

Machine Translation is Not One Size Fits All



Tweet this: [@lexworksMT](#)



Email me: Lori@LexWorks.com



Connect: [linkedin.com/in/lorithicke](https://www.linkedin.com/in/lorithicke)

Contents

- What is technology agnostic?
- Who is technology agnostic?
- Why is it important?
- Best practices for post-editors
- Rules of thumb for RBMT, SMT & Hybrid
- Case studies

First, a word about LexWorks

- MT services branch of Paris-based Lexcelera
- “T-shaped” expertise
- MT services provider
- Consulting, metrics, engine training, post-editing
- Technology agnostic



Current R&D projects

AC+OCEPT

- EU funded, Moses
- Developing new technologies to help MT work better with community content

acrolinx 

 Symantec™  Lexcelera

 UNIVERSITÉ
DE GENÈVE



TRANSLATORS
WITHOUT BORDERS



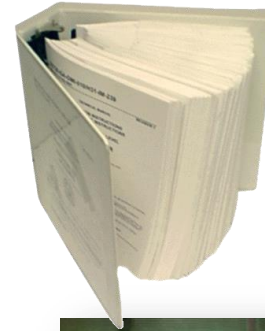
content rules®
the global content experts™

- Impact of pre-editing on post-editing productivity

acrolinx 

 Lexcelera

We already match the tool to the task

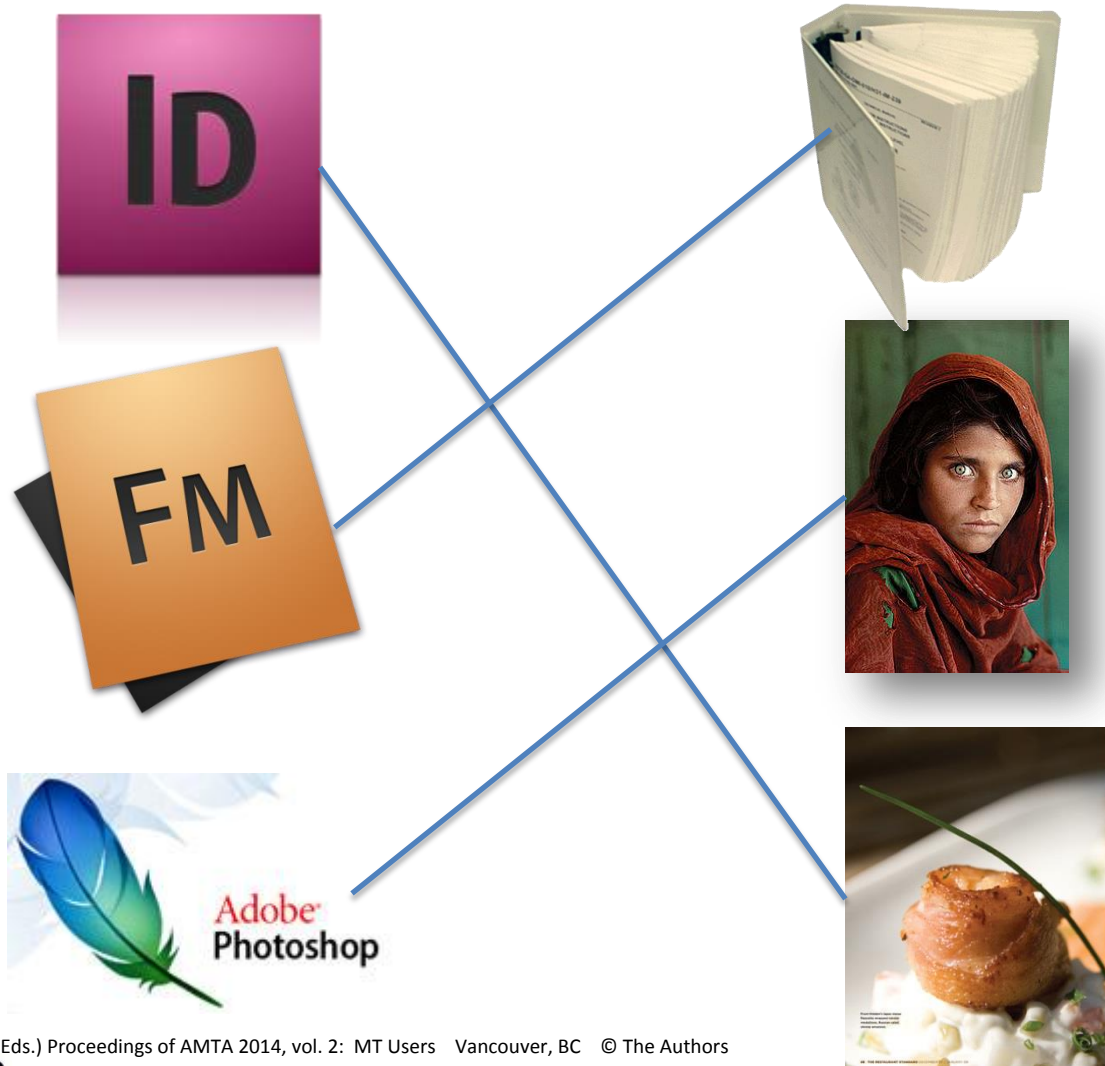


A PASSION FOR FOOD

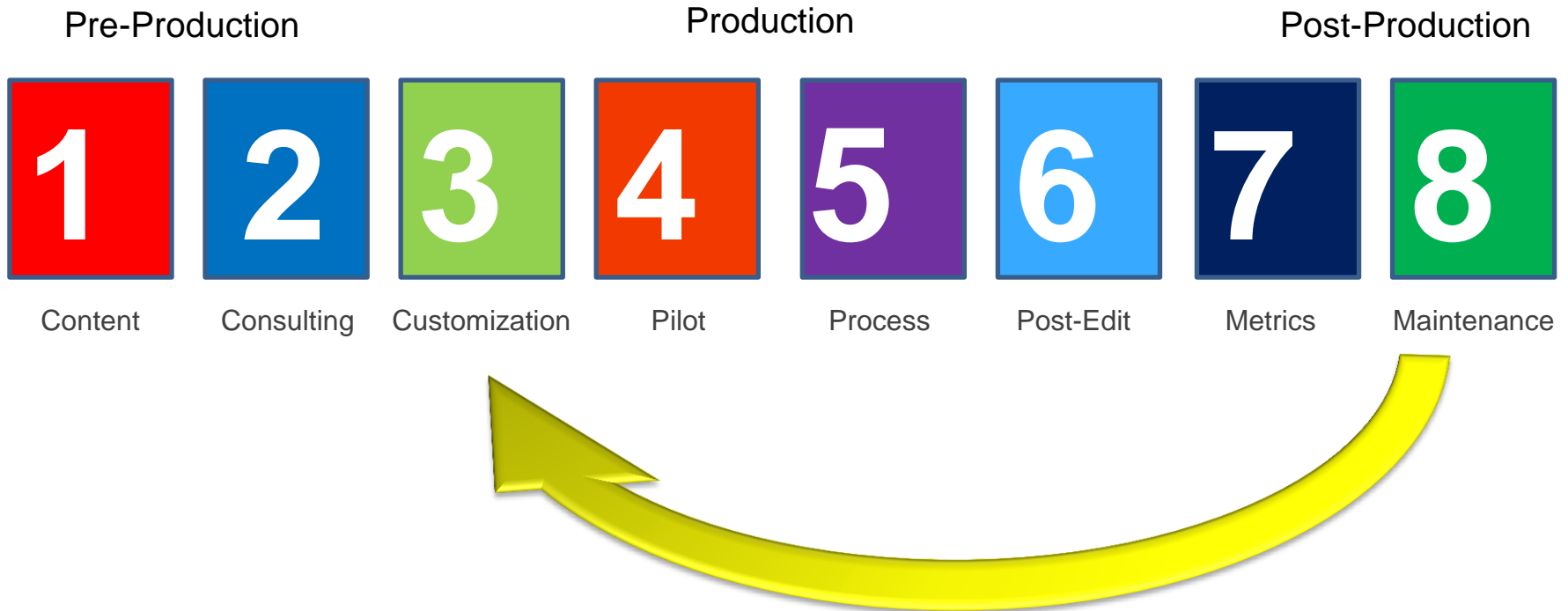
Modern restaurant cuisines: Four cuisines, two cultures and one restaurateur's love of the business

Modern restaurant cuisines: Four cuisines, two cultures and one restaurateur's love of the business

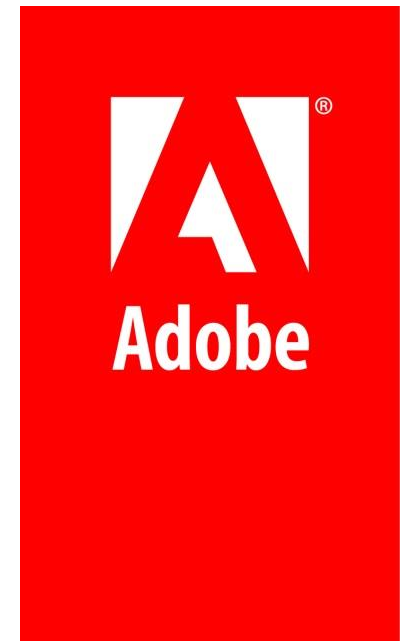
We already match the tool to the task



MT is more than a tool. MT is a process.



