

# PEMT for the Public Sector

## Evolution of a Solution

**Konstantine Boukhvalov and Sandy Hogg**  
ManpowerGroup Public Sector, Inc



ManpowerGroup™

# ManpowerGroup Public Sector

- **25+ years** supporting commercial and government clients around the world
  - Translation and I10n, transcription, interpretation, tech support, linguist placements, intel analysis and reporting
- Uniquely positioned to observe **language technology pain points for both clients and vendors**
- Developed a **universal TMS implementation procedure**
  - Advise and support **client TMS implementations**
  - Deploy and manage **SaaS TMS** for our clients
  - Provide client and vendor **TMS training**

## Presentation Objective

- Data protection concerns and knowledge of how to implement a practical HLT solution that delivers meaningful value have historically prevented public sector clients from using MT, CAT, and TMS
- We will show how our team successfully addressed data protection and client objectives to develop a practical, domain-specific PEMT solution to a public sector client who is now transforming how they use HLT

### Key Takeaways:

- How to develop a customized PEMT solution for public sector
- How to build and optimize TM corpora for statistical MT training
- How to gauge technological and procedural efficiencies for overall program success and scalability

# Historical HLT Challenges with MGPS PS Client Base

- **Limited HLT use due to various contract constraints**
  - No co-mingling data, no data in cloud
  - TM/TB corpora destruction
  - CONUS resources with citizenship, various clearance levels
    - HLT use still not widespread among PS linguist base (freelance)
- **No process automation**
  - Longer production timelines
  - Project-based translation



## Early Steps

- Secure isolated IT infrastructure
- Dedicated **enterprise-level CAT** setup
- **Centralized TM/TB**
- Resource requirements, e.g. **CAT-trained linguists/project managers**
- **TM/TB corpora** included as a **deliverable**



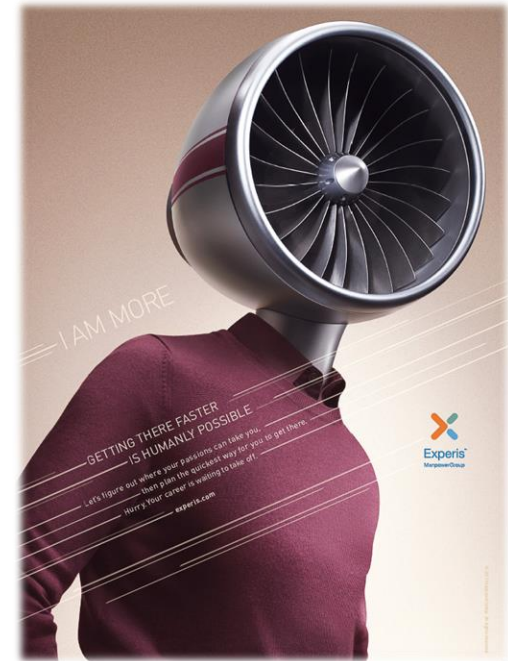
## Case Study

- **Objective:** Translate multiple domain-specific content streams with more automation and increased speed
- Large-volume legacy material alignment
- Geographically dispersed workforce of up to 120 participants in 3 continents/time zones
  - MGPS
  - Client stakeholders
  - Client linguists



# Program Requirements

- Centralized Globally Accessible HLT Resources
  - Projects
  - Integrated Domain-Specific Machine Translation
  - Translation Memories/TermBases
  - Tech Support
- Integrated Project Management
- Data and Personnel Security
  - Dedicated HLT resource instance
  - Controlled human access (US Citizen only)
- Continuous MT improvement cycle
- Process automation
- Seamless integration of cloud and local-install HLT solutions



ANSWER?

