

The PsyMine Corpus - A Corpus annotated with Psychiatric Disorders and their Etiological Factors



University of Zurich^{UZH}

OntoGene
www.ontogene.org

the cogito foundation

Tilia Renate Ellendorff, Simon Foster, Fabio Rinaldi

University of Zurich, Institute of Computational Linguistics

University of Zurich, Institute of Epidemiology

ellendorff@cl.uzh.ch

Motivation: Supporting Research in the Domain of Mental Health

Mental Disorders:

- Among main contributors to contemporary burden of disease
- Major group of disorder causing disability

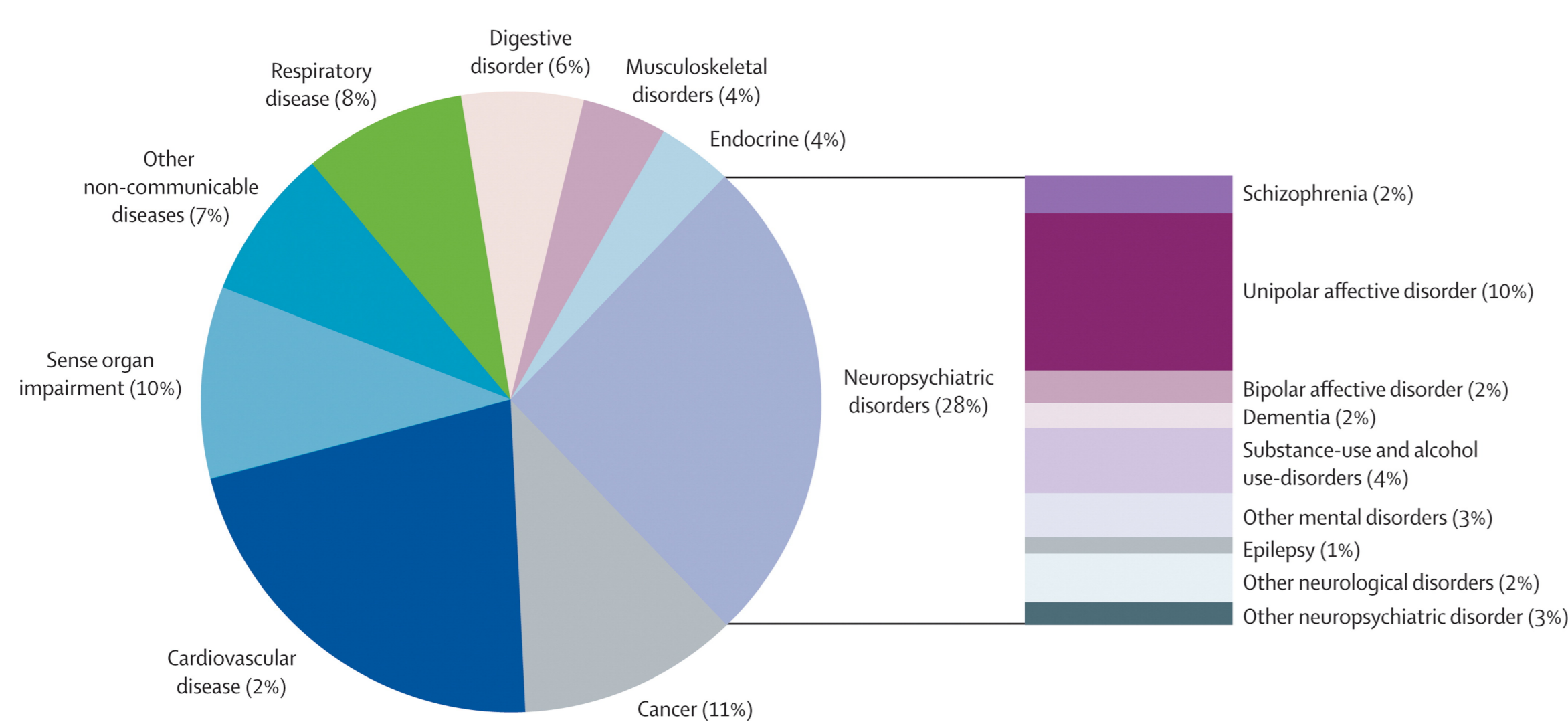


Figure 1: Contribution by different non-communicable diseases to disability-adjusted life-years worldwide in 2005 [3]

