

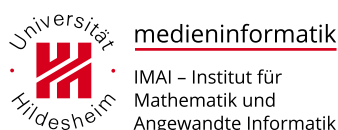
Topics Seminar Media Informatics

Updated: December 13, 2018

Jörg Cassens

Seminar Medieninformatik

Winter Term 2018/2019



Topics

A list of texts not assigned yet can be found on page 14.

1. Natural Language: Understanding and Generation

1.1. Survey: Sentiment Analysis

☒ **Presented by: Marvin Wolf**

▷ Paper: Maite Taboada. Sentiment Analysis: An Overview from Linguistics. Annual Review of Linguistics 2016 2:1, 325-347

[Download paper](#)

Abstract: Sentiment analysis is a growing field at the intersection of linguistics and computer science that attempts to automatically determine the sentiment contained in text. Sentiment can be characterized as positive or negative evaluation expressed through language. Common applications of sentiment analysis include the automatic determination of whether a review posted online (of a movie, a book, or a consumer product) is positive or negative toward the item being reviewed. Sentiment analysis is now a common tool in the repertoire of social media analysis carried out by companies, marketers, and political analysts. Research on sentiment analysis extracts information from positive and negative words in text, from the context of those words, and from the linguistic structure of the text. This brief review examines in particular the contributions that linguistic knowledge can make to the task of automatically determining sentiment.

1.2. Predicting Quality and Interestingness of Texts

☐ **Topic not assigned**

▷ Paper: Debasis Ganguly, Johannes Leveling, and Gareth J.F. Jones. Automatic prediction of text aesthetics and interestingness. In Proceedings of the International Conference on Computational Linguistics. 2014.

[Download paper](#)

Abstract: This paper investigates the problem of automated text aesthetics prediction. The availability of user generated content and ratings, e.g. Flickr, has induced research in aesthetics prediction for non-text domains, particularly for photographic images. This problem, however, has yet not been explored for the text domain. Due to the very subjective nature of text aesthetics, it is difficult to compile human annotated data by methods such as crowd sourcing with a fair degree of inter-annotator agreement. The availability of the Kindle "popular highlights" data has motivated us to compile a dataset comprised of human annotated aesthetically pleasing and interesting text passages. We then undertake a supervised classification approach to predict text aesthetics by constructing real-valued feature vectors from each text passage. In particular, the features that we use for this classification task are word length, repetitions, polarity, part-of-speech, semantic distances; and topic generality and diversity. A traditional binary classification approach is not effective in this case because non-highlighted passages surrounding the highlighted ones do not necessarily represent the other extreme of unpleasant quality text. Due to the absence of real negative class samples, we employ the MC algorithm, in which training can be initiated with instances only from the positive class. On each successive iteration the algorithm selects new strong

negative samples from the unlabeled class and retrains itself. The results show that the mapping convergence (MC) algorithm with a Gaussian and a linear kernel used for the mapping and convergence phases, respectively, yields the best results, achieving satisfactory accuracy, precision and recall values of about 74%, 42% and 54% respectively.

- ▷ Paper: Tong Wang, Ping Chen, Boyang Li: Predicting the Quality of Short Narratives from Social Media. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Main track. Pages 3859-3865.

[Download paper](#)

Abstract: An important and difficult challenge in building computational models for narratives is the automatic evaluation of narrative quality. Quality evaluation connects narrative understanding and generation as generation systems need to evaluate their own products. To circumvent difficulties in acquiring annotations, we employ upvotes in social media as an approximate measure for story quality. We collected 54,484 answers from a crowd-powered question-and-answer website, Quora, and then used active learning to build a classifier that labeled 28,320 answers as stories. To predict the number of upvotes without the use of social network features, we create neural networks that model textual regions and the interdependence among regions, which serve as strong benchmarks for future research. To our best knowledge, this is the first large-scale study for automatic evaluation of narrative quality.

1.3. **Headline Generation**

- **Topic not assigned**

- ▷ Paper: Enrique Alfonseca, Daniele Pighin, and Guillermo Garrido. HEADY: news headline abstraction through event pattern clustering. In ACL 2013, pages 1243–1253, 2013.

[Download paper](#)

Abstract: This paper presents HEADY: a novel, abstractive approach for headline generation from news collections. From a web-scale corpus of English news, we mine syntactic patterns that a Noisy-OR model generalizes into event descriptions. At inference time, we query the model with the patterns observed in an unseen news collection, identify the event that better captures the gist of the collection and retrieve the most appropriate pattern to generate a headline. HEADY improves over a state-of-the-art open-domain title abstraction method, bridging half of the gap that separates it from extractive methods using human-generated titles in manual evaluations, and performs comparably to human-generated headlines as evaluated with ROUGE.

