

Assessing the post-editor's effort to correct MT

David Canek, MemSource Technologies

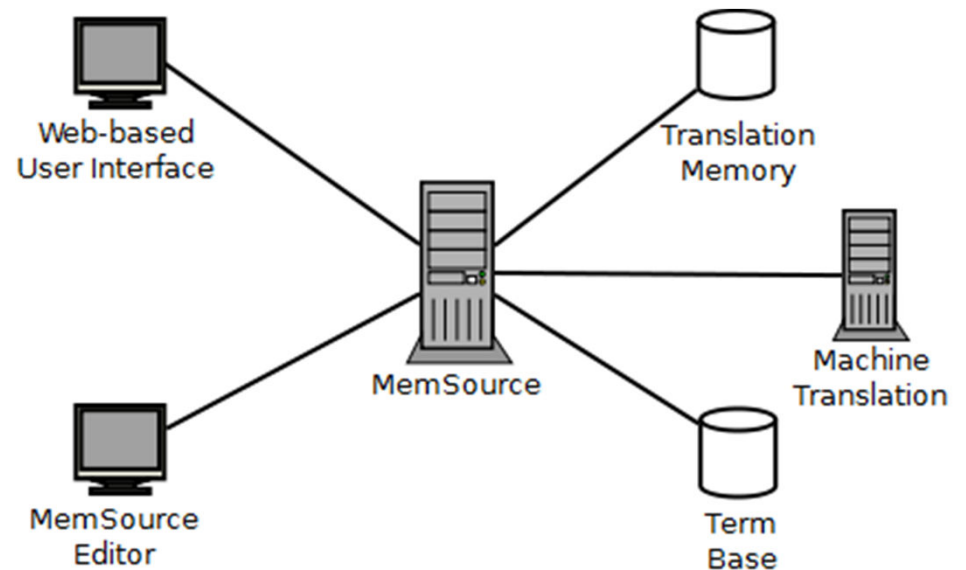
Getting Machine Translation Under Control

