



## Automated Translation 2008: Science Meets Solution

Mark Tapling  
President & CEO, Language Weaver  
[www.languageweaver.com](http://www.languageweaver.com)



### Questions?

- How is our industry unique as compared to other industries?
  - What's working, what's not?
  - What are the gating variables for growth?
- How are MT requirements different for commercial users, LSPs, government users, and digital content providers?
- How do we collectively as an industry prioritize needs to these markets?
  - What are the business and in technical requirements?



## First Question?

- What “Market” Are We Talking About?
- Opening impression is the market defined by:
  - The intersection of human translators
  - With technology
  - And professional project management
- How you frame the problem influences the perception of available solutions



## Definitions: Assets or Obstacles?

- Translation vs. Localization
- Assimilation vs. Dissemination vs. Communication
- Machine Translations vs. Statistical Translations
- Word vs. Phrase vs. Syntax
- Internationalization vs. Globalization
- Collaboration vs. Publication
- Research vs. Transaction
- Information vs. Technical Documentation



## Observations About our Market

- Our market tends to promote science; as opposed to the solution value
- We lack a tieretsu of industry partnerships
- A large opportunity is gated by human productivity issues
  - LSP's *own mindshare* on market definitions
- LSP market highly fragmented profile
  - Less than 20 vendors over \$20M
- Our customers have been incentivized to leverage vendors against each other
- Our customers have been coached to buy in low unit cost at high volumes (e.g. price per word/project)
  - Creates an incentive to do less as cost control strategy



## What's Different About our Market?

- Nobody is solving the "big problem"
  - "How do I translate everything?"
- Few or no product "solutions" for customers
  - Market to date has been tools and services
  - LSP's deliver solution as a service
- The customer's represent the big market
  - ...and customer's own "the big" problem
  - We have to define and address it for them
- Biggest players today are straddling the customer and translators
  - Razor thin margins
  - Hard to maintain loyalty



## Customers in Our Market

- Segmentation has been on vendor side
  - We've not segmented by customer need to date
- Customer is used to managing lots of vendor relationships
  - Switching costs are low
- Government customers are also starting to look for solutions
- More people want to try than buy
  - Value proposition lacks trust
- Corporate organizations for translations are fragmented in many cases



## Conclusions on Current Market

- Human Translators should be revered as the highest value add contributors in the publication space
  - Liberate human translators for the highest value add work
- Translation is not localization
  - Translation conveys meaning
- Translation is a non-negotiable requirement of global business (enterprise and government)
- *Translation is not a core competency of any end using customer*
- Deploy technology as vehicle to expand market on multiple fronts



## Obstacles

- Proposition lacks “trust”
  - Many believe MT does not work well enough to be commercially deployed
- No MT comparables with a working commercial strategy
- SMT runs on data; and the world has dirty data
- Risk of Indigestion: too many things to consider
- Communications challenges on the correct application use cases for demand generation
- Many human translators feel current MT positioning is a threat
- No successful established business models for fast growth; interesting margins
  - Limited fresh capital being invested by private equity investors



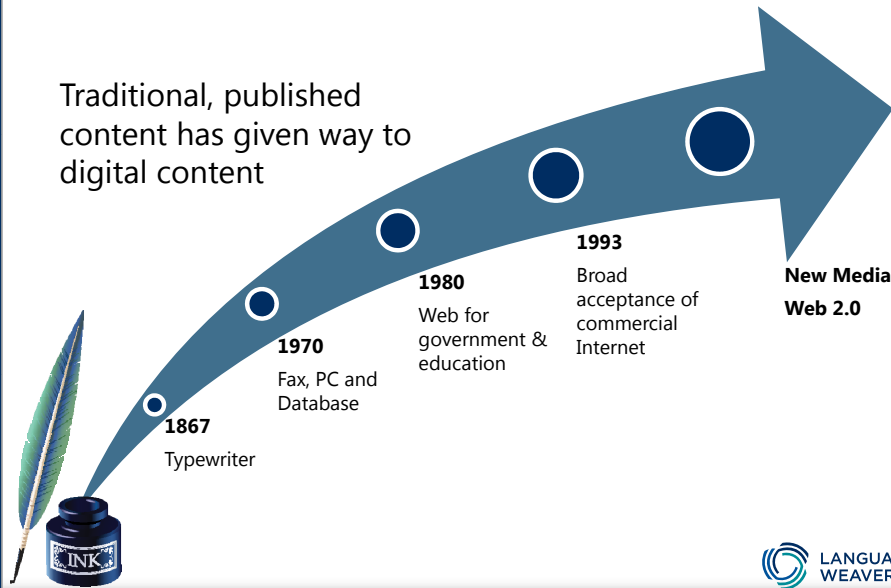
## What's Unique About the MT Proposition?

