



# **U.S. Army Machine Foreign Language Translation System (MFLTS)**



**October 2012**

**Michael J. Beaulieu  
Product Director, 703-704-0304**

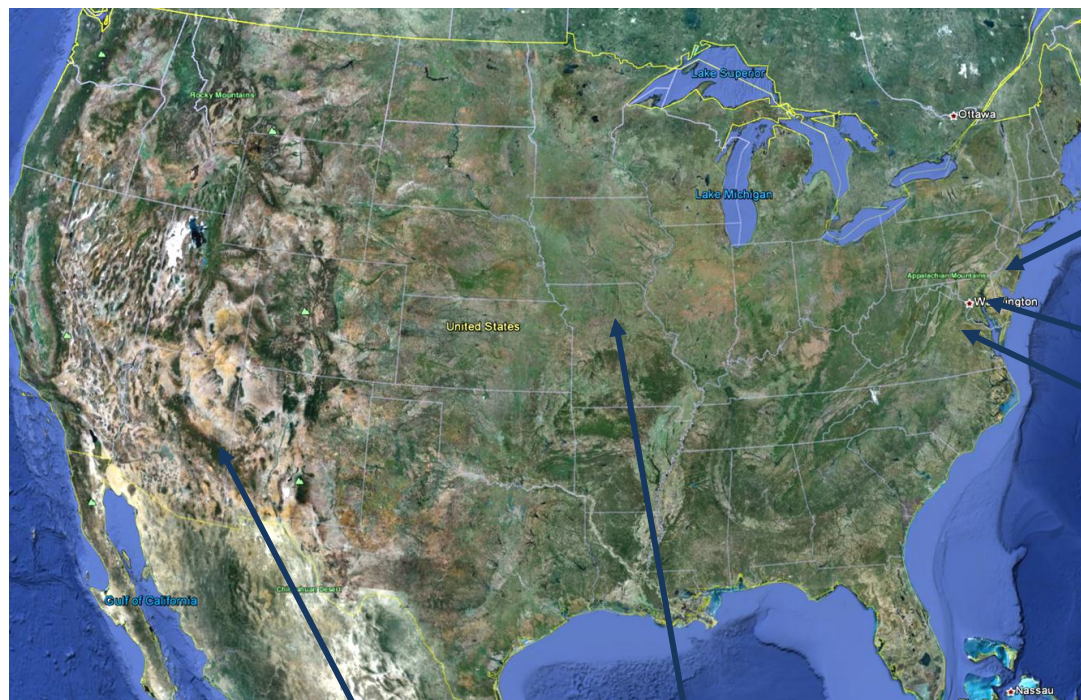


# Outline

- **United States Army Machine Foreign Language Translation System (MFLTS) – Who We Are**
- **A view of our Capability**
- **Types of Development**



# United States Army Machine Foreign Language Translation System (MFLTS) – Who We Are



**Arizona**  
Ft Huachuca

**Missouri**  
St Louis

**New Jersey**  
Eatontown, NJ

**Maryland**  
Aberdeen, MD

**VIRGINIA**  
Fort Belvoir

*Mission: Develop, field, and sustain a basic automated foreign speech and text translation capability into Army tactical systems in order to augment and complement limited human linguistic resources across all echelons in all environments*



## Description

MFLTS is a software-only program that will develop, acquire, field, and sustain a basic automated foreign speech and text translation capability into U.S. Army systems. The language development design allows for technical insertion of software language modules performing translation as required by the U.S. Army.



## Characteristics

- Portable (wearable or handheld), mobile (laptop), and web-enabled configurations
- Interoperable with commercial-off-the-shelf (COTS), or government-off-the-shelf (GOTS) automation equipment
- Plug-n-play language modules - a unit that translates between specific languages
- Rapid development and deployment of new language modules to meet warfighter operational needs
- Capable of deploying on multiple operating systems (e.g. Windows, Linux, Solaris)

*Not a "device" ... but rather an innovative Software Intensive Program employing modular, open system architecture*

*Vision: To be a high performance and learning organization that delivers world class products to the Warfighter and takes care of people.*

## Capabilities

2WS2S = Two-Way Speech-to-Speech  
 2WT2T = Two-Way Text-to-Text  
 COTS = Commercial Off-the-Shelf  
 GOTS = Government Off-the-Shelf  
 MSA = Modern Standard Arabic

### Initial Capability:

Speech	Text
<u>2WS2S</u> translation	<u>2WT2T</u> translation
3 languages (Pashto, Iraqi Arabic, & MSA)	2 languages (Dari & MSA)
portable & mobile configurations	mobile & web-enabled configurations
<u>2 domains</u> (checkpoint operations & base security)	<u>2 domains</u> (checkpoint operations & base security)

- **Follow-on Capability:** modules for additional languages, dialects (for already developed languages), OCR for handwriting, and technological upgrades.

