SciLex
RE Systematic Literature Review
ANNOTATION GUIDELINE

Version 6
Author: Célian RINGWALD
01/07/2025

ANNOTATION GUIDELINES V6

I. Basic Description

Theses fields are coming from annotations imported from https://linkedpaperswithcode.com/

1. Task [S,D,**M**]

Induced from Linked Paper With Code : https://paperswithcode.com/sota
Detail : Which task is it ? Please give here all the substask the paper is talking about
Where to find the information: Directly explained since the introduction

Borderline examples :

- **Relation Classification / Relation Extraction / EndToEndRE**
  must be clearly defined and differentiated
  To do it please consider the division of the Relation extraction task defined in [What Do You Mean by Relation Extraction? A Survey on Datasets and Study on Scientific Relation Classification](https://aclanthology.org/2022.acl-srw.7/) (Bassignana & Plank, ACL 2022)

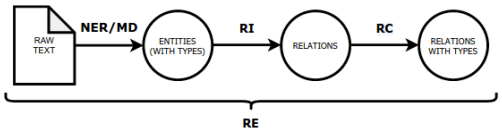


Figure 2: Relation Extraction pipeline. NER: Named Entity Recognition; MD: Mention Detection; RI: Relation Identification; RC: Relation Classification.

Possible values :

Could be extended if new Task found
Task:RelIdentification
Task:NER (named entity recognition)
Task:RelClassif (Relation classification)
Task:EndToEndRE (do everything in one model without dividing the process into substeps)
Task:EntityLinking (entity linking)
Task:Coref (Coreference resolution)
Task:EntityTyping (Type entity)
Task:NLU (use a model filter step as in rebel)
Task:SlotFilling (try to guess only the object)

2. Archi [S,D,**M**]

Induced partially from Linked Paper With Code: https://paperswithcode.com/methods
Details: Which architecture component is used? We exclude here the pretrained models that are annotated into PTM, pretrained models are instances of architectures. But we also keep Embeddings fields: that could in our case integrate Positional Embeddings as well as Relation or Entity based embeddings.
Where to find the information: Generally found in the model presentation part

Could be extended if new Archi found
Potential problem: could be difficult to list if many
Could be extended if new found
Possible values :
Archi:CNN
Archi:LSTM
Archi:GCN
Archi:BILSTM
Archi:RNN
Archi:RuleSystem
Archi:CRF
Archi:GRU
Archi:SVM
Archi:Encoder
Archi:GAN
Archi:PointerNet
Archi:Decoder
Archi:BIRNN
Archi:Encoder-Decoder
Archi:MarkovNet
Archi:KNN
Archi:CharacterEmbed (Character Embedding)
Archi:EntittyEmbed
…

3. Dataset_created [**M**]

Here we add the dataset created for the experiments
Induced partially from Linked Paper With Code
https://paperswithcode.com/datasets
A dataset here is supposed to be unique and never used before.
Details: Which datasets were created ?
Where to find the information: Generally found in the Dataset part.

4. Dataset [S,**M**]

Induced partially from Linked Paper With Code
https://paperswithcode.com/datasets
Could be extended if new found
Details: Which datasets are used (for model papers) or discussed (in survey papers)?
A dataset related to a scientific publication KB used is here excluded and must be given as a source
Where to find the information: Generally found in the Evaluation part of the Training process one.

Borderline Examples :
- could be difficult to list if many

Existing values :
Dataset:ACE2005, Dataset:Docred, Dataset:CDR, Dataset:TACRED, Dataset:ADE, Dataset:FewRel, Dataset:ScieRC, Dataset:NYT, Dataset:Wiki-ZSL, Dataset:WebNLG, Dataset:ReTaCRED, Dataset:SemEval, Dataset:GenIA, Dataset:Atis, Dataset:SNIPS, Dataset:COLA, Dataset:MRPC, Dataset:CONLL, Dataset:Rebel, Dataset:RAMS, Dataset:Wiki-Events, Dataset:Maven, Dataset:FB15k-237, Dataset:Atomic, Dataset:WikiMovies, Dataset:Penn-100, Dataset:Web-500, Dataset:DDI, Dataset:PERLEX, Dataset:KBP

5. PreTrainedModel (PTM) [D,S,**M**]

Induced partially from Linked Paper With Code
Details : Which pretrain models are cited / or used ?

