Which of the following is the annotation project (where you used AL) you are reporting on? (Select one alternative)

What kind of annotation had to be done? (Multiple answers allowed)

02/09/2009	298730	disambiguation

From which domain(s) where the annotated instances? (Multiple answers allowed)

02/06/2009	274047	web pages
02/09/2009	298730	patents
02/10/2009	311756	general
02/16/2009	347483	medical
02/20/2009	389148	WordNet synsets

What was the principal language? (Select one alternative)

How many person months were spent on annotation?

02/05/2009	270115	24
02/06/2009	274047	36
02/06/2009	274516	2
02/06/2009	277993	6
02/09/2009	298730	1
02/10/2009	310840	1
02/10/2009	311756	3
02/16/2009	347483	6
02/20/2009	386472	4
02/20/2009	389148	12
02/26/2009	444514	0.5

How many annotators have been involved?

02/05/2009	270115	12
02/06/2009	274047	5
02/06/2009	274516	4
02/06/2009	277993	4
02/09/2009	298730	1
02/10/2009	310840	1
02/10/2009	311756	6
02/16/2009	347483	4
02/20/2009	386472	1
02/20/2009	389148	10
02/26/2009	444514	1

What was the size of the resulting, labeled corpus in terms of annotated instances (please provide the annotation instance type in brackets, e.g., "1000 (tokens)") ?

02/05/2009	270115	65,000 (noun tokens)
02/06/2009	274047	thousands of phrases
02/06/2009	274516	63000 (tokens)
02/06/2009	277993	5000 named entities
02/09/2009	298730	100
02/10/2009	310840	100 features
02/10/2009	311756	3000 (words)
02/16/2009	347483	100K
02/20/2009	386472	200000 tokens
02/20/2009	389148	120K WordNet synsets
02/26/2009	444514	187 medline abstr.

How did you assemble the corpus of unlabeled annotation instances (i.e., the pool
from which your AL mechanism selected)? (Select one alternative)

02/05/2009	270115	Subsample based on sources
02/20/2009	389148	All WN synsets

What was the size of this unlabeled corpus in terms of annotatation instances
(please provide the annotation instance type in brackets, e.g., "1000 (tokens)") ?

02/05/2009	270115	40,000 (noun tokens)
02/06/2009	274047	7000 web pages
02/06/2009	274516	63000 (tokens)
02/06/2009	277993	1000 reports
02/09/2009	298730	10000
02/10/2009	310840	1000000 sentences
02/10/2009	311756	cant remember
02/16/2009	347483	10K
02/20/2009	386472	8000 tokens
02/20/2009	389148	100 mio
02/26/2009	444514	100000 medline abs

What kind(s) of active learning approach have you used? (Multiple alternatives allowed)

02/05/2009	270115	Unusual variant: see paper
02/06/2009	274047	human assessment
02/06/2009	277993	increase in variability (words and combinations of words we hadn't seen before but played similar semantic roles as others in our training set)
02/10/2009	310840	batch mode
02/10/2009	311756	informal

What was the base learning algorithm used (e.g., decision trees, Maximum Entropy)?

02/05/2009	270115	MaxEnt
02/06/2009	274047	SVM
02/06/2009	274516	ME
02/06/2009	277993	simple log score
02/09/2009	298730	svm
02/10/2009	310840	svm
02/10/2009	311756	informal
02/16/2009	347483	svm
02/20/2009	386472	Maximum Entropy
02/20/2009	389148	boosting
02/26/2009	444514	crf

How did you decide when to stop the annotation process? (Select one alternative)

02/05/2009	270115	Various; see paper

What was your primary motivation to use AL? (Select one alternative)

02/09/2009	298730	more trainings documents to improve svm based classification

If not, please describe the problems/shortcomings.

02/05/2009	270115	
02/06/2009	274047	
02/06/2009	274516	
02/06/2009	277993	bad experimental design; couldn't show it helped—needed to start over and didn't have time/resources to do that
02/09/2009	298730	
02/10/2009	310840	We could not perform anything other than batch mode labeling because of the computational complexity of ALL incremental active learning methods.
02/10/2009	311756	What we did in WASPS probably can't properly be called Active Learning but shares many characteristics
02/16/2009	347483	
02/20/2009	386472	
02/20/2009	389148	
02/26/2009	444514	Needed to hand-select more abstracts to fullfull our needs

Please estimate the reduction of annotation effort in terms of *annotation units in percent* (leave blank if you cannot give an estimate).

