

# TAUS – *NUNC EST TEMPUS* – MAPPING A FAST-CHANGING INDUSTRY

Andrew Joscelyne

OpenNMT Workshop Paris, March 2018



# Headlines 2017



**Nearly Indistinguishable From Human Translation — Google Claims Breakthrough** (*September 26, 2016*)

**Systran Goes Neural, Lets You Try It** (*October 18, 2016*)

**Baidu Sells Human Translations at Rock-Bottom Rates to Gain Edge in AI** (*March 16, 2017*)

**A Convolutional Encoder Model for Neural Machine Translation - @ Facebook** (*July 25, 2017*)

**How SDL's Secure Neural Machine Translation Helps Enterprises Reduce Risk** (*August 14, 2017*)

**Sogou Invests USD 6-7m in Shanghai-Based UTH in Scramble for Translation Data** (*August 22, 2017*)

**Linguee's Founder Launches DeepL in Attempt to Challenge Google Translate** (*August 30, 2017*)

**Amazon Pits Neural Machine Translation Framework Against Google, Facebook, Others**  
(*December 20, 2017*)

January 1954: “In Five Years... FAHQT”



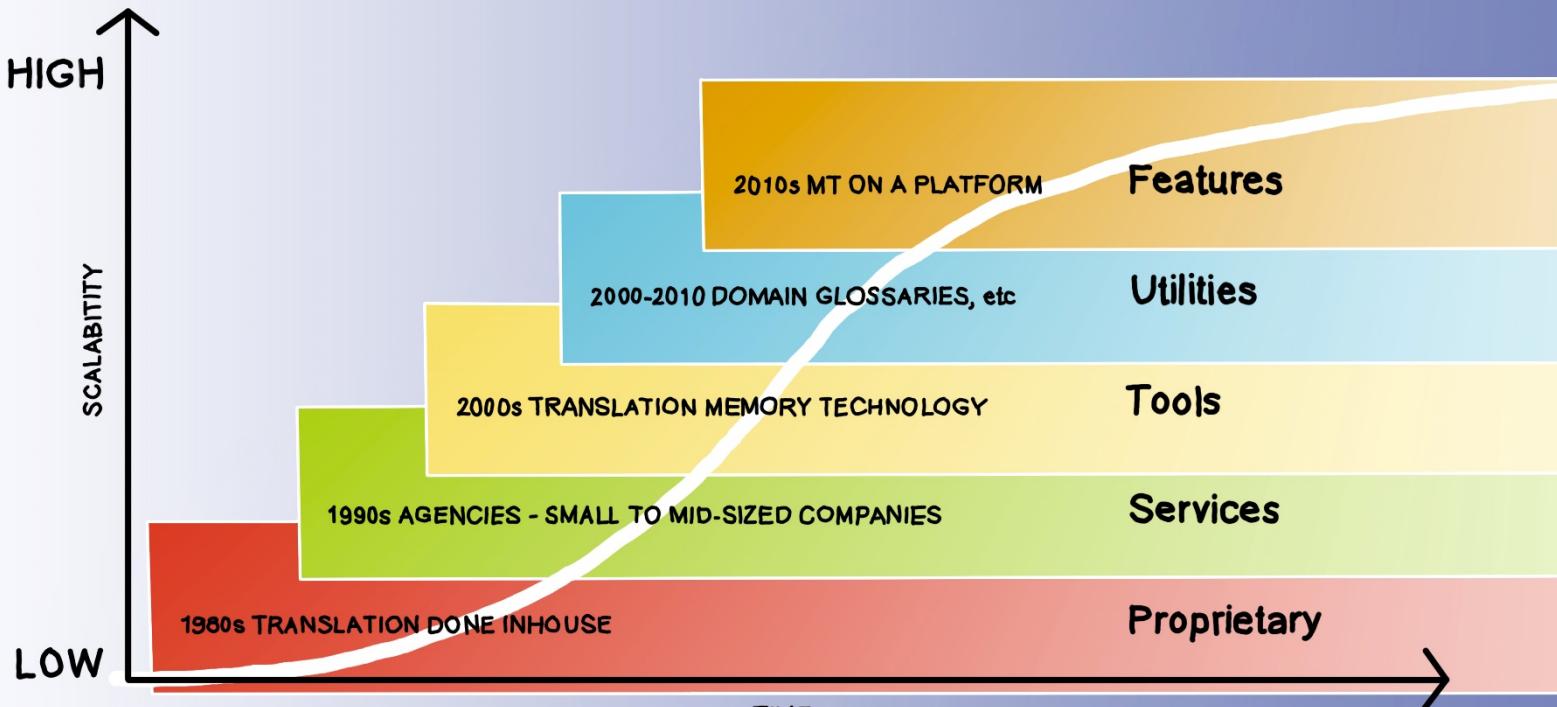
This slide may not be used or copied without permission from TAUS