Where to find the information : Generally found in the model presentation part
If transformer from scratch > we precise if this one is decoder/encoder/encoder-decoder
Nb identified values :77 values

Could be extended if new PTM found
Existing values :
PTM:XLNET, PTM:XLM, PTM:Word2Vec, PTM:VICUNA, PTM:TransformerXL, PTM:T5, PTM:T0, PTM:SpanBERT, PTM:SciBERT, PTM:Roberta, PTM:Rebel, PTM:PuMedBERT, PTM:OpenIE, PTM:OllIE, PTM:MT5, PTM:MBERT, PTM:MBART, PTM:LUKE, PTM:Longformer, PTM:LLAMA, PTM:LamDA, PTM:KGBART, PTM:KEPLER, PTM:KBERT, PTM:GPT4, PTM:GPT3, PTM:GPT2, PTM:GPT, PTM:Gopher, PTM:GLM, PTM:FlanT5, PTM:ERNIE, PTM:ELMO, PTM:Electra, PTM:DistillBERT, PTM:DeBERTa, PTM:COMET, PTM:CokeBERT, PTM:ClausIE, PTM:CharacterBERT, PTM:BioBERT, PTM:BERT, PTM:BART, PTM:BARD, PTM:ALPACA, PTM:AlBERT

6. Lang [D,S,M]

Induced partially from Linked Paper With Code
Where to find the information: Since introduction or in the title if other than english. In other case in the dataset presentation

Could be extended if new Lang found

Existing values :
Lang:Chinese, Lang:Multi, Lang:Arabic, Lang:German, Lang:French, Lang:Spanish, Lang:Italian, Lang:Russian, Lang:Dutch, Lang:Portuguese, Lang:Sloven, Lang:Polish, Lang:Finish, Lang:Turkish, Lang:Persian, Lang:Danish, Lang:Greek

II. New annotations

1. Binary dimensions

1. UseNegativeExample_Bin [DM]

Is the model taking in consideration counter examples, or is the dataset containing it ?
Where to find the information: In the dataset paragraph
TRUE >UseNegativeExample_Bin:1
FALSE - Default > Default UseNegativeExample_Bin:0

2. ObjectProperties_Bin [DM]

Does the dataset contain relation implying objects ? We can consider a triplet containing as a third element a unique Id as an object property.
Where to find the information: In the dataset paragraph
TRUE > ObjectProperties_Bin:1
FALSE - Default > ObjectProperties_Bin:0
Borderline Examples : In the fact this is sometimes hard to decide because models and datasets generally rely on label of the object rather than an ID

3. LossUpdate_Bin [M]

Does the model use Other Loss than the one defined in PTM (which is by default the Cross-Entropy)
Where to find the information: In the experimental set-up description part
TRUE > LossUpdate_Bin:1
FALSE - Default > LossUpdate_Bin:0

4. DecodingMethod_Bin [M]

Is the model using a specific method for decoding the output of the LLM, this is only related to generative models. (could be for example Beam Search / Grammar constrained / Constraint sampling)
Where to find the information: In the experimental set-up description part
Borderline Examples : is encoder classifier a decoding method ? No > we prefer to dedicate this dimension to encoder-decoder model
TRUE > DecodingMethod_Bin:1
FALSE - Default > DecodingMethod_Bin:0

5. CostEval_Bin [MDS]

Is the cost of the method evaluated ?
Where to find the information: In the experimental set-up description part
> Time Based / Carbon based
TRUE > CostEval_Bin:1
FALSE - Default > CostEval_Bin:0

6. Linearized_graph_Bin [MDS]

Do the model need to transform graph in a linearized sequence of text
Where to find the information: In the experimental set-up description part
TRUE > Linearized_graph_Bin:1
FALSE - Default > Linearized_graph_Bin:0

7. SynthGeneration_Bin [MD]
8.

Do the Data or model is using synthetic data ?
Where to find the information: Depends to the paper content and structuration
TRUE > SynthGeneration_Bin:1
FALSE - Default > SynthGeneration_Bin:0

9. Survey_Methodo_Bin [S]

Do the Survey follow a specific and defined methodology ?
Where to find the information: Depends to the paper content and structuration, but generally in the methodology part
TRUE > Survey_Methodo_Bin:1
FALSE - Default > Survey_Methodo_Bin:0

2. CATEGORICAL

1. MANUALANNOTATION_BIN [DM]

Was the dataset annotated manually?
Where to find the information: In the dataset description part
Possible values : 0/1 and partial

2. DatatypeProp: [D]

Details: Is the paper integrating the extraction of objects that are related to datatype properties?
https://www.w3.org/TR/owl-ref/#Datatype

Where to find the information: in the dataset presentation
Borderline example : hard to answer if not explicitly told in the paper, dates for examples are considered as object in WD
CUSTOM > (LIST of datatype used)

