




The world of metadata

How to create and where to use
metadata in memoQ



30 May – 1 June 2018 | Budapest, Hungary

Agenda

- What is metadata
 - Examples of some kinds of metadata available in memoQ
 - Where metadata can be used and how
 - Examples of metadata exchange – where it works well or not so well
-
- 

What is metadata and what is it used for?

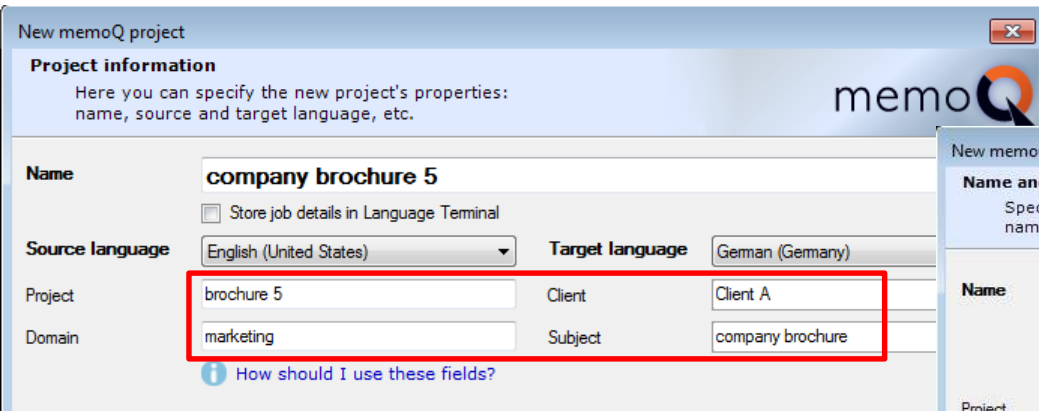
Data about data

- In the case of translation tools, metadata is additional information on entries in a database, on a project or on a database container, like a TM or on the bilingual file.
- It is used for
 - **Categorizing** entries, like segments, terms, LiveDocs documents
 - **Searching and Sorting** projects, database containers, entries...
 - **Filtering** the content of a database for specific entries
 - **Decision** making
 - Helping the translator to decide whether the match is a useful one or how much they can trust the translation (name of reviewer).

Categorization

Project level

- Pre-defined fields: Project, Domain, Client, Subject



New memoQ project

Project information
Here you can specify the new project's properties: name, source and target language, etc.

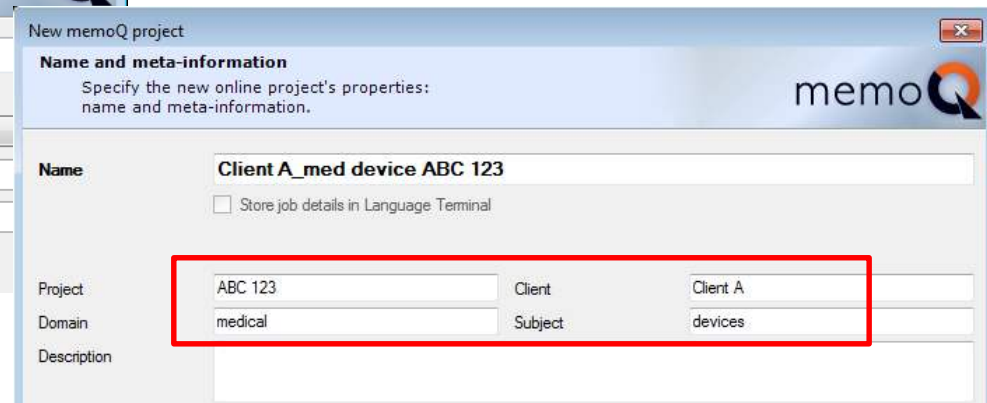
Name

Store job details in Language Terminal

Source language **Target language**

Project	<input type="text" value="brochure 5"/>	Client	<input type="text" value="Client A"/>
Domain	<input type="text" value="marketing"/>	Subject	<input type="text" value="company brochure"/>

[How should I use these fields?](#)



New memoQ project

Name and meta-information
Specify the new online project's properties: name and meta-information.

Name

Store job details in Language Terminal

Project	<input type="text" value="ABC 123"/>	Client	<input type="text" value="Client A"/>
Domain	<input type="text" value="medical"/>	Subject	<input type="text" value="devices"/>

Description

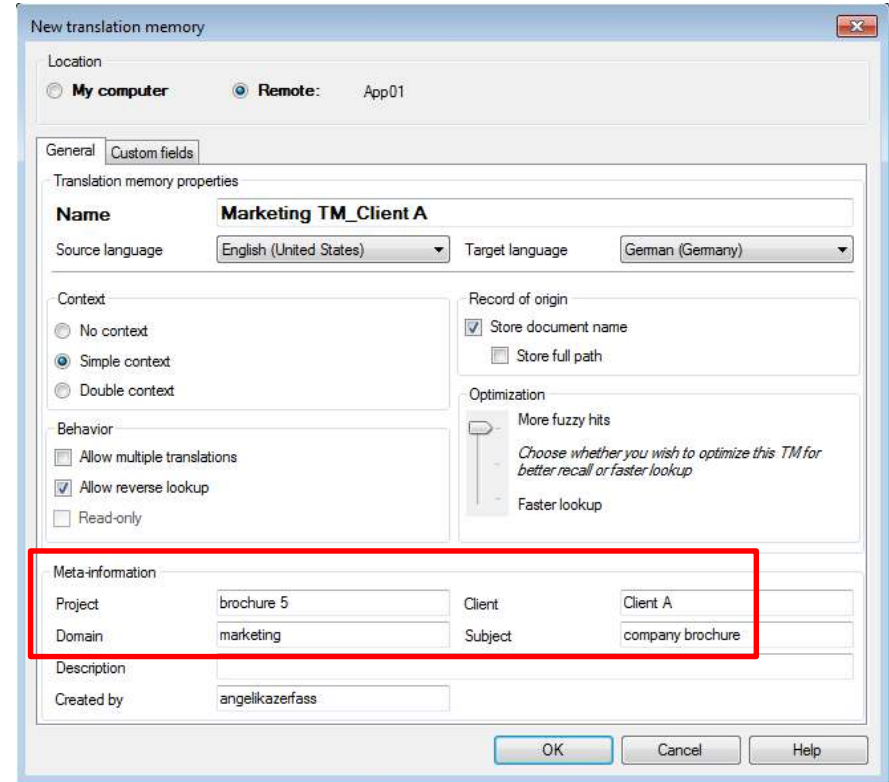
-> **Entry level**

- Information will be saved with each term and segment

Categorization

Project level

- Pre-defined fields: Project, Domain, Client, Subject are filled in when you create a resource like a TM or term base in the project.