# The S-Curve “is stronger than all of us...”

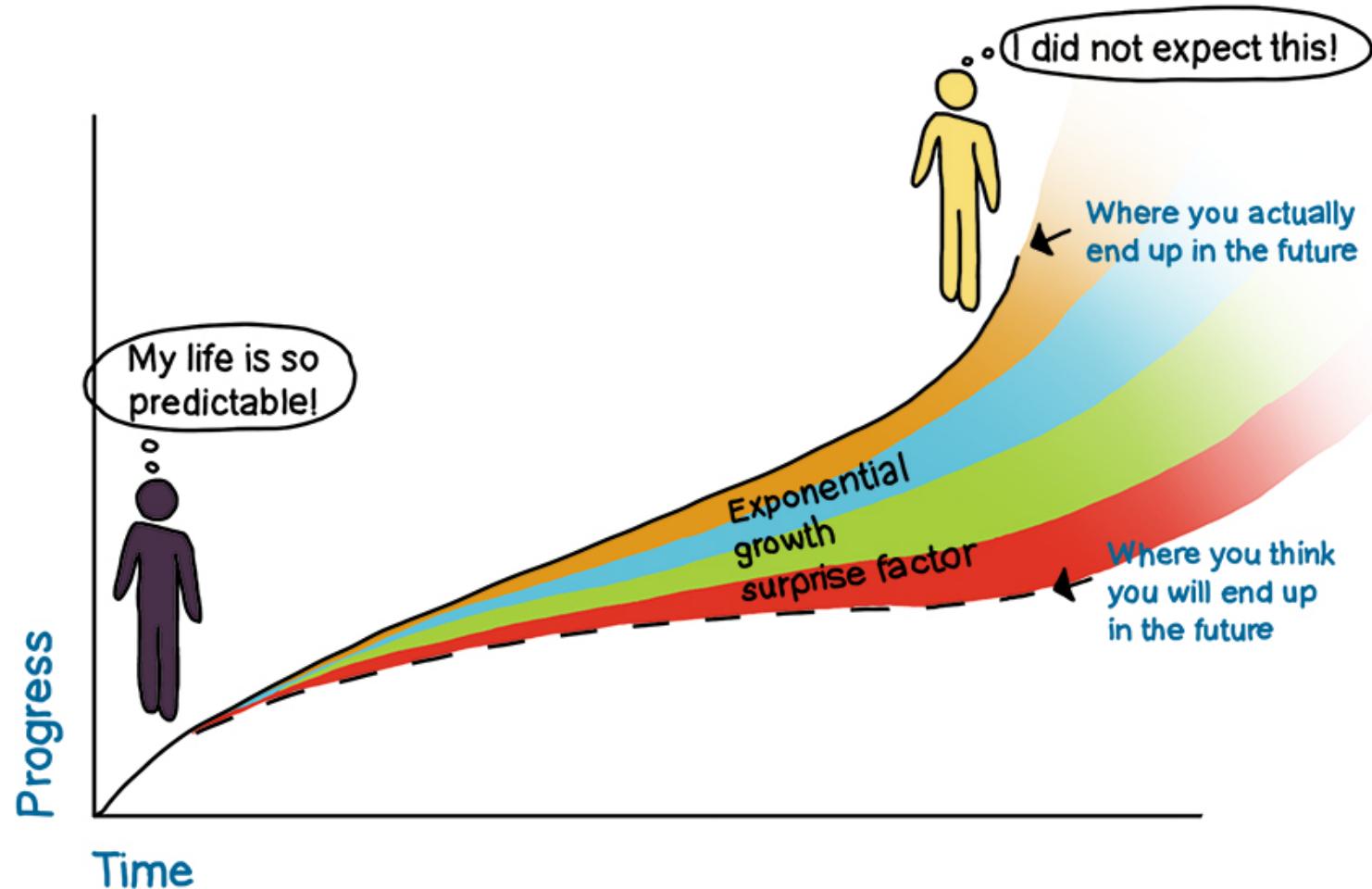


## The S-Curve Rules Technology:

