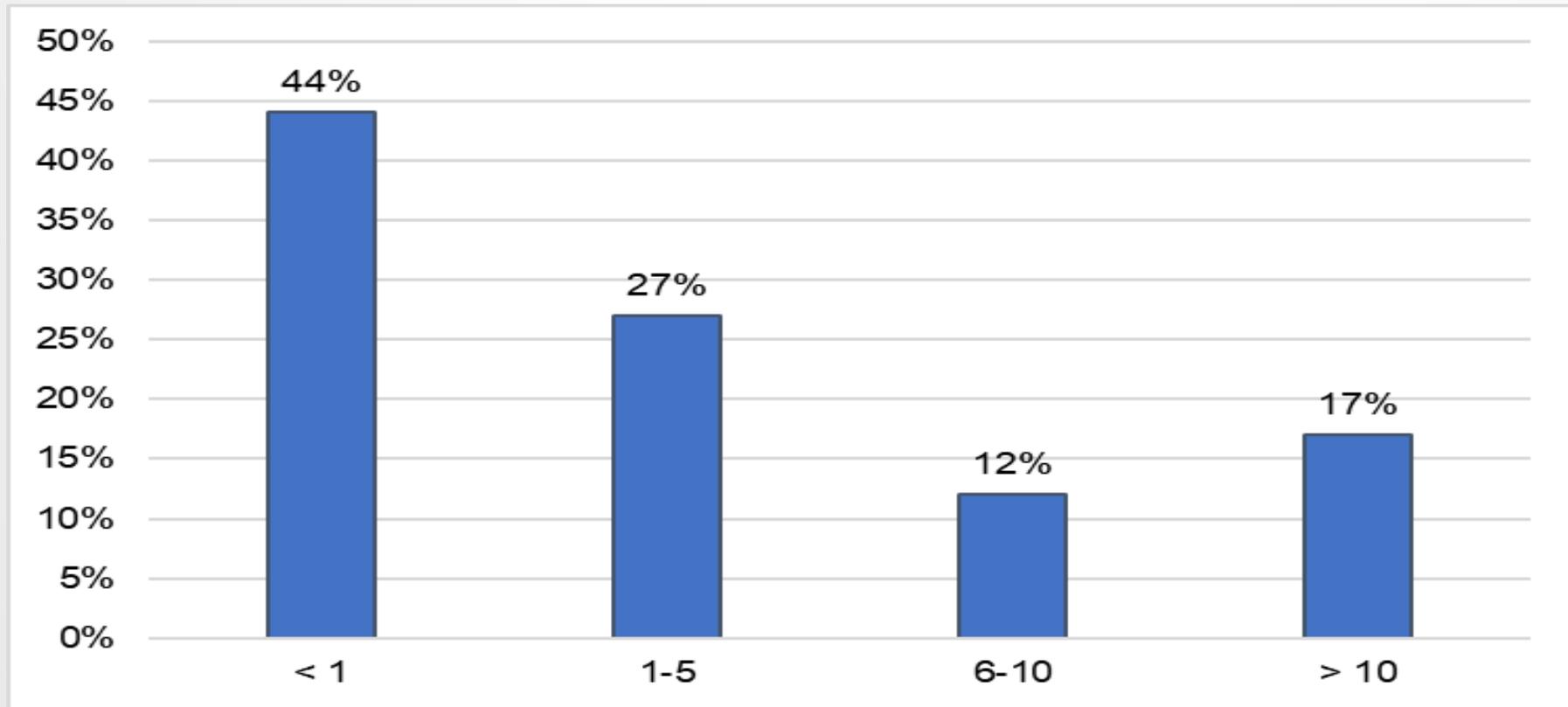
	2018	2019	2020
Germany	197	217	240
United Kingdom	189	209	232
France	88	96	105
Netherlands	55	60	66
Rest of EU 28	249	277	305
	778	859	948

