Introduction to OpenTM2 – An Open Source Solution for Translators

Gerhard Fetz & Michael Schneider

October 25, 2012     Version 1.0
Agenda

- General **Overview**
- **Strategy** & **Vision**
- **Objectives** & **Benefits**
- The **Global Architecture of OpenTM2**
- **2013** time line
- **Supporting Teams**
- How to **contribute to the project**
- Sources for **More Information**

---- Backup Slides ----

- Future **Core Functions**
- Future **Core Modules** & **Additional Modules**
- **Development Schedule** (key milestones)
### General Overview – What is OpenTM2, and what not ...

**OpenTM2 is** ...  
- An **open source** software project
- Based on **IBM TranslationManager/2** code
- A translator's **workbench** and **translation editor environment**
- An **enterprise-level** translation memory CAT software.
- Intended as the **reference implementation** platform for translation asset standards.

**OpenTM2 is not** ...  
- A **globalization management** system
- An open translation management **data** project
- A **machine translation** tool or environment
- A **complete** version of IBM TranslationManager/2
- **Free** of implementation **costs**.
General Overview – The Related Communities

Steering Committee

› **IBM**
  - Strong supporter, contributor of the OpenTM2 source code
  - Contributor of management resources
  - Provider of development resources

› **OSB / FOLT**
  - Strong supporter, contributor of management resources
  - Contributor of OpenTMS, a shared translation memory component
  - Contributor of a translation proofreading environment
  - Contributor of server hosting

 › **GALA**
  - Strong supporter, and contributor of management resources

Community

› Interaction is managed through discussion groups

› **TRAC** is used to report bugs or request new features.
Strategy and Vision

- **Develop a reference implementation for translation asset exchange standards**
  - Think of a lossless exchange of translation memories using **TMX**
  - Think of a lossless exchange of translation dictionaries/glossaries using **TBX**

- **Encourage the development of open standards across the entire content management chain**
  - Think of a lossless exchange of source files to be translated using **XLIFF**
  - Think of a lossless segmentation of source files using **SRX**

- **Deliver choice to translators and enable them to work on projects without tool lock-in**
  - Think of OpenTM2 as the open source (enterprise) translation environment
Objectives and Benefits

› **Objectives:** Restructuring of the OpenTM2 code
  • Build **modules** and **plug-ins**
  • **Re-design existing features** and **build new features** based on open standards
  • Compile it for **multi-platform usage** (WIN, OSX, AIX, LINUX, UNIX)
  • Build a **better architecture**, but re-use existing services
  • Make it **scalable**

› **Benefits:**
  • An **open source** solution that will **produce high quality** localization results **without high cost**
  • Generate an excellent **reuse of assets** (translation memories and terminology)
  • Provide hooks to **project management** software
  • Provide hooks to various **translation/globalization systems**
  • Provide ability to perform **real time collaboration** across multiple users
The Global Architecture of OpenTM2

The diagram illustrates the architecture of OpenTM2, showing how different components and plug-ins interact. The core engine includes modules such as Base GUI, Base Editor, Base API, Base Analysis, Base Segmentation, Base Markups, Base Memory, and Base Counting. These components are managed by the OpenTM2 "Plug-In" Manager. Additional plug-ins from various sources, such as "xx's Plug-In" and "yy's Plug-In," are integrated into the architecture.
2013 Time Line (high level)

- Further Modularization & New Plugins
- More Code Fixing & Implement Requirements
- Test of all Functions and all Modules
- Data Migration (Transl. Memories, Dictionaries, Documents etc.)
- Production Readiness

OpenTM2 Overview
How to contribute to the Project?

› **Subscribe** to the mailing lists

› **Help testing** OpenTM2, **report bugs** and **request new features**

› **Review** the **documentation** and help improving it (WiKi style)

› **Help** with OpenTM2 **release** and **maintenance tasks**

› **Help fixing bugs** and **offer patches**

› **High-level contributors** may be invited to take on greater responsibilities (e.g. Steering Committee)

› Provide **professional services** for interested customers
Sources for more information

› **The OpenTM2 home page:**
  • http://www.opentm2.org

› **The source code repository (SVN):**
  • http://145.253.107.23/svn/opentm2/

› **The OpenTM2 WiKi:**
  • http://www.beo-doc.de/opentm2wiki/index.php/Main_Page