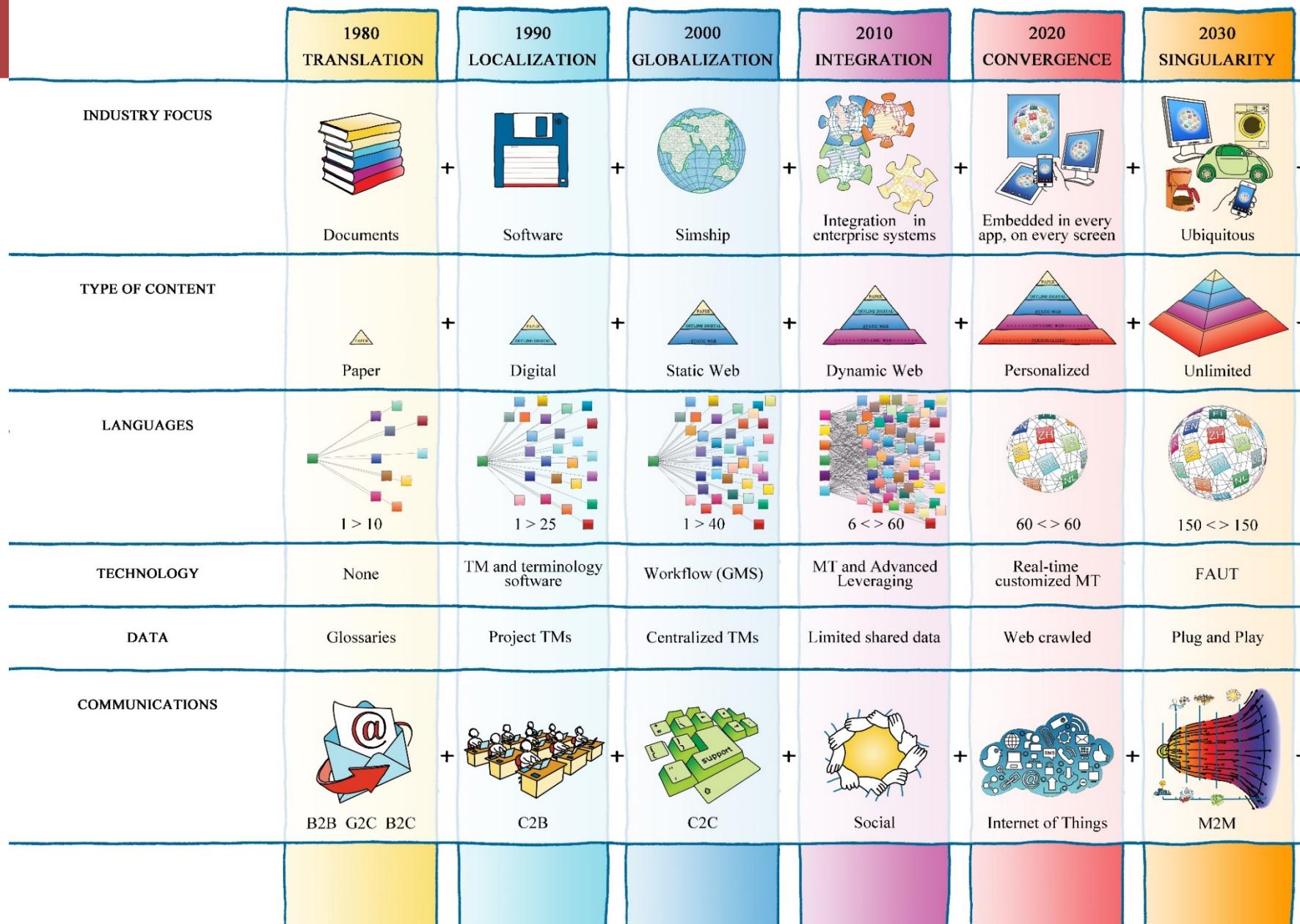
Adoption In The Translation Industry; Value Migrates as Technology infrastructures evolve.