In EUR million

LT VENDORS / EMPLOYEE NUMBERS

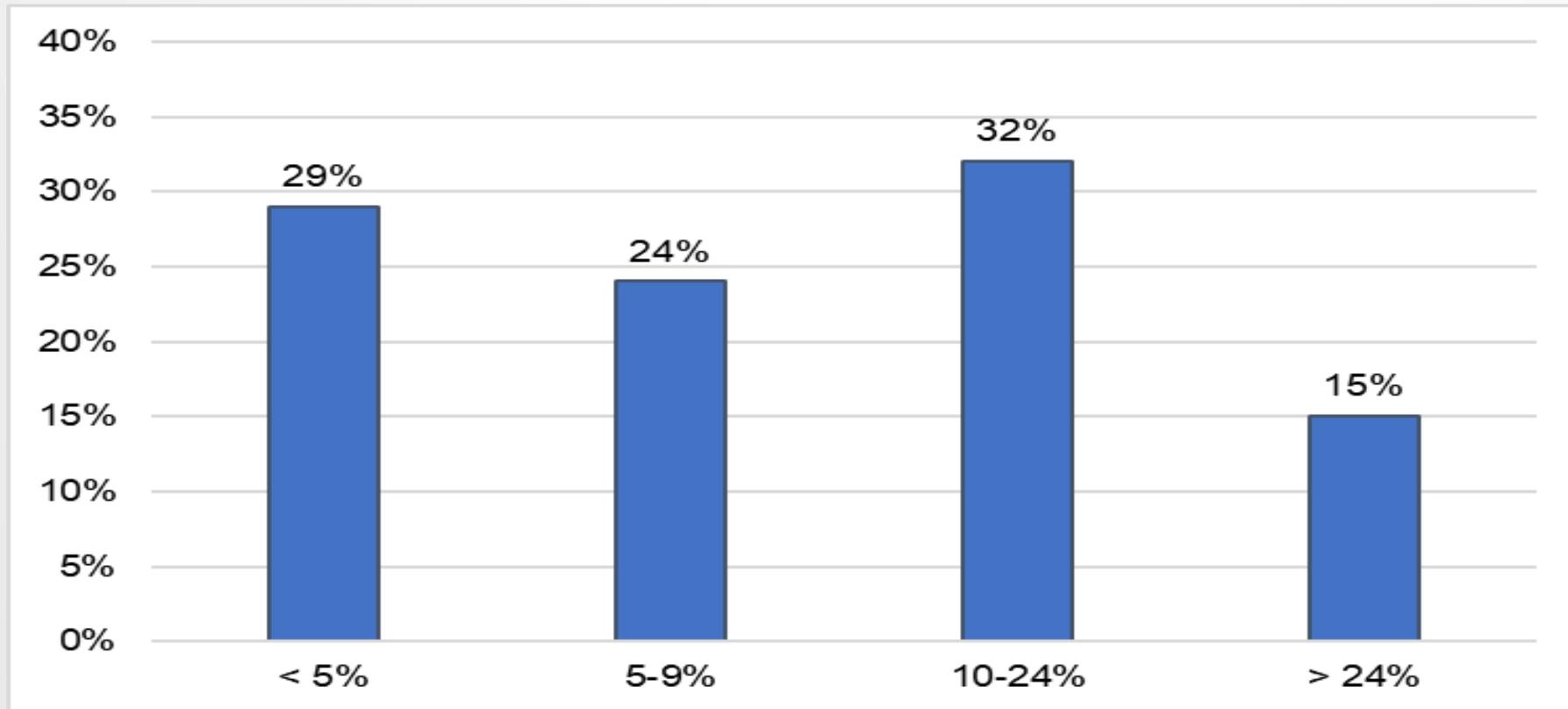


LT VENDORS / TURNOVER

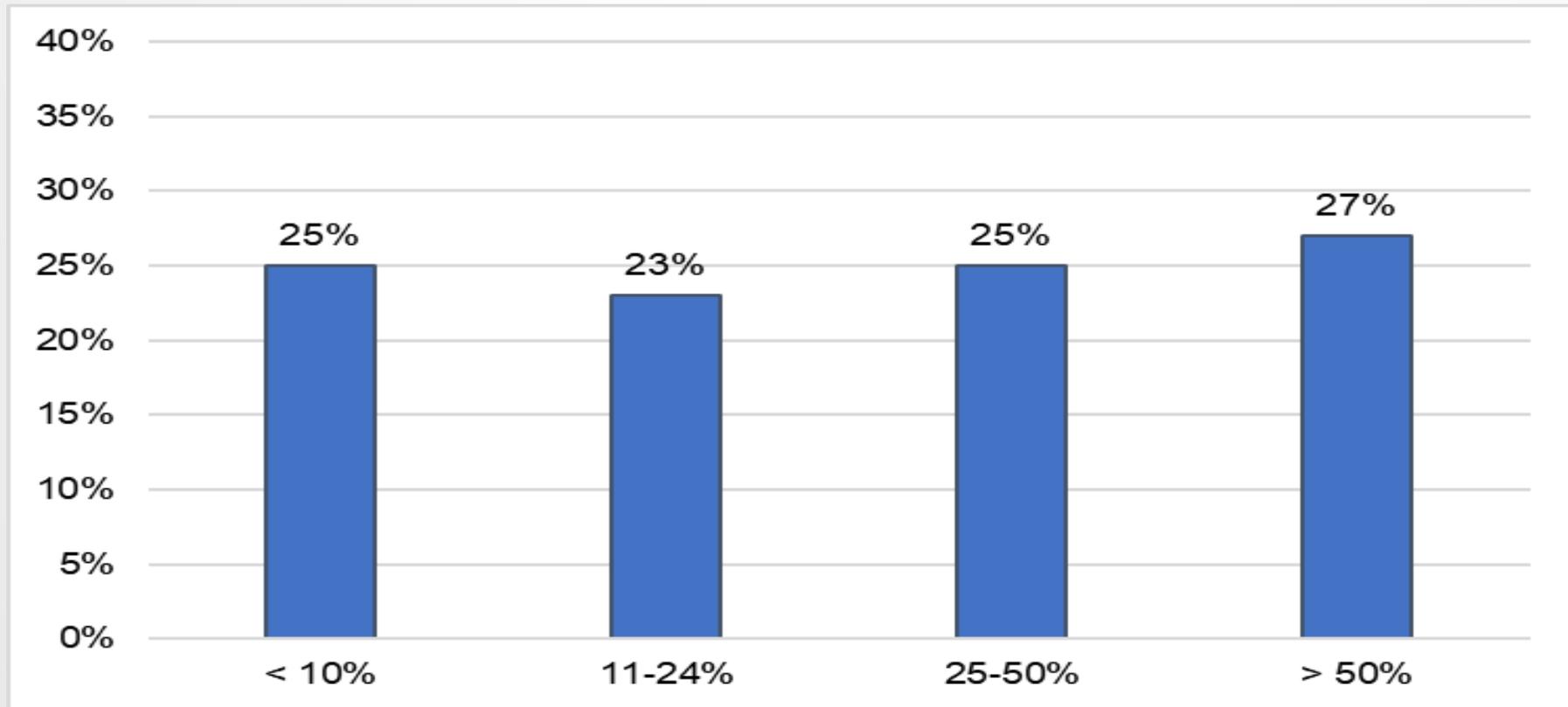


In EUR million

LT VENDORS / PROFITABILITY



LT VENDORS / REVENUE GROWTH 2019 - 2020



MOST CITED EMERGING TECHNOLOGIES

- Chatbots
- Machine Translation
- Automatic Summarization
- Natural Language Processing
- Automatic Speech Recognition

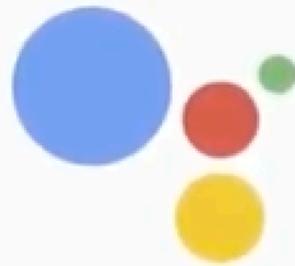


Multilingual Chatbots: The Conversation Has Yet to Get Longer

Just over 20 percent of search queries are made via voice now, and experts are predicting 50 percent of search to be done via voice by 2020.

Chatbots are relatively new to the market, but Gareth Dunlop, CEO and founder of user experience consultancy Fathom, points to a prediction by Gartner, forecasting that 25 percent of customer support and services will be integrated with virtual assistants by 2020.

GOOGLE I/O 2018



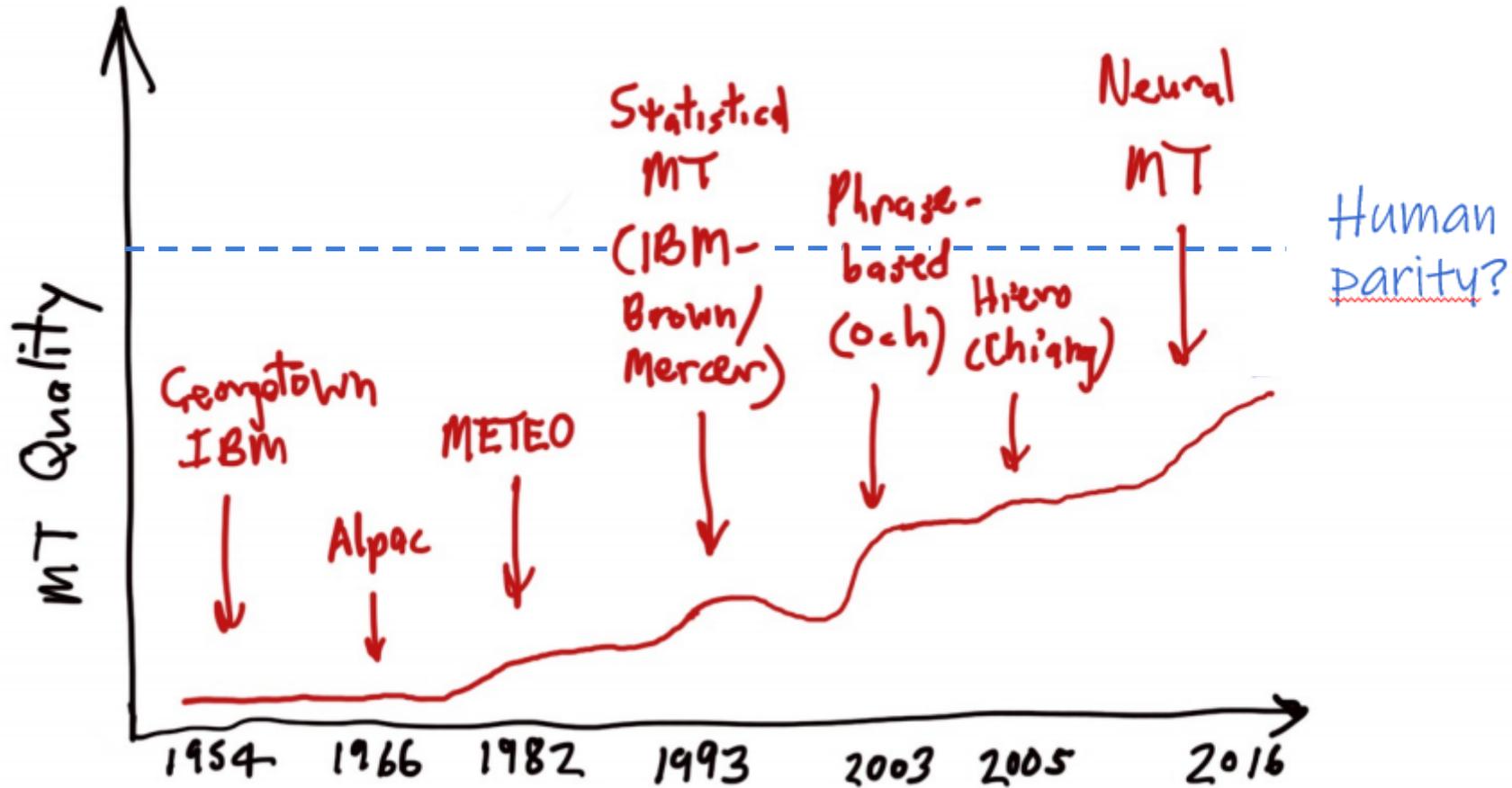
"Hi, I'm calling to book a women's haircut for a client."



