

ARMY RESEARCH LABORATORY



**Expanding the Toolkit and Resource Environment to Assist
Translation (TREAT) and Its User Base**

by Reginald L. Hobbs, Clare R. Voss, and Jamal Laoudi

ARL-TR-5563

June 2011

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Adelphi, MD 20783-1197

ARL-TR-5563

June 2011

Expanding the Toolkit and Resource Environment to Assist Translation (TREAT) and Its User Base

Reginald L. Hobbs and Clare R. Voss
Computational and Information Sciences Directorate, ARL

Jamal Laoudi
ARTI, Inc
1555 King Street Suite 400, Alexandria, VA 22314

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE (DD-MM-YYYY) June 2011		2. REPORT TYPE Final		3. DATES COVERED (From - To)	
4. TITLE AND SUBTITLE Expanding the Toolkit and Resource Environment to Assist Translation (TREAT) and Its User Base				5a. CONTRACT NUMBER	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Reginald L. Hobbs, Clare R. Voss, and Jamal Laoudi				5d. PROJECT NUMBER	
				5e. TASK NUMBER	
				5f. WORK UNIT NUMBER	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: RDRL-CII-T 2800 Powder Mill Road Adelphi, MD 20783-1197				8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-5563	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)				10. SPONSOR/MONITOR'S ACRONYM(S)	
				11. SPONSOR/MONITOR'S REPORT NUMBER(S)	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT Our initial aim in building the Toolkit and Resource Environment to Assist Translation (TREAT) was to provide translators with a hands-on framework as a single access point for learning about, using, and sharing a wide variety of online tools to support their task needs. Our extreme programming approach to the software engineering of this framework has enabled our in-house senior translator to participate fully in the software design, evaluation, and iterative modification process. In this report, we describe our ongoing work with extensions to TREAT and the unexpected result that we now see simple incremental changes to this framework introducing valuable side effects. The extensions support a wider range of users who provide more feedback to the tool and framework developers. With the inclusion in TREAT of new software tools to support these users, our translator discovered that the new tool combinations—though intended to support others—have had the side effect of helping him; he reports that he can now find more phrases that he used to miss in his translations, enabling him to post-edit his own work and boost the quality of his translations.					
15. SUBJECT TERMS Language translation, software development, user-centered design					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 14	19a. NAME OF RESPONSIBLE PERSON Reginald L. Hobbs
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) (301) 394-1981

Contents

List of Figures	iv
1. Introduction	1
2. Approach	2
3. Result and Ongoing Work	4
4. Conclusion	6
5. References	7
Distribution List	8

List of Figures

Figure 1. Characterization of TREAT users by source language level (ILR) and training level (tool use).....	3
Figure 2. Screenshot of TREAT (translation of Arabic source into English target) and two corresponding markup tool windows on Arabic source and English target, respectively (yellow = who, purple = when, blue = where).....	5
Figure 3. Wh-element markup on English translations output by BBLT.	6
Figure 4. Wh-markups on machine-translated texts output from seven MT engines into spreadsheet by MTriage.	6

1. Introduction

Our initial aim in building the Toolkit and Resource Environment to Assist Translation (TREAT) was to provide translators with a *hands-on framework* as a single access point for learning about, using, and sharing a wide variety of online tools to support their task needs. Our extreme programming approach to the software engineering of this framework has enabled our in-house senior translator—as subject matter expert and experienced software user—to participate fully in the software design, evaluation, and iterative modification process. As a result, TREAT incorporates principles of user-centered design and feedback from regular user sessions that guide ongoing development decisions (Hobbs et al., 2009).