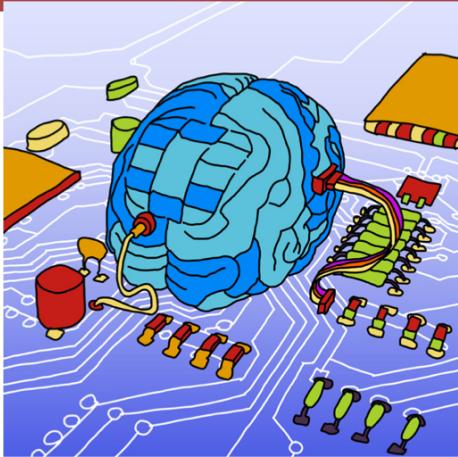
# Law of Accelerating Returns



# Evolution of the Translation Technology Landscape



# The Neural Wave



## Neural MT

- Five years of NMT = twenty years of SMT
- 5,000 lines of code – 200,000 lines of code
- Zero shot translation
- Thousand(s) decisions taken by machines
- Every week we see improvements in quality
- Enables ambient intelligence



**“The progress in what we can call the consumer translation tech industry is of much larger relevance than the classic corporate, outbound targeted, translation business.”**

*Chris Wendt: Going Vocal with Ambient Translation, Nunc Est Tempus, December, 2017*

# Democratization of Algorithms



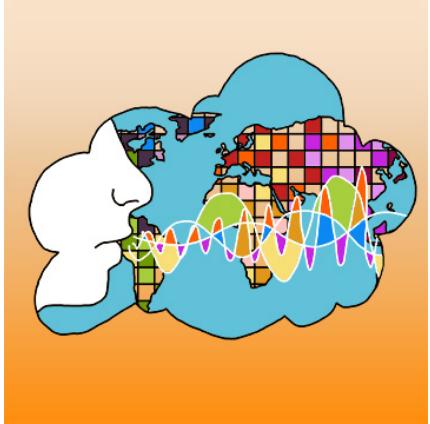
## Open source algorithms

- From SPSS and SAS Institute to TensorFlow, Open NMT, MXNet, Nematus, ConvNet
- 4 million R and Python users vs. 600 million Excel users
- The democratization of AI has only just begun

**“Once you have found your best service, you are less likely to switch. That’s how the circle of open algorithms and customer lock-in gets closed.”**

*Nunc Est Tempus, December, 2017*

# Follow the Data



**“Our language data travel freely across all borders to economies driven by free enterprises and long-term big data strategies.”**

*Nunc Est Tempus, December, 2017*



**“We are moving from a people-driven to a data-driven industry.”**

*Adolfo Hernandez, Nunc Est Tempus, December, 2017*



**“Cross-language big data inspires our future.”**

*Eric Yu, CEO of GTCOM, Nunc Est Tempus, December, 2017*

# Not Business as Usual



## Parallels to advertising, travel and banking industries:

- Standardization and automation lead to mass market adoption
- Ubiquitous, *always-on* translation triggers new business and pricing models
- The ease and omnipresence of translation will be the foundation for a plethora of new business opportunities.
- *Razor-sharp* localization



**“Fifty percent of the services we produce today, we were not doing five years ago, and they are not translation.”**

*Smith Yewell: In the Business of Global Talent Search, Nunc Est Tempus, December, 2017*

# Going Vocal



## Ambient intelligence

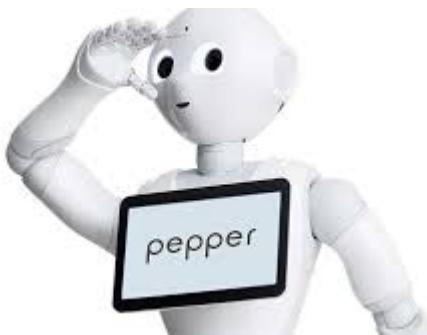
- Our environment automatically adjusts to the language and best means of communication.
- Spoken translation in wearable technology, robots, smart home devices
- Confluence of technologies, business models and markets