DatatypeProp:String / DatatypeProp:Date / DatatypeProp:Number

3. BenchmarkType [S]

CUSTOM > (quanti or quali)
Where to find the information: in the benchmark part
Details : A paper containing a table with quantitative or/and qualitative comparison of models
BenchmarkType:Quanti OR BenchmarkType:Quali

4. Input [D,S,**M**]

What is used as input of the models ?
Input:IndexedSpan
Input:EntityLink
Input:DependancyParsing
Input:POSTAG
Input:Embedding

5. LEARNING [S,**M**]

Learning method used to produce a model
Where to find the information: In the methodological part

Could be extended if new Lang found
Possible values :
LearningMethod:PromptBased:
LearningMethod:Distant
LearningMethod:Active
LearningMethod:FewShot
LearningMethod:Finetuning
LearningMethod:Contrastive

LearningMethod:Continual
LearningMethod:Reinforcement
LearningMethod:ZeroShot
LearningMethod:UnsupervisedPTM
LearningMethod:PromptTuning
LearningMethod:Instruct
LearningMethod:NO
LearningMethod:PrompTuning
LearningMethod:Unsupervised

6. GRANULARITY [D,S,**M**]

Is the model made to process Document or just sentences ?
Where to find the information: Generally since the introduction, or could be inferred in regards of the dataset used
Values :
Granularity:Document OR Granularity:Sentence

7. SOURCE [D]

Details : Which is the original source of the dataset ?
Where to find the information: In the methodological part
Could be internet corpora / a KB, an other Dataset…

Could be extended if new Lang found
Example of values :
Source:Wikidata, Source:Dbpedia, Source:Web, Source:Freebase, Source:PubMed, Source:NYT, Source:NCBI, Source:NLM-Gene, Source:GnormPlus, Source:C4, Source:Twitter, Source:WebDataCommon, Source:GoogleNews, Source:MLNLeague, Source:WikiNews, Source:Reuters, Source:DeutscheWelle, Source:TextBook, Source:Umls, Source:WordNet, Source:ConceptNet, Source:Atomic

8. DOMAIN [SDM]

Focused domain of application of the dataset/survey/model
Where to find the information: In the dataset presentation, the introduction or the other works
CUSTOM

Could be extended if new Lang found
Example of values :
Domain:Encyclo, Domain:Science, Domain:BioMedical, Domain:Politic, Domain:News, Domain:Litterature, Domain:General, Domain:Multi, Domain:Bio, Domain:Medical, Domain:Web, Domain:Finance, Domain:Sport, Domain:weather, Domain:Military, Domain:Music, Domain:Humanities, Domain:Economy, Domain:Legal, Domain:ScienceLitt

9. DATASET SPLIT [MD]

How the dataset was splitted into Train/Test/eval split
Where to find the information: In the dataset presentation

Could be extended if new Lang found

Example of values :
SelectionMethod:BalancedSampling / SelectionMethod:Random / SelectionMethod:ChallengingEx

4. STATS

> OPTIONNAL FOR MODEL > MANDATORY FOR DATASET
All the stats are reduced to Scale Base Notation 10^n

1. NbDoc [D,M]

Details: Number of document in the dataset or used by the model

2. NbSent [D,M]

Details: Number of sentences in the dataset or used by the model

3. NbDataset [D,M]

Details: Number of entity types integrated into the dataset or used by the mode

4. NbModel [SM]

Details : Number models compared or developed
Where to find the information: in methodological part/benchmark part or evaluation part

5. NbTypeEntity [D,M]

Details: Number of entity types integrated into the dataset or used by the model

6. NbEntity [D,M]

Details: Number of entities integrated into the dataset or used by the model

7. NbTypeRel [D,**M**]

Details : NB of properties or relation in the dataset or learn by a given model
Where to find the information: Generally related to the paper structure

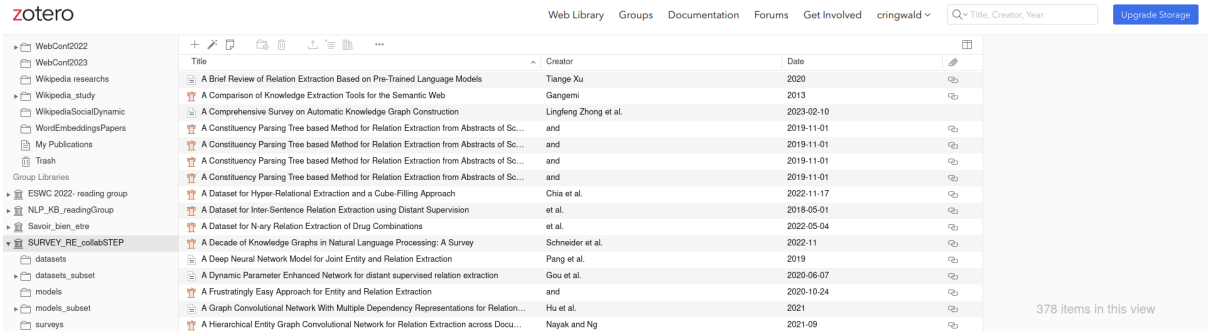
8. NbTriples [D,M]