02/05/2009	270115	35%
02/06/2009	274047	
02/06/2009	274516	
02/06/2009	277993	
02/09/2009	298730	
02/10/2009	310840	
02/10/2009	311756	70
02/16/2009	347483	30%
02/20/2009	386472	
02/20/2009	389148	
02/26/2009	444514	

Please estimate the reduction of annotation effort in terms of *time saved in percent* (leave blank if you cannot give an estimate).

02/05/2009	270115	35%
02/06/2009	274047	
02/06/2009	274516	
02/06/2009	277993	
02/09/2009	298730	100
02/10/2009	310840	20%
02/10/2009	311756	50
02/16/2009	347483	50%
02/20/2009	386472	
02/20/2009	389148	
02/26/2009	444514	

Please describe any other experiences regarding your practical use of AL

02/05/2009	270115	Our use of AL is nonstandard — might not even be considered AL, since we don't rerank annotation instances to be done after each step, but only in batches. See papers.
02/06/2009	274047	
02/06/2009	274516	
02/06/2009	277993	
02/09/2009	298730	
02/10/2009	310840	Never been able to build an AL interface which was easy to convince annotators to use it.
02/10/2009	311756	
02/16/2009	347483	
02/20/2009	386472	I will use it when the kind of entities annotated warrant a machine-learning based NER approach in the first place. Also, it only makes sense doing this (for entities) at a sufficiently abstract entity level (i.e. Person, Gene, Cell, etc.) Still, an unsolved problem in practice is the extensibility of AL-derived training data with new labels/cases, a scenario that often arises in practical systems
02/20/2009	389148	
02/26/2009	444514	

Would you use AL again in an annotation project of this type and scope? (Select one alternative)

Which of the following is the annotation project you are reporting on? (Select one alternative)

02/05/2009	269783	Only one

What kind of annotation had to be done? (Multiple alternatives allowed)

02/05/2009	269459	Coreference
02/05/2009	269783	Multimodal behaviour accompanying speech
02/05/2009	270094	discourse connectives
02/06/2009	274503	WePS2 attribute extraction
02/06/2009	274900	dialogue acts
02/06/2009	275845	discourse structure
02/06/2009	278622	Named entity normalization
02/07/2009	290122	Coreference relations
02/08/2009	294403	error
02/09/2009	299359	CCG category
02/11/2009	320896	discourse topic boundaries
02/12/2009	327471	entailment
02/12/2009	327680	factuality degree of events, sources
02/12/2009	330590	segmentation of definitions in genus and specifics, and annotation respectively with conceptual categories and conceptual relations
02/13/2009	334014	alignment
02/14/2009	340515	genre classification
02/16/2009	348166	error
02/16/2009	349556	coreference
02/16/2009	352374	normalization
02/19/2009	379494	coreference

From which domain(s) where the annotated instances? (Multiple alternatives allowed)

02/05/2009	269783	Custom for the application
02/05/2009	270094	Penn Treebank
02/05/2009	270477	general text
02/05/2009	271884	Academic literature
02/06/2009	273660	Chemistry
02/06/2009	273837	General (BNC)
02/06/2009	274503	Web document
02/06/2009	274900	dialogue
02/06/2009	274935	reference corpus
02/06/2009	275845	meeting dialogues
02/08/2009	294403	writing
02/09/2009	299359	questions
02/10/2009	307028	Education
02/10/2009	309907	Query logs
02/11/2009	319755	movie reviews
02/11/2009	320896	my own dataset of conversations
02/11/2009	324964	novel
02/12/2009	325736	politics news - financial news - tourism
02/12/2009	327471	educational assessment
02/12/2009	330590	various domains
02/12/2009	333405	literary text, blogs
02/13/2009	333898	instruction manuals
02/13/2009	334014	prose
02/13/2009	334304	encyclopedia, literature
02/14/2009	340515	web genres
02/16/2009	348166	error annotated corpora
02/19/2009	379494	newspaper

What was the principal language of the annotated texts? (Select one alternative)

02/05/2009	270477	danish
02/06/2009	274900	Korean
02/06/2009	274935	Slovene
02/11/2009	321971	Greek
02/12/2009	325736	Greek
02/12/2009	333405	Bengali
02/13/2009	334014	multilingual

How many person months were spent on annotation?