GOOGLE I/O 2018



Progress in MT



From: <https://nlp.stanford.edu/projects/nmt/Luong-Cho-Manning-NMT-ACL2016-v4.pdf>

What's so Massive About Google's Massively Multilingual Neural Machine Translation?

6 Days Ago

Technology · By Esther Bond On July 18, 2019



Google's AI team recently unveiled a new research paper on neural machine translation that has been five years in the making. The research paper, entitled "Massively Multilingual Neural Machine Translation in the Wild: Findings and Challenges," published on July 11, 2019.

TASK 2. COMPETITIVENESS ANALYSIS

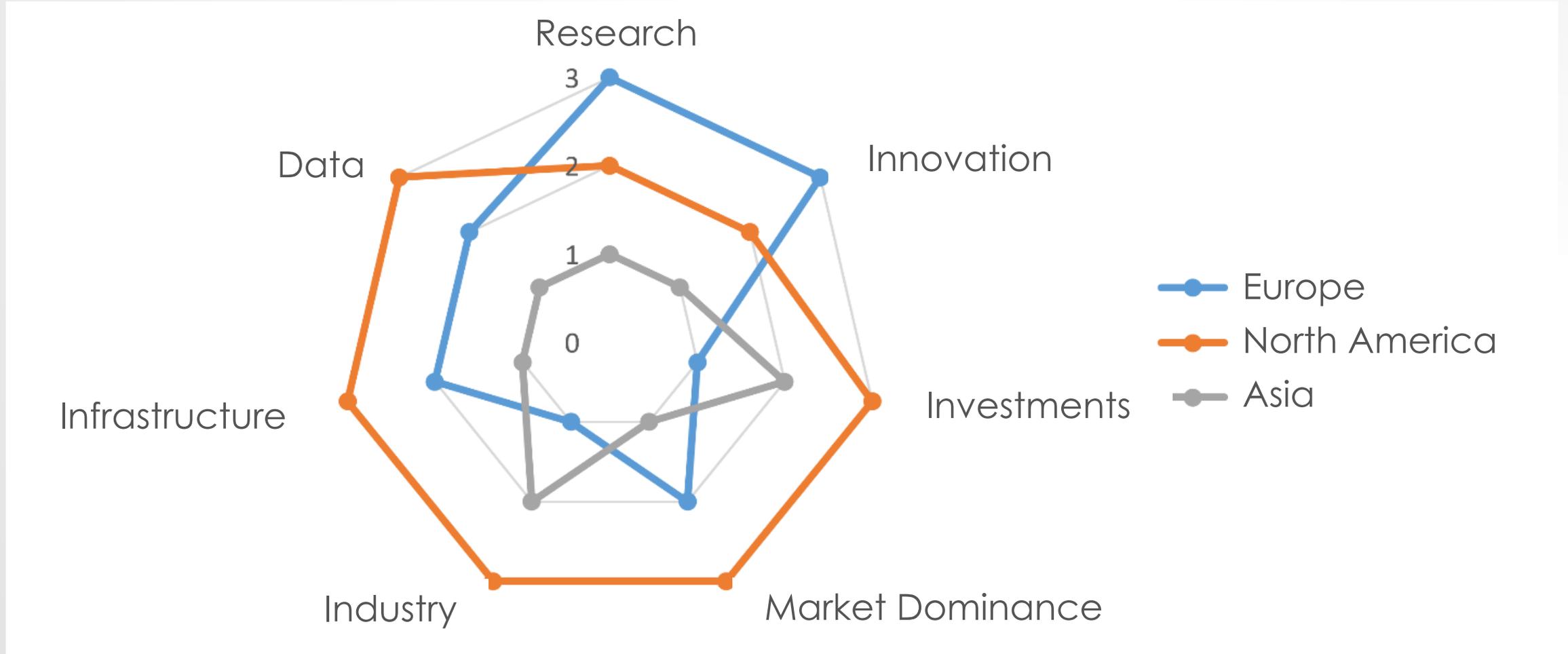
Objectives

- Conduct a competitiveness analysis
- Identify possible shortcomings and opportunities for the European LT market

Approach

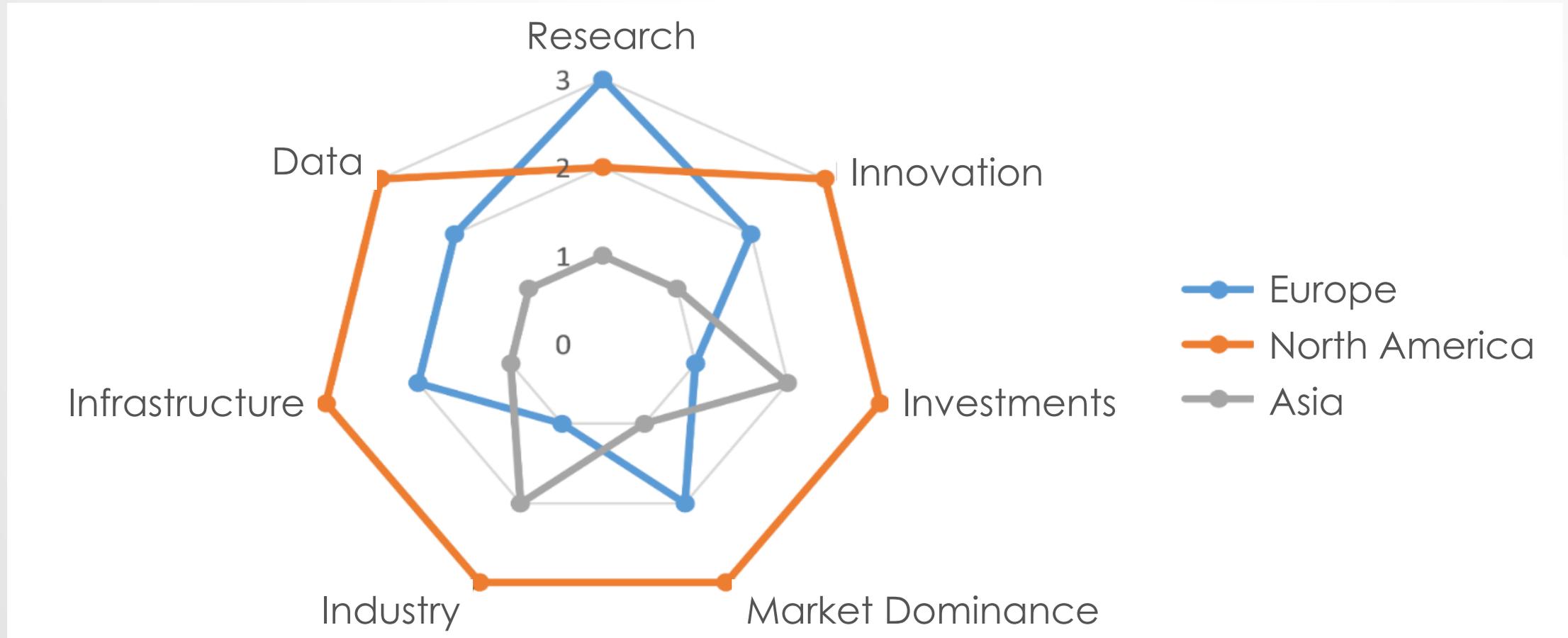
- An analysis of 7 dimensions of the LT market within each of the three domains (score card)
- SWOT analysis from European perspective

