



*Sue Ellen Wright
Kent State University
SDL TRADOS Webinar,*



20 September, 2006

The Role of Terminology Management in Localization



Acronyms



- I18n = Internationalization
- JIT = Just-in-Time
- KOS=Knowledge Organization System
- L10n = Localization
- OWL=Web Ontology Language
- QA=Quality Assurance
- ROI = Return on Investment
- SL=Source Language / TL=Target Language
- TMM=Terminology Management (TMgmt)
- TWF=Terminology Workflow



The Role of Terminology



- **Units of thought**, focusing on the psychological aspect of recognizing objects as part of reality
- **Units of knowledge**, focusing on the ways that information is collected, stored and retrieved
- **Units of communication**, stressing the fact that concepts are the prerequisite for knowledge transfer in specialized discourse

Voices of Experience



- Medtronic has long managed its terminology as a ***strategic asset*** and achieved significant savings not only in translation costs, but also accelerated time to market by improving the efficiency of the author-translate-review process.
- Terminology is developed very early in the product development cycle in order to allow for a consistent use of terms in the whole spectrum of communications and documentation.

Inna Geller
Translation manager
Medtronic





Why Manage Terminology?



- Managing terminology
 - ◆ Supports your ***corporate brand image***
 - ◆ Makes your products easier to use, easier to translate, and easier to adapt to global markets.
- The information business is shifting from a fixed product model, where information products had a long shelf life and a common customer base, to a dynamic product model, where information is customized “on-the-fly,” based on users’ particular needs

Kara Warburton
Corporate Terminologist
IBM





Facing New Challenges



- Integrating terminology management into the product development process
- ***Managing terminology on a global, rather than a product scale***
- Finding appropriate tools that can handle a mixed bag of monolingual, bilingual, and multilingual terminology activities
- Addressing the diverse needs of both authors and translators
- Dealing with employee resistance and the barriers that result from corporate culture

*Kara Warburton
Corporate Terminologist
IBM*





Consistent Corporate Language



- A significant component of globalization costs is attributable to rework
- The primary cause of rework is inconsistent terminology
- Inconsistencies in corporate language compound across projects and throughout stakeholder groups
- Inconsistent terminology, especially in multilingual environments, causes costly delays and downstream misunderstandings



Terminology Management Spectrum



- Continuum of practice in a range defined by multiple factors:
 - ◆ Enterprise type and mission
 - ◆ Criticality of terminology for core business
 - ◆ Criticality of quality and branding concerns
 - ◆ Recognition of tangible and intangible ROI
 - ◆ Terminology user groups
 - ◆ Buy-in by stakeholders



Localization Stakeholders



- Companies who own knowhow & who generate information
- Management at all levels
- Primary generators of information products
 - ◆ In-house & out-house tech writers, programmers, designers, etc.
- Secondary generators of localized products
 - ◆ Translation and localization service providers
- Marketing
- In-country partners
- End users



Enterprise Type



- Nature of enterprise
 - ◆ Government
 - ◆ Industry
 - ◆ Research institutes
 - ◆ Localization/translation bureaus
 - Nature of the client
 - Text type
 - Negotiated client/vendor agreements
 - ◆ Web content management providers
 - ◆ Freelance & in-house technical writers, translators
 - ◆ Librarians & knowledge organization environments



TMM Mission & Strategic Position



- Support for national language policy
- Support for global enterprise activity
- Support for technical writing & translation
 - ◆ Ongoing, stable subject fields
 - ◆ Frequent one-off, unrelated, unanticipated subject fields
- TMM approaches
 - ◆ Prescriptive vs. descriptive
 - ◆ Ad hoc vs. systematic



TWF Task Issues for L10n



- Specific task types
 - ◆ Source language product design
 - GUI
 - ◆ Source language documentation
 - Manuals
 - Help files
 - ◆ Controlled vocabulary & controlled language for document production
 - ◆ Just-in-time documentation of TL equivalents
 - ◆ Support for machine vs. human-oriented translation
 - ◆ Multilingual content management in dynamically changing Web environments
 - ◆ Consistent localized marketing, training, and user support



Knowledge-oriented TWF



- ✦ Enterprise-related information and knowledge management
 - ✦ Terminology as a function of taxonomy, ontology, and information retrieval (knowledge organization systems: KOS)
 - ✦ KOS interaction with:
 - Monolingual technical writing & product development
 - Translation and localization
 - Multilingual technical writing
 - E-business solutions
 - Inventory control and logistics
 - General information retrieval and processing
 - Standards issues (SKOS and OWL)



Cost Effectiveness



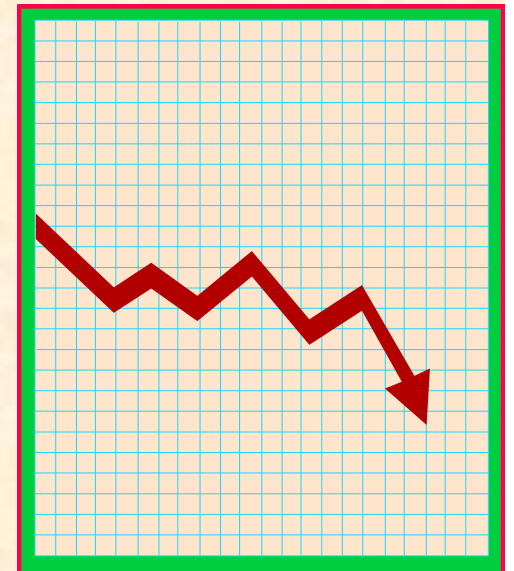
- Terminology management activity costs time and money. Is it economically feasible:
 - ◆ For individual translators?
 - ◆ For groups working together?
 - ◆ For bureaus and translation services in government and industry?