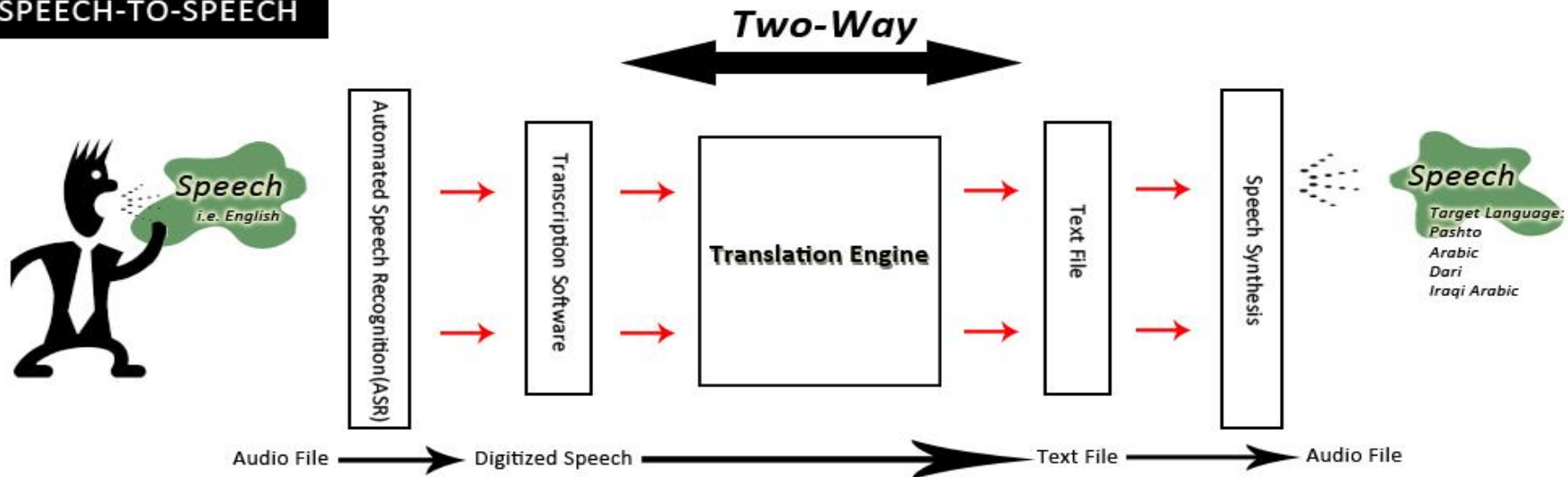
## Benefits

- Greatly increases points of presence for foreign language translation capabilities on the battlefield
- Enables communications with foreign speakers when human linguists are not available, enhancing mission success with:
  - Host Nations
  - Coalition partners
  - Indigenous peoples
- Enables rapid translations and triage of documents and media focusing analysts on the most important documents/media for exploitation