The screenshot shows the 'New translation memory' dialog box. The 'Location' is set to 'Remote: App01'. The 'Name' is 'Marketing_TM_Client A'. The source language is 'English (United States)' and the target language is 'German (Germany)'. The 'Context' is set to 'Simple context'. The 'Behavior' section has 'Allow reverse lookup' checked. The 'Record of origin' section has 'Store document name' checked. The 'Optimization' section has 'More fuzzy hits' selected. The 'Meta-information' section is highlighted with a red box and contains the following fields:

Project	brochure 5	Client	Client A
Domain	marketing	Subject	company brochure
Description			
Created by	angelikazerfass		

-> **Container level**

- Information describes the container

Levels

Different levels of metadata

- Project data
 - Data to be saved with terms and segments
 - Find projects in the project list
 - Entry data
 - Find segments, terms in the databases
 - Apply penalties on TM segments
 - Container data
 - Attach correct container to project via template
-

Levels

Sorting and finding projects

- Type meta data field name... (pop-down list of categories)
- Sort the project list with the metadata columns

Dashboard

App01 Forget login Manage logins Vendor availability

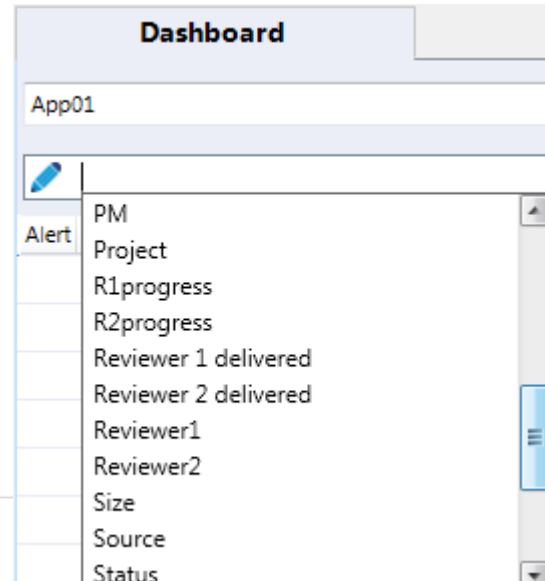
cl All projects

Alert Type	Project	Domain	Client	Subject	Size
px-cc-2018-03-19	px	cc			58
multilingual xml-2017-08	multilingual...				16
demo-2017-02-10	demo				621
Client A_med device ABC 123	ABC 123	medical	Client A	devices	0
P1	12345				198
Server-Demoprojekt	123			Training	735
Demo Project	123			software...	324
Demo	123			Training	232

Levels

Finding projects

- Type space and select the category from the list (for online projects the list is more extensive)



Levels

Finding segments or terms in the databases

Filtering and sorting

Filter - Entry contents

Source words

Target words

Filter - Modification date

After 22.05.2018 11:19

Before 22.05.2018 11:19

Filter - Meta-info

Project brochure 5 Client

Domain Subject

Author Modified by

Roles TR R1 R2 No role

Entries With tags Without tags All

Sort by

ID

Source text

Target text

Modification date

Sort order

Ascending

Descending

• Clear

OK

Cancel

Help

Filter and sorting

Filter - Text criteria

Language

Text

Exact match

Language	Text	Exact match
----------	------	-------------

Add

Remove

Sort by

Source text

Target text

Modifier

Modification date

Domain

Subject

Needs moderation

Internal

Sort order

Ascending

Descending

Reset

Accept

No filtering/sorting

Help

Filter - Modification date

After 22.05.2018 11:23

Before 22.05.2018 11:23

Filter - Meta-info

Project Client

Domain marketing Subject

Author

Note Show all

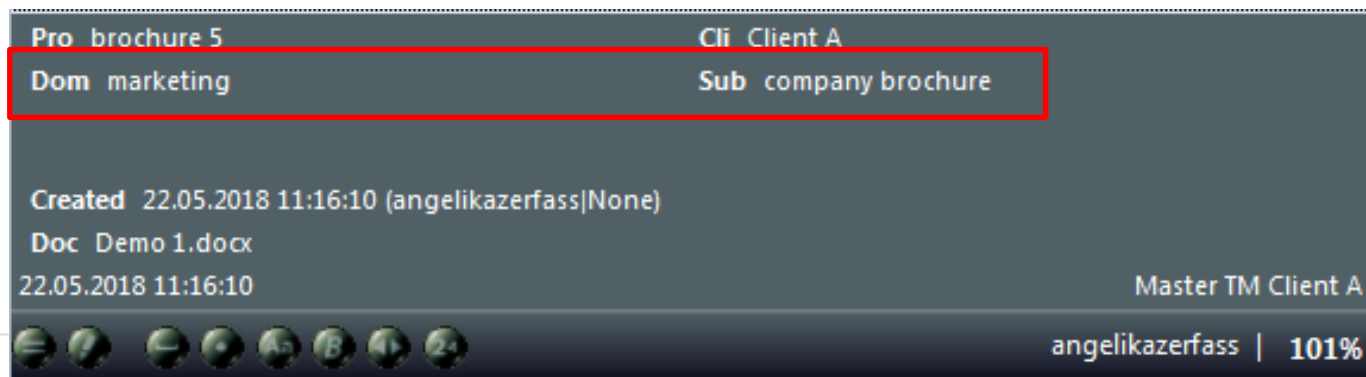
Needs moderation

Has forbidden term(s)

Levels

Deciding on the right match

- Categories can help a translator to find the most useful match.
- For example: a suggestion coming from the translation of a marketing piece might have to be re-phrased for the translation of a contract – even though the source segment is a 100% match.