# Post-Edited Machine Translation

MT

CAT

TMS

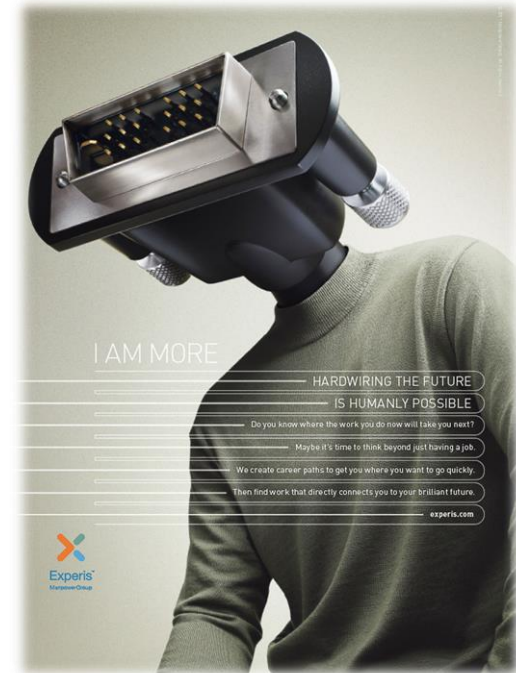
...Talent





## Define Stakeholders and Budget

- Dedicate a representative team of production experts – include the client!
  - Get early buy-in from the future production team
  - Start building the TMS operations culture
  - Let the production-side stakeholders define a business case and the best solution
- Align Budget and HLT options
  - Define Scope and Level of Effort - manage budget and expectations
  - HLT costs (CAT/TMS/MT)
  - IT setup (local install vs. SaaS)



# Challenge

**There are a growing number of strong HLT solutions. How do you select the right one, and how do you implement effectively?**



# Choosing/Validating the Right Solution

- Perform preliminary research
  - *What do I need?*
  - *What are my options (commercial/custom/open source)?*
  - *What are my community peers saying?*
- Choose solution candidates
- Set up orientation calls with solution developers
  - *Identify dedicated contacts for technical and contractual questions*
  - *Explore data security options for data and support*

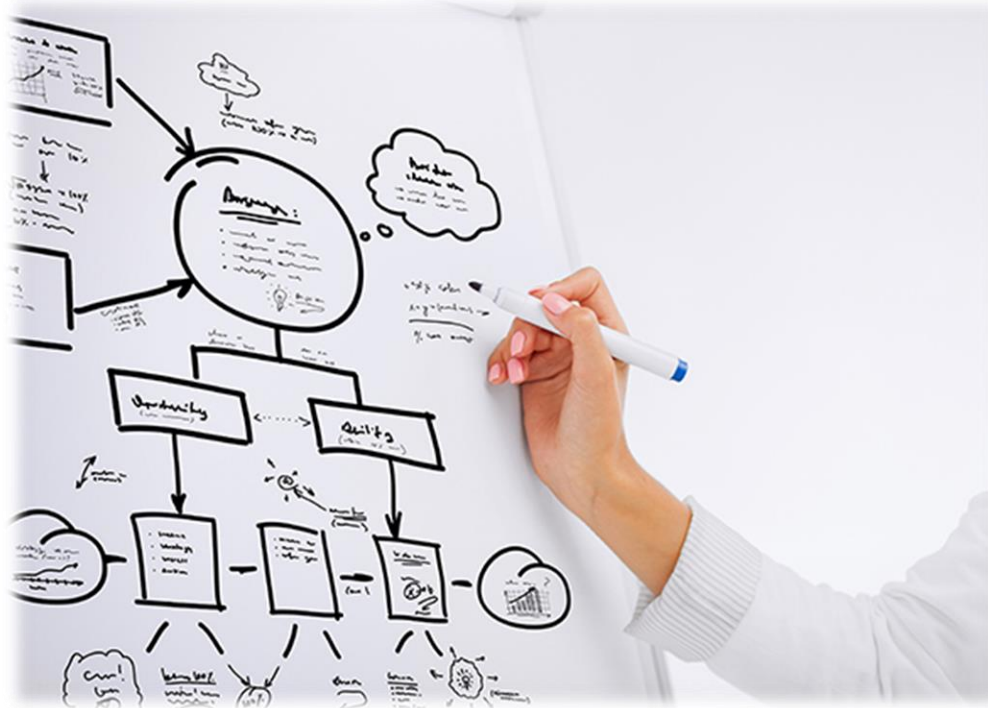


# Choosing/Validating the Right Solution (cont.)

- Create evaluation matrix
  - *Evaluate all products with the same criteria*
  - *Standard criteria include:*
    - Key features
    - Benefits
    - Shortcomings
    - Technical and contract support
    - Deployment options
    - Costs