The design of TREAT assumes that translators will come to this framework with different technical support needs. For new users, TREAT is set up by default with settings that give them direct access to a simple Translation Page (screen) for immediate task work upon opening the framework. For users with prior experience using the tool, TREAT has “any-time” options so that they can adjust, as needed or preferred, the built-in settings. TREAT permits three progressively more complex “levels of use” of options:

1. *Configure* the framework and screen layouts via checkboxes on a Configuration Page (screen) by selecting resources from among the available data sets, tools, and settings, appropriate for each task.
2. *Extend* the framework from either the Configuration Page using the Browse/Upload buttons by resource type or the Translation Page using a right-click menu (pop-up panel, at cursor) with a reconfigurable list of context-sensitive calls sending selected text to available or other new applications.
3. *Build* their own applications (mashups) in the Toolbar Window of the Translation Page (bottom of screen) to automate frequently repeated sequences of steps by combining two or more resources into a new service.

In this report, we describe our ongoing work with extensions to TREAT and the unexpected result that we now see simple incremental changes to this framework introducing valuable side effects. The extensions support a wider range of users including language learners, their instructors, non-translators, as well as the original users and translators, who provide more feedback to the tool and framework developers. With the inclusion in TREAT of new software tools to support these users, our translator discovered that the new tool combinations—though intended to support others—have had the side effect of helping him; he reports that he can now find more phrases that he used to miss in his translations, enabling him to post-edit his own work and boost the quality of his translations.

2. Approach

With the initial framework in place, we decided to expand TREAT to provide support to two new groups of users: students learning to be Arabic-language translators and teachers training them. The students and the teachers are native English speakers, so the training includes learning how to read Arabic script, understand Arabic text, and translate Arabic text into English.

We began by taking a first guess at the tools that the students would find most useful in supplementing their lessons. To that end, we reviewed our own in-house patterns of tool use, identifying the following:

- The software tools that the non-Arabic speakers and one student of Arabic in our lab *selected* and used effectively when needing to understand Arabic texts in the context of a project (when the translator was not available to assist them).
- The extent to which people *configured* a tool by changing the default settings as needed.
- The extent to which people *extended* a tool either by modifying the underlying code or by going back to the developer with a request to make the modifications.
- The combinations and regular sequences in which these tools were used as each person pieced together possible meanings of their texts. The “levels of use” analysis—from selecting to configuring to extending, and then to building effective, new sequences—by analogy with the design breakout in TREAT (see the three TREAT options outlined previously) also defined a new role for team members in developing TREAT: they would be non-translators in the user base of the TREAT framework.

The three software tools that we identified were the following:

- *For short passage or single segment translations:* MTriage, a front-end desktop application with numerous configuration settings, preprocesses and sends a source language text through multiple, back-end machine translation (MT) engines and displays the results in a spreadsheet or table with one source language segment per row and corresponding MT outputs horizontally aligned across columns (Hobbs et al., 2008).
- *For word sense disambiguation of one or more Arabic tokens:* the Buckwalter-based Lookup Tool (BBLT), a front-end application (desktop and Web) with a configuration page, displays all Buckwalter analyses of each input token in a table with one sense per cell beneath that token and lets users determine cell content in the display with lemma forms, parts-of-speech, translations, and other options (Micher and Voss, 2008).

- *For token or segment markup for downstream processing:* The item markup tool (IMT), a desktop application, enables users to build tagged texts (as ground-truth data sets for machine learning in other applications, such as named entity or wh-element annotation (in MT evaluation), by “highlighting” the relevant strings in a textbox window and labeling them by given categories (Tate and Voss, 2006).

While conducting our review, we were also asked by non-Arabic speaking archivists and analysts for assistance in identifying and learning to use software tools to triage Arabic language texts that they could not be understood, for possible follow-on translation by experts. Given all the possible users of TREAT, we realized that to continue with a user-centered design for an expanded user base, we would need to identify the levels of source language expertise and tool training of TREAT users (figure 1). Doing this would enable us to track and ultimately assess when the framework is effective for different types of users.*