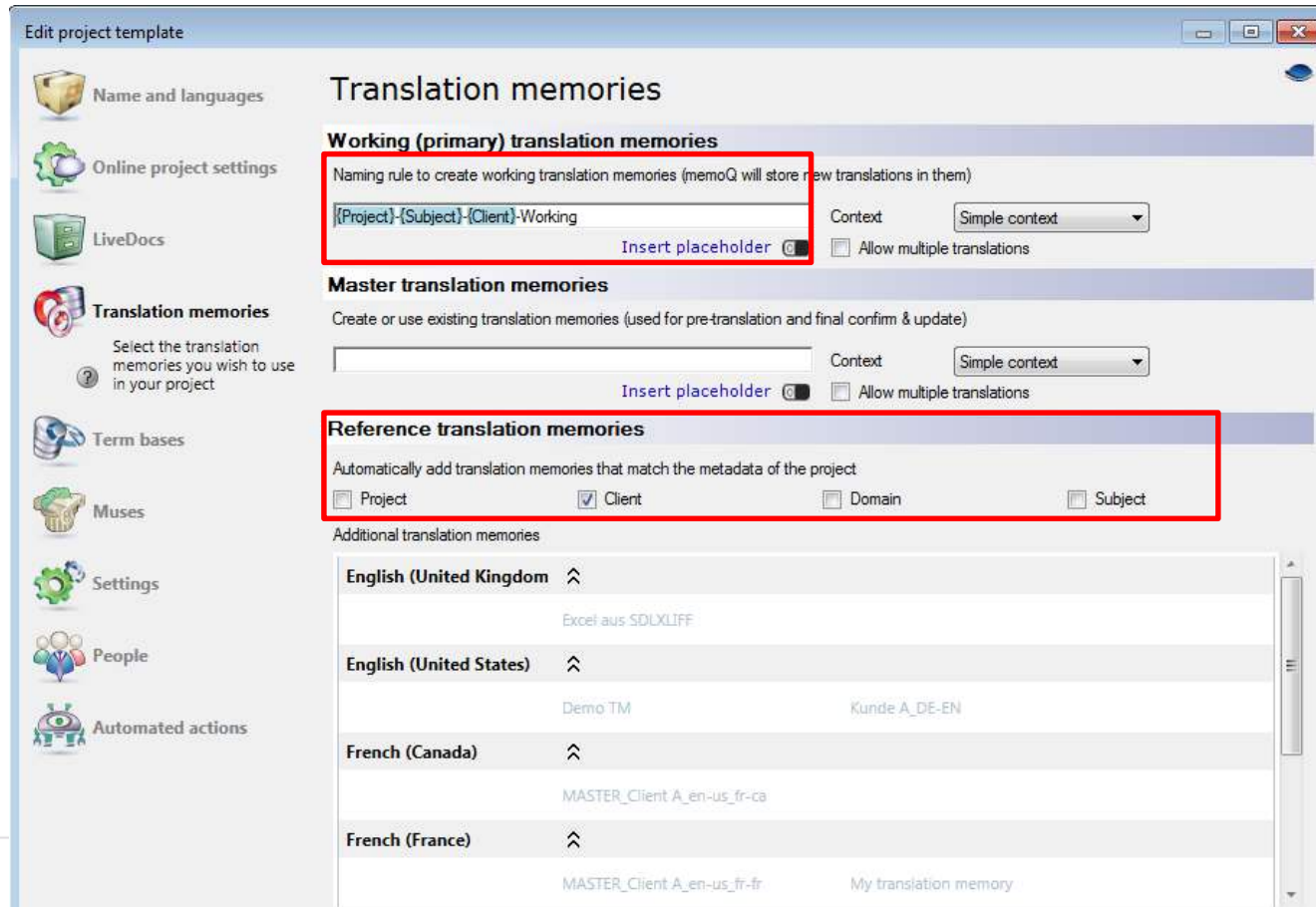


Naming projects or containers

- Categories can be used to name the projects or containers consistently.

Attaching containers

- If the metadata of the project corresponds to the metadata of the container!



Edit project template

Translation memories

Working (primary) translation memories

Naming rule to create working translation memories (memoQ will store new translations in them)

{Project}-{Subject}-{Client}-Working

Context: Simple context

Allow multiple translations

Master translation memories

Create or use existing translation memories (used for pre-translation and final confirm & update)

Context: Simple context

Allow multiple translations

Reference translation memories

Automatically add translation memories that match the metadata of the project

Project Client Domain Subject

Additional translation memories

English (United Kingdom)	Excel aus SDLXLIFF
English (United States)	Demo TM Kunde_A_DE-EN
French (Canada)	MASTER_Client_A_en-us_fr-ca
French (France)	MASTER_Client_A_en-us_fr-fr My translation memory