- ▷ Paper: Jiwei Tan, Xiaojun Wan, Jianguo Xiao: From Neural Sentence Summarization to Headline Generation: A Coarse-to-Fine Approach. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. Main track. Pages 4109-4115.

[Download paper](#)

Abstract: Headline generation is a task of abstractive text summarization, and previously suffers from the immaturity of natural language generation techniques. Recent success of neural sentence summarization models shows the capacity of generating informative, fluent headlines conditioned on selected recapitulative sentences. In this paper, we investigate the extension of sentence summarization models to the document headline generation task. The challenge is that extending the sentence summarization model to consider more document information will mostly confuse the model and hurt the performance. In this paper, we propose a coarse-to-fine approach, which first identifies the important sentences of a document using document summarization techniques, and then exploits a multi-sentence summarization model with hierarchical attention to leverage the important sentences for headline generation. Experimental results on a large real dataset demonstrate the proposed approach significantly improves the performance of neural sentence summarization models on the headline generation task.

1.4. **Time Aware Knowledge Extraction for Summarization**

- **Topic not assigned**

- ▷ Paper: Carmen De Maio, Giuseppe Fenza, Vincenzo Loia, Mimmo Parente, Time Aware Knowledge Extraction for microblog summarization on Twitter, In Information Fusion, Volume 28, 2016, Pages 60-74, ISSN 1566-2535

[Download paper](#)

Abstract: Microblogging services like Twitter and Facebook collect millions of user generated content every moment about trending news, occurring events, and so on. Nevertheless, it is really a nightmare to find information of interest through the huge amount of available posts that are often noisy and redundant. In the era of Big Data, social media analytics services have caught increasing attention from both research and industry. Specifically, the dynamic context of microblogging requires to manage not only meaning of information but also the evolution of knowledge over the timeline. This work defines Time Aware Knowledge Extraction (briefly TAKE) methodology that relies on temporal extension of Fuzzy Formal Concept Analysis. In particular, a microblog summarization algorithm has been defined filtering the concepts organized by TAKE in a time-dependent hierarchy. The algorithm addresses topic-based

summarization on Twitter. Besides considering the timing of the concepts, another distinguishing feature of the proposed microblog summarization framework is the possibility to have more or less detailed summary, according to the user's needs, with good levels of quality and completeness as highlighted in the experimental results.

1.5. Cross-Lingual Event Tracking

□ Topic not assigned

- ▷ Paper: Jan Rupnik, Andrej Muhič, Gregor Leban, Primož Škraba, Blaž Fortuna, and Marko Grobelnik. 2016. News across languages – cross-lingual document similarity and event tracking. *J. Artif. Int. Res.* 55, 1 (January 2016), 283-316.

Download paper

Abstract: *In today's world, we follow news which is distributed globally. Significant events are reported by different sources and in different languages. In this work, we address the problem of tracking of events in a large multilingual stream. Within a recently developed system Event Registry we examine two aspects of this problem: how to compare articles in different languages and how to link collections of articles in different languages which refer to the same event. Taking a multilingual stream and clusters of articles from each language, we compare different cross-lingual document similarity measures based on Wikipedia. This allows us to compute the similarity of any two articles regardless of language. Building on previous work, we show there are methods which scale well and can compute a meaningful similarity between articles from languages with little or no direct overlap in the training data. Using this capability, we then propose an approach to link clusters of articles across languages which represent the same event. We provide an extensive evaluation of the system as a whole, as well as an evaluation of the quality and robustness of the similarity measure and the linking algorithm.*

1.6. Annotator Rationales for Relevance Judgment

□ Topic not assigned

- ▷ Paper: Tyler McDonnell, Matthew Lease, Mucahid Kutlu, and Tamer Elsayed. Why Is That Relevant? Collecting Annotator Rationales for Relevance Judgments. In *Proc. of the 4th AAAI Conference on Human Computation and Crowdsourcing (HCOMP)*, pages 139–148, 2016.