Details : Number of triples used for the model training or available in the dataset (based on the models/dataset studied)
Where to find the information: in the dataset or methodological part

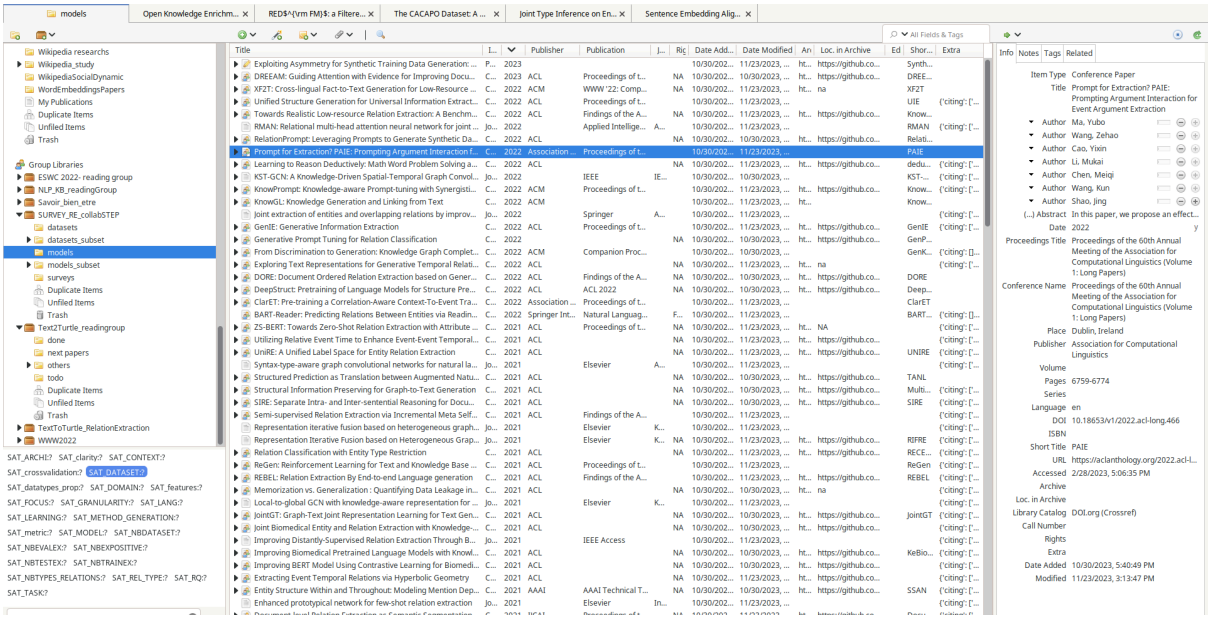
III. Using Zotero for the annotation process