Custom categories

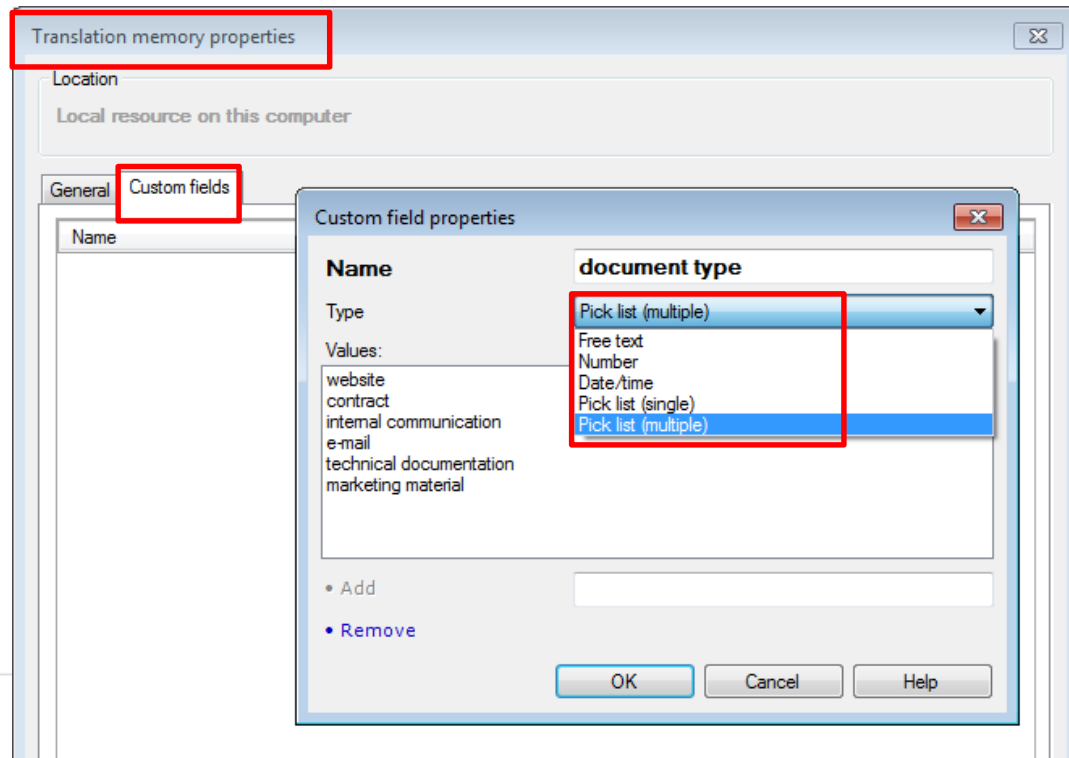
Additional categories in TMs

- The TM allows the user to specify customized metadata fields in addition to the 4 pre-defined project fields.
- The customized fields can be used for
 - Information on segments in the TM
 - Default setup of new TMs

Custom categories

Additional categories in TMs

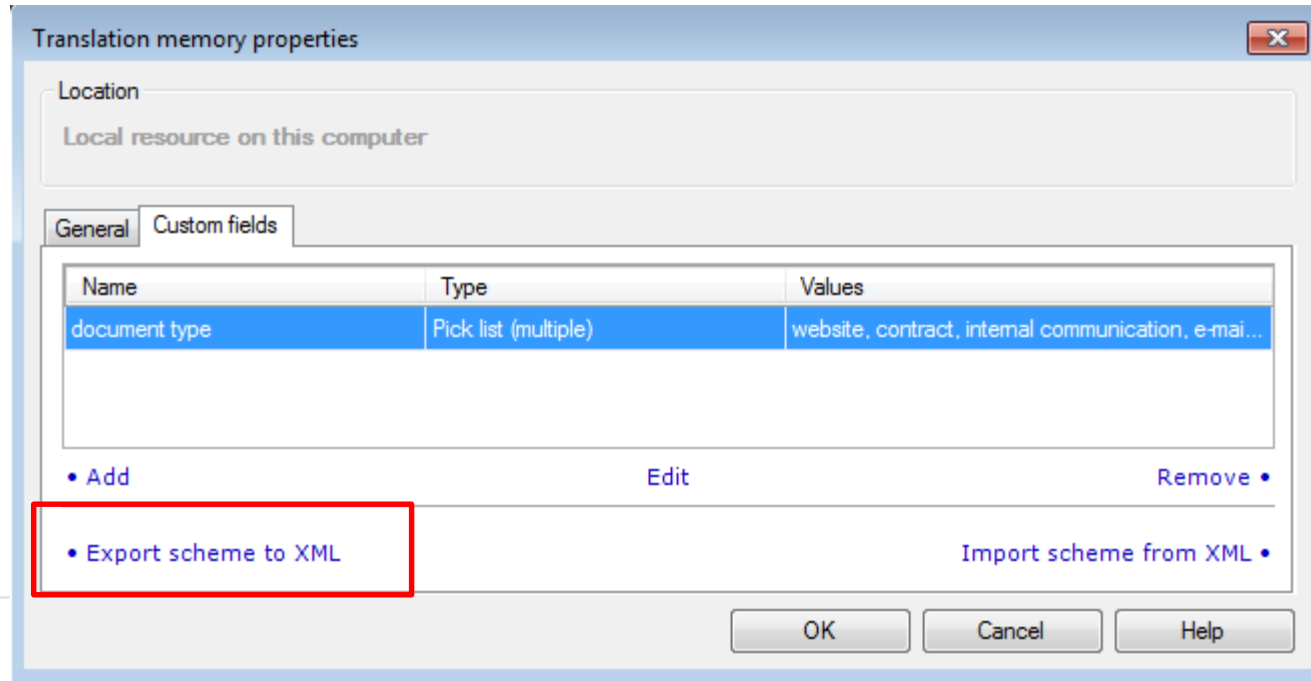
- Different types of fields with free text field, list of pre-defined values or pre-defined content (number/date)



Custom categories

Additional categories in TMs

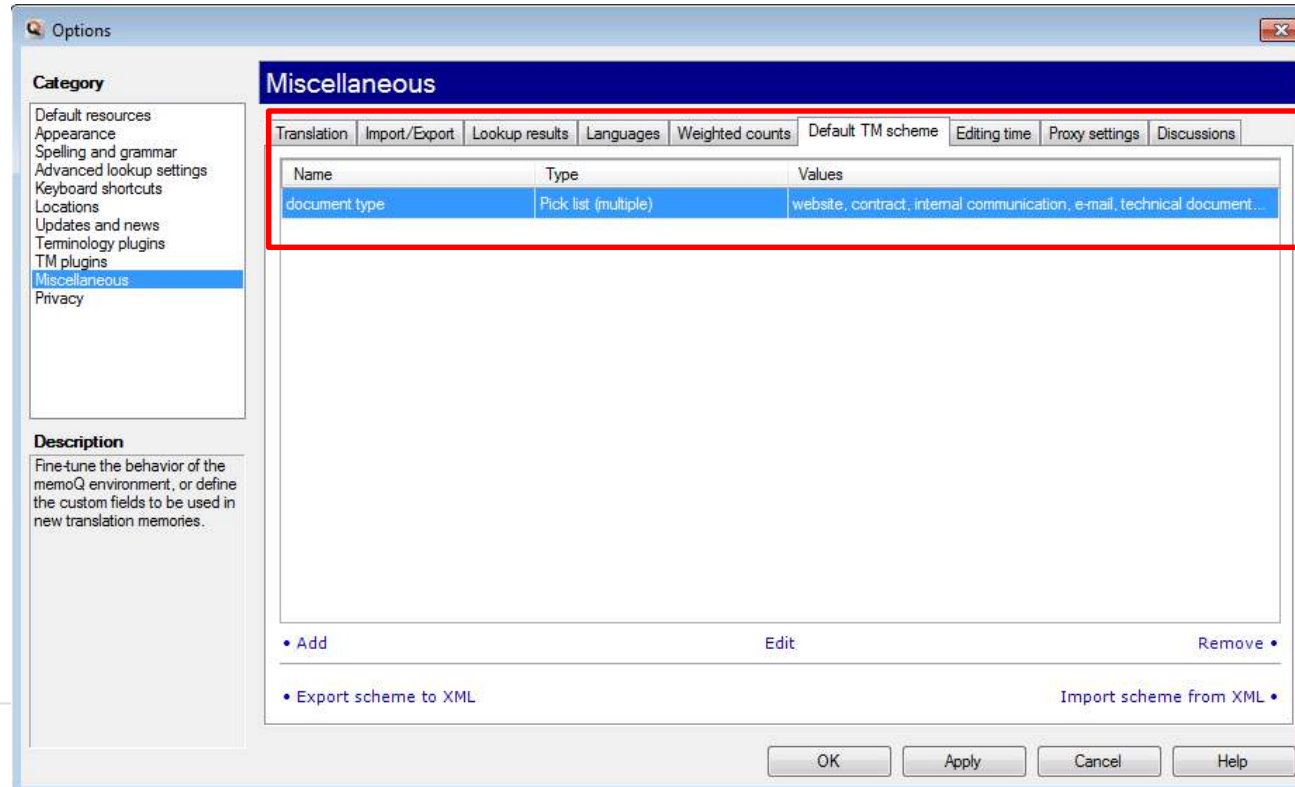
- Export the scheme (list of custom-defined fields)