MACHINE TRANSLATION

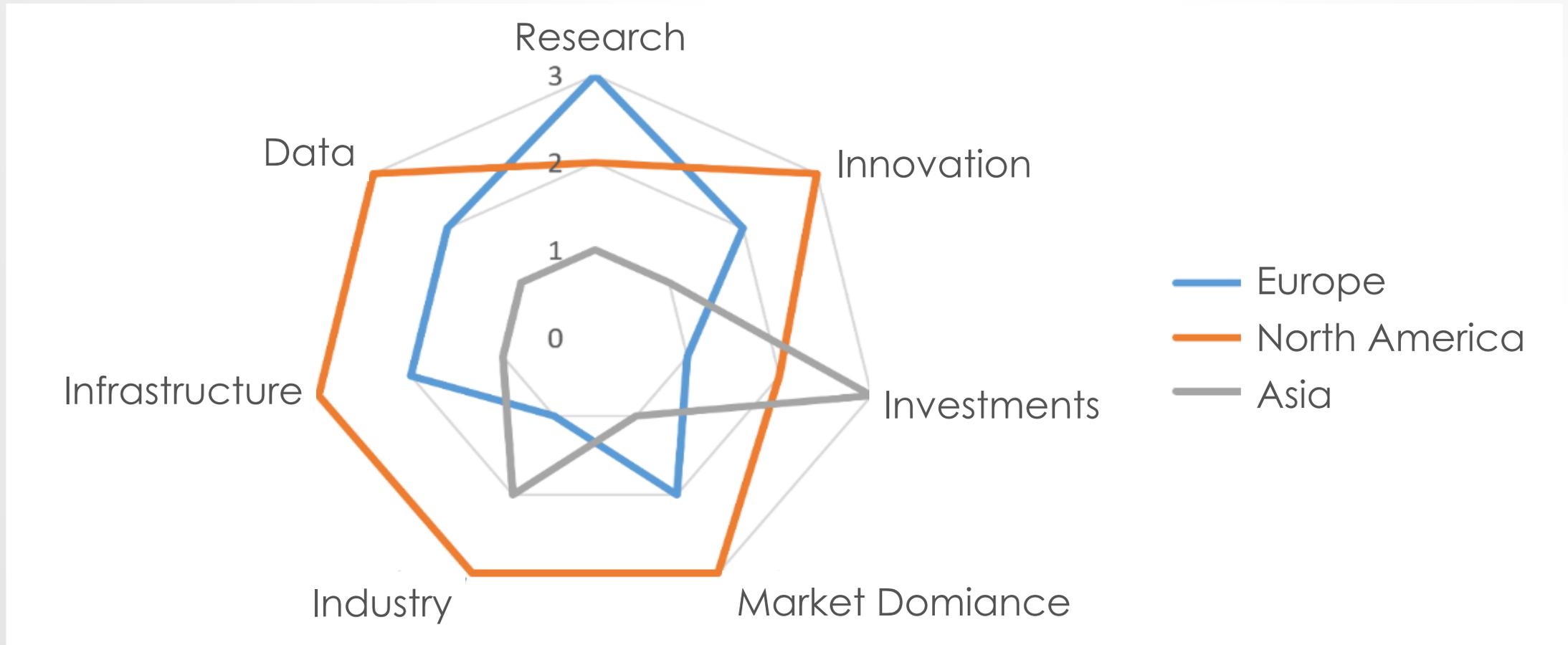


- Europe
- North America
- Asia

SPEECH TECHNOLOGY

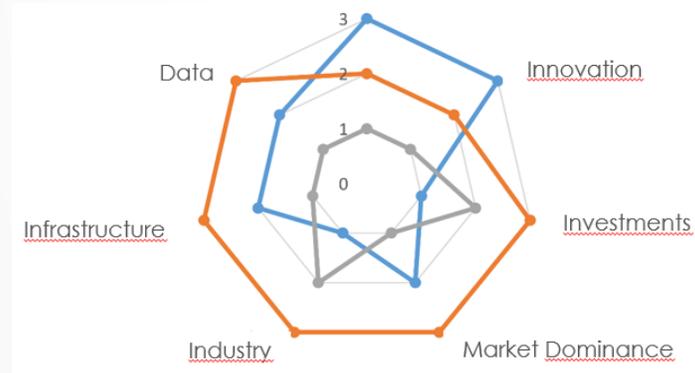


SEARCH TECHNOLOGY



- Europe
- North America
- Asia

RESEARCH



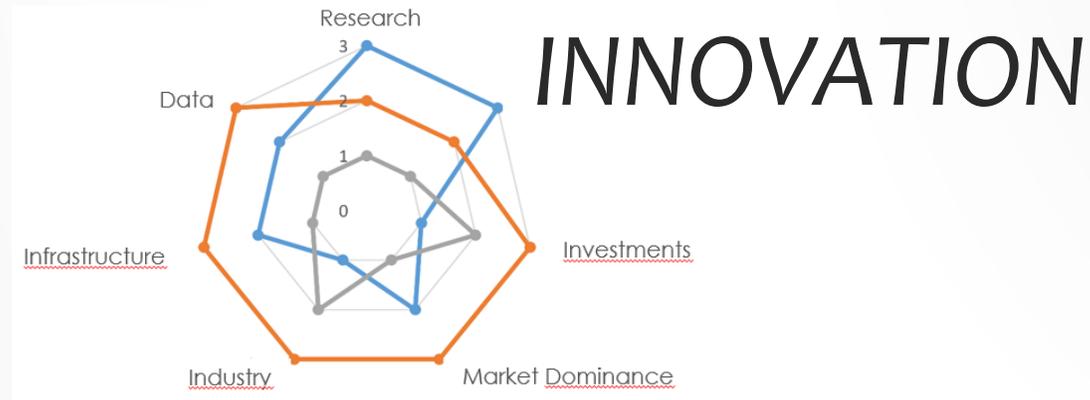
- Research is based on :
 - Research centres
 - Research Publications (Scopus database – industry & academia)
 - Organisational infrastructure (eg. Associations, networks, infrastructures)