Download paper

Abstract: *When collecting subjective human ratings of items, it can be difficult to measure and enforce data quality due to task subjectivity and lack of insight into how judges' arrive at each rating decision. To address this, we propose requiring judges to provide a specific type of rationale underlying each rating decision. We evaluate this approach in the domain of Information Retrieval, where human judges rate the relevance of Webpages to search queries. Cost-benefit analysis over 10,000 judgments collected on Mechanical Turk suggests a win-win: experienced crowd workers provide rationales with almost no increase in task completion time while providing a multitude of further benefits, including more reliable judgments and greater transparency for evaluating both human raters and their judgments. Further benefits include reduced need for expert gold, the opportunity for dual-supervision from ratings and rationales, and added value from the rationales themselves.*

- ▷ Paper: Tyler McDonnell, Mucahid Kutlu, Tamer Elsayed, Matthew Lease: The Many Benefits of Annotator Rationales for Relevance Judgments. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence Best Sister Conferences*. Pages 4909-4913

Download paper

Abstract: *When collecting subjective human ratings of items, it can be difficult to measure and enforce data quality due to task subjectivity and lack of insight into how judges arrive at each rating decision. To address this, we propose requiring judges to provide a specific type of rationale underlying each rating decision. We evaluate this approach in the domain of Information Retrieval, where human judges rate the relevance of Webpages. Cost-benefit analysis over 10,000 judgments collected on Mechanical Turk suggests a win-win: experienced crowd workers provide rationales with no increase in task completion time while providing further benefits, including more reliable judgments and greater transparency.*

2. Affective Computing

2.1. Affective Loop

□ Topic not assigned

- ▷ Paper: Höök K. (2008) Affective Loop Experiences – What Are They?. In: Oinas-Kukkonen H., Hasle P., Harjumaa M., Segerstahl K., Øhrstrøm P. (eds) *Persuasive Technology. PERSUASIVE 2008. Lecture Notes in Computer Science*, vol 5033. Springer, Berlin, Heidelberg

Download paper

Abstract: *A research agenda for bodily persuasion through a design approach we name affective loops is outlined. Affective loop experiences draw upon physical, emotional interactions between user and system.*

2.2. Psychophysiology in Games

□ **Topic not assigned**

- ▷ Paper: Yannakakis G.N., Martinez H.P., Garbarino M. (2016) Psychophysiology in Games. In: Karpouzis K., Yannakakis G. (eds) Emotion in Games. Socio-Affective Computing, vol 4. Springer, Cham

[Download paper](#)

Abstract: *Psychophysiology is the study of the relationship between psychology and its physiological manifestations. That relationship is of particular importance for both game design and ultimately game-playing. Players' psychophysiology offers a gateway towards a better understanding of playing behavior and experience. That knowledge can, in turn, be beneficial for the player as it allows designers to make better games for them; either explicitly by altering the game during play or implicitly during the game design process. This chapter argues for the importance of physiology for the investigation of player affect in games, reviews the current state of the art in sensor technology and outlines the key phases for the application of psychophysiology in games.*

2.3. Survey emotional body gesture recognition

□ **Topic not assigned**

- ▷ Paper: Noroozi, Fatemeh, Ciprian Adrian Corneanu, Dorota Kamińska, Tomasz Sapiński, Sergio Escalera, and Gholamreza Anbarjafari. "Survey on emotional body gesture recognition." arXiv preprint arXiv:1801.07481 (2018).

[Download paper](#)

Abstract: *Automatic emotion recognition has become a trending research topic in the past decade. While works based on facial expressions or speech abound recognizing affect from body gestures remains a less explored topic. We present a new comprehensive survey hoping to boost research in the field. We first introduce emotional body gestures as a component of what is commonly known as "body language" and comment general aspects as gender differences and culture dependence. We then define a complete framework for automatic emotional body gesture recognition. We introduce person detection and comment static and dynamic body pose estimation methods both in RGB and 3D. We then comment the recent literature related to representation learning and emotion recognition from images of emotionally expressive gestures. We also discuss multi-modal approaches that combine speech or face with body gestures for improved emotion recognition. While pre-processing methodologies (e.g. human detection and pose estimation) are nowadays mature technologies fully developed for robust large scale analysis, we show that for emotion recognition the quantity of labelled data is scarce, there is no agreement on clearly defined output spaces and the representations are shallow and largely based on naive geometrical representations.*

2.4. Affective Computing in Game Design

☒ **Presented by: Beritan Seven**

- ▷ Paper: Nalepa, Grzegorz J., Barbara Gizycka, Krzysztof Kutt, and Jan K. Argasinski. "Affective design patterns in computer games. Scrollrunner case study." In Communication Papers of the 2017 Federated Conference on Computer Science and Information Systems, FedCSIS, vol. 2017, p. 345. 2017.