1. Acces to the library

From the website

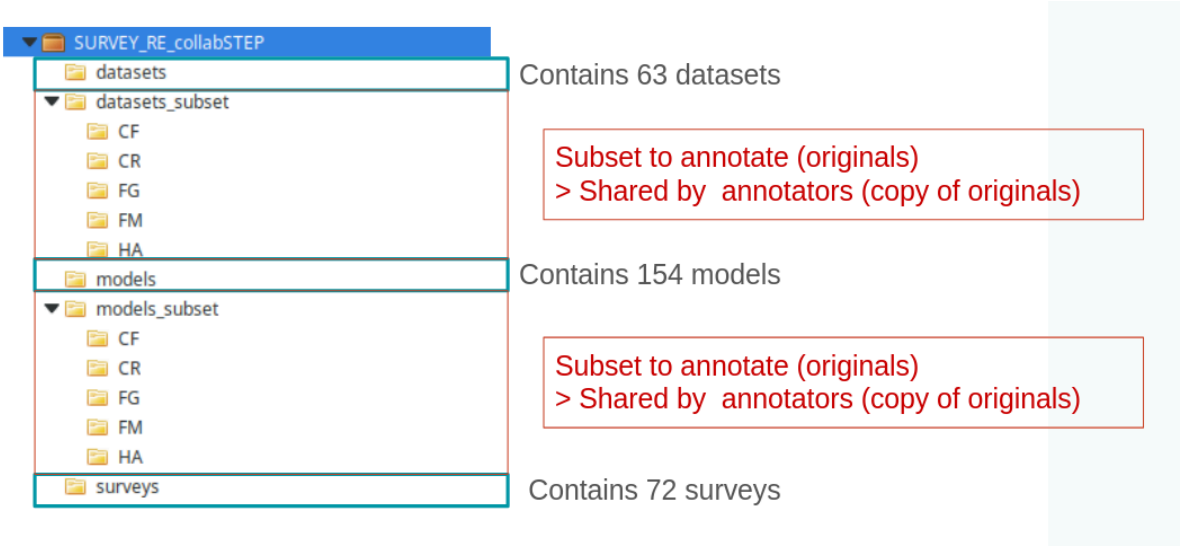


or from the app :



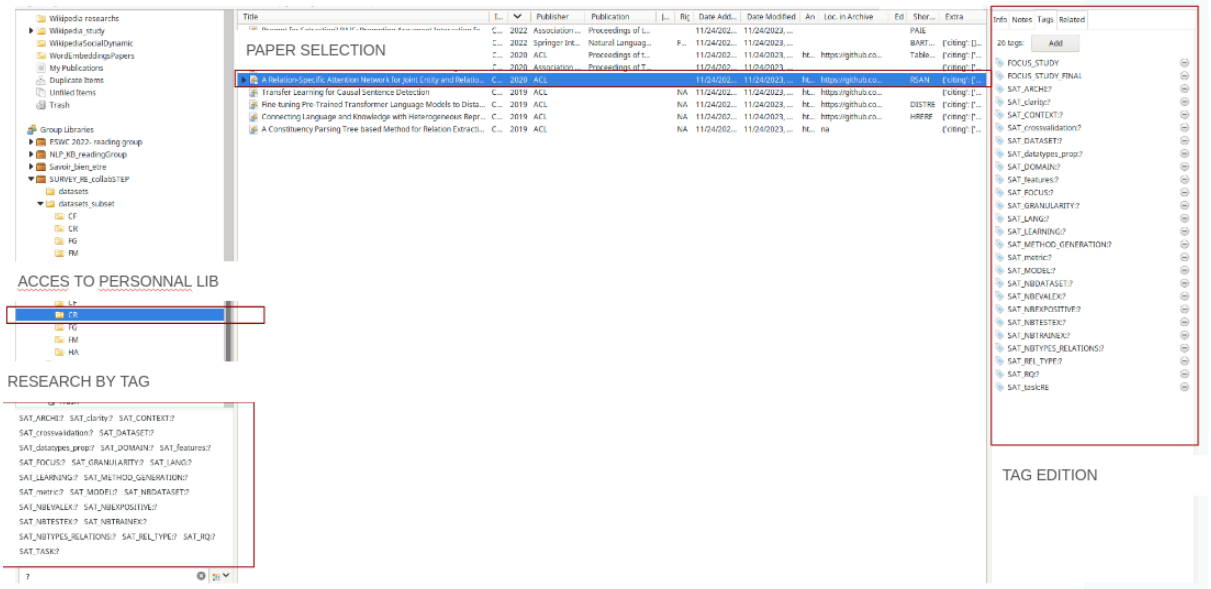
I advice to download and use the local app
> it synchronize the content and allows to work without internet
> allows to download and annotate with PDF in the same windows

2. Content of the lib :



3. Usage

Reminder : the pdf could be downloaded and watched via the local app



ACCES TO PERSONNAL LIB :
each paper is checked by 3 annotators, and each of them receives the same copy of the paper to annotate.

RESEARCH BY TAG :

- Researching only “?” filter the papers by the fields still blank
- Filtering by only giving the annotation dim name allows to check the values already given in the subset

/!\ Using the research on the entire set (ex: “modele” lib) allows us to get the values already given at the corpora level - > could be also useful to use

TAG EDITION :
> A missing tag must be completed or deleted
> That is possible to add multiple values by dimension
> A tag could be deleted or edited