RESEARCH IN MT

- Research Centres in Europe : 54 (vs. 23 in US and 14 in Asia)
- Publications in Europe : 1752 (vs. 975 in US and 1932 in Asia)

Per Country:		Per institute/university:	
Spain	293	DCU	162
Germany	266	Politecnica de Valencia	65
UK	266	Edinburgh	58
Per Author:		Per private organization:	
Way A.	75	DFKI	54
Casacuberta F.	45	FBK	54
Specia L.	44		

- Organisational Infrastructure : EAMT/ELRC/LT Innovate/TAUS/META-NET/CLARIN/ELEXIS/DARIAH/Toolkits (Moses, Nematus, Open NMT, Marian)



- Innovation is based on :
 - Market of origin of most popular tools
 - Eg. CAT tools (Across, Déjà Vu, Matecat, memoQ, Memsource, ..)
 - Implementation of latest techniques
 - Eg Adoption of NMT : European companies were quick to follow Google and Facebook
 - Emergence of start-ups
 - <https://angel.co>



Regional distribution of translation technology start-up companies

TECHNOLOGY



TECHNOLOGY

June 11 – 13 2020, Paris

The world's rendezvous
for startups & leaders

[PRE REGISTER](#)

124,000
Attendees

13,000
Startups

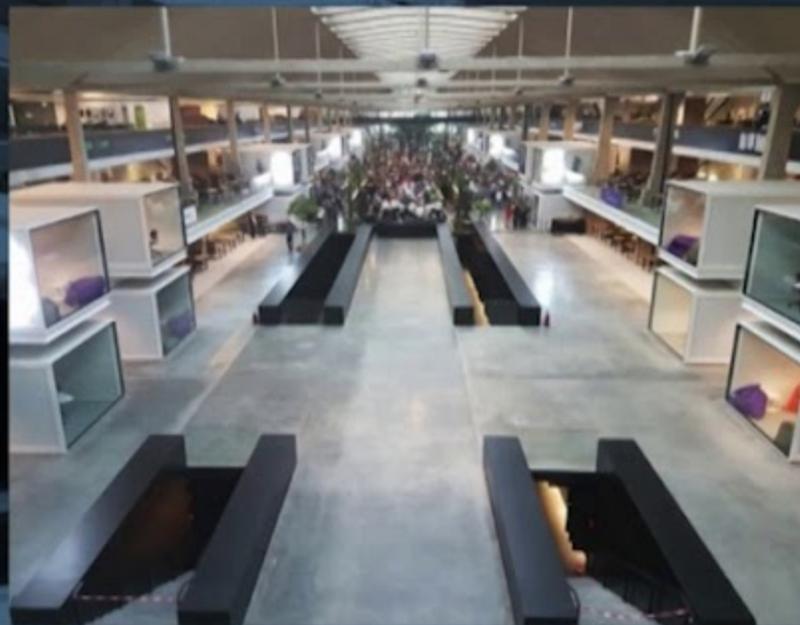
3,300
Investors

2,500
Journalists

125
Countries

STATION F

The largest startup campus

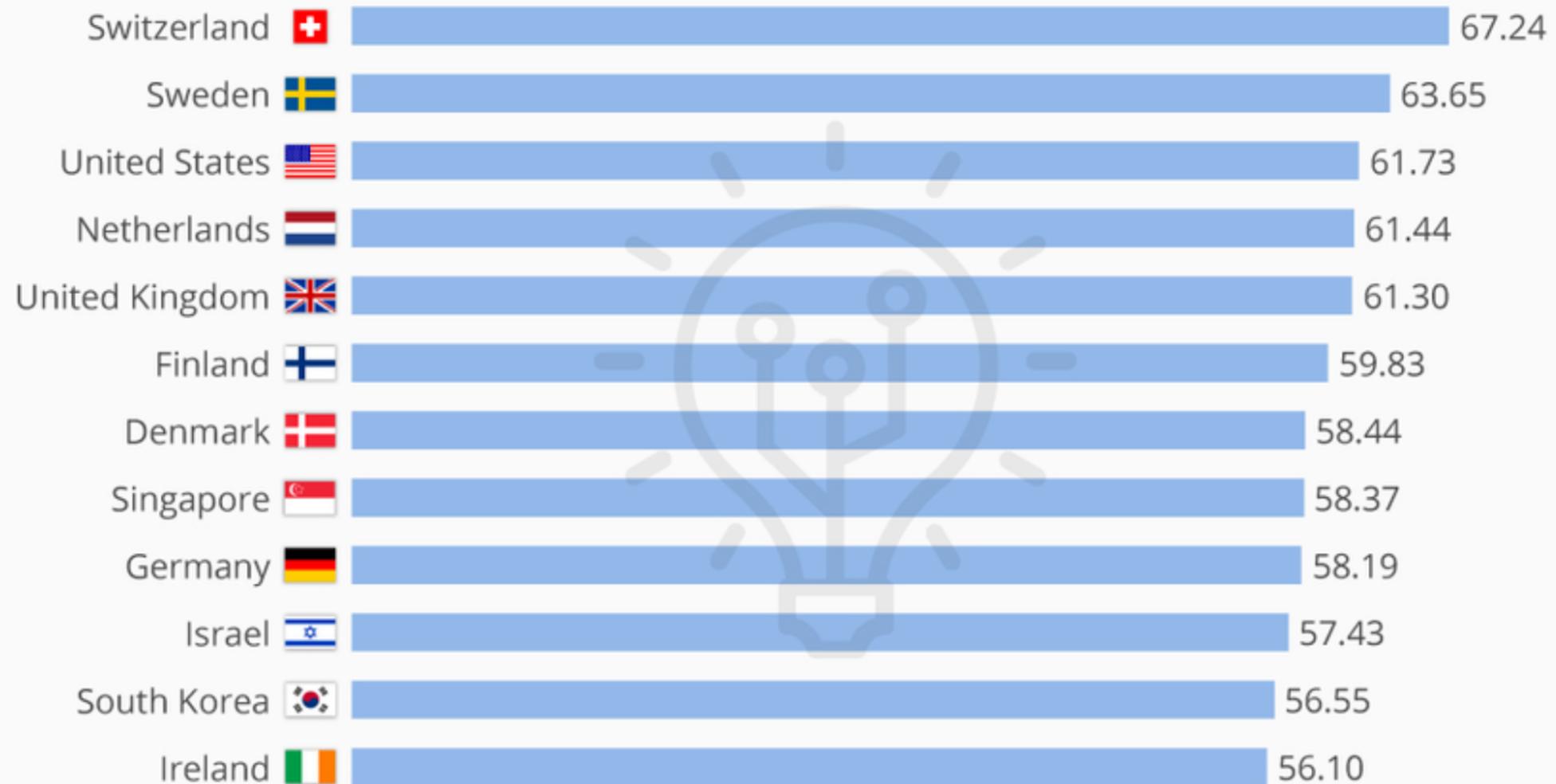


1,000
startups



The World's Most Innovative Countries

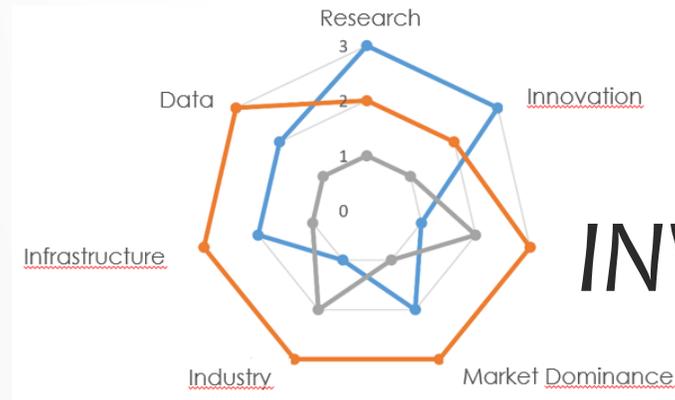
2019 rankings of the Global Innovation Index (100=most innovative)



@StatistaCharts

* Scores determined through various factors such as business sophistication, level of human & capital research and creative outputs.