Custom categories

New TMs with custom fields

- Options – Miscellaneous – Default TM scheme
- Works also for server TMs (or you can clone an existing TM or import the XML scheme)



The screenshot shows the 'Options' dialog box with the 'Miscellaneous' tab selected. A red box highlights the table of custom fields. The table has three columns: Name, Type, and Values. One entry is visible: 'document type' with type 'Pick list (multiple)' and values 'website, contract, internal communication, e-mail, technical document...'. Below the table are buttons for 'Add', 'Edit', and 'Remove'. At the bottom of the dialog are 'OK', 'Apply', 'Cancel', and 'Help' buttons.

Name	Type	Values
document type	Pick list (multiple)	website, contract, internal communication, e-mail, technical document...

System metadata

Date and time / user name / role

- Date and time is used for
 - Creation time: Time of creation of a segment pair, term, project
 - Creation user: User who created the segment pair, term, project
 - Change time: Time when a segment or term was changed
 - Change user: User who saved the existing unit again – not necessarily with a change
 - The actual changes are not saved.
 - Except for
 - the history on changes in an entry of qTerm
 - the Online project history
 - the row history of a segment
-

History data

Project history | Notification history

Who: Document: Language: Refresh history

Who	When	Document	Action	Details
PM1	24.05.2018 15:08:53	Demo 1.docx (g		
revDE	24.05.2018 15:07:36	Demo 1.docx (g		
revDE	24.05.2018 15:05:36	Demo 1.docx (g		
PM1	24.05.2018 15:03:07	Demo 1.docx (g		
trans1	24.05.2018 15:01:43	Demo 1.docx (g		
trans1	24.05.2018 14:57:48	Demo 1.docx (g		
PM1	24.05.2018 14:56:37	Demo 2.docx (g		
PM1	24.05.2018 14:56:37	Demo 2.docx (g		
PM1	24.05.2018 14:56:37	Demo 2.docx (g		
PM1	24.05.2018 14:56:37	Demo 1.docx (g		
PM1	24.05.2018 14:56:37	Demo 1.docx (g		
PM1	24.05.2018 14:56:37	Demo 1.docx (g		
PM1	24.05.2018 14:56:37	Demo 1.docx (g		
PM1	24.05.2018 14:55:31	Demo 2.docx (g		
PM1	24.05.2018 14:55:31	Demo 1.docx (g		

Entry history			
Modified at	User name	Change	Value
5/25/2018 11:43:53 AM	pmaz	+field (eng-US / screen / Abteilung)	Vertrieb
5/25/2018 11:43:53 AM	pmaz	@field (eng-US / screen / Abteilung)	Einkauf => Vertrieb
5/25/2018 11:43:53 AM	pmaz	@term (ger-DE)	Blds. => Bs.
5/25/2018 11:43:53 AM	pmaz	@field (ger-DE / Blds. / Case sensitivity)	Near => Strict
5/25/2018 11:43:53 AM	pmaz	@field (ger-DE / Bildschirm / Abteilung)	Einkauf => Vertrieb
7/1/2016 10:33:44 AM	admin	+field (ger-DE / Blds. / Abteilung)	Einkauf
7/1/2016 10:33:44 AM	admin	+field (ger-DE / Bildschirm / Abteilung)	Einkauf

Row history			
C...	St...	S...	Target segment
<input checked="" type="checkbox"/>			Dies ist ein kurzer, netter Satz.
<input checked="" type="checkbox"/>			Dies ist ein kurzer schöner Satz.
<input checked="" type="checkbox"/>			Dies ist ein kurzer schöner Satz.
<input checked="" type="checkbox"/>			Dies ist ein kurzes schönes Segment.

	Current (modified)		
	24.05.2018 15:14	revDE2	
	24.05.2018 15:14	revDE2	
	24.05.2018 15:07	revDE	
	24.05.2018 14:59	trans1	
	24.05.2018 14:55	PM1	

System metadata

Date and time / user name / role

- Use it for
 - Filtering the entries in a database to export only the filtered part of the database.
 - Exclude matches or apply a penalty to all entries in a TM that were created before a certain date, or by a specific user.
 - Show or hide matches because of the role of the user who saved the segment.
 - On a server, configure the system so that entries that are older than ... are deleted automatically.
-

Filtering/excluding Penalty in the TM (on any metadata)

- Filters to exclude matches altogether or to apply a penalty to a match with specific metadata.

Strict filters

Summary
Subject is different from the project ()

• Add Edit Remove •

Soft filters

Summary	%
Created by is any of "Joe"	6




• Add Edit
• Move up
 Use first applicable penalty
 Add up penalties

OK Cancel Help

If the content of the Subject field in the project differs from that of the match, the match will not be shown.

If the metadata of the match equals the filter, the match will receive a penalty.