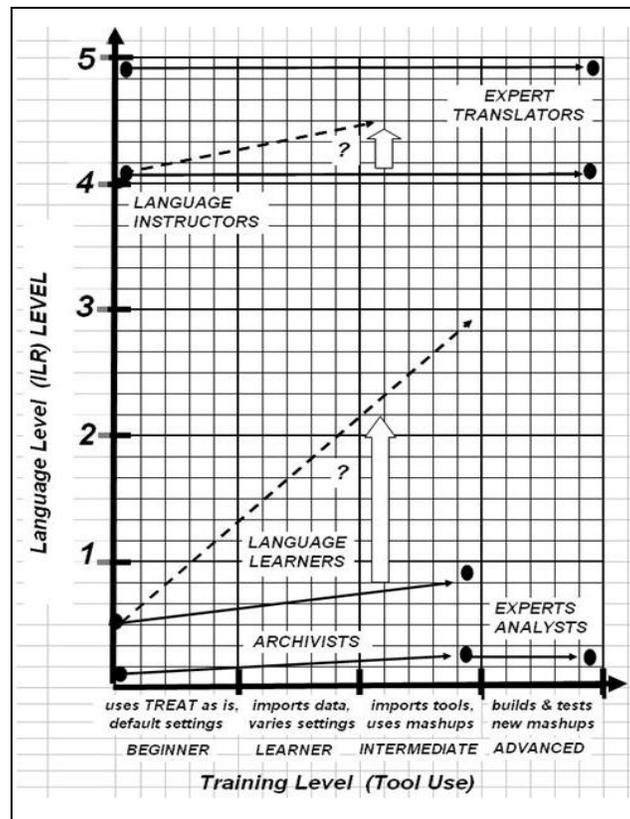


Figure 1. Characterization of TREAT users by source language level (ILR) and training level (tool use).

Note: Dots indicate users; straight lines show training paths of TREAT users, and dashed lines are notional training paths for TREAT users as they progress in language learning.

*We expect that methodological insights from Koeling et al., 2004; Koehn, 2009; and forthcoming experimental results by Day et al., 2006, together with our own rapid evaluations (per Kirkpatrick, 1994) will guide future development choices. While non-Arabic speakers have been found to perform as high as Interagency Language Roundtable (ILR) level 3 on MT test materials (Jones et al., 2007), we are not aware of research on the impact of translation software availability on rate of language learning.

3. Result and Ongoing Work

One result of our approach for the student/teacher users of TREAT has been to develop an activity with these tools that gives our targeted new students, in particular, both immediate and training-relevant exposure to a variety of Arabic language sources on the internet. The activity was selected to enable the students to practice a range of document exploitation tasks as diverse as triage, information extraction, and short summary-report writing. The students are asked to find and write short gist of a daily news story that is relevant to a current news event described by the teacher. The teacher may conduct a search beforehand and store these as a collection as an answer set, or choose to search at the same time as the students, encouraging everyone's curiosity and giving the class a sense of shared discovery.

By copying and pasting of an Arabic news story or perhaps just a title or a photo caption into the MTriage window, the students can quickly generate multiple parallel English translations. The students can scan these translations to get an initial sense of whether the article might be relevant to the designated event. They can then use BBLT to lookup the meanings of individual phrases in the stories that were oddly translated by the MT engines. They can also back-translate English words into Arabic with other engines in MTriage to trace the source of odd word choices. BBLT also can be set with diacritics on, so that words will appear in their table cell with their English translation and disambiguated in the Arabic spelling.

Both these tools immerse the students in an active process of looking for equivalent Arabic and English phrases, a key part of training for text translation that necessarily entails choosing equivalent expressions in English that match the meaning intended in the given Arabic text. To annotate the textual evidence for their gist of the Arabic story, students can upload the source text into IMT and use it to highlight the essential elements of information, the wh-elements, such as the people and organizations (who), the times and dates (when), and the locations and spatial relations (where).

Current work with TREAT involves building the interfaces to these three tools, so that they are properly called from within TREAT, with results from the tools returned as needed for TREAT. The TREAT Configuration Page now indicates the option to select these tools. Figure 2 shows a TREAT Translation Page window with an Arabic sentence in the top source language textbox and a translator's English target language version of that sentence. By selecting the Text Markup icon from the Toolbox at the bottom of the TREAT window, the user can launch IMT (markup tool) windows and then markup the source and target texts, as shown in figure 2.