Source: World Intellectual Property Organization



INVESTMENTS

- Investments are measured through:
 - M&A activity
 - Venture Capital / Financing of companies



- *“The European HLT industry is mainly made up by innovative smaller companies (..). Although most of them have been established in the market for several years, the fragmentation of the LT market in Europe (..) hamper their growth. The transformation into global players (..) requires financing in all stages of business life cycle, not only in research activities.” (European Parliament, 2017, p. 104)*

State-Sponsored Speech Recognition: China's iFlytek Raises USD 400m in 'Private' Placement

1 Week Ago

M&A And Funding · By [Marion Marking](#) On July 22, 2019

Unbabel raises \$23M for its 'AI-powered, human-refined' translation platform

How we can protect innovative tech companies from being sold and moving away

12 June 2019



#VentureEU

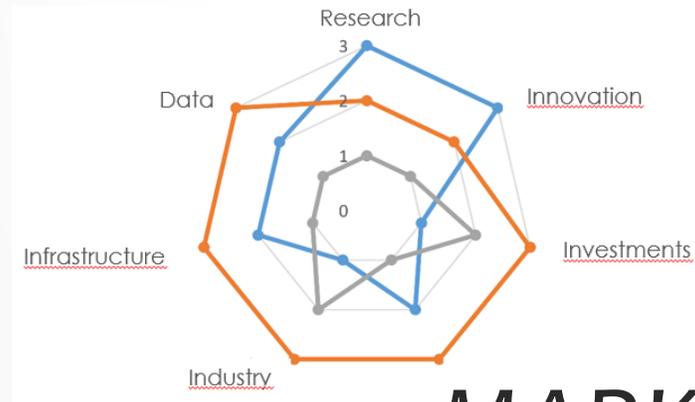


VentureEU

**BOOSTING VENTURE
CAPITAL INVESTMENT
IN EUROPE'S INNOVATIVE
START-UPS**



European
Commission



MARKET DOMINANCE

- ‘Market Dominance is measured by strength of a brand, product, service or firm’
- Comparing total web traffic of largest providers of LT services



Karen Tso 

@cnbcKaren



The world's most valuable global brand...  is ...[#Amazon!](#) Here is the list from [#Brandz](#) [@WPP](#) [@Kantar](#)

1. Amazon
2. Apple
3. Google
4. Microsoft
5. Visa
6. Facebook
7. Alibaba
8. Tencent
9. McDonalds
- 10 AT&T

 5 7:47 AM - Jun 11, 2019



 [See Karen Tso's other Tweets](#)

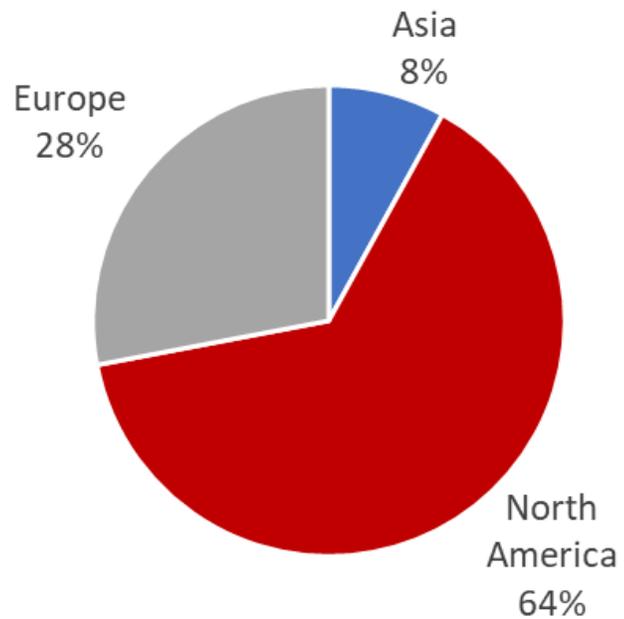


MARKET DOMINANCE

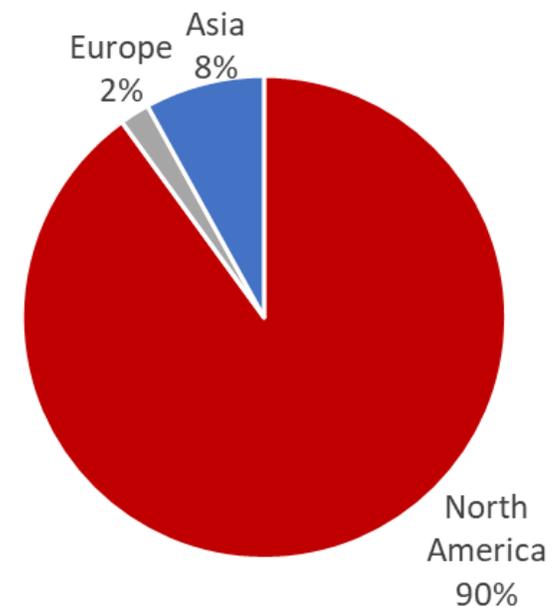
MT

SPEECH

Web traffic share by region



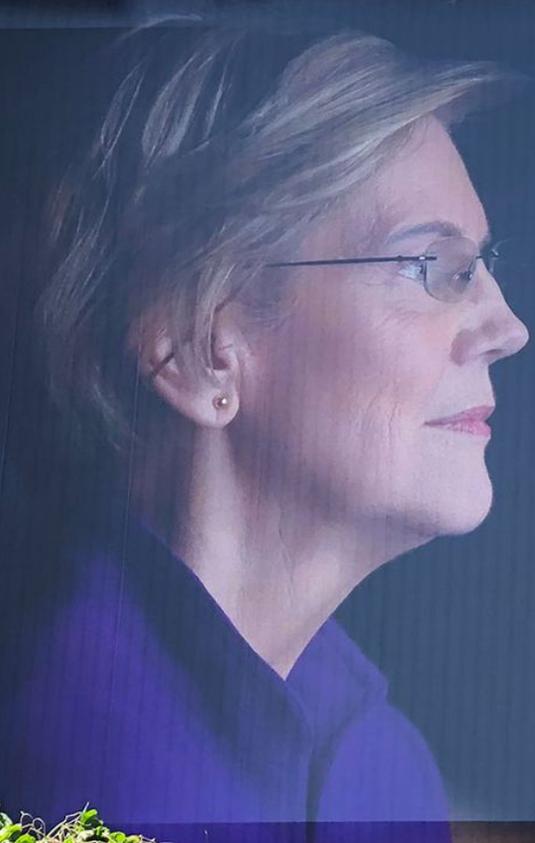
Web traffic share by region



LA / OC / SF

HDR CLIP.COM

#503688



WARREN

BREAK UP BIG TECH

TEXT TECH TO 24477
TO JOIN OUR FIGHT

BY PROVIDING YOUR CELL PHONE NUMBER YOU CONSENT TO RECEIVE PERIODIC CAMPAIGN UPDATES FROM WARREN FOR PRESIDENT.
TEXT HELP FOR HELP, STOP TO END. MSG & DATA RATES MAY APPLY. [HTTPS://WWW.ELIZABETHWARREN.COM/PRIVACY-POLICY](https://www.elizabethwarren.com/privacy-policy).