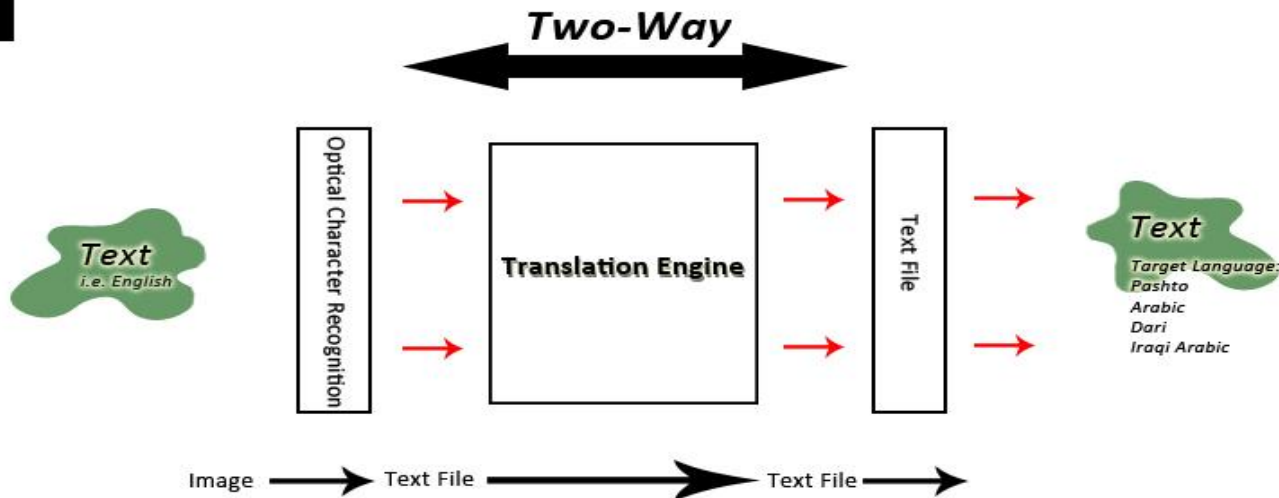
**President Obama "... in the 21st century, military strength will be measured not only by weapons our troops carry, but by the languages they speak and the cultures that they understand"**



## SPEECH-TO-SPEECH



## Text-to-Text







# Speech-to-Speech (S2S)

**Raytheon** Iraqi  
**BBN Technologies**

You: Hello how are you doing



هو : مر حبا شلونك



Sleep 



# Text-to-Text (T2T)

SDL T2T UI Demo 0.1

File View Tools Help

Translate

New Translation Document Queue Document View

بدأ في جدة اجتماع وزراء خارجية منظمة التعاون الإسلامي، قبيل انطلاق قمة مكة المرتقبة غدا لزعماء العالم الإسلامي التي من المقرر أن تركز على الوضع في سوريا وعلى "المشكلات التي تعوق مسيرة العمل الإسلامي"، بحسب عيد الحزير بن عبد الله نائب وزير الخارجية السعودي. وفي كلمة له بالمناسبة، قال الأمين العام للمنظمة أكمل الدين إحسان أوغلو إن منظمة التعاون "تؤكد ضرورة الحفاظ على وحدة سوريا ووقف العنف ضد المدنيين فوراً".

وأضاف أوغلو أن خطورة ما يجري على مستوى العالم الإسلامي "يتطلب توحيد الجهود".

وجدد الأمين العام استنكار المنظمة لما تتعرض له أقلية الروهينغا المسلمة في ميانمار من سلب للحقوق وانتهاك للحرمان والحريات، مشيراً في هذا المضمار إلى حرمانها من حقوقها الأساسية في الجنسية، حيث تعتبرهم سلطات ميانمار "لاجئين" من بنغلاديش.

وعدد أوغلو عددا من الخطوات التي قامت بها المنظمة مؤخرا في سبيل نصرة أقلية الروهينغا، ومن ذلك الإشراف على اجتماع تنسيقي في العاصمة الماليزية كوالالمبور للجمعيات الخيرية والإغاثية لصالح مسلمي ميانمار، وإيفاد مبعوث من المنظمة إلى ميانمار التقى برئيسها، كما قام وفد من المنظمة بزيارة إقليم أركان، و"هي أول مرة نزر فيها هذا الإقليم".

وبحسب أوغلو فإن منظمة التعاون تعمل الآن على توقيع اتفاق مع حكومة ميانمار يسمح بوجود دائم للمنظمة في إقليم أركان.

In Jeddah began a meeting of foreign ministers of the Organization of [Islamic cooperation](#), such as the launching of Mecca summit expected tomorrow of the leaders of the Islamic world, which is expected to focus on the situation in Syria, and on "problems that impede the process of Islamic work", according to Abdul Aziz bin Abdullah, Deputy Saudi foreign minister.

In his speech, he said the secretary-general of the Organization Ekmeleddin Ihsanoglu, the organization of cooperation, "stresses the need to preserve the unity of Syria and stop violence against civilians immediately."

Cakmakoglu added that the seriousness of what is happening on the level of the Islamic world, "requires the consolidation of efforts."

The Secretary-General renewed condemnation of the Organization to [Rohingya](#) Muslim minority in Myanmar of robbery and the violation of rights and freedoms sacrilegious, referring in this regard to deprive them of their fundamental rights in nationality, where it considers the Myanmar authorities "refugees" from Bangladesh.