Discovering Underlying Factors of Mental Disorders

- Underlying factors spread over many research areas, e.g. genetics, sociology, chemistry, ...
 - Information is encoded mostly in research articles (= written text)
 - Very large amounts of text (Table 1)
- ⇒ Obtaining a “complete picture” is not easy

Automatic Extraction of Causal Factors from Written Text

- Fast and efficient
- Inclusion of all relevant research areas
- Aim: overview of all known causal factors for a given disorder
- Helps to provide a complete picture of etiological factors

Disorder Type	Relevant MeSH Term	Topic	Major Topic
Depression	Depressive Disorder	84,513	64,588
	Dysthymic Disorder	1,029	620
Anxiety	Anxiety Disorder	72,032	56,049
	Obsessive-Compulsive Disorder	11,839	9,280
	Phobic Disorder	9,524	6,745
	Agoraphobia	2,412	1,396
	Panic Disorder	6,133	4,661

Table 1: Number of Abstracts per Disorder over the whole of PubMed¹

Data Selection

Focus Psychiatric Disorders

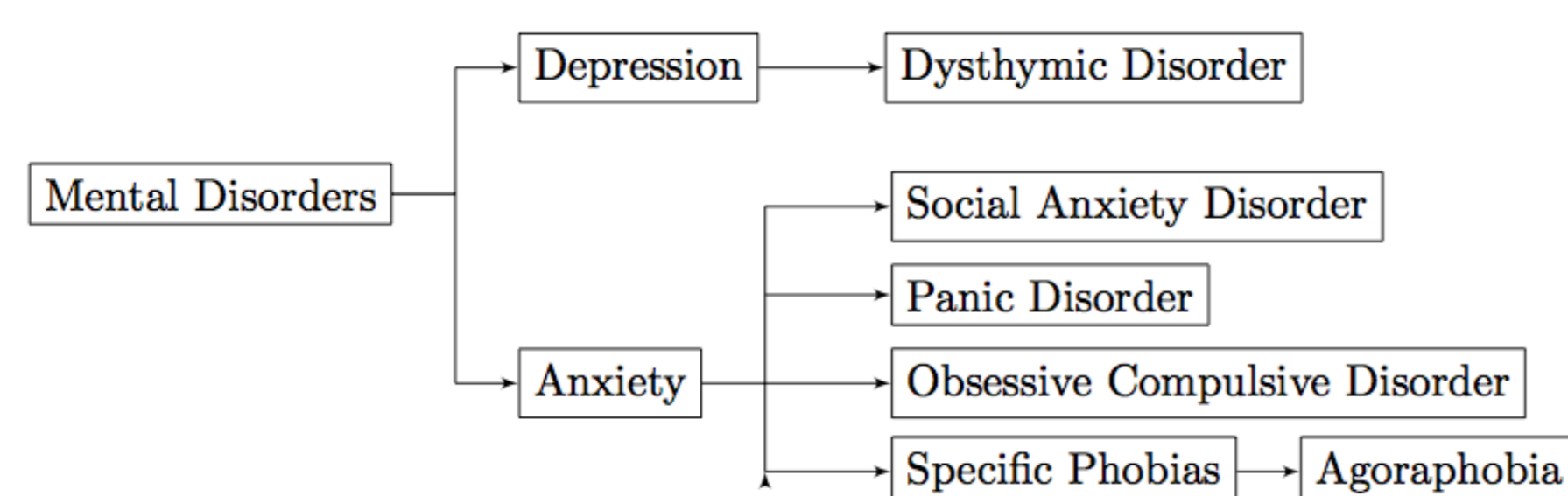


Figure 2: Hierarchy of included Mental Disorders

Categories of Etiological Factors

Etiological Factors	Examples
Genetic Factors	5HTTLPR
Neurotransmitters	Serotonin
Hormones	Estrogen
Neuronal Growth Factors	Brain derived neurotrophic factor
Brain Physiology	Amygdala, Brain Network, Neuroplasticity
Physiologic Stress Response	HPA axis, Cortisol, glucocorticoids
Autonomic Nervous System	Anxious Arousal, Autonomic bodily symptoms
Inflammation	Pro inflammatory Markers, Interleukin 6
Cardiovascular/Vascular Factors	Heart rate, Blood pressure, Hypertension
Disorder related/Other Psychiatric Disorders	Disorder severity, post traumatic stress disorders
Personality Factors	Neuroticism
Self Identity	Core perceptions of the self, self-acceptance associations