 VOTE DEMOCRATIC

PAID FOR BY WARREN FOR PRESIDENT

WARREN HAS A
PLAN
FOR THAT

1446 x 25

POLITICS

Justice Department to Open Broad, New Antitrust Review of Big Tech Companies

Inquiry signals Barr's deep interest in tech sector, poses threat to companies such as Facebook, Google, Amazon, Apple

By *Brent Kendall*

Updated July 23, 2019 5:34 pm ET

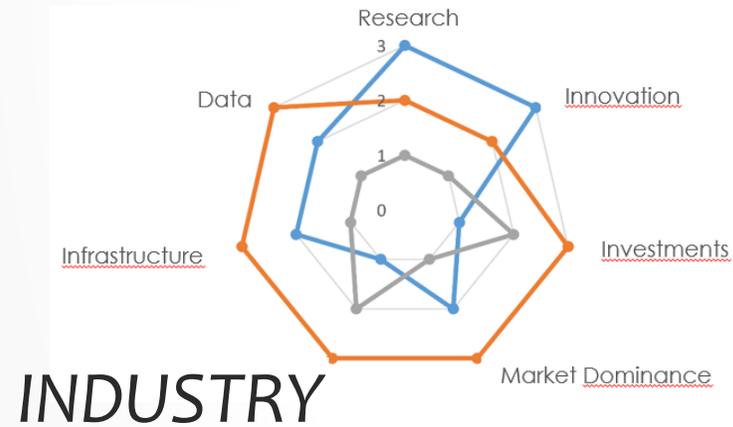


THE WALL STREET JOURNAL.

Margrethe Vestager scares the tech giants.

Trump says the competition commissioner hates the US, but what she really hates is tax avoidance





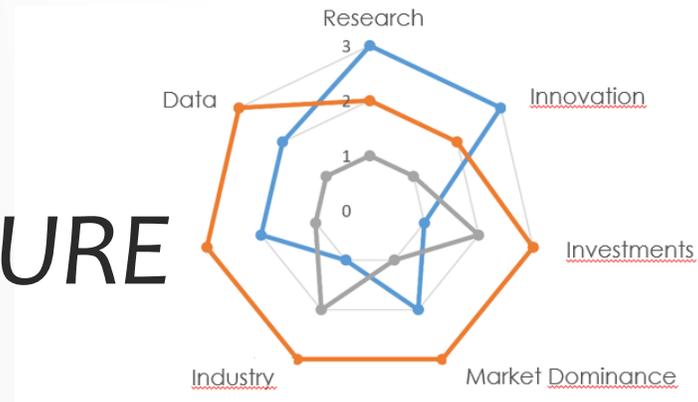
- Market capitalization / Market revenues
- Number and size of LT companies

“Like Google, Microsoft, Baidu, and Yandex, these Internet giants do not primarily make money from the MT technology itself but their investments and offerings have of course a disruptive and innovative effect on the MT market.” (TAUS, 2017, p. 27)

Table 32 Top 20 global companies by market capitalisation and their MT activities, as of March 31, 2018

COMPANY NAME	NATIONALITY	INDUSTRY	MARKET CAP 2018 (\$B)	IN-HOUSE MT
Apple	United States	Technology	851	MT
Alphabet	United States	Technology	719	MT
Microsoft	United States	Technology	703	MT
Amazon	United States	Consumer Services	701	MT
Tencent	China	Technology	496	MT
Berkshire Hathaway	United States	Financials	492	
Alibaba	China	Consumer Services	470	MT
Facebook	United States	Technology	464	MT
JPMorgan Chase	United States	Financials	375	
Johnson & Johnson	United States	Health Care	344	
ICBC	China	Financials	336	
Exxon Mobil	United States	Oil & G a s	316	
Bank of America	United States	Financials	307	
Samsung Electronics	South Korea	Consumer Services	298	MT
Walmart	United States	Consumer Services	264	
Royal Dutch Shell	United Kingdom	Oil & G a s	263	
China Construction Bank	China	Financials	259	
Wells Fargo	United States	Financials	256	
Nestle	Switzerland	Consumer Goods	246	
Visa	United States	Financials	246	

INFRASTRUCTURE



- Availability of technical infrastructure
- Access to infrastructure

Europe needs \$800 billion investment in digital infrastructure

3 MIN READ

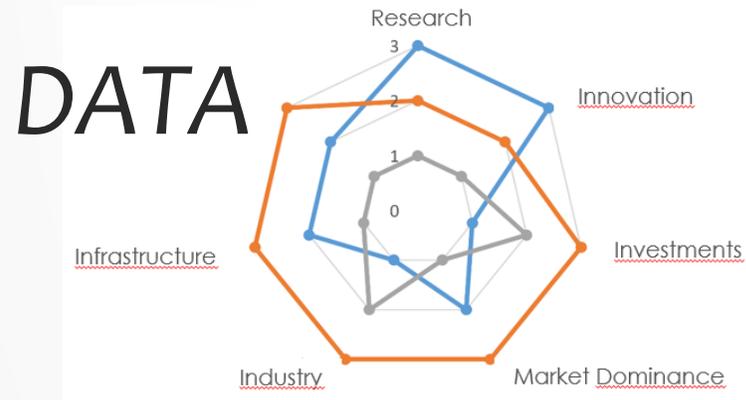


ALPBACH, Austria (Reuters) - Europe needs to invest close to \$800 billion in its digital infrastructure to catch up with the United States and China, the European Union commissioner responsible for the issue said on Thursday.



Table 40 Top 20 countries by Network Readiness Index (NRI)

Rank	Country	NRI	Region
1	Singapore	6	Asia
2	Finland	6	Europe
3	Sweden	5.8	Europe
4	Norway	5.8	Europe
5	United States	5.8	North America
6	Netherlands	5.8	Europe
7	Switzerland	5.8	Europe
8	United Kingdom	5.7	Europe
9	Luxembourg	5.7	Europe
10	Japan	5.6	Asia
11	Denmark	5.6	Europe
12	Hong Kong SAR	5.6	Asia
13	Korea, Rep.	5.6	Asia
14	Canada	5.6	North America
15	Germany	5.6	Europe
16	Iceland	5.5	Europe
17	New Zealand	5.5	Pacific
18	Australia	5.5	Pacific
19	Taiwan, China	5.5	Asia
20	Austria	5.4	Europe



- Availability of open data
 - EU institutions, European Open Data Portal
 - Large national corpora (Bulgaria, Croatia, Slovenia, ..)
 - Corpora developed by European universities (Opus corpora, ..)
 - ELRC
- Legal regulation of data usage
 - GDPR regulation is more restrictive
- Access to proprietary data resources
 - Global dominance allows big tech to harvest unmatched amounts of data

Smart talking: are our devices threatening our privacy?