Penalty or hiding matches

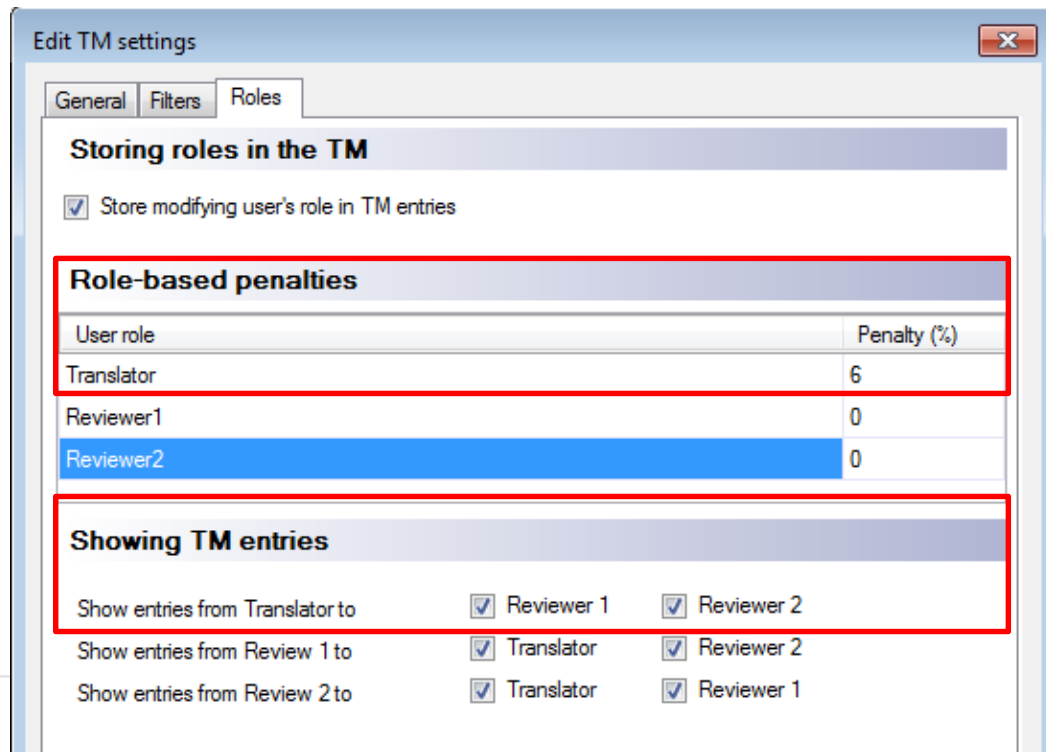
1.	This·is·a·new·sentence.	Dies·ist·ein·neuer·Satz.	0% 
2.	This·is·a·short·new·sentence.	Dies·ist·ein·kurzer,·neuer·Satz.	73% 
3.	This·is·a·short·nice·sentence.	Dies·ist·ein·kurzer,·netter·Satz.	70% 
4.	This·is·a·new·sentence.	Dies·ist·ein·neuer·Satz.	0% 

- Segments saved by users with different roles to the TM
- If the role is saved together with the segment (TM Settings), penalties can be applied depending on the user's role.

System metadata

Penalty in the TM on the user role

- Depending on who saved a segment, the segment will be visible or invisible to certain other users or might be shown only with a penalty.



Edit TM settings

General | Filters | Roles

Storing roles in the TM

Store modifying user's role in TM entries

Role-based penalties

User role	Penalty (%)
Translator	6
Reviewer1	0
Reviewer2	0

Showing TM entries

Show entries from Translator to Reviewer 1 Reviewer 2

Show entries from Review 1 to Translator Reviewer 2

Show entries from Review 2 to Translator Reviewer 1

Category

- Connection
- User management
- Server connections
- Group management
- Language Terminal
- Windows single sign-on
- Default server resources
- CAL licenses
- Configuration and logging
- Web Service interface
- Corpus indexing
- Omniscien Technologies
- Archiving
- Performance counters
- Background tasks
- Broadcast message
- Discussions
- Spelling in memoQWeb
- Weighted counts
- Audit log for TMs**
- Customer Portal
- CMS connections

Audit log for TMs

Configuration

Create audit log for translation memories

<input checked="" type="checkbox"/> TM entry deletion	<input type="checkbox"/> Creation
<input checked="" type="checkbox"/> Export	<input checked="" type="checkbox"/> Deletion
<input type="checkbox"/> Import	<input checked="" type="checkbox"/> Renaming
<input checked="" type="checkbox"/> Modification of TM permissions	<input checked="" type="checkbox"/> Modification of TM properties