**“Machine Translation will become more acceptable not because it miraculously becomes perfect, but because it can act as a utility for people with reasonable expectations.”**

*Chris Wendt: Going Vocal with Ambient Translation, Nunc Est Tempus, December, 2017*

# Going Vocal

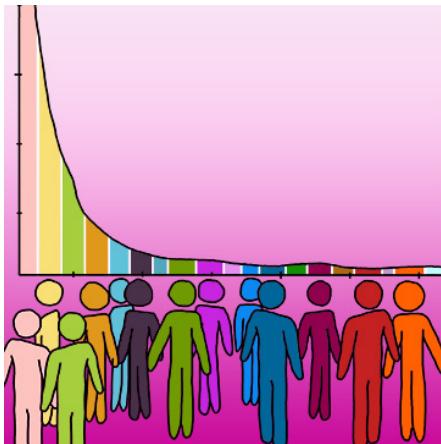


Megaphonyaku is translation service that uses a megaphone-shaped translator to automatically translate Japanese into English, Chinese, and Korean. The device is programmed with 300 preset phrases and can generate 1,800 patterns of phrases

Image: Panasonic



# The Long Tail of Languages



## Geolinguistic shifts

- We are in a semi-state of globalization: 30% international trade
- China One-Belt One-Road Initiative (OBORI)
- China Next Generation AI Development Plan
- Zero shot translations adds MT for more languages
- Stimulates interest in new cultures and languages

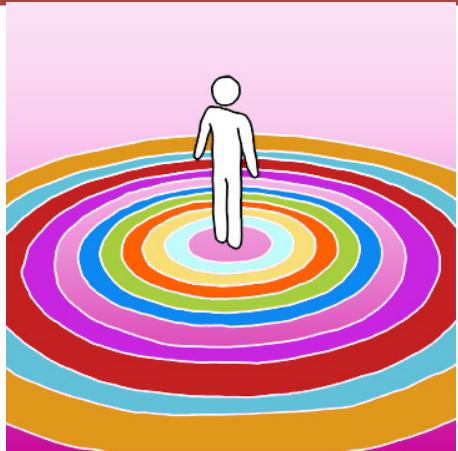
## Chinese Government Going All in to Back Language Industry

(Slator, January 10, 2018)



**“The web, in the mentality of the Chinese, has always been a key tool to discover the world outside.”**

*Eric Liu: Translation World Shifting from English-Centric to Chinese-Centric, Nunc Est Tempus, December, 2017*



## Boutique style translations:

- Don't be blinded by technology
- 'Dumb translations' are not enough
- Stories need to be told and imagination needs to be triggered

# Reading the Charts



**Dashboards become commonplace:**

- Can't manage, if you can't measure
- ROI on a per segment basis
- Data help to decide where to invest

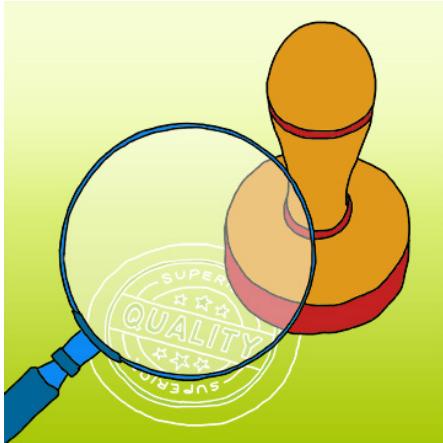
**"Localization on the whole has been more of a shotgun approach until now for most companies."**