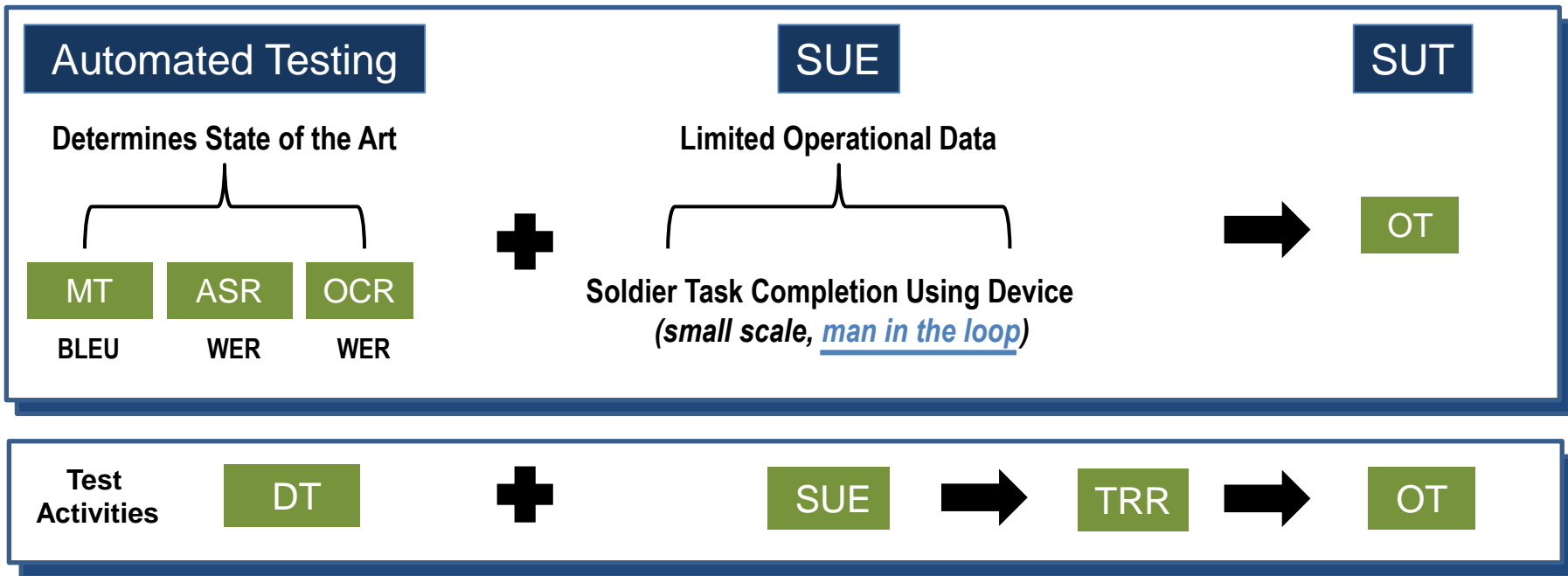
The number of Ihsanoglu a number of the steps taken by

Finished Auto Translation | utf-8 | Words/minute: 0, 100%



# Open Issues

## ➤ Metrics



1.  $ILR^* \approx ILR + (\text{Definition of Applicability of Machine Translation})$

2. Manual Protocol for Evaluating  $ILR^*$

- Soldier knows how to do job
- Soldier knows how to use machine



Manual Scoring Protocol  
**(Developed)**

3. Automated Metric Which Correlates to ILR

... **Poor Correlation**

ASR = Automatic Speech Recognition  
BLEU = Bilingual Evaluation Understudy  
DT = Developmental Testing  
ILR = Interagency Language Roundtable  
MT = Machine Translation  
OCR = Optical Character Recognition  
OT = Operational Testing  
SUE = System Under Evaluation  
SUT = System Under Test  
TRR = Test Readiness Review  
WER = Word Error Rate





# Open Issues (continue)

## ➤ Data Collection

- Expensive
- Time Consuming
- Need Industry / Academia / S&T / Community Collaboration

## ➤ Domain Development

- Expensive
- Time Consuming
- Need Industry / Academia / S&T / Community Collaboration



# Conclusions

## ➤ Initial Capability

- On schedule, within budget and performance parameters
- Need Industry / Academia / S&T / Community Collaboration with Open

Issue:

- Metrics

## ➤ Follow-on Capability

- Need Industry / Academia / S&T / Community Collaboration with Open

Issues:

- Metrics
- Data Collection
- Domain Development