Criticality of Terminology



- Potential for market losses
- Potential for communicative losses
- Risk of product failure, human injury
- Adverse effects on branding efforts (marketing issues)
- Relative significance of terminology
 - ◆ To the process
 - ◆ To the product
 - ◆ Example: Terminology is more critical if you are selling software than if you are selling wheat.





ROI: Variable vs. Fixed Costs



- ❖ TMM tasks hidden in general overhead
- ❖ Cost of *not agreeing to uniform terminology* hidden in overhead costs
- ❖ Costs of correcting mistakes or recouping damage to branding difficult to document
- ❖ Compounded costs due to the persistence of defective communications



Fixed Costs



- ✦ The costs of doing systematic terminology work, on the other hand, are up front costs that can be calculated using traditional means.





Return on Investment



- ✦ The greater general applicability of specific terminological units (greater frequency) = greater the return on terminology management costs.
- ✦ The greater the quality, safety, or branding-related criticality = greater the return.
- ✦ The greater the degree of integration between controlled authoring, CAT, TM and MT applications, the greater the payback in leverageable data.



Terminology Management Spectrum



- These factors result in varying:
 - ◆ Integrated and non-integrated workflow models
 - ◆ Complexity with respect to the data model
 - ◆ Investment of human resources in the form of
 - Term extraction
 - Context documentation
 - Concept definition (ideally by SL & TL subject experts)
 - Terminology product delivery



Data Model Complexity



- A continuum of practice in a range varying from/to:
- Creation of non-systematic tab-delimited, text and job-oriented glossaries and spreadsheets
 - ◆ SL term + TL term
 - ◆ + POS, Gender, and/or a note
- Theoretically rich complex entry input models
 - ◆ Subject field classification
 - ◆ Client designation
 - ◆ Concept-oriented definition
 - ◆ SL term + TL term + POS + Term Type + Register + Context, etc.
 - ◆ Documentation through source citations, graphics, etc.

Excel Spreadsheet Glossary

	A	B	C	D	E	F	G	H	I	J	K	L
1		Spanish Term	POS	Gender	English Term	POS	Subject Field					
2	12	árbol	noun	masculine	tree	noun	Forestry					
3	1	biodiversidad	noun	feminine	biodiversity	noun	Forestry					
4	24	blanqueo	noun	masculine	bleaching	noun	Forestry					
5	26	caducifolio	adjective		deciduous	adjective	Forestry					
6	4	conífera	noun	feminine	conifer	noun	Forestry					
7	17	contaminación	noun	feminine	contamination	noun	Forestry					
8	20	corta	noun	feminine	felling	noun	Forestry					
9	18	descortezar	verb	verb	debark	verb	Forestry					
10	21	desechos	noun	plural	waste	noun	Forestry					
11	5	ecosistema	noun	feminine	ecosystem	noun	Forestry					
12	6	erosión	noun	feminine	erosion	noun	Forestry					
13	19	eucalipto	noun	masculine	eucalyptus	noun	Forestry					
14	7	forestación	noun	feminine	forestry	noun	Forestry					
15	22	hacha	noun	masculine	axe	noun	Forestry					
16	13	injerto	noun	masculine	graft	noun	Forestry					
17	28	licor blanco	noun	masculine	white liquor	noun	Forestry					
18	23	licor negro	noun	masculine	black liquor	noun	Forestry					
19	27	licor verde	noun	masculine	green liquor	noun	Forestry					
20	11	madera	noun	feminine	timber	noun	Forestry					
21	8	micropropagación.	noun	feminine	micropropagation	noun	Forestry					
22	2	motosierra	noun	feminine	chainsaw	noun	Forestry					
23	3	noun	noun	feminine	chlorination	noun	Forestry					
24	9	pesticida	noun	feminine	pesticide	noun	Forestry					
25	10	poda	noun	feminine	pruning	noun	Forestry					
26	16	polución	noun	feminine	pollution	noun	Forestry					
27	15	sustentable	adjective		sustainable	adjective	Forestry					
28	14	topografía	noun	feminine	topography	noun	Forestry					
29	25	tronco	noun	masculine	bole	noun	Forestry					
30												
31												
32												
33												
34												

forestrytab

start

3 Microso...

Gmail - Pre...

60011_2006

Search Res...

Microsoft E...