02/05/2009	269459	64
02/05/2009	269783	3
02/05/2009	269985	36
02/05/2009	270094	50
02/05/2009	270477	24
02/05/2009	271884	1
02/06/2009	273573	2
02/06/2009	273660	2
02/06/2009	273837	6
02/06/2009	274072	12
02/06/2009	274396	6
02/06/2009	274503	3
02/06/2009	274900	3
02/06/2009	274935	12
02/06/2009	275006	1
02/06/2009	275845	6
02/06/2009	277427	0.05
02/06/2009	278622	5
02/06/2009	279020	2
02/06/2009	279527	360
02/07/2009	290122	18
02/08/2009	294403	12
02/09/2009	299359	1
02/09/2009	303860	3
02/10/2009	307028	2
02/10/2009	309907	1
02/11/2009	319755	3
02/11/2009	320896	.5
02/11/2009	321971	8
02/11/2009	324964	6
02/12/2009	325736	10
02/12/2009	327471	6
02/12/2009	327680	6
02/12/2009	330590	0,5
02/12/2009	333405	6
02/13/2009	333898	1
02/13/2009	334014	3
02/13/2009	334304	2
02/14/2009	340515	0.25

02/16/2009	348166	24
02/16/2009	349556	170
02/16/2009	352374	2
02/19/2009	379494	27.8

How many annotators have been involved?

02/05/2009	269459	4
02/05/2009	269783	2
02/05/2009	269985	3
02/05/2009	270094	20
02/05/2009	270477	5
02/05/2009	271884	1
02/06/2009	273573	4
02/06/2009	273660	3
02/06/2009	273837	2
02/06/2009	274072	1
02/06/2009	274396	3
02/06/2009	274503	4
02/06/2009	274900	3
02/06/2009	274935	6
02/06/2009	275006	2
02/06/2009	275845	6
02/06/2009	277427	1
02/06/2009	278622	2
02/06/2009	279020	2
02/06/2009	279527	48
02/07/2009	290122	4
02/08/2009	294403	1
02/09/2009	299359	2
02/09/2009	303860	2
02/10/2009	307028	1
02/10/2009	309907	25
02/11/2009	319755	1
02/11/2009	320896	2
02/11/2009	321971	3
02/11/2009	324964	2
02/12/2009	325736	2
02/12/2009	327471	6
02/12/2009	327680	4
02/12/2009	330590	1
02/12/2009	333405	3
02/13/2009	333898	1
02/13/2009	334014	10
02/13/2009	334304	10
02/14/2009	340515	1

02/16/2009	348166	1
02/16/2009	349556	20
02/16/2009	352374	2
02/19/2009	379494	20

What was the size of the resulting, labeled corpus in terms of annotated instances
(please provide the annotation instance type in brackets, e.g., "1000 (tokens)") ?

02/05/2009	269459	Still building it
02/05/2009	269783	1000 facial displays
02/05/2009	269985	5000 (tokens)
02/05/2009	270094	>40.000 (tokens)
02/05/2009	270477	120000 words
02/05/2009	271884	50 (documents)
02/06/2009	273573	35000
02/06/2009	273660	7297 (NEs)
02/06/2009	273837	1443 (relations)
02/06/2009	274072	3000(sentences)
02/06/2009	274396	1250 (patient rec's)
02/06/2009	274503	about10000 attribute
02/06/2009	274900	10000(sentences)
02/06/2009	274935	30000 (words)
02/06/2009	275006	5000 (tokens)
02/06/2009	275845	30 (meetings)
02/06/2009	277427	200 documents
02/06/2009	278622	3000 (sentences)
02/06/2009	279020	25000 (sentences)
02/06/2009	279527	274,167 (tokens), 52
02/07/2009	290122	800000 (tokens)
02/08/2009	294403	ongoing
02/09/2009	299359	2800 (sentences)
02/09/2009	303860	3000 sentences
02/10/2009	307028	30000 (tokens)
02/10/2009	309907	25000 (queries)
02/11/2009	319755	10000 sentences
02/11/2009	320896	100 topic segments
02/11/2009	321971	163 (documents)
02/11/2009	324964	500 (tokens)
02/12/2009	325736	200.000 tokens
02/12/2009	327471	16K student answers
02/12/2009	327680	22454 (events)
02/12/2009	330590	1022 (segments)
02/12/2009	333405	1,20,00o tokens
02/13/2009	333898	500 (sentences)
02/13/2009	334014	3000 (sentences)
02/13/2009	334304	800 (sentences)
02/14/2009	340515	1400 (web pages)

02/16/2009	348166	8000
02/16/2009	349556	"500000"
02/16/2009	352374	2000 (sentences)
02/19/2009	379494	36.000 (sentences)

How did you assemble the corpus of annotation instances to be labeled? (Select
one alternative)

02/06/2009	273660	Random subsample, 3 papers per available journal
02/06/2009	274935	Subsample of all available annotation instances based on disagreements between two automatic taggers
02/06/2009	277427	random documents
02/06/2009	279527	defined our corpus
02/12/2009	327680	Used existing corpora already annotated with information complementary to what I needed
02/14/2009	340515	from web archives, web portals, keywords