› **The OpenTM2 problem reporting database (TRAC):**
  • http://source.opentm2.org:8000/opentm2/report

› **The OpenTM2 Google Group:**
  • https://groups.google.com/forum/#!forum/opentm2-support
Backup Slides
Future Core Functions (1 of 2)

> **Core FUNCTIONS** supporting translation processes and tools:
  * Globalization Management support
  * Translation Work Flow support
  * Translation Project Management support

> **“Core ENGINE” functionality:**
  * Run on various platforms such as WIN32, LINUX, AIX, UNIX, MAC
  * Translation Workbench component
  * Folder Management component (ZIP technology)
  * Translation Memory component
    * Open technology to support “third party” translation memory plug-ins for:
      * MySQL, PostgreSQL, other (relational) DBs.
  * **Shared Translation Memory component:**
    * Server based shared translation memory
    * Local translation memories with replication & synchronization (offline/online capabilities)
Future Core Functions (2 of 2)

› **Language support:**
  - Support major languages including emerging markets (e.g. India).

› **Translation Dictionary component:**
  - A shared, server based dictionary component
  - Local dictionary with replication & synchronization function (offline/online capabilities)

› **Translation Editor component:**
  - Editor for text-based files (e.g. *.TXT, *.XML, *.HTML etc.)
  - Editor for GUI-files (e.g. *.DLG, *.RC)
  - WYSIWYG capabilities (e.g. Symphony, Open Office, Libre Office, MS Office, …)

› **Word Count component:**
  - Generates consistent and reliable word count independent from source file formats

› **API / Connector component:**
  - Allow to easily plug-in “third-party“ components
  - Allow to control the entire environment via API for automated work-flow support

› **File Parsers / Markup Tables:**
  - Supports all leading file formats
Future Core Modules & Additional Modules (1 of 2)

- Plug-In Manager with Plugins representing the OpenTM2 CORE ENGINE:
  - Analysis Engine
  - Segmentation Engine
  - Markup Tables (file parser)
  - Translation Memory Component
  - Word Count
  - Workbench
  - Translation Editor
  - API / Batch-commands
Future Core Modules & Additional Modules (2 of 2)

- ADDITIONAL MODULES (Plug-Ins developed by supporting dedicated company needs):
  - Language Component (e.g. HUNSPELL)
  - Translation quality checking component (e.g. TRANSLATE 5)
  - Dictionary Component / Term List Component
  - Machine Translation Component
  - Pre- and Post translation component
  - Counting component
  - Segmentation component (Markupo Tables as well as segmentation rules).
  - Additional Translation Editors component
Development Schedule (key milestones) (1 of 3)

**Priority 1:**

- **Translation Editor** Plugin (performs interactive translation task)
- **Translation Memory Editor** Plugin (allows editing of OpenTM2 integrated translation memory)
- **Translation Dictionary Editor** Plugin (allows editing of OpenTM2 integrated translation dictionary)
- **Translation Memory** Plugin (the OpenTM2 internal translation memory/-ies)
- **Dictionary** Plugin (the OpenTM2 internal dictionary/-ies)
- **Markup Table** Plugin (support for file parsers)
- **Shared Translation Memory** Plugin (located on a server accessed via Web-Services).
- "Abstract Document Layer" Plugin (access translation documents in a controlled way)
- "Abstract Folder Layer" Plugin (access OpenTM2 in a controlled way)
- **Word Count** Plugin (static & dynamic word counting)
- Plug-in for **Language Support & Linguistic Processing** (e.g. HUNSPELL)
- **API** Plugin (programmatically control OpenTM2)
- **Q-Environment** (web based application for Translation Quality Insurance)
Development Schedule (key milestones) (2 of 3)

.Priority 2:
• OpenTM2 Workbench Plugin
• Markup Table Identification Plugin
• Automated OpenTM2 Version Update
• Plug-In for ITM (Initial Translation Memory)
• Plug-in for Synchronization & for Online & Offline Capabilities
• Analysis Engine Plugin
• Plugin-Manager GUI
• MTI Plugin (Machine Translation Integration & Support)
Development Schedule (key milestones) (3 of 3)

> **Priority 3:**
  - Markup Table Development Environment
  - Translation Workflow Management Plugin
  - User-Interface Rework
  - Globalization Management Plugin
  - Translation Project Management Plugin
  - Terminology Management plugin
  - Content Management Support Plugin
  - Improve Usability