*Nunc Est Tempus, December, 2017*

**"Using our dashboards we will be able to switch to a razor-sharp approach to localization."**

*Nunc Est Tempus, December, 2017*

# Making Translation Invisible



## Towards Self-Driving Autonomous Workflows:

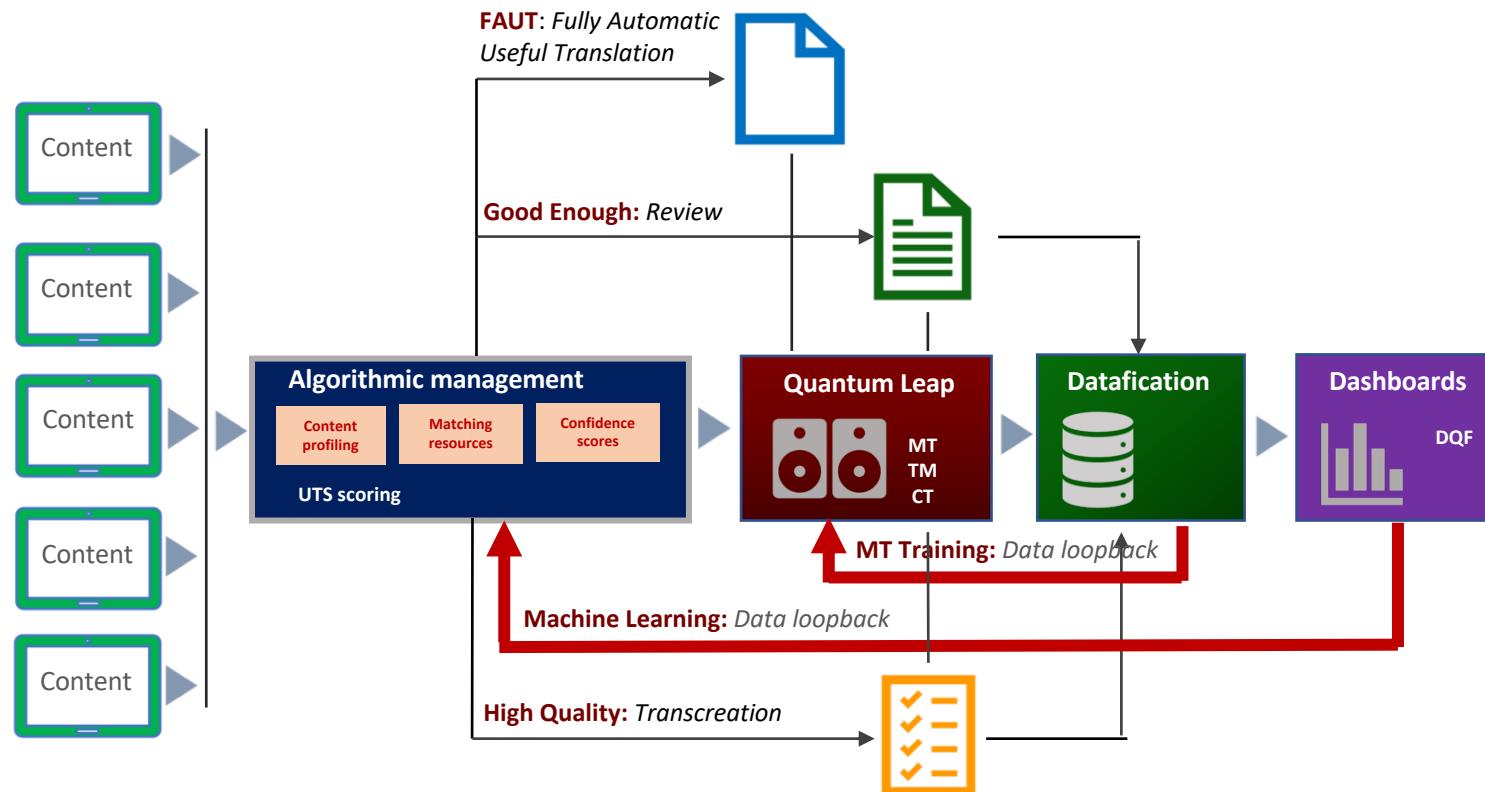
- Automate dozens of human touchpoints in a multi-layer multi-company supply chain
- Microservices and SOA have helped solve the challenges of interoperability
- Resolve diverging issues in localization: quality levels, nomenclature, metrics, how we measure