E-mail the audit log of the previous day every day

Time: 14:00

E-mail: levente.galbats@kilgray.com

Delete records older than 60 days every month

Day of the month: 5

Scheduled time: 00:00

Save

Description

Export

Export audit log into CSV

From: 22.04.2018 00:00

Until: 22.05.2018 13:13

Export

Delete

Delete records created before: 23.03.2018 00:00

Delete

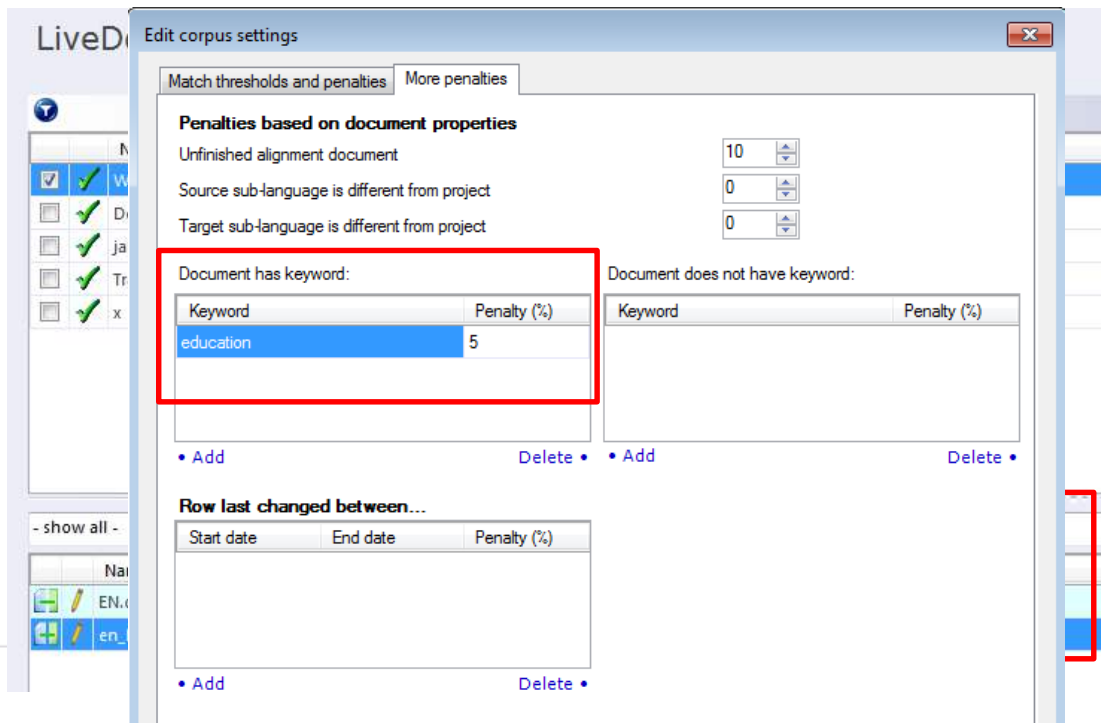
Close

Help

Custom metadata in LiveDocs

Penalty on matches from LiveDocs

- Depending on the keyword that was associated with the material in the LiveDocs, penalties can be applied to matches in the LiveDocs settings.



LiveDocs Edit corpus settings

Match thresholds and penalties More penalties

Penalties based on document properties

Unfinished alignment document 10

Source sub-language is different from project 0

Target sub-language is different from project 0

Document has keyword:

Keyword	Penalty (%)
education	5

• Add Delete •

Document does not have keyword:

Keyword	Penalty (%)
---------	-------------

• Add Delete •

Row last changed between...

Start date	End date	Penalty (%)
------------	----------	-------------

• Add Delete •

Exchanging metadata

Success depends on the specific metadata, the tool combination and the direction

- User name and date/time can be exchanged without issues.
 - Custom fields can be exchanged either directly or with a little tweaking (search&replace) in the TMX files.
 - Tool specific information (segment comes from an alignment, source text has been edited, role of the user who saved the segment...) usually cannot be exchanged between tools.
-

Exchanging metadata

Context information

- Context information will not be usable by other tools because they use a different system or don't save that kind of metadata themselves.

How memoQ saves context information to the TMX

```
<prop type="x-context-pre">&lt;seg&gt;previous sentence&lt;/seg&gt;</prop>  
<prop type="x-context-post">&lt;seg&gt;following sentence&lt;/seg&gt;</prop>
```

How Studio saves context information to the TMX

```
<prop type="x-Context">2977040540754490337, -2182033961215568804</prop>
```

Exchanging metadata

XLIFF

- XLIFF offers a wide range of options to add metadata, like comments, history of a segment, alternative translations from other sources, where the first match came from...)
- Also, XLIFF allowed a lot of customization so that every tools provider created their own flavor of XLIFF.
 - An exchange of metadata like comments etc. is only possible if those XLIFF flavors are specifically supported by the receiving tool.
 - Hopefully XLIFF 2.0 will help with this situation a bit, as it prescribes a core of functionality that needs to be handled in the same way by all tools that produce XLIFF.

Exchanging metadata

XLIFF – Examples for metadata exchange

- Locked segments
 - Comments
 - Status of a segment
 - Match values

 - memoQ to SDL Trados Studio
 - SDL Trados Studio to memoQ
-

memoQ

- Segment with comment (yellow bubble)
- Locked segment (gray line / lock symbol)
- Rejected segment (status field red)
- Segment confirmed as reviewer (double checkmark)



Project home Demo 1.docx x

Path to source: Demo 1.docx












Source Target

Sort No sorting

1.	This is a new sentence.	Dies ist ein neuer Satz.	0%	🔒
2.	This is a short new sentence.	Dies ist ein kurzer, neue Satz.	101%	✅
3.	This is a short nice sentence.	Dies ist ein kurzer, netter Satz.	101%	💬
4.	This is a new sentence.	Dies ist ein neuer Satz.	0%	🔒
5.	Sentence 5 is similar to the next sentence.	Satz 5 ist ähnlich zum folgenden Satz.	100%	✅
6.	Sentence 6 is similar to the previous sentence.	Satz 6 ist ähnlich zum vorherigen Satz.	100%	❌









When opening this XLIFF file from memoQ in Studio:

- All segments have match values and status as in memoQ
- Locked segment appears as locked in Studio as well
- Reviewer 2 confirmed segments are locked as well (depends on the filter settings)
- Comment is not visible
- Rejected segment does not appear as rejected, but as not edited (there is no mapping for this status in the filter)

Demo 1.docx_ger-DE.maxl2 (Translation en-US-de-DE)	
1 This is a new sentence.	  Dies ist ein neuer Satz.
2 This is a short new sentence.	  CM  Dies ist ein kurzer, neue Satz.
3 This is a short nice sentence.	  CM Dies ist ein kurzer, netter Satz.
4 This is a new sentence.	 100%  Dies ist ein neuer Satz.
5 Sentence 5 is similar to the next sentence.	 100% Satz 5 ist ähnlich zum folgenden Satz.
6 Sentence 6 is similar to the previous sentence.	 100% Satz 6 ist ähnlich zum vorherigen Satz.

SDL Trados Studio

- Segment with comment (highlighted text)
- Locked segment (segment grayed out / lock symbol)
- Rejected segment (reject symbol)
- Segment confirmed by reviewer (reviewer confirmed symbol)

Demo 1.docx_en-US_de-DE.sdlxliff [Translation]		
1	This is a new sentence.	  Dies ist ein neuer Satz.
2	This is a short new sentence.	 Dies ist ein kurze, neuer Satz.
3	This is a short nice sentence.	 Dies ist ein kurze, netter Satz.
4	This is a new sentence.	 100%  Dies ist ein neuer Satz.
5	Sentence 5 is similar to the next sentence.	 Satz 5 ist ähnlich zum folgenden Satz.
6	Sentence 6 is similar to the previous sentence.	 Satz 6 ist ähnlich zum vorherigen Satz..