# PUT THE DATA ASIDE – TIME FOR WHITEBOARDING!



# Whiteboard Your Workflows

- Define and document/update existing production processes
- Do *\*not\** adjust workflows based on the solutions' limitations
  - *If it doesn't fit, it's not right for you*
- Generate a master workflow that addresses the variations
  - *Define production steps as “required” or “optional”*
- Whiteboard other business requirements/expectations
  - *Manage expectations*



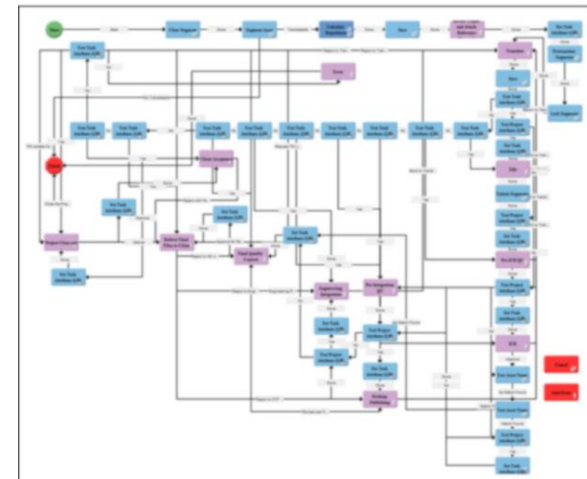
## See What “Fits” – Select and Acquire

- Combine the research with the master workflow and business requirements
- Identify the solution that provides the most value
- Generate TMS/MT Selection Report:
  - Fund the acquisition and deployment
  - Maintain technology knowledgebase
  - Validate your decision
- Finalize the deployment plan
- Minimize the time between the acquisition and production deployment



# Initial Configuration

- Master production workflow
- Sample business rules
- Sample linguistic resources (TMs, TBs, baseline MT)
- Optional/custom components and workflow steps (forms, fields, etc.)
- Production pilot
  - *Test the workflow, not just the filters*





# Document TMS/PEMT Production Procedures – Role-Specific Instructions

- Project Managers
  - *Production*
  - *Offline procedures*
- Linguist users
  - *Production*
- Client users
  - *Portal access/request*
  - *Production*
- Other production roles (as applicable)



# Develop and Implement Training

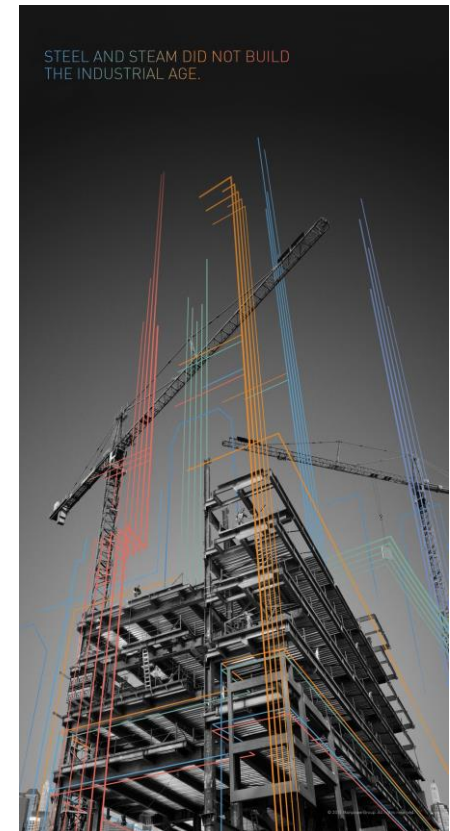
*Client stakeholder participation and buy-in is key to project success.*

- Develop reusable curriculum
- Provide general system overview
- Provide role-specific training



# Production Deployment

- Configure and deploy client portal and other auxiliary components
- Align legacy content
- Optimize TM corpora for MT training
  - Segmentation
  - Markup
- Perform initial training of Domain-Specific MT engines/language pairs
- Perform first automated and human evaluation of MT – start the measuring