- MT is like a new hire (...or whole team)
  - If trained well; it can deliver unmatched productivity
- The MT market must deliver against a triad of need
  - Volume
  - Speed
  - Accuracy
- If the volume and speed requirements are not sufficient to incent a desire to train; then the value proposition can be diluted or polluted
  - Baseline system deployments need to be considered carefully against application use case



## The Problem: Speed of Innovation

Traditional, published content has given way to digital content



## Market Drivers

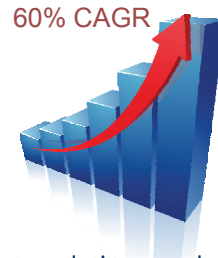
- Internet growth has exponentially driven digital content volume across languages
  - **1.48B people on line**
  - **CAGR 40%**
- Industry shortage of human linguists
  - Historical solutions have focused on minimizing the number of words to be translated to meet capacity limits
    - 350K-500K human translators
- Most content falls below the cost threshold for human translation (**average 21 cents per word**)
- The demand for digital content is significantly greater than the currently identified market of **\$14 billion for publication oriented translations**



Sources: IDC, Internet World Stats, CI Almanac, NUA Ltd., Milestone Group, Common Sense Advisory

## Digital Content Growth\*

- 2007
  - 281B total Gigabytes of Digital Content
  - 25% newly created (70B G)
    - 75% backed up or repurposed
- 2011
  - 1,800B Gigabytes of Digital Content
  - 30% created in the workplace (540B Gigabytes)
    - 1/100 %X540BG=54MGigabytes digital text with translation need
- Opportunity Assumption
  - 54MGigabytes of new text
  - 6.75MGWords to be translated over next 3 years
  - 10% of 6.75MGWords=675KGwords

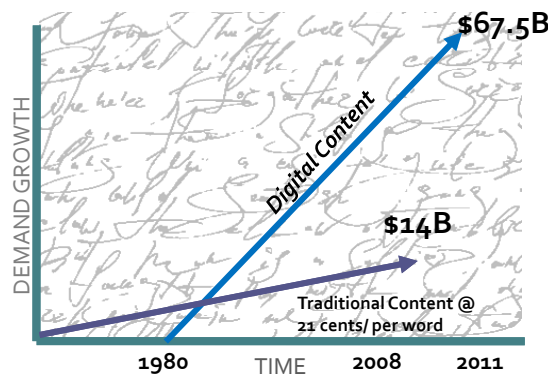


\*IDC-The Expanding Digital Universe 2007/2008



## Accelerating the Way World Communicates...

... because good communication is good for business



**Everyone wants to communicate,  
few want to translate**

Assumptions: .01 cents per word  
2,000 times more cost effective than current market



## Radical Changes in Communication Requirements

Regional Economic Strategy	Global Economic Strategy
Internally driven deadlines MarCom, manuals	Rapid/real time communication
Limited Business Languages	Variety of languages for emerging markets
Project based	Market driven demand e.g. Blogs, UGC
Objective to simplify by reducing the amount of content translated	Continuous demand
Objective to simplify by reducing the amount of content translated	Objective to translate everything
Quality Measured by defect rate	Quality measured by communication & meaning