EN

Norton

9:00 AM

Highly Granular Input Model

The screenshot displays the SDL MultiTerm software interface. The main window is titled "SDL MultiTerm [Project Untitled.xdp]". The interface is divided into several panes:

- Left Pane (List):** A list of terms including "axe", "biodiversity", "black liquor", "bleaching", "bole", "chainsaw", "chlorination", "conifer", "contamination", "debark", "deciduous", "ecosystem", "erosion", "eucalyptus", and "felling". The term "chlorination" is selected.
- Bottom-Left Pane (Project Structure):** Shows a folder named "Project: Untitled.xdp" containing a sub-folder named "Forestry".
- Top-Right Pane (Entry Details):** Displays the entry level, entry number (3), and subject (Forestry). It shows the English entry for "chlorination" with its definition, source, part of speech (noun), usage (technical), and context. The Spanish entry for "clorinación" is also shown with its definition, source, part of speech (noun), grammatical gender (feminine), status (standardized), usage (technical), and context.
- Bottom-Right Pane (Navigation):** Shows the current entry path: "\chlorination /".

The Windows taskbar at the bottom shows the Start button, several open applications (Internet Explorer, File Explorer, Microsoft Word, Adobe Reader, Microsoft Office Word), and the system tray with the time 3:44 PM.

SDL MultiTerm [Project Untitled.xdp]

Termbase Project Entry Search View Help

English Spanish Flags layout2

Forestry Input Model

axe
biodiversity
black liquor
bleaching
bole
chainsaw
chlorination
conifer
contaminati
debark
deciduous
ecosystem
erosion
eucalyptus

Project: Ur
Forestr

Entry level:
Entry number: 3
Subject: Forestry
Graphic:
G_Source:

EN English
Definition: The application of chlorine to drinking water, sewage, or industrial waste to disinfect or to oxidize undesirable compounds.
D_Source: http://teachmefinance.com/Scientific_Terms/Chlorination.html
Graphic:
G_Source:

• **chlorination**

Part of Speech: noun
Usage Register: technical
Context: Traditionally, the bleaching of pulp is performed with the help of elemental chlorine. The bleaching is essential for the improvement in brightness, cleanliness and removing impurities. The advantages of chlorination are that it is most effective delignifying agent and at the same time least expensive of all the bleaching chemicals. It is also excellent for the shive and dirt removal. But, in the recent years, chlorine is viewed as enemy of the environment, as it destroys the stratospheric ozone layer, and produces dioxin and organic chlorides. Its low consistency produces a large volume of acidic effluent, which has to be neutralized before waste treatment. It degrades pulp to some extent and is corrosive in nature.
C_Source: <http://www.cpcb.nic.in/sept97ii.htm>
Grammatical Gender:
Grammatical Number:
Status:
Type:
Note:
N_Source:
Related Term:

Spanish
Definition: Aplicación de cloro al agua potable, a las aguas negras, o al drenaje industrial para desinfectar u oxidar componentes indeseables.
D_Source: <http://www.ine.gob.mx/ueajej/publicaciones/libros/109/glosario.html>
Graphic:
G_Source:

• **clorinación**

Part of Speech: noun
chlorination /

Entry Level

start | 2 Int... | tlpask... | 5 Mic... | Adobe... | 3 Mic... | SDL M... | EN | Norton | 4:21 PM

SDL MultiTerm [Project Untitled.xdp]

Termbase Project Entry Search View Help

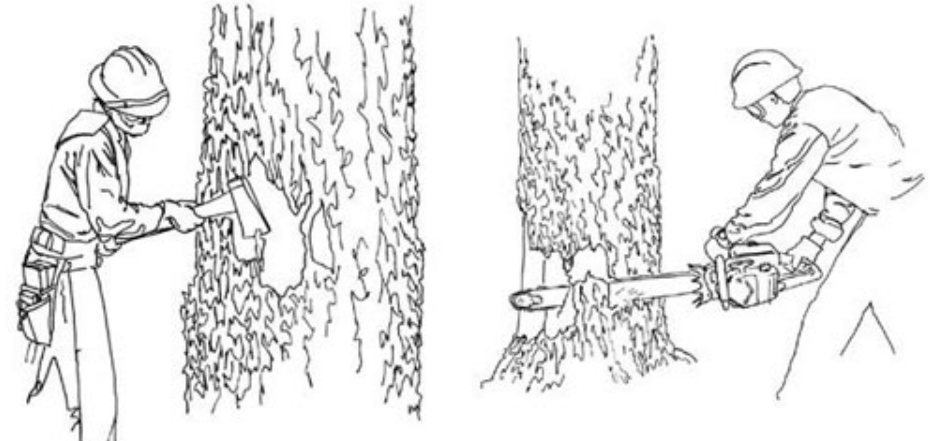
English Spanish Flags layout2

Forestry Input Model

axe
biodiversity
black liquor
bleaching
bole
chainsaw
chlorination
conifer
contaminati
debark
deciduous
ecosystem
erosion
eucalyptus

Project: Ur
Forestr

Entry level
Entry number: 18
Subject Forestry
Graphic



G_Source http://www.coloradofirecamp.com/s-212-chainsaws/glossary_D-F.htm

EN English
Definition To remove bark from trees or logs.
D_Source <http://forestry.about.com/library/glossary/blloggld.htm>

debark
Part of Speech verb
Context There are regional differences on when to debark logs, but it is usually best to leave the bark on as long as possible because it protects the wood from damage in the log yard caused by cranes, loaders and other handling equipment.
C_Source <http://www.timberlinemag.com/articledatabase/view.asp?articleID=1215>

Spanish
Definition Eliminar la corteza de los árboles o troncos.
D_Source <http://www.cohdefor.hn/glosario/de.shtml>

descortezar
Part of Speech verb
Context Aunque un árbol puede tardar muchos años en alcanzar un tamaño considerable, con los métodos de la industria forestal moderna se puede cortar, desmochar y descortezar un árbol recto, como el pino, en pocos minutos.
C_Source http://www.asturtalla.com/Curso_de_talla_02.htm

Done

start 2 Int... tlpask... 5 Mic... Adobe... 3 Mic... SDL M... EN Norton 4:22 PM



What is workflow?