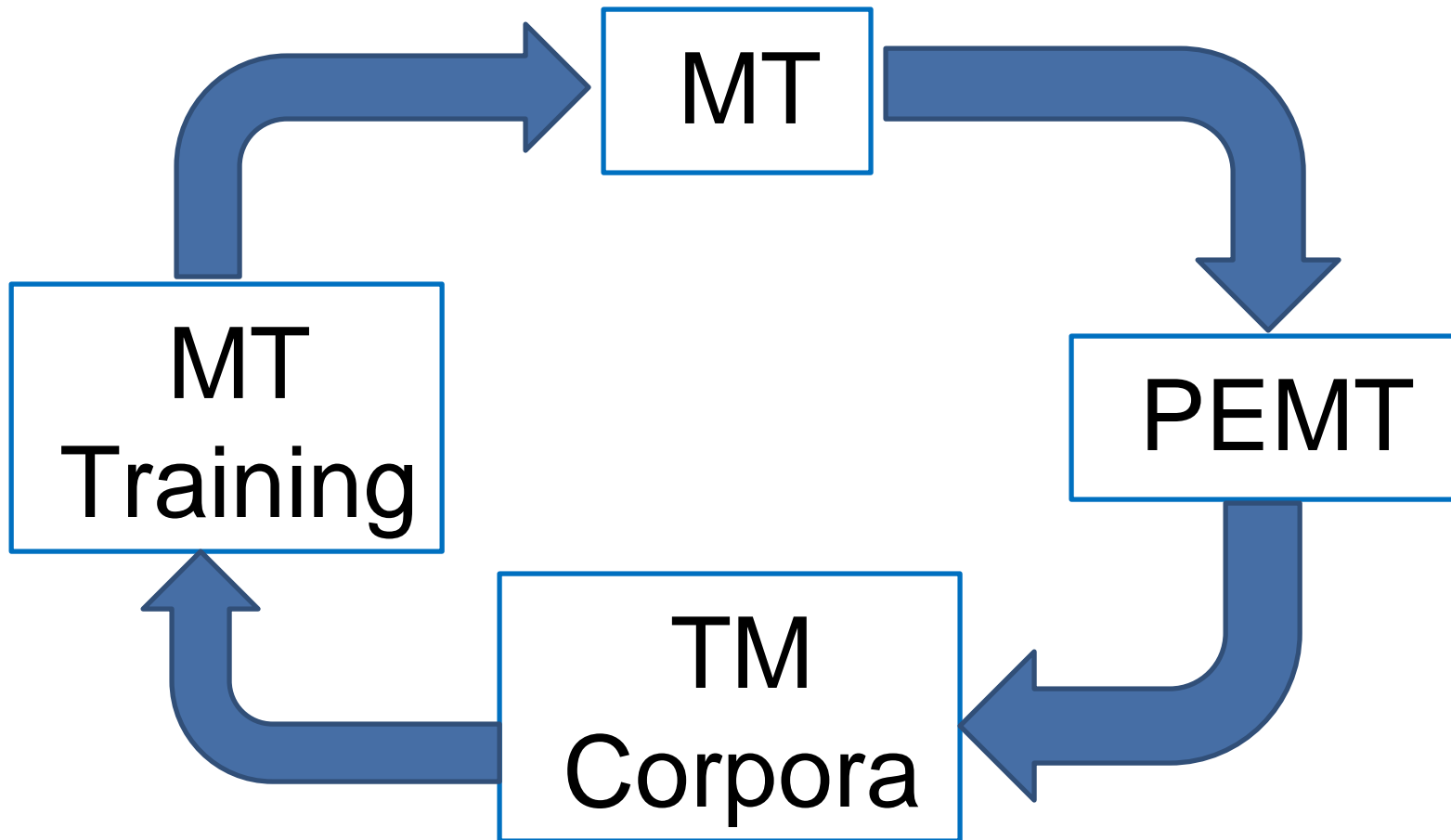


## Start PEMT!!!

- Start production for the selected programs/projects
- Adjust configuration, procedures, and documentation, as applicable
  - *Deliver the updates to the appropriate parties*