[Download paper](#)

Abstract: *The emotional state of the user is a new dimension in human-computer interaction, that can be used to improve the user experience. This is the domain of affective computing. In our work we focus on the applications of affective techniques in the design of video games. We assume that a change in the affective condition of a player can be detected based on the monitoring of physiological signals following the James-Lange theory of emotions. We propose the use of game design patterns introduced by Björk and Holopainen to build games. We identify a set of patterns that can be considered affective. Then we demonstrate how these patterns can be used in a design of a scroll-runner game. We address the problem of the calibration of measurements in order to reflect responses of individual users. We also provide results of practical experiments to verify our approach.*

- ▷ Paper: Gizycka, Barbara, Grzegorz J. Nalepa, and Paweł Jemioło. "Aided with emotions"-a new design approach towards affective computer systems. arXiv preprint arXiv:1806.04236 (2018).

[Download paper](#)

Abstract: *As technologies become more and more pervasive, there is a need for considering the affective dimension of interaction with computer systems to make them more human-like. Current demands for this matter include accurate emotion recognition, reliable emotion modeling, and use of unobtrusive, easily accessible and preferably wearable measurement devices. While AI methods provide many possibilities for better affective information processing, it is not a common scenario for both emotion recognition and modeling to be integrated in the design phase. To address this concern, we propose a new approach based on affective design patterns in the context of video games, together with summary of experiments conducted to test the preliminary hypotheses.*

2.5. Affective Level Design

□ **Topic not assigned**

- ▷ Paper: Balducci, F., Grana, C. and Cucchiara, R., 2017. Affective level design for a role-playing videogame evaluated by a brain-computer interface and machine learning methods. *The Visual Computer*, 33(4), pp.413-427.

[Download paper](#)

Abstract: Game science has become a research field, which attracts industry attention due to a worldwide rich sell-market. To understand the player experience, concepts like flow or boredom mental states require formalization and empirical investigation, taking advantage of the objective data that psychophysiological methods like electroencephalography (EEG) can provide. This work studies the affective ludology and shows two different game levels for Neverwinter Nights 2 developed with the aim to manipulate emotions; two sets of affective design guidelines are presented, with a rigorous formalization that considers the characteristics of role-playing genre and its specific gameplay. An empirical investigation with a brain-computer interface headset has been conducted: by extracting numerical data features, machine learning techniques classify the different activities of the gaming sessions (task and events) to verify if their design differentiation coincides with the affective one. The observed results, also supported by subjective questionnaires data, confirm the goodness of the proposed guidelines, suggesting that this evaluation methodology could be extended to other evaluation tasks.

2.6. Deep Learning for Missing Sensor Data

□ **Topic not assigned**

- ▷ Paper: Jaques, Natasha, Sara Taylor, Akane Sano, and Rosalind Picard. "Multimodal autoencoder: A deep learning approach to filling in missing sensor data and enabling better mood prediction." In *Affective Computing and Intelligent Interaction (ACII), 2017 Seventh International Conference on*, pp. 202-208. IEEE, 2017.

[Download paper](#)

Abstract: To accomplish forecasting of mood in real-world situations, affective computing systems need to collect and learn from multimodal data collected over weeks or months of daily use. Such systems are likely to encounter frequent data loss, e.g. when a phone loses location access, or when a sensor is recharging. Lost data can handicap classifiers trained with all modalities present in the data. This paper describes a new technique for handling missing multimodal data using a specialized denoising autoencoder: the Multimodal Autoencoder (MMAE). Empirical results from over 200 participants and 5500 days of data demonstrate that the MMAE is able to predict the feature values from multiple missing modalities more accurately than reconstruction methods such as principal components analysis (PCA). We discuss several practical benefits of the MMAE's encoding and show that it can provide robust mood prediction even when up to three quarters of the data sources are lost.

3. Multimodality

3.1. Multimodal Interaction

□ **Topic not assigned**

- ▷ Paper: M Turk: Multimodal interaction: A review. *Pattern Recognition Letters*, Volume 36, January, 2014, Pages 189-195. Elsevier

[Download paper](#)

Abstract: People naturally interact with the world multimodally, through both parallel and sequential use of multiple perceptual modalities. Multimodal human-computer interaction has sought for decades to endow computers with similar capabilities, in order to provide more natural, powerful, and compelling interactive experiences. With the rapid advance in non-desktop computing generated by powerful mobile devices and affordable sensors in recent years, multimodal research that leverages speech, touch, vision, and gesture is on the rise. This paper provides a brief and personal review of some of the key aspects and issues in multimodal interaction, touching on the history, opportunities, and challenges of the area, especially in the area of multimodal integration. We review the question of early vs. late integration and find inspiration in recent evidence in biological sensory integration. Finally, we list challenges that lie ahead for research in multimodal human-computer interaction.