- **Workflow** is the operational aspect of a work procedure:
 - ◆ how *tasks* are structured
 - ◆ what their relative order is
 - ◆ how they are synchronized
 - ◆ who performs them (& where)
 - ◆ how information flows to support the tasks
 - ◆ and how tasks are being tracked

<http://en.wikipedia.org/wiki/Workflow>



Macrostructures & Task Structures



- Terminology workflow (TWF) as a task-set in overall enterprise workflow
- Terminology management (TMM) mission within the enterprise
- Workflow parameters
- Task issues
- Input issues
- Output (terminology products)



Workflow Models



- ❖ TWF as non-detailed item in overall workflow
- ❖ TWF plotted as task set during project planning
- ❖ TWF plotted independently for individual projects
- ❖ TWF plotted for ongoing activities involving ongoing global TMM
- ❖ Project-oriented workflow vs. ongoing global enterprise-related terminology management



Critical Workflow Parameters



- **Dimensions:**
 - ◆ **Tasks (activities)**
 - **Time**
 - How much time does it take?
 - How much time do you have?
 - **Throughput**
 - ◆ **Resources**
 - **Human**
 - **Tools**
 - **“Raw Materials”**



Critical Workflow Parameters



✦ Dimensions:

- ✦ **Cost accounting as a function of workflow management**
 - **Difficult to calculate because of individual task differences**
 - **Balance between intensity & granularity of TMM vs. the investment costs incurred for “doing” TMM**
 - **Balance between investment and incidence of reuse**
 - **Balance between investment and criticality of terminological accuracy**



Project Inputs



- Raw source text
 - ◆ Frequently flawed (particularly in localization environments)
 - ◆ Subject to ambiguity: polysemy, synonymy
 - ◆ Term extraction from the source text
 - Human or automatic term recognition
 - Synonym identification and disambiguation
 - Co-occurrence issues with advanced automatic indexing software
 - ◆ Determining target language equivalents
 - Verification issues as a function of workflow
 - Issues with client-side terms in multilingual projects



Order & Synchronization



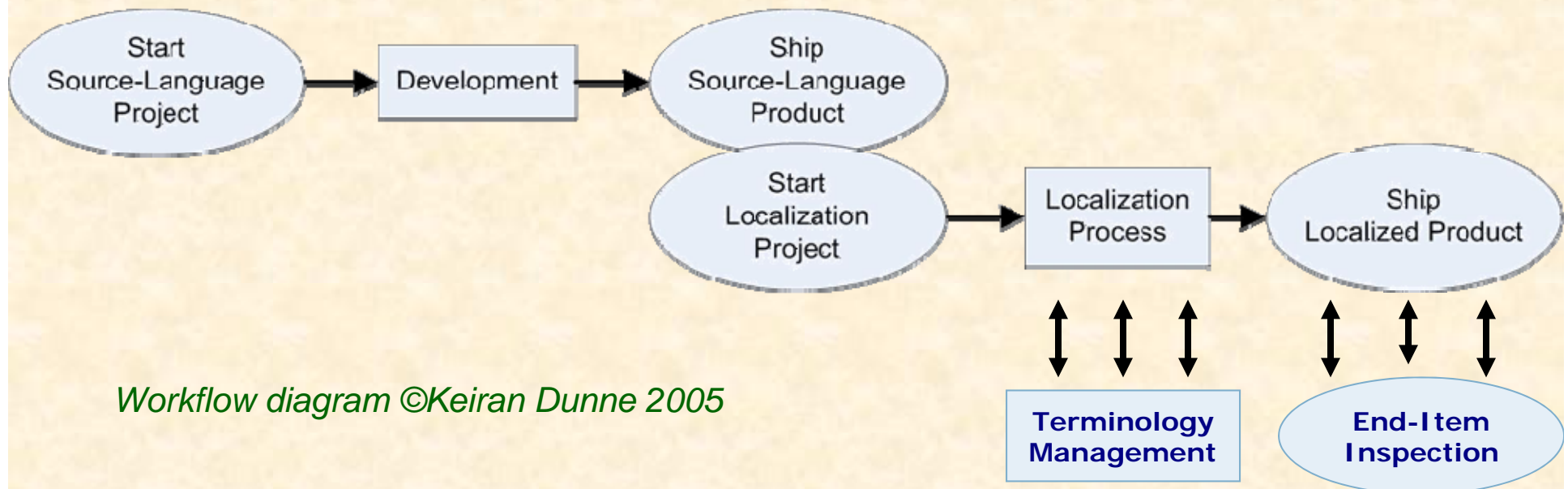
- Traditional TMM workflow position
- Rationalized workflow position
- Global TMM solutions
- Systematic TMM
- Information feedback loop



Traditional Workflow



- Ad hoc TMM
- Reactive project-specific TMM
- No influence on document production, i18n