Core Self Evaluation	Self esteem
Pregnancy/Birth factors	Prenatal maternal anxiety, Birth weight
Developmental Factors	Parenting style
Cognitive Factors	Dysfunctional attitudes
Lifestyle Factors	Physical activity, Alcohol consumption
Sociologic Factors	Social status, Income inequality, Discrimination
Socioeconomic Factors	Education, Income
Socio-demographic Factors	Age, Gender
Social Roles	Parenting
Chemicals/Recreational Drugs	Cigarette Smoke, Alcohol, Cocaine
Medical Drugs	Class SSRI, Escitalopram, Class SSNRI, Venlafaxine
Environmental Factors	Aircraft noise, Amount of sunlight

Table 2: Categories of Included Etiological Factors

Text Selection

- 175 abstracts of research papers (chosen based on focus disorders)
- Average abstract length: 250 words
- Automatic download from PubMed
- More text (also full text articles) will be included in the future

Corpus Annotation

- Manual annotation in brat rapid annotation tool (BRAT) [4]
- 3 human annotators (domain experts)

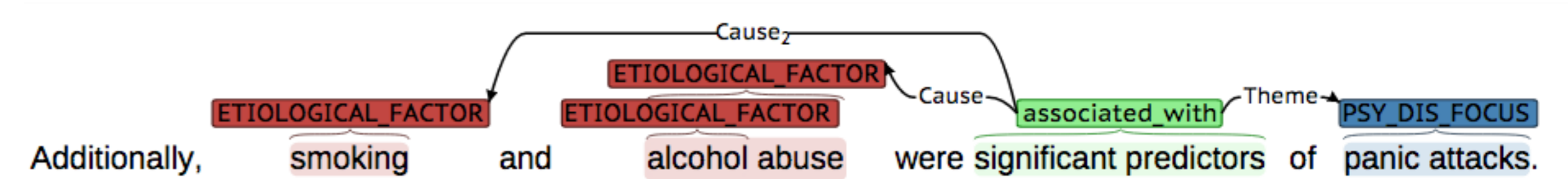


Figure 3: Annotation in Brat of an Etiological Event from [2]

Annotation Schema

Annotated components:

- Annotated Entities:
 - Focus Disorders (including patient groups, disorder scales and implicit disorders)
 - Other Psychiatric Disorders, Other Disorders
 - Etiological Factors (possibly double annotated disorders)
 - Symptoms and Treatments
- Annotated Events:
 - Etiological events: *associated with*
 - Treatment events: *treatment for*
 - Symptomatic events and other event types: *other relation*
- Other:
 - Annotation of implicit Entity Mentions
 - Anaphora Relations (e.g. for implicit entity mentions)
 - Equivalence Relations (e.g. for local abbreviations)