When opening this XLIFF file from Studio in memoQ:

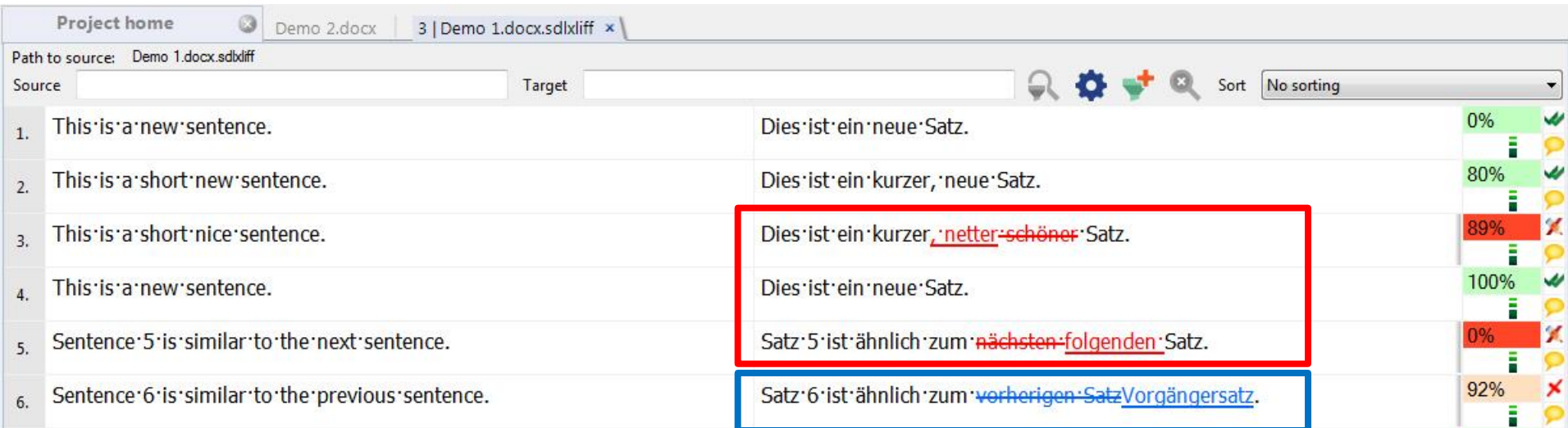
- Comment is visible
- Locked segments appear as locked
- Rejected segment status visible
- Reviewed segment status visible



Source	Target	Progress	Status
1. This is a new sentence.	Dies ist ein neuer Satz.	0%	Locked
2. This is a short new sentence.	Dies ist ein kurze, neuer Satz.	0%	Reviewed
3. This is a short nice sentence.	Dies ist ein kurzer netter Satz.	0%	Reviewed
4. This is a new sentence.	Dies ist ein neuer Satz.	100%	Locked
5. Sentence 5 is similar to the next sentence.	Satz 5 ist ähnlich zum folgenden Satz.	0%	Reviewed
6. Sentence 6 is similar to the previous sentence.	Satz 6 ist ähnlich zum vorherigen Satz.	0%	Rejected

- SDLXLIFF file with tracked changes in memoQ
- Red (tracked changes as they came from Trados Studio)
 - Blue (tracked changes added in memoQ)

With the right settings in the SDLXLIFF filter in memoQ, the tracked changes can be brought back to Trados Studio.







The screenshot shows the memoQ interface with the SDLXLIFF filter applied to a project. The interface displays a list of source sentences and their corresponding target translations. The target translations are color-coded to indicate tracked changes: red for changes from Trados Studio and blue for changes added in memoQ.

Source	Target	Similarity	Tracked Changes
1. This is a new sentence.	Dies ist ein neue Satz.	0%	None
2. This is a short new sentence.	Dies ist ein kurzer, neue Satz.	80%	None
3. This is a short nice sentence.	Dies ist ein kurzer, <u>netter</u> -schöner Satz.	89%	Red
4. This is a new sentence.	Dies ist ein neue Satz.	100%	None
5. Sentence 5 is similar to the next sentence.	Satz 5 ist ähnlich zum <u>nächsten</u> - <u>folgenden</u> Satz.	0%	Red
6. Sentence 6 is similar to the previous sentence.	Satz 6 ist ähnlich zum <u>vorherigen</u> -Satz <u>Vorgängersatz</u> .	92%	Blue

- SDLXLIFF file with tracked changes in memoQ
- All tracked changes have the same colors

Demo 1.docx.sdlxliff [Translation]

1	This is a new sentence.		Dies ist ein neue Satz.	P
2	This is a short new sentence.	 80%	Dies ist ein kurzer neue Satz	
3	This is a short nice sentence.		Dies ist ein kurzer <u>netter</u> schöner Satz.	
4	This is a new sentence.		Dies ist ein neue Satz.	
5	Sentence 5 is similar to the next sentence.		Satz 5 ist ähnlich zum nächsten <u>folgenden</u> Satz.	P
6	Sentence 6 is similar to the previous sentence.	 92%	Satz 6 ist ähnlich zum vorherigen Satz <u>Vorgängersatz</u> .	
7	This is a sentence with a		Dies ist ein Satz mit einer	P
8	line break.		Zeilenschaltung.	P
9	New sentence.		Neuer Satz.	P
a				
9	Next sentence.		Nächster Satz.	
b				
10	Several sentences already contained the fragment „This is a“.		Einige Sätze beinhalteten bereits das Fragment "Dies ist ein".	P












Demo 1.docx

MQXLIFF file with tracked changes in Trados Studio

Tracked changes color change is shown with MRK tags.

4	0%	Dies ist ein neuer Satz.	CXT+
5		Dies ist mrk noch mrk ein anderer Satz.	CXT+
		Schicken Sie eine mrk E- mrk Mail an	CXT+
6		info@company.com mrk mrk	
		info@company.com mrk mrk	
		info@company.com mrk mrk	
		um das Bonus-Material zu erhalten.	
7			CXT+

Tracked changes will show up again after importing the changed file back into memoQ

4.	This is a new sentence.	Dies ist ein neuer Satz.	0%	
5.	This is another sentence.	Dies ist <u>noch</u> ein weiterer Satz.	0%	
6.	Send a mail to   info@company.com  to receive the bonus material.	Schicken Sie bitte eine <u>E</u> -Mail an   info@company.com   um das Bonus- M <u>a</u> aterial zu erhalten.	0%	
7.	mailto:info@company.com		0%	

Thank you for your
attention!