Assessing the post-editor's effort to correct MT

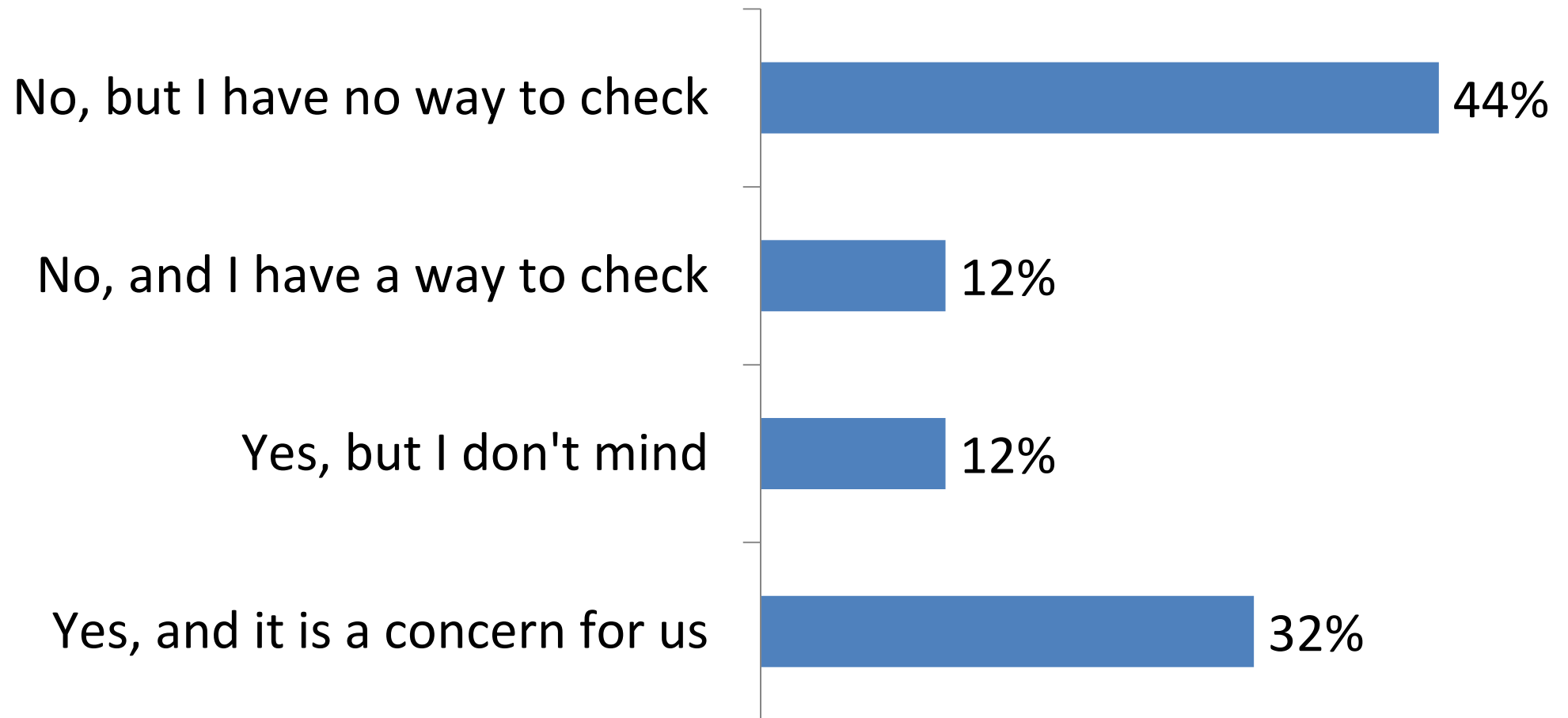
David Canek, MemSource Technologies

About MemSource

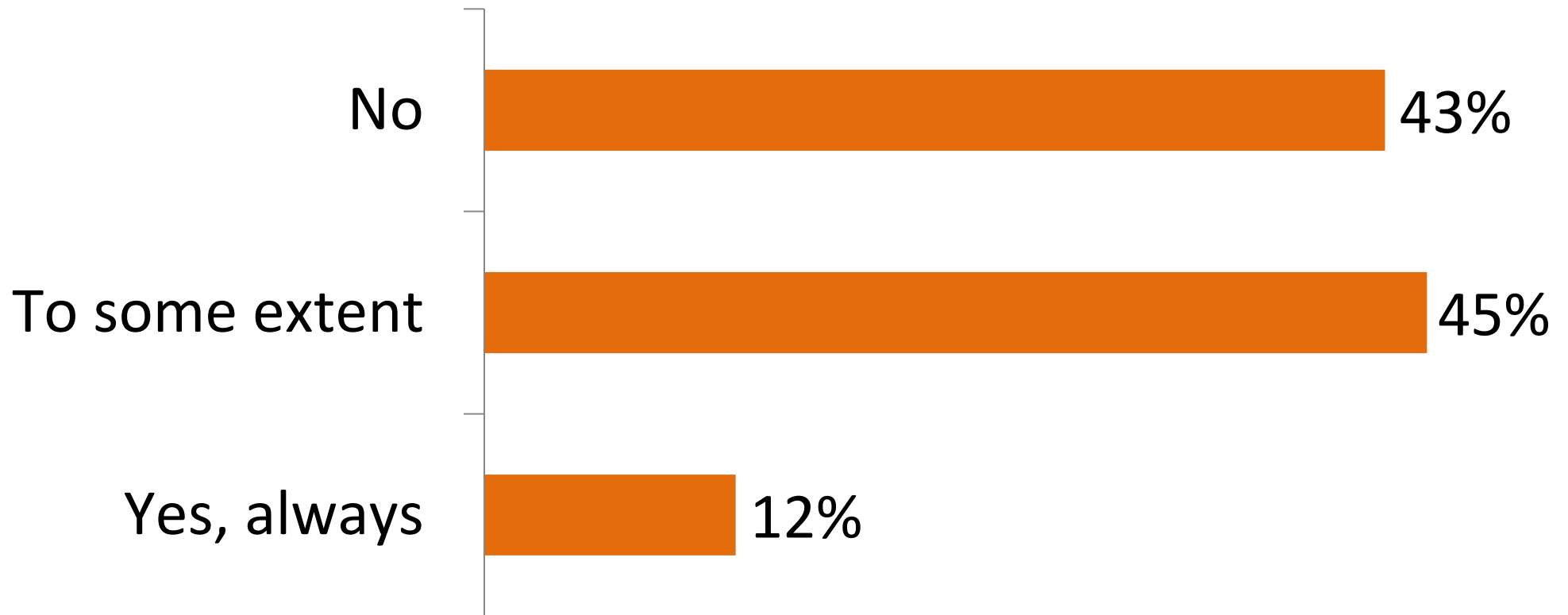
- A complete translation environment
- Founded 2010 in Prague
- Launched 2011
- 1,000+ users in 2012