ComplementaryValues :
NSP
DONE
EXCLUDED

IV. ANNOTATION SUMMARY:

tag	nb	datasets	surveys	models	type	new for models	detail
LearningMethod	168			X	quali		TO EXTRACT
Granularity	123	X	X	X	quali		TO EXTRACT
Source	84	X			quali		TO EXTRACT
DATASETSplit	63	X		X	quali		TO EXTRACT
Domain	61	X	X		quali		TO EXTRACT
Survey_Methodo_Bin	51		X		Bin		TO EXTRACT
Input	42			X	Quali		TO EXTRACT
BenchmarkType	28		X		Quali		TO EXTRACT
UseNegativeExample_Bin	12			X	Bin		TO EXTRACT
DatatypeProp	9	X		X	Quali		TO EXTRACT
SynthGeneration_Bin	8	X		X	Bin		TO EXTRACT
LinearizedGraph_Bin	4			X	Bin		TO EXTRACT
DecodingMethod_Bin	4			X	Bin		TO EXTRACT
LossUpdate_Bin	4			X	Bin		TO EXTRACT
CostEval_Bin	3			X	Bin		TO EXTRACT
ObjectProperties_Bin	0			X	Bin	1	TO EXTRACT
Task	362	X	X	X	quali		TO CHECK
PTM	139		X	X	quali		TO CHECK
Archi	119		X	X	quali		TO CHECK
Lang	118	X	X	X	quali		TO CHECK
Dataset	78		X	X	quali		TO CHECK
NbTypeRel	47	X		X	Scaled		OPTIONAL
NbTriples	32	X			Scaled		OPTIONAL
NbDoc	29	X			Scaled		OPTIONAL
NbModel	27		X		Scaled		OPTIONAL
NbTypeEntity	23	X			Scaled		OPTIONAL
NbEntity	21	X			Scaled		OPTIONAL
NbDataset	21		X	X	Scaled		OPTIONAL
NbSent	20	X			Scaled		OPTIONAL

V. MODEL CHECKLIST

tag	nb	type	detail	Model1	Model2	Model3	Model4	Model5	Model6
Task	362	quali	TO CHECK						
LearningMethod	168	quali	TO EXTRACT						
PTM	139	quali	TO CHECK						
Granularity	123	quali	TO EXTRACT						
Archi	119	quali	TO CHECK						
Lang	118	quali	TO CHECK						
Dataset	78	quali	TO CHECK						
DATASETSPLIT	63	quali	TO EXTRACT						
NbTypeRel	47	Scaled	OPTIONAL						
Input	42	Quali	TO EXTRACT						
NbDataset	21	Scaled	OPTIONAL						
UseNegativeExample_Bin	12	Bin	TO EXTRACT						
DatatypeProp	9	Quali	TO EXTRACT						
SynthGeneration_Bin	8	Bin	TO EXTRACT						
LinearizedGraph_Bin	4	Bin	TO EXTRACT						
DecodingMethod_Bin	4	Bin	TO EXTRACT						
LossUpdate_Bin	4	Bin	TO EXTRACT						
CostEval_Bin	3	Bin	TO EXTRACT						
ObjectProperties_Bin	0	Bin	TO EXTRACT						

VI. DATASET CHECKLIST

tag	nb	datasets	type	detail	Dataset1	Dataset2	Dataset3
Task	362	X	quali	TO CHECK			
Source	84	X	quali	TO EXTRACT			
Domain	61	X	quali	TO EXTRACT			
Granularity	123	X	quali	TO EXTRACT			
Lang	118	X	quali	TO CHECK			
DATASET SPLIT	63	X	quali	TO EXTRACT			
NbTypeRel	47	X	Scaled	OPTIONAL			
DatatypeProp	9	X	Quali	TO EXTRACT			
SynthGeneration_Bin	8	X	Bin	TO EXTRACT			
NbTriples	32	X	Scaled	OPTIONAL			
NbDoc	29	X	Scaled	OPTIONAL			
NbTypeEntity	23	X	Scaled	OPTIONAL			
NbEntity	21	X	Scaled	OPTIONAL			
NbSent	20	X	Scaled	OPTIONAL			

VII. SURVEY CHECKLIST

tag	nb	detail	Survey1	Survey2	Survey3
Task	362	TO CHECK			
PTM	139	TO CHECK			
Domain	61	TO EXTRACT			
Survey_Methodo_Bin	51	TO EXTRACT			
Granularity	123	TO EXTRACT			
BenchmarkType	28	TO EXTRACT			
Archi	119	TO CHECK			
Lang	118	TO CHECK			
Dataset	78	TO CHECK			
NbDataset	21	OPTIONAL			
NbModel	27	OPTIONAL			