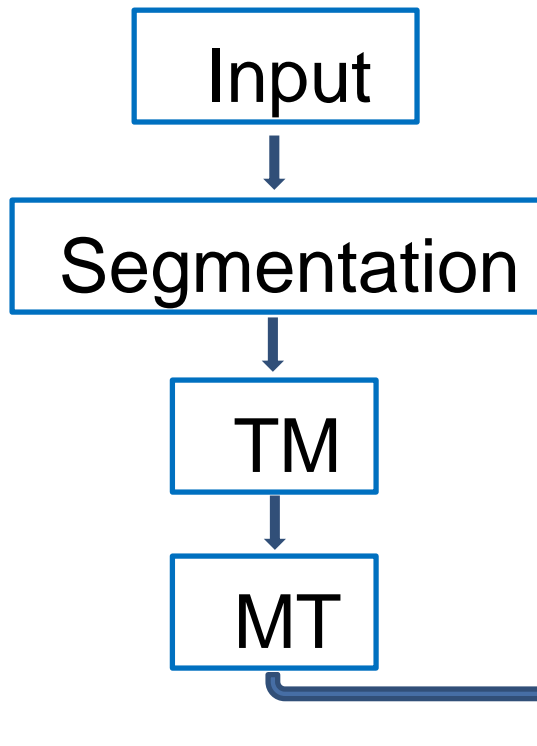


## Continuous MT Improvement Cycle

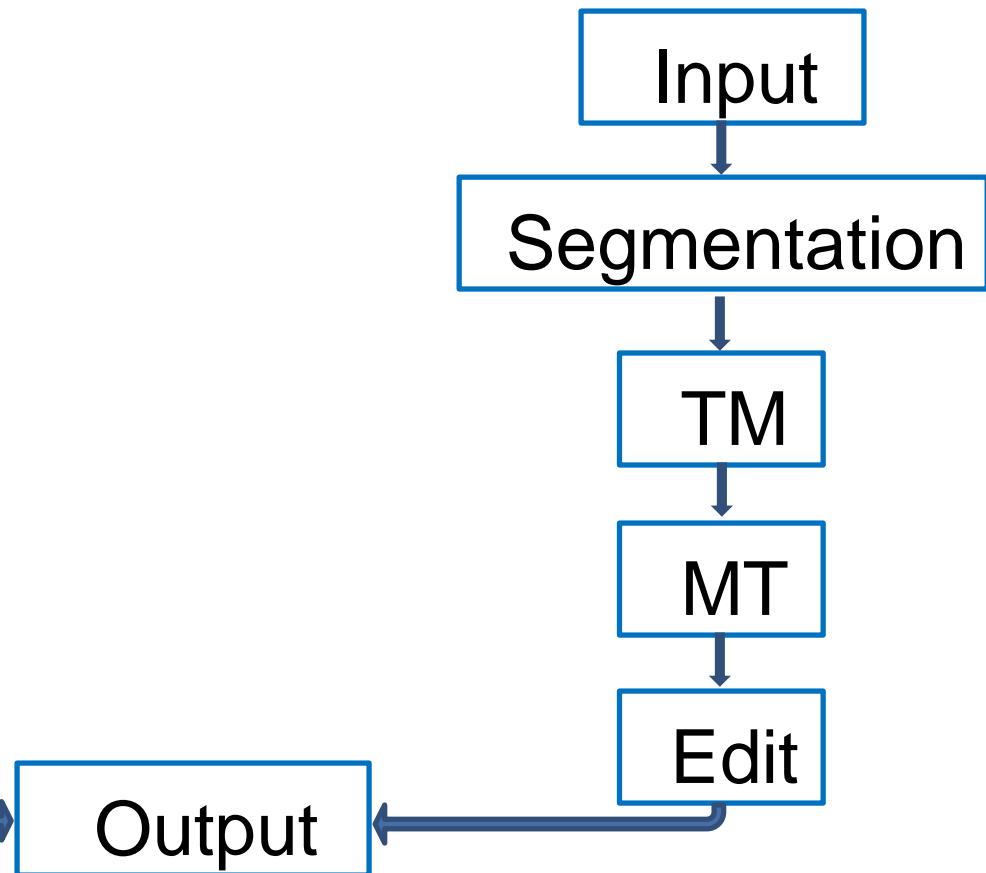


# Additional Automated Workflow Options

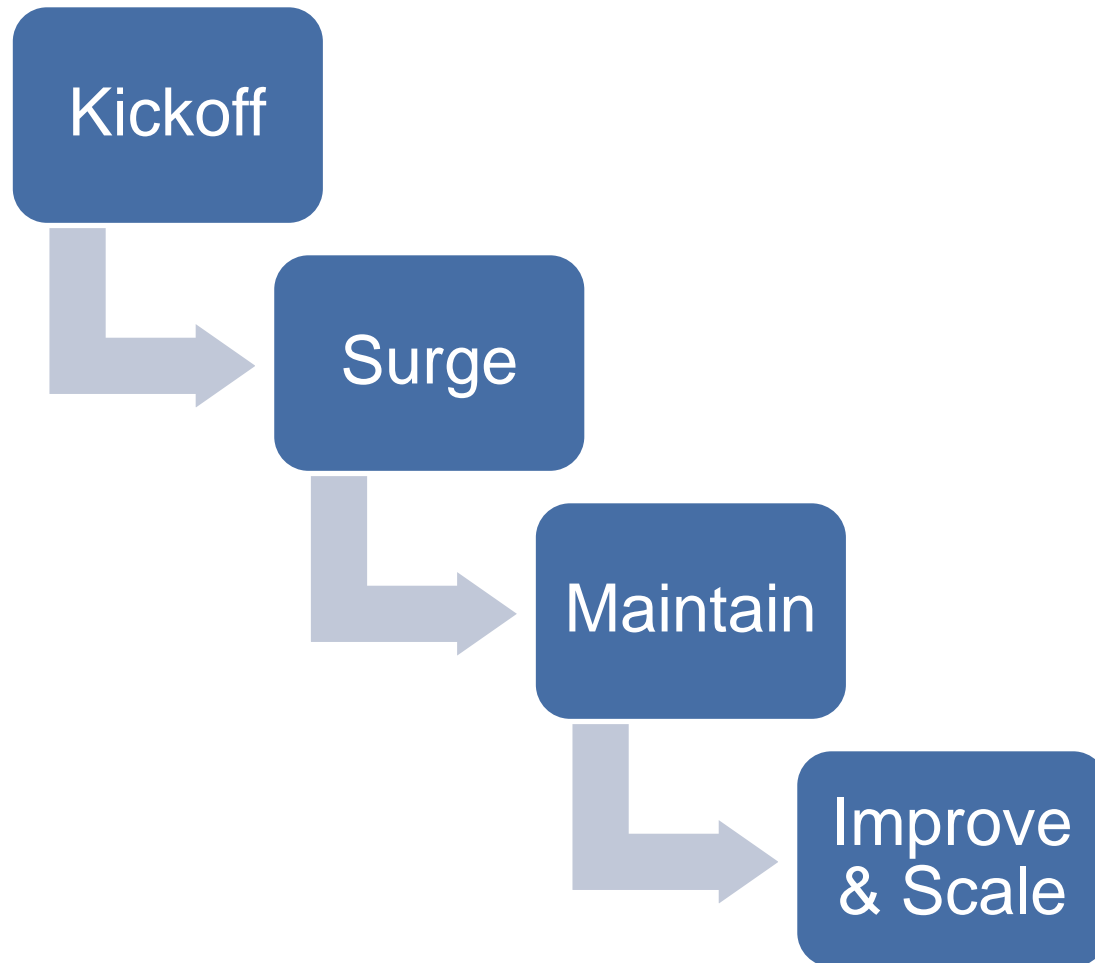
## TM-MT Only



## TM-MT with Edit

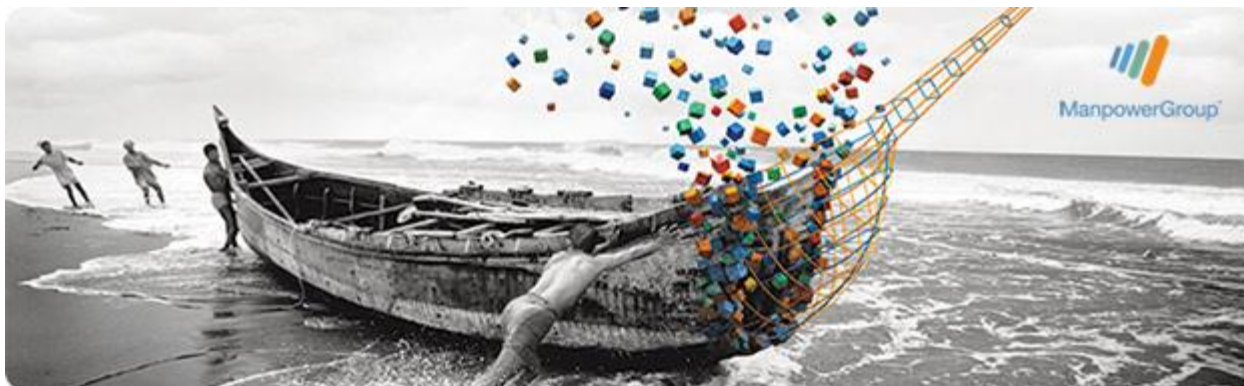


# Program Launch



# Kick Off

- **Review SOW**
  - Clarify parameters and assumptions at kickoff meeting with stakeholders
- **Inform Stakeholders – before, during, after kickoff!**
  - How to “engineer for success” with source selection, MT training corpus
  - Manage expectations for productivity, timeline
- **Set Goals and KPIs**
  - Linguist productivity
  - Tool effectiveness





# Surge

- **Build a Team**

- Linguists, Engineers, PM
- Success factors: full-time dedicated resources; native speakers, US citizens, tech-savvy; ability to handle breadth of subject matter domain

- **Baseline**

- Translate sample set of material (larger = better) outside of PEMT environment to gauge productivity sans HLT

- **Track Everything!**

- Client queries, client feedback, adjustments made to workflow
- Technology data: MT training corpus size and details, BLEU