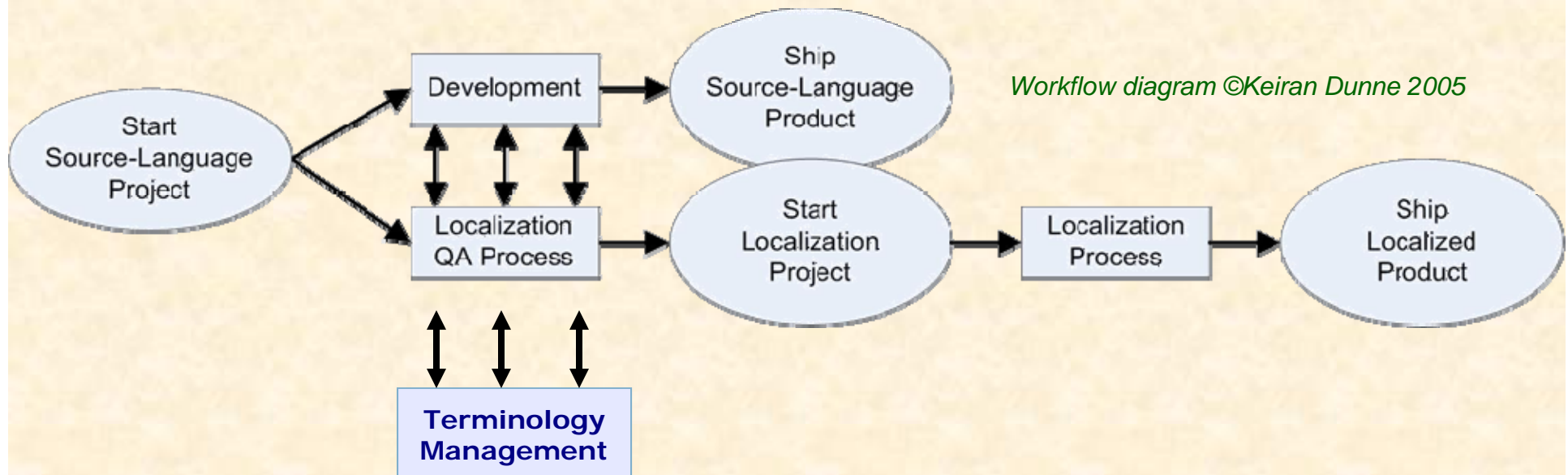
Workflow diagram ©Keiran Dunne 2005



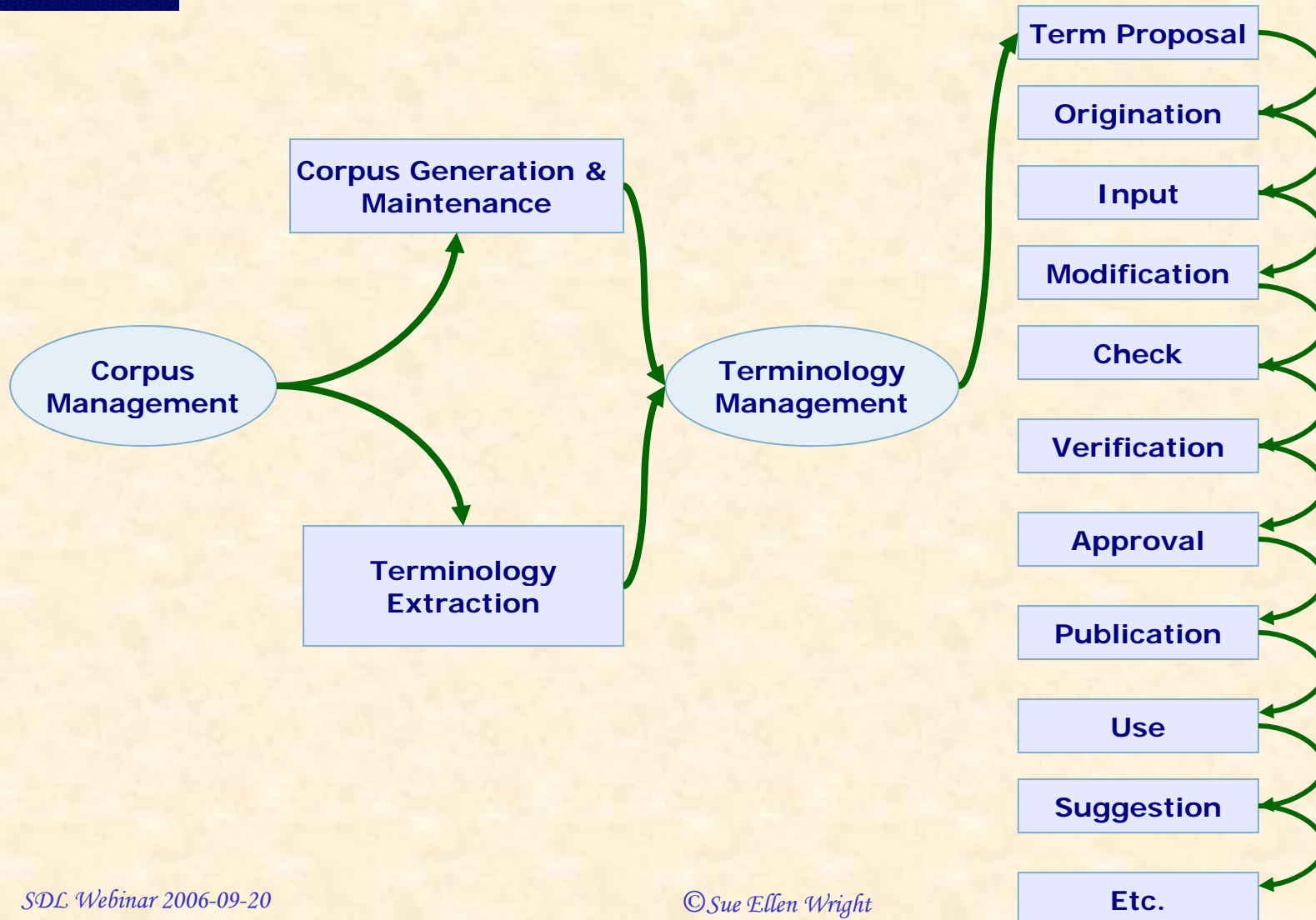
Rationalized Project-Oriented TMM



- TMM as a function of QA (Quality Assurance) management
- TMM and QA upstreamed to planning stage
- Proactive Just-in-Time (JIT) TMM