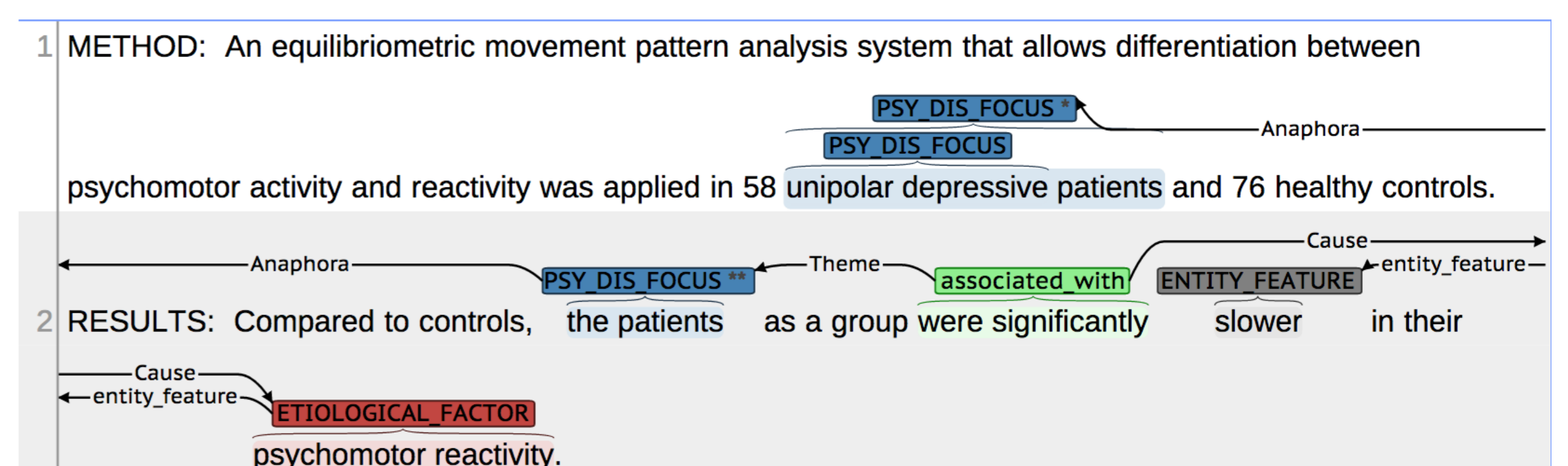


Figure 4: Event Annotation including the Annotation of an implicit Entity Mention and an Anaphoric Relation in [5]

Inter-Annotator Agreement

		annotator 1 - annotator 2	annotator 2 - annotator 3	annotator 1 - annotator 3
Focus	strict	0.37	0.35	0.77
Psychiatric Disorders	relaxed	0.81	0.81	0.88
Etiological Factors	strict	0.27	0.22	0.38
	relaxed	0.65	0.63	0.69
Both	strict	0.27	0.27	0.49
Entity Types	relaxed	0.71	0.70	0.77

Table 3: Overview of inter-annotator scores between annotator pairs measured in average observed agreement across coders and items

Format and Release

- Published online: <http://www.ontogene.org/current-pr/psymine>
- First version of PsyMine Corpus; Extension planned
- Available in two formats (converters provided):
 - BioNLP/Brat Format (<http://2013.bionlp-st.org/file-formats>)
 - BioC Format [1]

References

- [1] Donald C. Comeau, Rezarta Islamaj Dogan, Paolo Ciccarese, Kevin Bretonnel Cohen, Martin Krallinger, Florian Leitner, Zhiyong Lu, Yifan Peng, Fabio Rinaldi, Manabu Torii, Alfonso Valencia, Karin Verspoor, Thomas C. Wiegers, Cathy H. Wu, and W. John Wilbur. BioC: a minimalist approach to interoperability for biomedical text processing. *Database*, 2013, 2013.
- [2] Kristine Ria Hearld, Henna Badhwani, and Daniel Chavez-Yenter. Panic attacks in minority americans: The effects of alcohol abuse, tobacco smoking, and discrimination. *Journal of Affective Disorders*, 174:106 – 112, 2015.
- [3] Martin Prince, Vikram Patel, Shekhar Saxena, Mario Maj, Joanna Masello, Michael R Phillips, and Atif Rahman. No health without mental health. *The Lancet*, 370(9590):859 – 877, 2007.
- [4] Pontus Stenetorp, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. Brat: A web-based tool for nlp-assisted text annotation. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, EACL '12, pages 102–107, Stroudsburg, PA, USA, 2012. Association for Computational Linguistics.
- [5] Petya D Terzivanova and Svetlozar H Haralalov. Latent bipolarity in unipolar depression: Experimental findings, conceptual analysis and implications for treatment strategies. *Folia medica*, 56(4):282288, 2014.

Acknowledgements

We would like to thank the three annotators involved in the creation of the PsyMine Corpus for their efforts. TRE and SF are supported within the PsyMine Project by grand 15-115-R of the COGITO Foundation.