VIII. VALUES DICTIONNARY

level1	**tag**	nb
Archi	Archi:BILSTM	4
Archi	Archi:BILSTM	9
Archi	Archi:BIRNN	1
Archi	Archi:CNN	23
Archi	Archi:CNN	3
Archi	Archi:CRF	3
Archi	Archi:Decoder	2
Archi	Archi:Encoder	2
Archi	Archi:Encoder-Decoder	5
Archi	Archi:Encoder-Decoder	4
Archi	Archi:GAN	2
Archi	Archi:GCN	6
Archi	Archi:GCN	10
Archi	Archi:GraphTransformer	1
Archi	Archi:GRU	1
Archi	Archi:GRU	2
Archi	Archi:KNN	1
Archi	Archi:LSTM	15
Archi	Archi:MarkovNet	1
Archi	Archi:PointerNet	3
Archi	Archi:RNN	9
Archi	Archi:RuleSystem	6
Archi	Archi:SVM	3
Archi	Archi:Unet	1
BenchmarkType	BenchmarkType:?	1
BenchmarkType	BenchmarkType:Quali	12
BenchmarkType	BenchmarkType:Quanti	15
CostEval_Bin	CostEval_Bin:1	3
Dataset	Dataset:ACE2005	8
Dataset	Dataset:ACE2005	4
Dataset	Dataset:ADE	4
Dataset	Dataset:Atis	1
Dataset	Dataset:Atomic	1
Dataset	Dataset:CDR	4
Dataset	Dataset:CDR	1
Dataset	Dataset:CONLL	1
Dataset	Dataset:DDI	1
Dataset	Dataset:Docred	6

Dataset	Dataset:FB15k-237	1
Dataset	Dataset:FewRel	4
Dataset	Dataset:FewRel	1
Dataset	Dataset:GenIA	1
Dataset	Dataset:KBP	1
Dataset	Dataset:Maven	1
Dataset	Dataset:NYT	2
Dataset	Dataset:NYT	6
Dataset	Dataset:Penn-100	1
Dataset	Dataset:PERLEX	1
Dataset	Dataset:RAMS	1
Dataset	Dataset:Rebel	3
Dataset	Dataset:ReTaCRED	3
Dataset	Dataset:ScieRC	4
Dataset	Dataset:SemEval	3
Dataset	Dataset:SNIPS	1
Dataset	Dataset:TACRED	5
Dataset	Dataset:Web-500	1
Dataset	Dataset:WebNLG	3
Dataset	Dataset:Wiki-Events	1
Dataset	Dataset:Wiki-ZSL	2
Dataset	Dataset:WikiMovies	1
DatatypeProp	DatatypeProp:?	2
DatatypeProp	DatatypeProp:Date	3
DatatypeProp	DatatypeProp:Number	1
DatatypeProp	DatatypeProp:String	3
DecodingMethod_Bin	DecodingMethod_Bin:1	4
Domain	Domain:Bio	3
Domain	Domain:BioMedical	7
Domain	Domain:BioMedical	1
Domain	Domain:Economy	1
Domain	Domain:Encyclo	9
Domain	Domain:Finance	2
Domain	Domain:General	4
Domain	Domain:General	1
Domain	Domain:Humanities	2
Domain	Domain:Legal	2
Domain	Domain:Litterature	2
Domain	Domain:Medical	4
Domain	Domain:Military	1
Domain	Domain:Multi	2
Domain	Domain:Music	1

Domain	Domain:News	3
Domain	Domain:Politic	3
Domain	Domain:Science	7
Domain	Domain:Sport	1
Domain	Domain:weather	1
Domain	Domain:Web	4
FILTER_STEP	FILTER_STEP	1
Granularity	Granularity:Document	59
Granularity	Granularity:Document	9
Granularity	Granularity:Sentences	51
Granularity	Granularity:Sentences	4
Input	Input:?	1
Input	Input:Code	1
Input	Input:DependancyParsing	1
Input	Input:Embedding	1
Input	Input:EntityLink	1
Input	Input:Graph	12
Input	Input:Graph	5
Input	Input:Image	1
Input	Input:IndexedSpan	2
Input	Input:Multimodal	3
Input	Input:POSTAG	1
Input	Input:Text	13
Lang	Lang:Arabic	14
Lang	Lang:Chinese	17
Lang	Lang:Danish	1
Lang	Lang:Dutch	2
Lang	Lang:English	38
Lang	Lang:Finish	1
Lang	Lang:French	6
Lang	Lang:German	8
Lang	Lang:Greek	1
Lang	Lang:Italian	3
Lang	Lang:Multi	8
Lang	Lang:Multi	5
Lang	Lang:Persian	1
Lang	Lang:Polish	1
Lang	Lang:Portuguese	1
Lang	Lang:Russian	2
Lang	Lang:Sloven	1
Lang	Lang:Spanish	6
Lang	Lang:Turkish	2