Other technology agnostics



A Best-of-Breed approach

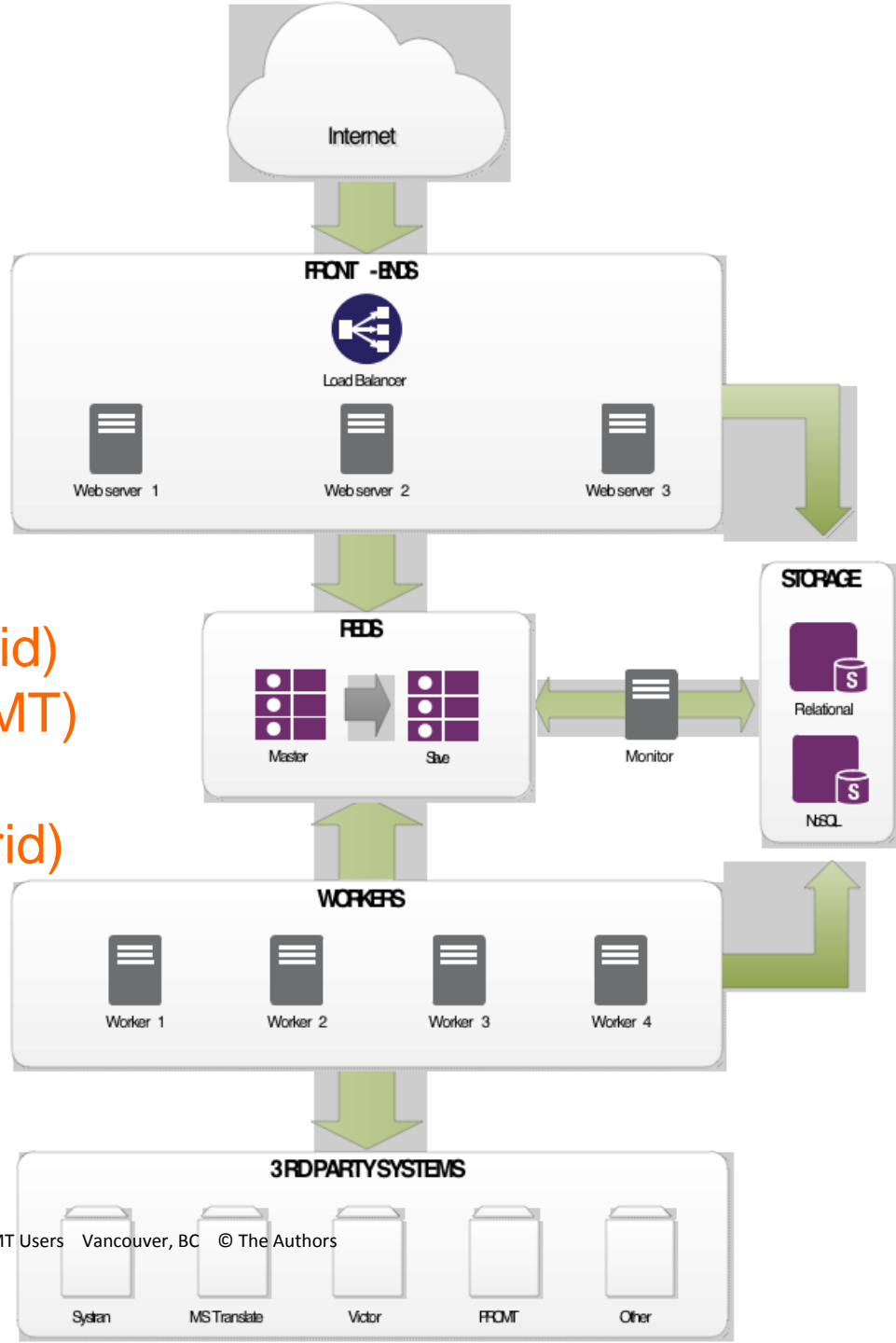
“A good MT strategy should be technology-agnostic and look for the most efficient solution on a case-by-case basis.”