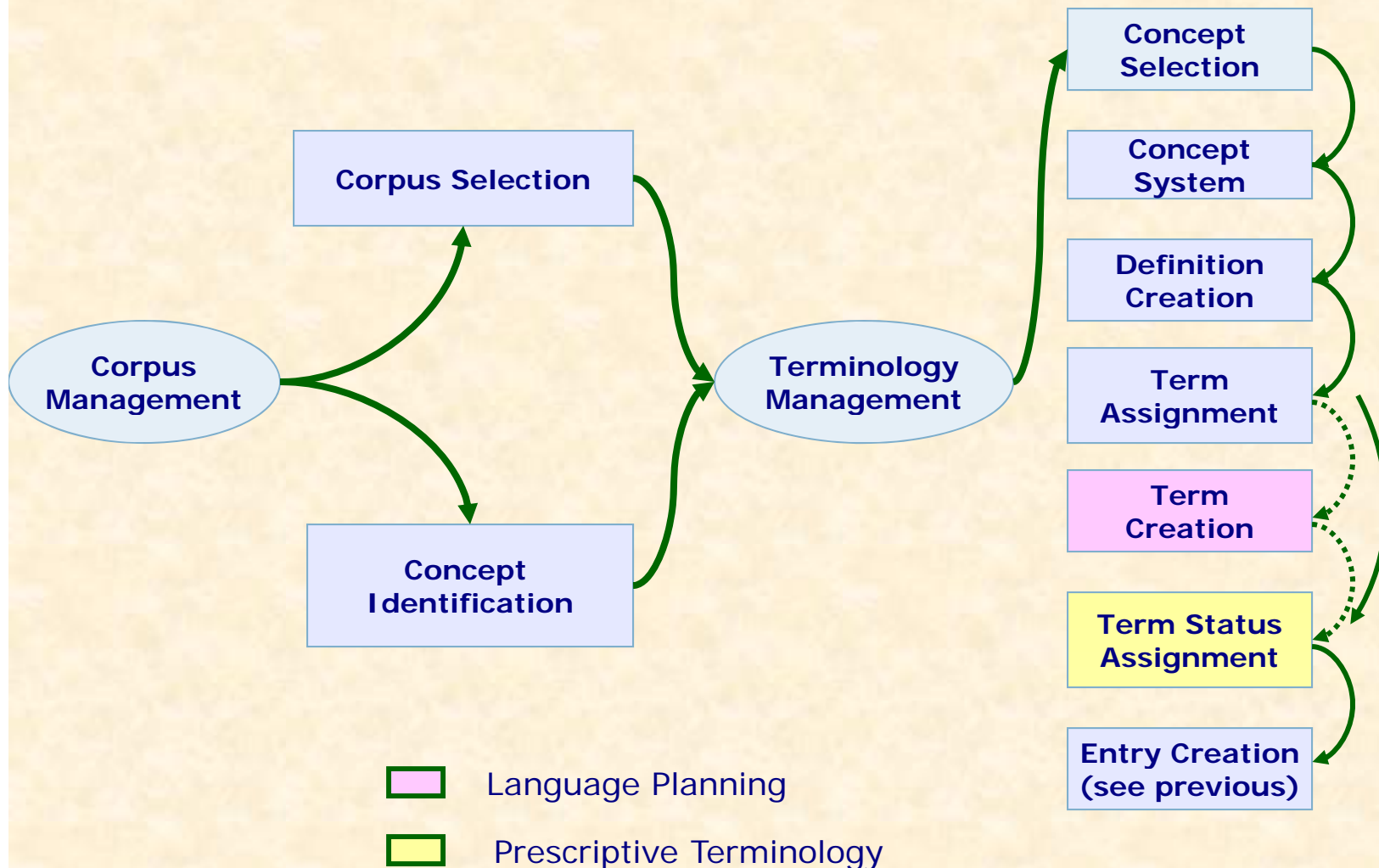


Global Enterprise Terminology Management Solutions



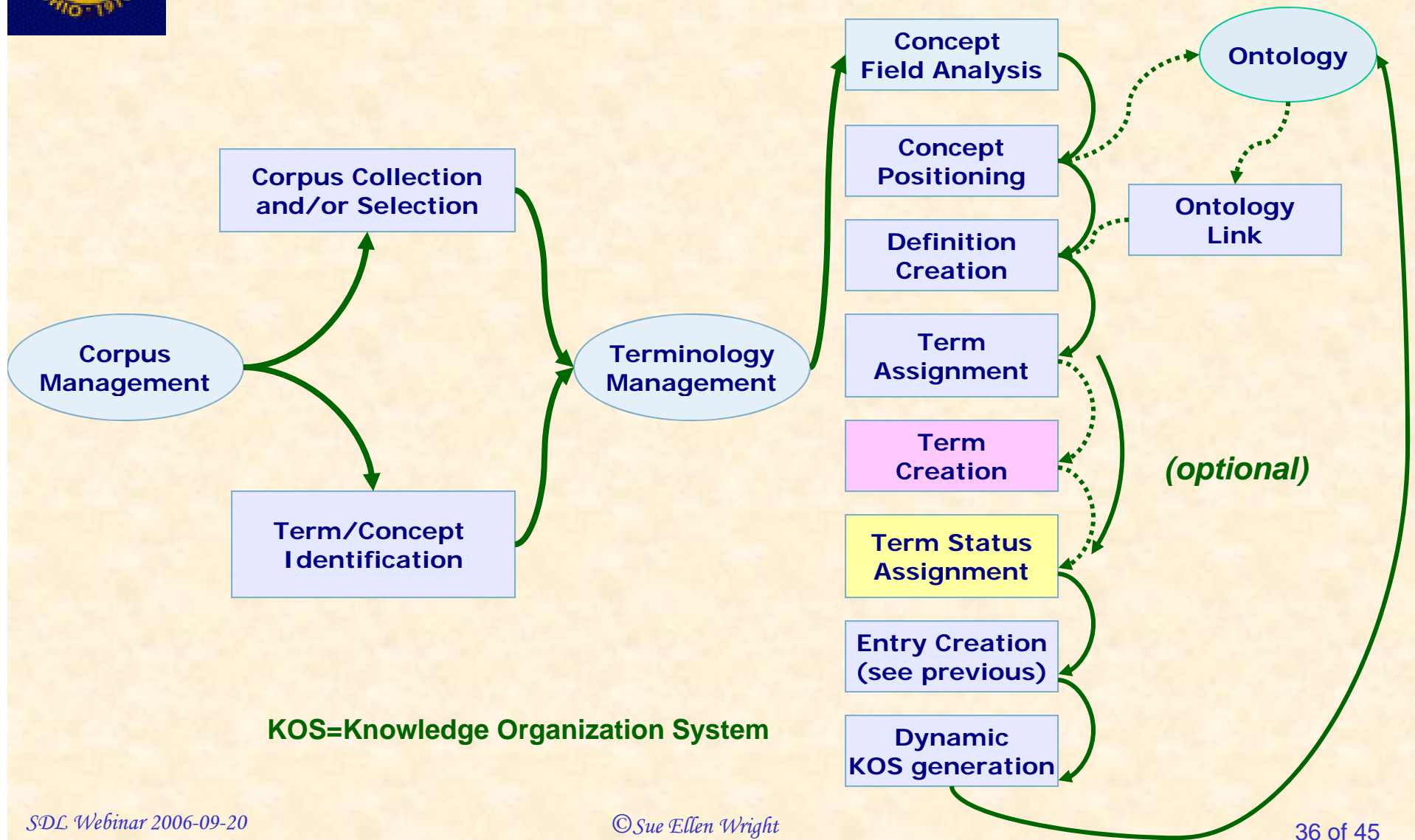


Systematic Term Management for Collections of Controlled Size



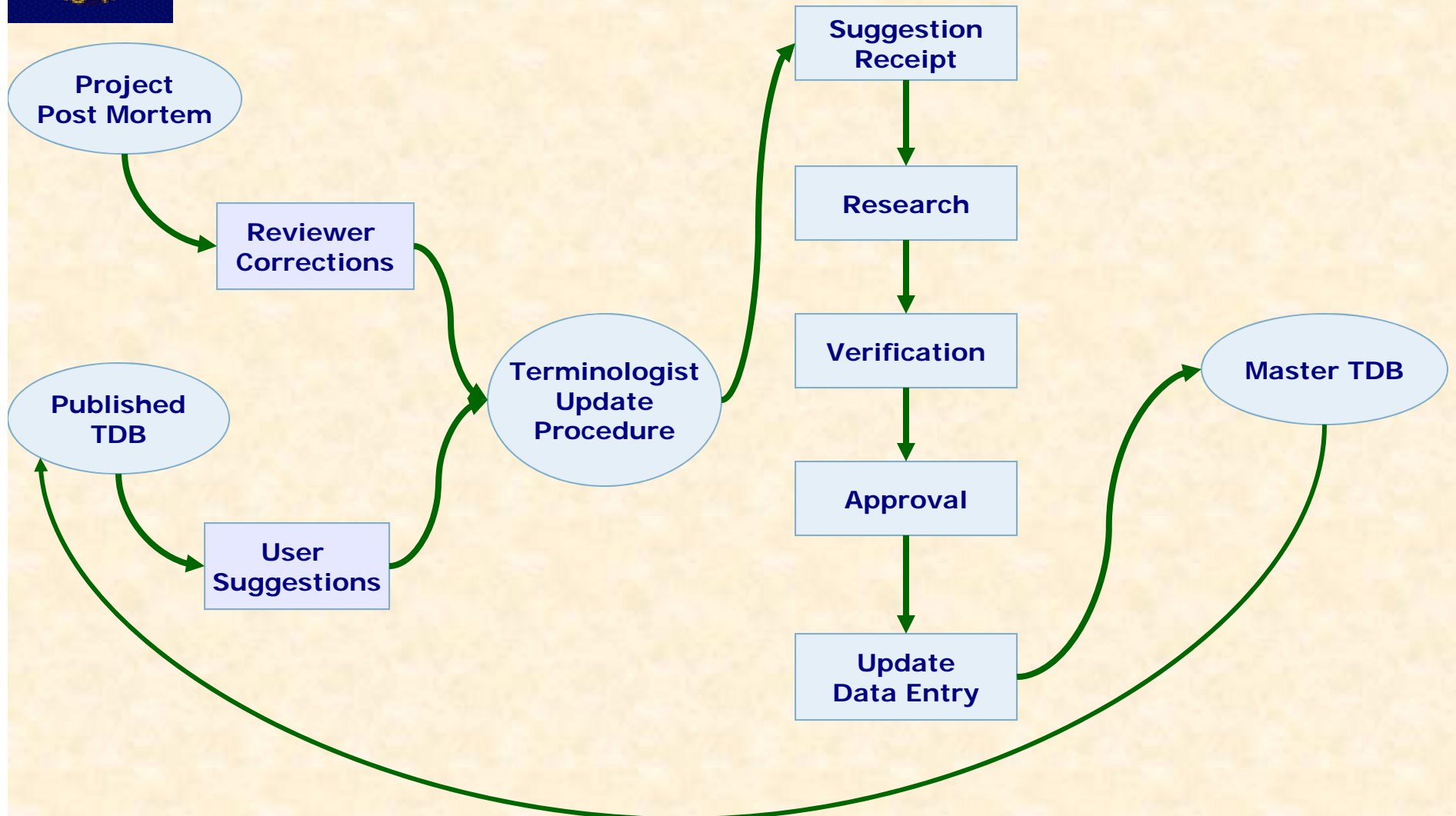


Dynamic Systematic Management





Info Feedback Loop





Who Does What



- Project developers (engineers, designers, other subject experts)
 - ◆ Terminological inconsistencies
 - ◆ Uncontrolled coinage of neologisms
- Terminology project group representative
 - ◆ Technical writers
 - Frequently unaware of multilingual issues
 - Language specific problems
 - ◆ Translators
 - ◆ Knowledge engineers
 - ◆ Terminologists



Who Does what



- Master terminologist
 - ◆ Authority to change records in master file
 - ◆ Authority for final approval
- Language-specific terminologist
 - ◆ Source and target language research
- Data input specialist
- Subject-field specialist
 - ◆ Source language
 - ◆ Target language
- Conflation and expansion of tasks according to need



Special Outputs



- Total resource output to Intranet or LAN