LearningMethod	LearningMethod:?	80
LearningMethod	LearningMethod:?	38
LearningMethod	LearningMethod:Continual	3
LearningMethod	LearningMethod:Continual	2
LearningMethod	LearningMethod:Contrastive	2
LearningMethod	LearningMethod:Contrastive	1
LearningMethod	LearningMethod:Distant	4
LearningMethod	LearningMethod:FewShot	5
LearningMethod	LearningMethod:FewShot	2
LearningMethod	LearningMethod:Finetuning	11
LearningMethod	LearningMethod:Instruct	1
LearningMethod	LearningMethod:Pretraining	1
LearningMethod	LearningMethod:PromptBased	8
LearningMethod	LearningMethod:PromptTuning	1
LearningMethod	LearningMethod:PrompTuning	1
LearningMethod	LearningMethod:Reinforcement	3
LearningMethod	LearningMethod:Unsupervised	2
LearningMethod	LearningMethod:UnsupervisedPTM	1
LearningMethod	LearningMethod:ZeroShot	2
LinearizedGraph_Bin	LinearizedGraph_Bin:1	4
LossUpdate_Bin	LossUpdate_Bin:1	4
PTM	PTM:AlBERT	3
PTM	PTM:ALPACA	1
PTM	PTM:BARD	1
PTM	PTM:BART	17
PTM	PTM:BERT	37
PTM	PTM:BioBERT	2
PTM	PTM:CharacterBERT	2
PTM	PTM:ClausIE	2
PTM	PTM:CokeBERT	1
PTM	PTM:COMET	2
PTM	PTM:DeBERTa	1
PTM	PTM:DistillBERT	1
PTM	PTM:Electra	1
PTM	PTM:ELMO	3
PTM	PTM:ERNIE	4
PTM	PTM:FlanT5	2
PTM	PTM:GLM	2
PTM	PTM:Gopher	1
PTM	PTM:GPT	5
PTM	PTM:GPT2	3
PTM	PTM:GPT3	3

PTM	PTM:GPT4	1
PTM	PTM:KBERT	1
PTM	PTM:KEPLER	1
PTM	PTM:KGBART	1
PTM	PTM:LamDA	1
PTM	PTM:LLAMA	1
PTM	PTM:Longformer	1
PTM	PTM:LUKE	1
PTM	PTM:MBART	3
PTM	PTM:MBERT	1
PTM	PTM:MT5	1
PTM	PTM:OllIE	1
PTM	PTM:OpenIE	1
PTM	PTM:PuMedBERT	1
PTM	PTM:Rebel	1
PTM	PTM:Roberta	9
PTM	PTM:SciBERT	2
PTM	PTM:SpanBERT	2
PTM	PTM:T0	1
PTM	PTM:T5	7
PTM	PTM:TransformerXL	2
PTM	PTM:VICUNA	1
PTM	PTM:Word2Vec	1
PTM	PTM:XLM	1
PTM	PTM:XLNET	2
SelectionMethod	SelectionMethod:Automatic	24
SelectionMethod	SelectionMethod:Manual	36
SelectionMethod	SelectionMethod:Random	3
Source	Source:Atomic	1
Source	Source:C4	1
Source	Source:ConceptNet	1
Source	Source:Dbpedia	8
Source	Source:DeutscheWelle	1
Source	Source:Freebase	3
Source	Source:GnormPlus	1
Source	Source:GoogleNews	1
Source	Source:MLNLeague	1
Source	Source:NCBI	1
Source	Source:NLM-Gene	1
Source	Source:NYT	1
Source	Source:PubMed	3
Source	Source:Reuters	1

Source	Source:TextBook	1
Source	Source:Twitter	2
Source	Source:Umls	1
Source	Source:Web	3
Source	Source:WebDataCommon	1
Source	Source:Wikidata	23
Source	Source:WikiNews	1
Source	Source:Wikipedia	25
Source	Source:Wikipedia	1
Source	Source:WordNet	1
Survey_Methodo_Bin	Survey_Methodo_Bin:0	43
Survey_Methodo_Bin	Survey_Methodo_Bin:1	8
SynthGeneration_Bin	SynthGeneration_Bin:1	8
Task	Task:Coref	15
Task	Task:EndToEndRE	26
Task	Task:EntityLinking	16
Task	Task:EntityLinking	1
Task	Task:EntityTyping	7
Task	Task:NER	51
Task	Task:NER	5
Task	Task:NLU	9
Task	Task:NLU	1
Task	Task:RelClassif	36
Task	Task:RelExtract	121
Task	Task:RelExtract	68
Task	Task:SlotFilling	6
UseNegativeExample_Bin	UseNegativeExample_Bin:0	5
UseNegativeExample_Bin	UseNegativeExample_Bin:1	7