**"In the past, the keyword was first cost, then quality, then efficiency.  
Today it's *agility*."**

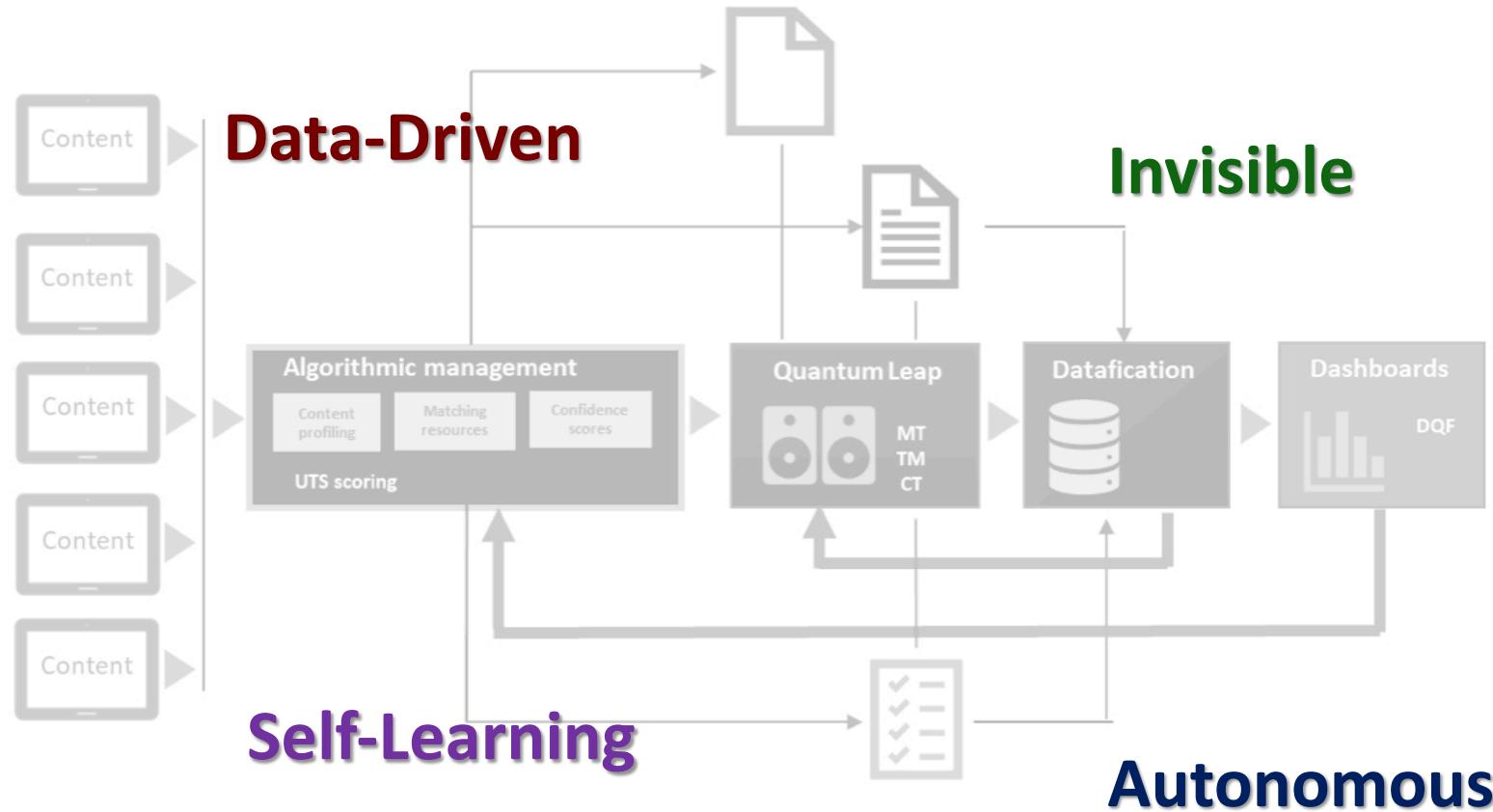
*Aiman Copty: Making Translation Invisible. Nunc Est Tempus, December, 2017*