Rubén Rodríguez de la Fuente, PayPal



Engines:
Systran (RBMT/Hybrid)
MSFT Hub (Online SMT)
Victor (SMT)
PROMT (RBMT/Hybrid)
Other

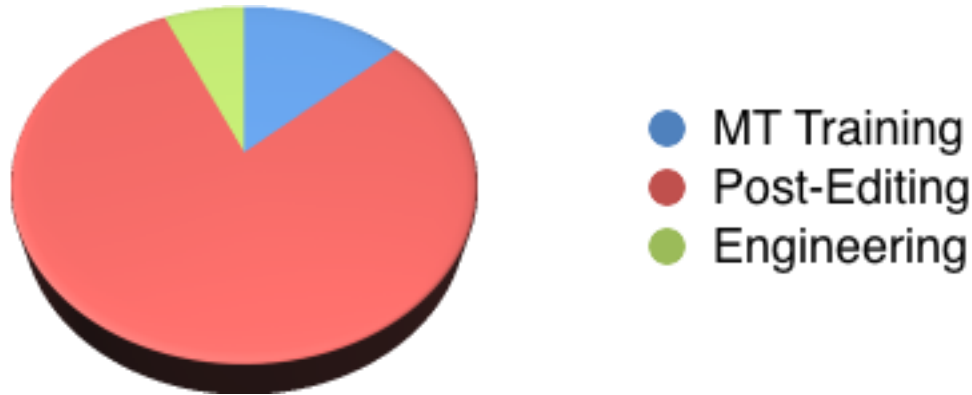


Post-editors really hate bad MT (and why we should care)

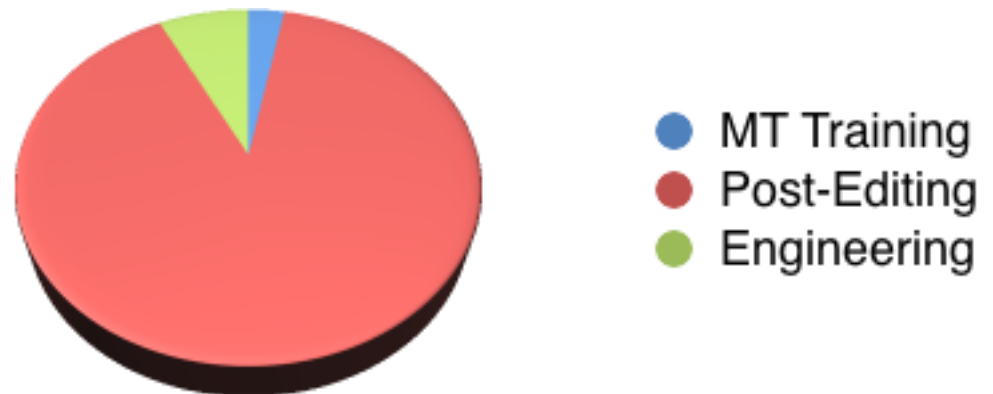


Post-editing as a share of total costs

Initial MT Project



Subsequent MT Project



7 best practices for post-editors

1. Don't give post-editors bad MT & expect them to fix it – for a discount
2. Choose the right engine for the content
3. Train it well
4. Test the output: retrain if necessary
5. Set post-editing expectations: what type of errors to look for and what quality is needed
6. Let post-editors be involved in improving the engine
7. Integrate changes quickly

4 Types: RBMT, SMT, Hybrid & Online SMT



BNP PARIBAS | The bank for a changing world

Français | Our local websites | Online banking | Sitemap

BNP Paribas 49.500 € +0.182 %
12-09-2013 - 09:48 (Paris time)



About us

Responsible Bank

BNP Paribas Foundation

News & Press

Investors & Shareholders

Careers

Type your search

Translate

ABOUT US

The bank for a changing world

Home > About us > Corporate Culture

Share:     

Corporate Governance

Compliance

Core Businesses

Our Corporate Social Responsibility

Corporate Culture

> History

> Core Values

> Management Principles

> Risk Management

> Spirit of Innovation

> Diversity

MANAGEMENT VALUES AND PRINCIPLES: CULTURAL FOUNDATIONS OF THE BNP PARIBAS GROUP

Over the past ten years, the BNP Paribas Group has embarked on a series of major transformations.