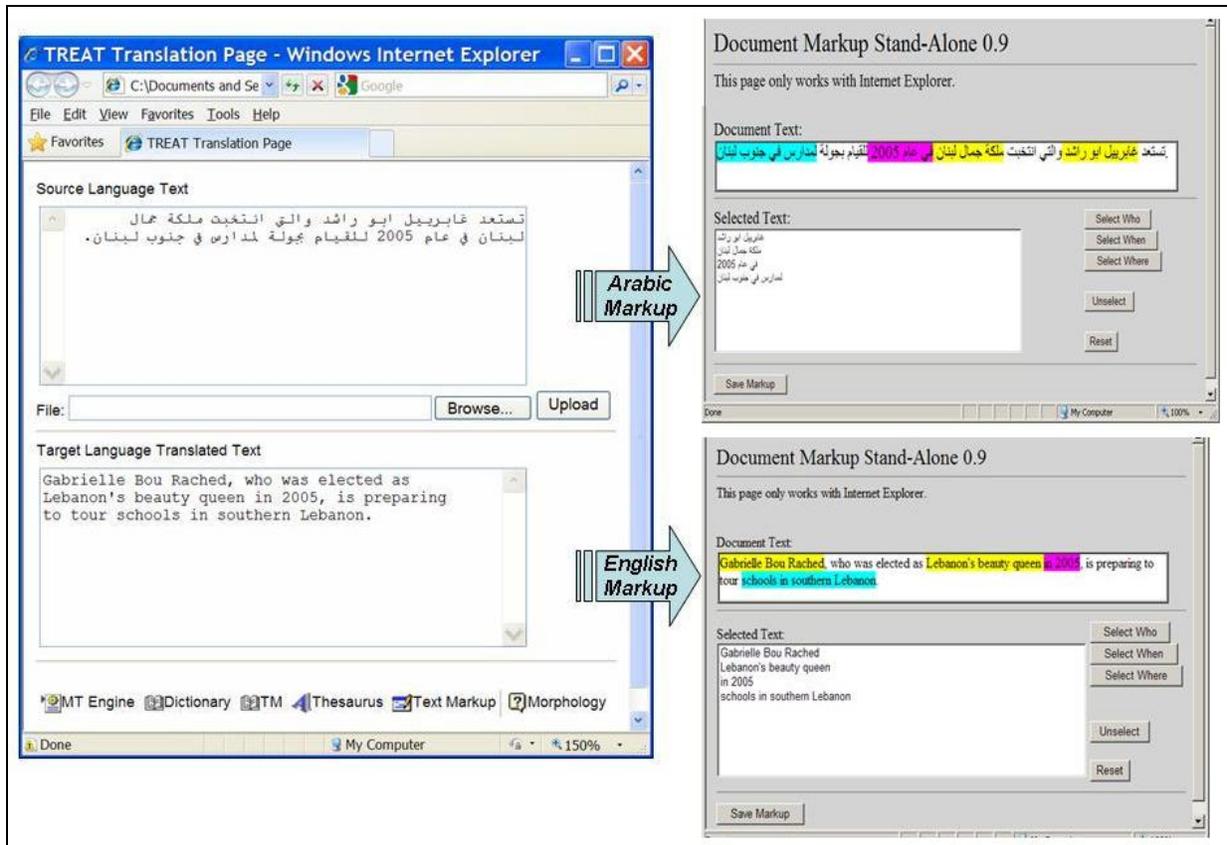


Figure 2. Screenshot of TREAT (translation of Arabic source into English target) and two corresponding markup tool windows on Arabic source and English target, respectively (yellow = who, purple = when, blue = where).

After numerous sessions testing the combination of producing his translations and doing the parallel wh-element markups with the IMT[†] in the process of another project whose texts contained lengthy sentences, our translator began to notice that the wh-element markup process provided him indirectly with an informal quality control tool. The highlighting gave him an easy way to quickly verify that he had captured the corresponding elements in the translations.