TASK 2. COMPETITIVENESS ANALYSIS

	Europe	North America	Asia
Machine Translation	14	19	9
Speech Technology	13	20	9
Search Technology	13	19	10
TOTAL	40	58	28

TASK 3. DEMAND SIDE ANALYSIS WITHIN PUBLIC SECTOR

- **Objectives**

- Analysis of LT-based Services and Solutions CURRENTLY in use in Public Administration in the EU
- Identification of needs and future requirements

- **Approach**

- Identify all different domains of public sector
- Obtain results through online questionnaire

TASK 3. DEMAND SIDE ANALYSIS WITHIN PUBLIC SECTOR

Results based on 80 respondents

- Low current demand for LT solutions
- Translation Technology is most frequently used technology...
- .. followed by OCR, Speech and Search
- Major vendors cited are US based (exception: CAT tools)
- Low collaboration with academia
- Optimistic views on new technologies to be incorporated from 2020 and beyond

TASK 4. IDENTIFICATION OF THE VALUE PROPOSITION

- **Objectives**

- Identify value proposition of CEF eTranslation
- Assess potential avenues (development, sustainability, ...)

- **Approach**

- Business model canvas
- Input from Tasks 1, 2 and 3
- Input from focused meetings with DSIs

TASK 4. IDENTIFICATION OF THE VALUE PROPOSITION

Today's situation

- Services offered :
 - Provide an (asynchronous) MT service: secure & integrated/stand-alone
 - Provide a customer Service Desk
 - Provide a language resource repository (ELRC-SHARE)
- Focus on DSIs and Public Administrations – help them make their service/content multilingual
- Key partners: DGT, DIGIT
- Fixed cost structure
- No revenue streams

TASK 4. IDENTIFICATION OF THE VALUE PROPOSITION

Potential avenues

- Extend current business model
 - Scale up service / Offer real-time translations
 - Customize MT with special attention to under-resourced languages
 - Offer LT services : NER/Anonymization, Quality Estimation, ..
- Focus on Governance
 - Provide data for training models : Euramis !!
 - Extra promotion through CEF Telecom calls
 - Provide MT models / LT models (API access)

CONCLUSIONS & RECOMMENDATIONS

- Market is disrupted by dominant global players
- MT market evolved in non-core environment
 - Google, Baidu, Yandex : MT to sell ads
 - eBay, Amazon, Alibaba : MT to drive e-commerce
 - SDL : MT to drive localization business
- Significant areas of market deficiency
 - Quality gap for smaller and complex languages
 - Domain and application specific MT
 - Security and privacy concerns

CONCLUSIONS & RECOMMENDATIONS

- Europe is strong in research & innovation, but not successful to scale innovations and capture the market
- Europe needs European alternatives to fill the gaps/deficiencies and to avoid reliance from market monopolies
- Multilingual DSM should be developed on own infrastructure
- Public procurement can be the major driver for European LT industry to avoid dependence on existing monopolies

CONCLUSIONS & RECOMMENDATIONS

- Review of IPR regulation of data (fair use of copyrighted data)
- Initiatives needed to avoid brain drain
- Closer collaboration between Industry - EC

We listened and worked hard to prepare the launch of the LTI Cloud project at our upcoming Summit.

The *LTI Cloud* is a SaaS wrapper around LT components with the goal of facilitating the plug 'n' play of language technology. It will allow you to market your capabilities, to easily build new LT components based on others... and it will be a great new customer acquisition channel.

We intend it to become a single point of access for companies wishing to implement language technology, empowering people and businesses by simplifying digital language technology solutions in Europe and globally. So far, we have architected the *LTI Cloud*. Influential politicians love the idea and will come to the LTI Summit to support us.

We would like to share this project with you. It is, therefore, very important that you come to the LT-Innovate Summit on 25-26 June in Brussels to show your interest and support. Only together can we push the project ahead and launch soon the first version of the *LTI Cloud*!

We count on you and are looking forward to seeing you soon in Brussels,



Jochen Hummel
CEO, ESteam and
Chairman, LT-Innovate



Robert E. Etches
CIO, TextMinded



Luc Meertens
CEO, CrossLang



Christoph Prinz
CEO, SailLabs

The LTI Cloud is the one-stop-shop platform for making available, discovering, assembling, testing and prototyping language technology components.



Knowledge Management



Natural Language Processing



Search



Translation

Welcome to the Catalogue of CEF eTranslation Services.

Multilingualism has become an essential part of our lives, which is why enabling seamless cross-border communication has never been more important. Therefore, the catalogue of services includes a wide range of tools and services which aim to help you overcome language barriers and to support your multilingual needs. Browse through our catalogue, explore the language tools (LT) and services offered by [CEF eTranslation](#) and discover other potentially useful third-party services* which can facilitate your multilingual operations.

Explore the services provided by CEF eTranslation [here](#)

Explore third-party services [here](#)

Browse the [ELRC-SHARE repository](#) for Natural Language Processing tools/services

Results 8 (1/1)

Filter By:

Service Function

- Translation (4)
- Language Technologies (3)

MORE

Intended User

- Any third party interested in training MT engines (1)
- CEF Automated Translatio (1)

MORE

Language Coverage

- English (6)
- French (5)

MORE

Service Provider

- Administrative Modernizaton Agency (1)
- Austrian Presidency of the Council of the European Union (1)

MORE

Service Ownership Status

- External (8)

Clear Filters

Translator of the Bulgarian Presidency of the Council of the EU

Function: Translation

Task: Machine Translation

By Ministry of the Presidency, Bulgaria

Translator of the Austrian Presidency of the Council of the EU

Function: Translation

Task: Machine Translation

By Austrian Presidency of the Council of the European Union

Latvian state administration language technology platform

Function: Language Technologies

Task: Speech Recognition , Text To Speech Synthesis , Machine Translation

By Culture Information Systems Centre, Latvia

ParaCrawl OpenSource Pipeline

Function: Language Technologies

Task: Crawling , Alignment

By Universitat d'Alacant

PLaTa: Spanish state administration Machine Translation Platform

Function: Translation

Task: Machine Translation

By Ministerio de Política Territorial y Función Pública, Secretaría de Estado de Función Pública, Secretaría General de Administración Digital

eTradução - Repositório Nacional de Recursos de Tradução

Function: Data And Data Management

By Administrative Modernizaton Agency



Cyril Hlakkache

@hlakkache



Als antwoord op [@hlakkache](#) en [@EVANVAERENBERGH](#)

L'UE est sur la voie et je pense que [@olesovhcom](#) (si j'ai bien interprété un podcast) a touché du doigt le bon concept qui pourrait nous être destiné : créer des alliances d'acteurs experts EU versus créer des Titans (GAFA) afin de préserver la privacy propre à la culture EU.



**EUROPEAN
LANGUAGE
GRID**

[About](#) [Grid](#) [Community](#) [Open Calls](#) [Events](#) ▼ [News](#) [Consortium](#) [Contact](#)

Towards the Primary Platform for Language Technologies in Europe





THANK YOU!