The first stage was the creation, in 2000, of BNP Paribas. The merger between BNP and Paribas was an industrial project of enormous scale, as well as a tremendous adventure in human terms. The new Group believed that in order to make this corporate effort a collective success, it was essential to base it on [strong values: Responsiveness, Creativity, Commitment and Ambition.](#)

These values are now shared by the 180,000 employees of the Group. They are a reflection of our employees' commitment to customer service.

In 2006, the adventure continued with the integration of BNL bc. In 2009, Fortis and BGL joined BNP Paribas, strengthening its presence in Europe. The size of the workforce more than doubled, with nearly 170 nationalities represented in the 75 countries in which the Group operates. Furthermore, this new dimension to BNP Paribas resulted in higher customer expectations.

For all these reasons, BNP Paribas has decided to incorporate a number of principles into its daily management practices in order to clarify its management culture: Customer Focus, Risk-Aware Entrepreneurship, Valuing People and Leading by Example.

Trained Online SMT

Rules of thumb: content type

	Online SMT	Hybrid	RBMT	SMT
Documentation, reports, online help, UI		✓	✓	
FAQs, forums, UGC	✓			✓
Patents, other broad domain	✓			✓
Marketing materials				
Insufficient in-domain/out-of-domain data	✓	✓	✓	
Poor grammar, spelling	✓			✓

Rules of thumb: languages

	Online SMT	Hybrid	RBMT	SMT
French, Spanish, Italian	✓	✓	✓	✓
Russian, Japanese, German		✓	✓	
Norwegian, Danish, Thai	✓			✓

A comparison of features

Area	Feature	RBMT	SMT
Capability	Number of languages handled out of the box	~20	~50
Capability	Ability to add language pairs (if training data available)		✓
Cost	Free or Open Source version exists (e.g. Apertium, Moses)	✓	✓
Cost	SaaS (Software-As-A-Service) models exist	✓	✓
Quality	Output is fluent		✓
Quality	Output is predictable	✓	
Quality	Pre-editing significantly improves output quality	✓	
Quality	Can handle poor grammar or spelling		✓
Quality	Uses specified terminology applying correct grammar	✓	
Quality	Handles software tags without special programming	✓	
Quality	Can be integrated with TM (Translation Memory) tools	✓	✓

Area	Feature	RBMT	SMT
Suitability	Better performance with online chat		✓
Suitability	Better performance with UGC and broad-domain content (e.g. patents)		✓
Suitability	Better performance for documentation/UI and other narrow domain content	✓	
Suitability	Suited to rare language pairs (where training data is available)		✓
Suitability	Suited to full post-editing/real-time improvement cycles	✓	
Training	Learns automatically		✓
Training	Rapid improvement cycle	✓	
Training	Can be trained by engineers		✓
Training	Can be trained by linguists	✓	
Training	Effective with limited training data	✓	

Testing assumptions is sometimes surprising

Translation engine	Engine Type	BLEU Score	GTM Score (SymEval)
<i>Systran</i>	<i>Hybrid</i>	69.74	72.69
<i>Moses</i>	<i>Statistical</i>	50.46	57.93
<i>Microsoft Translator</i>	<i>Statistical</i>	54.01	60.81

Case study



Alstom Transport Factory, Novochoerkassk Russia

Challenge:

- English to Russian
- 2 pages to 100 pages daily
- Over 3 years
- Technical specifications, contracts
- No bilingual data at project start

Technology Used:

- RBMT



Looking for that Smoking Gun

Challenge:

- Japanese to English
- 30,000 pages, mostly emails, technical reports, meeting minutes
- Poor grammar, colloquial

Technology Used:

- Hybrid



Case study



SNCF, Response to Technical RFP

Challenge:

- French & English to Braz. Portuguese
- 3400 pages, lots of different documents
- Multiple passes on same file
- Tender response
- Limited data at project start

Technology Used:

- Hybrid



Online Forums and FAQs

Challenge:

- 9 languages
- Dynamic content
- Poor grammar, colloquial
- Solve problems before help desk
- Need 24/7 availability

Technology Used:

- Online SMT



Case study



BNP PARIBAS

BNP's Self-Service MT Server

Challenge:

- World's 3rd largest bank
- Centralized MT server for 200,000 employees
- Multiple business units with their own terminology
- Self-Service
- Behind the firewall; we train & update remotely

Technology Used:

- Hybrid



Questions?



Tweet this: [@lexworksMT](#)



Email me: Lori@LexWorks.com



Connect: [linkedin.com/in/lorithicke](https://www.linkedin.com/in/lorithicke)