Ongoing work with TREAT now includes augmenting the interfaces to BBLT and MTriage so that users can open the XML files output by those tools directly within IMT, rather than as separate windows, as shown in figures 3 and 4, and work with the same IMT interface to do their markups.

[†]His work also entailed debugging other options in the markup tools hidden from these screenshots.

انتخبها	ابو غابرييل	والده	والتي	انتخبته	ملكة	جمال	لبنان	في	عام	2005	التجول	بجولة	في المدارس	في	جنوب	لبنان	
(she-it) gets ready	Gabriel	father	adult/rational/ rightly guided	and which/who/whom [fem. sg.]	I elected/chose	queen	camels	Lebanon		general/ common/ public	for to the undertaking/ carrying out/ setting up	by/with tour/ patrol/ round	for to schools		south/ South	Lebanon	PUNC
(she-it) appeals (for help)/ provokes/ incites	Abu	Rashid		and instrument/ apparatus/ appliance/ machine	you (m. s.) elected/ chose	aptitude/ faculty	camel driver	for to finger tips	in me				indeed/ truly/ schools	in me	south/ southern	for to finger tips	
(she-it) recovers/ regains/ reclaims		reluctant/ unwilling		and my instrument/ apparatus/ appliance/ machine	you (f. s.) elected/ chose	(he-it) had/ held/ controlled (ham it)	Jamal/ Gamal	indeed/ truly/ finger tips	V.					V.		indeed/ truly/ finger tips	
(she-it) is recovered/ is regained/ is reclaimed					(she-it) elected/ chose	(he-it) made owner (ham it)	beauty		and, so Y./10th					and, so Y./10th			

Figure 3. Wh-element markup on English translations output by BBLT.

MT 1 Output	MT 2 Output	MT 3 Output
English	English	English
an adult and which a beauty queen elected Lebanon 2005 for rising in a year Abu (ghabreeeel) be ready in South of Lebanon for schools in a tour...	Gabriel Abu Rashed and who elected Lebanon beauty queen in year 2005 for carrying out of a tour to schools in south Lebanon gets ready .	gets GhABRIIL father adult that selected beauty queen Lebanon in general 2005 for the standing tour for the schools in the south of Lebanon .
MT 4 Output	MT 5 Output	MT 6 Output
English	English	English
Gabriel is preparing Abu Rashid, who was elected Miss Lebanon in 2005 for a tour to schools in southern Lebanon.	Gabriel Abu prepares grown-up and which queen of beauty elected yogurts in year 2005 for the rising in tour for schools in south of yogurts.	Preparing Gabriel Abu Rashid, which was elected Miss Lebanon in 2005 for a tour to schools in south Lebanon.
		MT 7 Output
		English
		Prepares Gabriel Abu Rashid, which elected miss Lebanon in the year 2005 to carry out the round of the schools in south Lebanon.

Figure 4. Wh-markups on machine-translated texts output from seven MT engines into spreadsheet by MTriage.

4. Conclusion

The purpose of the TREAT framework is to enable translators to access multiple components for performing translation tasks. The most effective method for developing (and extending) the framework involves an iterative, user-centered approach, where specific user tasks impact the design, as opposed to software functionality driving what the user can do with the tools. The extreme programming paradigm enabled the translator (as both subject matter expert and potential end-user) to participate fully in the software design, evaluation, and iterative modification process. The resulting framework has a more direct impact on translator productivity as well as being an excellent basis for training inexperienced users on core translation tasks by supporting multiple levels of usage.