## Customers are Drowning in Digital Content

### Content served up by LSPs

- User manuals
- Contracts
- Technical documentation
- Sales & marketing material

\$14 Billion

**Success Measured by:**  
Minimizing Translations  
Defect Rate

21 cents

### Content liberated by automated translation software (non LSP)

- Knowledge bases
- User generated content
- Search results
- Blogs
- Email/chat

Over  
\$67.5 Billion

**Success Measured by:**  
Value of human communication  
Market Penetration  
Business Growth





## What's Required?

- High speed, cost effective translation for communication
- Translations that convey meaning
- Accommodate virtually unlimited authoring
- Respond to very fluid content
- Accommodate unstructured content
- No limits on volume



## Digital Industry Needs

Business	Technical
Wide language coverage	Speed
Ad hoc access	Ability to handle continuous volumes
Low cost per article	Ability to handle unstructured content
Application based pricing	SaaS backbone access
Enduring partnership approach	Lightweight training with heavyweight results

## Government Needs

Business	Technical
Stability in Vendor	Extremely stable products
Patient partners	Ease of integration
Willingness to partner as a lifestyle	Standards in architecture
Perpetual licenses to match regulatory requirements	Multiple deployment footprints Large Server Single User on PC
Willingness to pursue low density languages	Extremely high speeds



## Commercial Enterprise Needs

Business	Technical
Ability to handle SaaS or on site deployments	Quality in high density languages
Application driven packaging	Ability to do domain training on company specific basis
Solution oriented delivery directed towards empirical operating improvements	Open architectures for integrations
Willingness to adjust to accommodate operating models and culture	Non-controversial platform stack
Willingness to be measured on the same metrics as your customer	Build into operational workflow



## Customer Case Study: Community Content

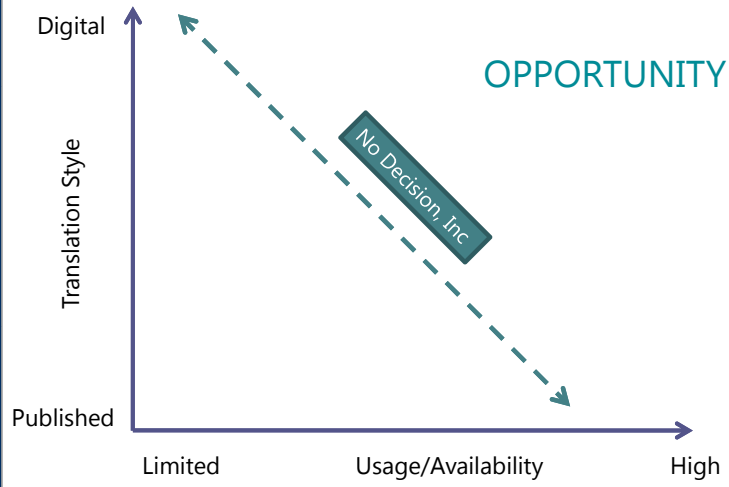
- **Business Situation**
  - Industry leading, travel focused web property has desire to expand user-generated contributions from English to EU based local markets
  - Unable to afford cost or time required for human translation
- **Desired Solution**
  - High speed, cost effective and accurate translations for unstructured digital content
  - Ability to handle continuous volumes of millions of user-generated articles
- **Language Weaver Solution**
  - SaaS based subscription to cost effectively train software and translate millions of user-generated articles into multiple European languages
- **Result**
  - Translations delivered successfully as measured by:
    - Site Hits
    - Dwell Times
    - Page Views
    - Increased number of EU contributions from EU customer base



“Flat” world has pushed the language problem down the org chart in daily operations

*We must deliver application solutions that enable and enhance human communications*

## Market Landscape



*"...accelerating the way the world communicates"*



**Thank you**