- Project data: word count, TM leveraging, subject matter, time spent

# Maintain

- **Prioritize Knowledge Share**
  - Training materials, lessons learned
  - Meet regularly
- **Monitor, Report, Adjust**
  - Provide reports and recommendations monthly
  - Metrics
- **Review client level of engagement**
  - Client involved too much or too little?
  - Client requests within contract scope?



## Improve & Scale

- **Monitor technology developments and provide recommendations as necessary**
  - Raise questions/issues to software developers as needed
- **Evaluate MT output monthly; experiment and make adjustments as needed**
  - Capture qualitative and quantitative data
- **Communicate success stories and lessons learned**
  - Continually demonstrate ROI
- **Scale with additional domains and locales**
  - Ensure HLT solution can accommodate growth

# Recap

- **Historical challenges with MGPS PS client base**
- **Early steps toward HLT culture**
- **Case Study Success Factors**
  - Client buy-in
  - Customized HLT solution -- one size does not fit all
  - Documentation for reusability and scalability
  - Talent development through training
  - Clear KPIs to evaluate success
  - Continuous improvement

## Future Enhancements

- Neural MT
- Adaptive MT/Augmented Translation
- Substring tokenization
- Integrated speech-to-text supported by TMS/CAT/MT
- Multilingual Redaction Database (MRD)



# Thank you

**Konstantine G. Boukhvalov**

Operations Manager, Engineering  
ManpowerGroup Public Sector  
703-245-9372

[Konstantine.Boukhvalov@Experis.com](mailto:Konstantine.Boukhvalov@Experis.com)

**Sandra Hogg**

Deputy Program Manager  
ManpowerGroup Public Sector  
571-358-7671

[sandra.hogg@manpowergroup.com](mailto:sandra.hogg@manpowergroup.com)

# Appendix

<b>Acronym</b>	<b>Expansion</b>
BLEU	Bilingual evaluation understudy
CAT	Computer-assisted translation
CONUS	Contiguous United States
HLT	Human language technology
KPI	Key performance indicator
I10n	Localization
MGPS	ManpowerGroup Public Sector
MT	Machine translation
PEMT	Post-edited machine translation
PM	Project manager
PS	Public sector
ROI	Return on investment
SaaS	Software as a service
SOW	Statement of work
TB	TermBase
TMS	Translation management system