5. References

- Day, E. et al. A Platform for the Empirical Analysis of Translation Resources, Tools and Their Use. *Proceedings of Fifth International Conference on Language Resources and Evaluation (LREC06)*. Third International Workshop on Language Resources for Translation Work, Research & Training, Genoa, Italy, 2006.
- Hobbs, R.; Laoudi, J.; Voss, C. MTriage: Web-enabled Software for the Creation, Machine Translation, and Annotation of Smart Documents. *Proceedings of Sixth Language Resources and Evaluation Conference (LREC08)*, Marrakech, Morocco, 2008.
- Hobbs, R.; Voss, C.; Laoudi, J. On Beyond TM: When the Translator Leads the Design of a Translation Support Framework. *In Proceedings of the Twelfth MT Summit*, Ottawa, Canada, 2009.
- ILR Skill Level Descriptions for Translation Performance (adopted 2006)
www.govtilr.org/Skills/Adopted_ILRTranslationGuidelines.htm (accessed March 2010).
- Jones, D.; Herzog, M.; Ibrahim, H.; Jairam, A.; Shen, W.; Gibson, E.; Emonts, M. ILR-Based MT Comprehension Test with Multi-Level Questions. *Proceedings of the NAACL-HLT07*, Rochester, NY, 2007.
- Kirkpatrick, D. *Evaluating Training Programs: The Four Levels*; Berrett-Koehler: San Francisco, CA, 1994.
- Koehn, P. A Web-based interactive Computer Aided Translation Tool. *Proceedings of ACL/IJCNLP, Software Demonstrations*, Suntec, Singapore, 2009.
- Koeling, R.; Kilgarriff, Adam; Tugwell, David; Evans, Roger. An Evaluation of a Lexicographer's Workbench: Building Lexicons for Machine Translation. *Proceedings of 7th EAMT Workshop, Improving MT through Other Language Technology Tools*, Hungary, 2003.
- Micher, J.; Voss, C. Buckwalter-based Lookup Tool as Language Resource for Arabic Language Learners. *Proceedings of the ACL08 Workshop on Software Engineering, Testing, and Quality Assurance for NLP*, Columbus, Ohio, 2008.
- Voss, C.; Tate, C. Task-based Evaluation of MT Engines. Measuring How Well People Extract Who, When, Where-type Elements in MT Output. *Proceedings of 11th Annual Conference of the European Association for MT (EAMT06)*, Oslo, Norway, 2006.

NO. OF COPIES	ORGANIZATION
1 ELEC	ADMNSTR DEFNS TECHL INFO CTR ATTN DTIC OCP 8725 JOHN J KINGMAN RD STE 0944 FT BELVOIR VA 22060-6218
1 CD	OFC OF THE SECY OF DEFNS ATTN ODDRE (R&AT) THE PENTAGON WASHINGTON DC 20301-3080
1	US ARMY RSRCH DEV AND ENGRG CMND ARMAMENT RSRCH DEV & ENGRG CTR ARMAMENT ENGRG & TECHNLOGY CTR ATTN AMSRD AAR AEF T J MATTS BLDG 305 ABERDEEN PROVING GROUND MD 21005-5001
1	PM TIMS, PROFILER (MMS-P) AN/TMQ-52 ATTN B GRIFFIES BUILDING 563 FT MONMOUTH NJ 07703
1	US ARMY INFO SYS ENGRG CMND ATTN AMSEL IE TD A RIVERA FT HUACHUCA AZ 85613-5300
1	COMMANDER US ARMY RDECOM ATTN AMSRD AMR W C MCCORKLE 5400 FOWLER RD REDSTONE ARSENAL AL 35898-5000
1	US GOVERNMENT PRINT OFF DEPOSITORY RECEIVING SECTION ATTN MAIL STOP IDAD J TATE 732 NORTH CAPITOL ST NW WASHINGTON DC 20402
13	US ARMY RSRCH LAB ATTN IMNE ALC HRR MAIL & RECORDS MGMT ATTN RDRL CII B BROOME ATTN RDRL CII T C VOSS ATTN RDRL CII T C SANDERS ATTN RDRL CII T J LAUDI ATTN RDRL CII T R HOBBS (5 HCS) ATTN RDRL CII T V M HOLLAND ATTN RDRL CIO LL TECHL LIB ATTN RDRL CIO MT TECHL PUB ADELPHI MD 20783-1197

TOTAL: 20 (1 ELEC, 1 CD, 18 HCS)