How did you decide when to stop the annotation process? (Select one alternative)

02/06/2009	273660	Size of corpus decided in advance, based on available time
02/06/2009	274503	the evaluation decide to use 30 people names and 100 documents for each name
02/08/2009	294403	ongoing
02/12/2009	330590	Annotation isn't finished yet
02/14/2009	340515	comparing my collection size to the size of other existing collections

Are there specific reasons to why you did not apply Active Learning in this project?
(Multiple alternatives allowed)

02/06/2009	273660	Wanted to do cross-validation; was concerned that material selected by AL would be unsuitable for evaluation
02/06/2009	274503	This is an evaluation data. it must be near 100% accurate
02/06/2009	279527	project was to create manual annotations; AL not developed when project began (2001)
02/09/2009	299359	AL may be involved in next stage
02/12/2009	327471	perceived timeline issues
02/12/2009	327680	The phenomenon to annotate had never been annotated before, and so I wanted only manually annotated data
02/12/2009	330590	The annotated corpus extract is meant for machine learning. (Is that "Active Learning"?)

Please elaborate on the following two aspects if they apply to you: Why did Active Learning not fit your specific requirements? What were your doubts concerning the applicability of Active Learning to your scenario?

02/05/2009	269783	
02/05/2009	269985	
02/05/2009	270094	There were too many types of annotation decisions that would be hard for active learning to 'learn', such as marking implicit connectives, offering an explicit connective as annotation, identifying span of arguments of connectives etc. We judged that the task was too hard for active learning to be really useful.
02/05/2009	270477	We did dependency annotation of small discourses. While active learning may reduce the total number of required annotated instances, I worry that it may actually increase the unit cost of annotating an instance, since the annotator needs to consider an entire sentence/discourse anyway in order to determine whether it is correct. That is, the annotator must basically construct the entire analysis mentally, and it may then be wasteful to prevent he/she from annotating only the chosen instance.
02/06/2009	273573	Wanted to build evaluation data and AL would result in an un-natural dataset biased towards some selection criterion.
02/06/2009	274072	
02/06/2009	274396	
02/06/2009	274503	It must be near 100% accurate for the evaluation project purpose
02/06/2009	275845	
02/06/2009	277427	most natural way to present material for annotation is BY DOCUMENT (not by instance)
02/09/2009	299359	Not part of the remit for the first stage of the project.
02/10/2009	307028	First I want to work on unsupervised learning i.e. LSA for semantics.
02/11/2009	320896	the point was to build an evaluation set, not a training set. so active learning not applicable. also, not a straightforward classification problem, but rather a question of structure induction.
02/12/2009	327680	Already answered in previous question.
02/13/2009	334304	I wanted to get some statistics on the annotated corpus that might be biased with this approach. But I'm not very familiar with AL, so I might be wrong.
02/16/2009	352374	
02/19/2009	379494	corpus was created for a multitude of purposes/ applications

Would you consider applying active learning in future annotation projects? (Select one alternative)

| 02/06/2009 | 274503 | If I create a machine learning training data or discovering semantic knowledge, maybe yes. But I found Amazon's Mechanical Turk so attractive... |
| 02/13/2009 | 333898 | if it helps the annotators, reduces the overall annotation time |

Where are you currently working? (Select one alternative)

| 02/06/2009 | 277427 | not-for-profit |
| 02/12/2009 | 327471 | research lab |

What is your educational background? (Select one alternative)

02/06/2009	273660	Chemistry
02/06/2009	274047	Computer Science and Lingusitics
02/06/2009	277993	Biomedical Informatics
02/06/2009	278622	Bioinformatics
02/09/2009	298730	computation physics

Which field are you currently working in? (Select one alternative)

02/05/2009	269783	Dialogue systems
02/05/2009	270094	Anaphora resolution, Discourse coherence
02/05/2009	270477	dependency parsing / machine translation
02/06/2009	273660	Chemistry Text Mining
02/06/2009	274503	NLP
02/06/2009	274900	Machine Translation
02/06/2009	275845	computational linguistics for dialogue
02/06/2009	279527	corpus development and improvement
02/08/2009	294403	ESP
02/09/2009	299359	Bio parsing
02/10/2009	311756	corpus linguistics
02/11/2009	320896	nlp
02/11/2009	324964	morphology
02/12/2009	330590	Terminology
02/13/2009	333898	treebanks
02/13/2009	334014	corpus linguistics
02/14/2009	340515	genre classification
02/26/2009	444514	bio tm and ml