Do your translators use MT without you knowing about it?



When you use MT, are you able to measure its quality?



MemSource gets MT under control

1. Control **which** MT will be used for a project or **whether** MT will be used at all
2. Measure **post-editing effort** in an automated way at a segment level
3. Determine **MT quality** for any language and any MT engine
4. Produce a **fair compensation** scheme for MT post-editors

The Post-editing Analysis

File: Agreement.docx

	Segments			Words			Characters			%		
	TM	MT	All	TM	MT	All	TM	MT	All	TM	MT	All
After discount	43	1	44	1	18	19	3	124	127			
All	45	3	48	8	32	40	34	213	247	20	80	100
Repetitions	-	-	0	-	-	0	-	-	0	-	-	0
100%	2	2	4	8	10	18	34	64	98	20	25	45
95%–99%	0	0	0	0	0	0	0	0	0	0	0	0
85%–94%	0	1	1	0	22	22	0	149	149	0	55	55
75%–84%	0	0	0	0	0	0	0	0	0	0	0	0
50%–74%	0	0	0	0	0	0	0	0	0	0	0	0
0%–49%	43	0	43	0	0	0	0	0	0	0	0	0

The Post-editing Analysis

File: Agreement.docx

	Segments			Words			Characters			%		
	TM	MT	All	TM	MT	All	TM	MT	All	TM	MT	All
After discount	43	1	44	1	18	19	3	124	127			
All	45	3	48	8	32	40	34	213	247	20	80	100
Repetitions	-	-	0	-	-	0	-	-	0	-	-	0
100%	2	2	4	8	10	18	34	64	98	20	25	45
95%–99%	0	0	0	0	0	0	0	0	0	0	0	0
85%–94%	0	1	1	0	22	22	0	149	149	0	55	55
75%–84%	0	0	0	0	0	0	0	0	0	0	0	0
50%–74%	0	0	0	0	0	0	0	0	0	0	0	0
0%–49%	43	0	43	0	0	0	0	0	0	0	0	0

The Post-editing Analysis

File: Agreement.docx

	Segments			Words			Characters					
	TM	MT	All	TM	MT	All	TM	MT	All	TM	MT	All
After discount	43	1	44	1	18	41	3	5	124			
All	45	3	48	8	32	40	34	213				
Repetitions	-	-	0	-	-	0	-	-				
100%	2	2	4	8	10	18	34	64				
95%–99%	0	0	0	0	0	0	0	0				
85%–94%	0	1	1	0	22	22	0	149				
75%–84%	0	0	0	0	0	0	0	0	0	0	0	0
50%–74%	0	0	0	0	0	0	0	0	0	0	0	0
0%–49%	43	0	43	0	0	0	0	0	0	0	0	0

18 words for translation instead of 32