- Project-specific term entry subsetting and exportation (electronic)
- Selective subset delivery, especially to freelance developers, tech writers, and translators
- Output of SL→TL glosses in hard copy
 - ◆ For subject area specialists in multilingual environments
 - ◆ For interpreters



Technical Writers and Translators



- Translators: search for equivalent terms based on source language terms and concepts
- Technical writers: search for source language terms based on developers' concepts
 - ◆ Frequently flawed
 - ◆ Frequently newly coined terms and concepts
- Support to technical writers using controlled authoring tools



Technical Writers & Program Developers



- Ontology-enabled
- Synonym classification
- Push-technology to deliver the correct terms
- Style sheets
- Automatic term use checkers
- Issues:
 - ◆ Resistance to synonymy control
 - ◆ Coinage of translation-unfriendly terms
 - ◆ Coinage of “cute” unmotivated terms



Translation-oriented Solutions



- Terminology database integrated with translation memory and text production software
- Concordance features for access to terms in context
- Multiple layers of information
 - ◆ Data category subsetting for rapid information retrieval
 - ◆ Optional view of full terminological entries



Issues of Ownership



- Client based
 - ◆ Client control over strategic knowledge assets
 - ◆ Client control over quality and information input
 - ◆ Client flexibility to change vendors
- Service-provider based
 - ◆ Locking in the client
 - ◆ Content knowledge on the client side, language knowledge on the vendor side



Term Management Standards



- ISO 12620:2009, *Computer applications for terminology – Data Categories*
 - ◆ Governs <http://www.isocat.org>, an online Data Category Registry
- ISO 26162:2009, *Systems to manage terminology, knowledge and content — Design, implementation and maintenance of Terminology Management Systems*
 - ◆ Designing your own system.



For More Information



- ✦ Sue Ellen Wright
Institute for Applied Linguistics
Kent State University
109 Satterfield Hall
Kent, Ohio 44242, USA

sellenwright@gmail.com