# The Modern Translation Pipeline

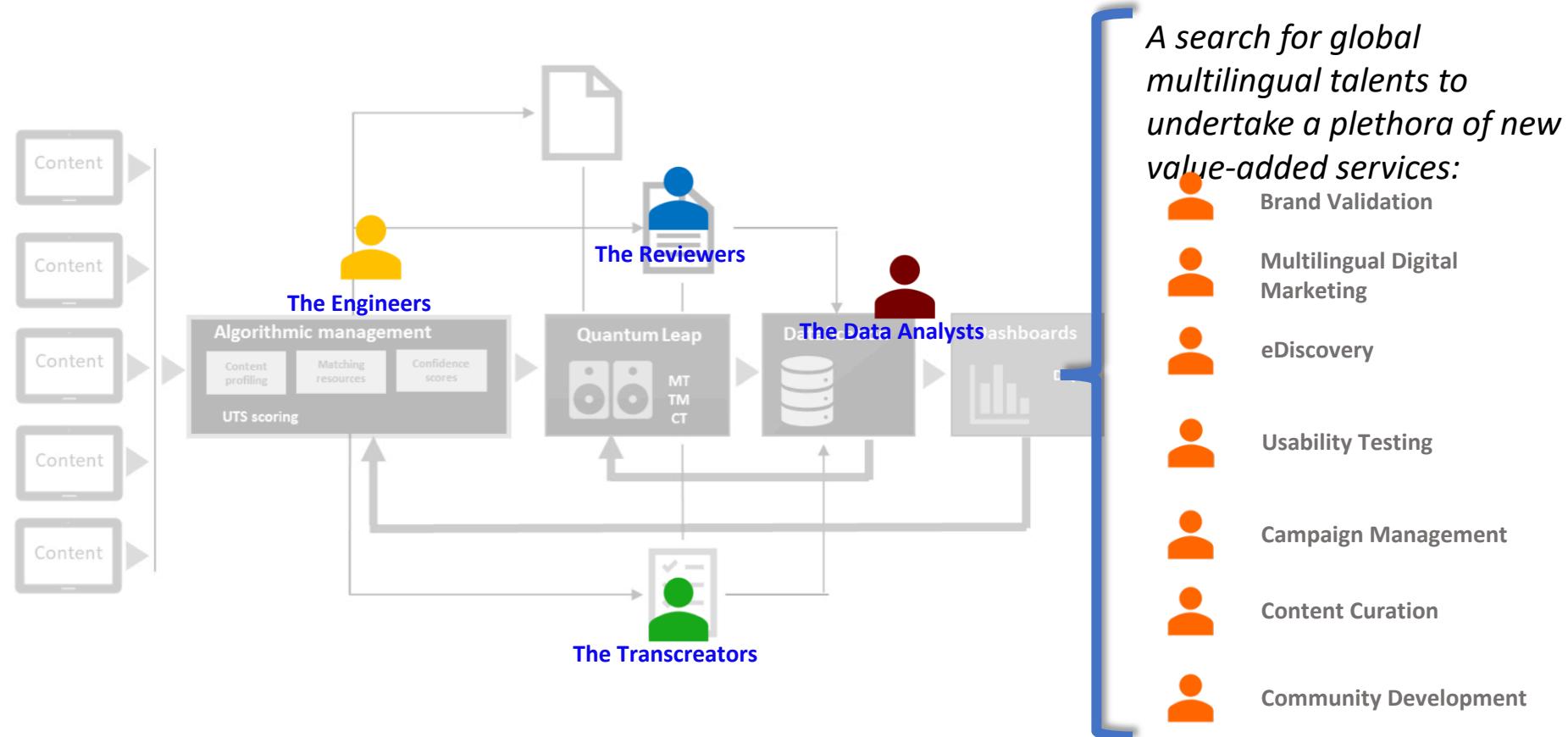


# The Modern Translation Pipeline ...

1. made great progress in interoperability
2. is adopting DQF as a quality standard



# ... and where are the Jobs?!



# Quality Management Becomes So Vital

1. Jobs of post-editor/translator, quality evaluator, editor and **reviewer** are blurring
2. They feed the data that train the translation pipeline, to **fork content** in three levels:
  - FAUT
  - Good enough
  - High Quality
3. Their annotation and corrections train the **machine translation engines**.



# Key Messages of Nunc Est Tempus:



1. The consumer is driving the change
2. Democratization of algorithms has only just begun
3. Follow the data
4. Not business as usual
5. Going vocal
6. Going vertical
7. Geolinguistic shifts
8. Quality DOES matter
9. Reading the charts
10. Making translation invisible

➤ **And now is the time!**



1954

2017

2022

## *Key Take-Aways*

1. Change – jobs, processes, technology
2. Control – measure, dashboard
3. Analyze – use your data

# TAUS Customer Experience Journey



We guide adoption of technologies and data-driven processes, normalized quality metrics and future-proof business models.