- ▷ Paper: MR Morris: Web on the wall: insights from a multimodal interaction elicitation study. *Proceedings of the 2012 ACM international conference on Interactive tabletops and surfaces*, Pages 95-104, 2012. ACM

[Download paper](#)

Abstract: New sensing technologies like Microsoft's Kinect provide a low-cost way to add interactivity to large display surfaces, such as TVs. In this paper, we interview 25 participants to learn about scenarios in which they would like to use a web browser on their living room TV. We then conduct an interaction-elicitation study in which users suggested speech and gesture interactions for fifteen common web browser

functions. We present the most popular suggested interactions, and supplement these findings with observational analyses of common gesture and speech conventions adopted by our participants. We also reflect on the design of multimodal, multi-user interaction-elicitation studies, and introduce new metrics for interpreting user-elicitation study findings.

3.2. Gesture & Speech

☒ **Presented by: Nastja Heinrich**

▷ Paper: P Wagner, Z Malisz, S Kopp: Gesture and speech in interaction: An overview. In: Journal Speech Communication, Volume 57, February, 2014. Pages 209-232. Elsevier

[Download paper](#)

Abstract: *Gestures and speech interact. They are linked in language production and perception, with their interaction contributing to felicitous communication. The multifaceted nature of these interactions has attracted considerable attention from the speech and gesture community. This article provides an overview of our current understanding of manual and head gesture form and function, of the principle functional interactions between gesture and speech aiding communication, transporting meaning and producing speech. Furthermore, we present an overview of research on temporal speech-gesture synchrony, including the special role of prosody in speech-gesture alignment. In addition, we provide a summary of tools and data available for gesture analysis, and describe speech-gesture interaction models and simulations in technical systems. This overview also serves as an introduction to a Special Issue covering a wide range of articles on these topics. We provide links to the Special Issue throughout this paper.*

3.3. Multimedia Summarization

☐ **Topic not assigned**

▷ Paper: F Metze, D Ding, E Younessian, A Hauptmann: Beyond audio and video retrieval: topic-oriented multimedia summarization. In: International Journal of Multimedia Information (2013) Vol. 2, pages 131-144. Springer

[Download paper](#)

Abstract: *Given the deluge of multimedia content that is becoming available over the Internet, it is increasingly important to be able to effectively examine and organize these large stores of information in ways that go beyond browsing or collaborative filtering. In this paper we review previous work on audio and video processing, and define the task of Topic-Oriented Multimedia Summarization (TOMS) using natural language generation: given a set of automatically extracted features from a video (such as visual concepts and ASR transcripts) a TOMS system will automatically generate a paragraph of natural language (“a recounting”), which summarizes the important information in a video belonging to a certain topic area, and provides explanations for why a video was matched and retrieved. We see this as a first step towards systems that will be able to discriminate visually similar, but semantically different videos, compare two videos and provide textual output or summarize a large number of videos at once. In this paper, we introduce our approach of solving the TOMS problem. We extract visual concept features and ASR transcription features from a given video, and develop a template-based natural language generation system to produce a textual recounting based on the extracted features. We also propose possible experimental designs for continuously evaluating and improving TOMS systems, and present results of a pilot evaluation of our initial system.*

3.4. Multimodal Ensembles

☐ **Topic not assigned**

▷ Paper: Maskey, S., & Hirschberg, J. (2005). Comparing lexical, acoustic/prosodic, structural and discourse features for speech summarization. In Ninth European Conference on Speech Communication and Technology.

[Download paper](#)

Abstract: *We present results of an empirical study of the usefulness of different types of features in selecting extractive summaries of news broadcasts for our Broadcast News Summarization System. We evaluate lexical, prosodic, structural and discourse features as predictors of those news segments which should be included in a summary. We show that a summarization system that uses a combination of these feature sets produces the most accurate summaries, and that a combination of acoustic/prosodic and structural features are enough to build a “good” summarizer when speech transcription is not available.*

▷ Paper: Sameer Maskey and Julia Hirschberg. 2006. Summarizing speech without text using hidden Markov models. In Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers (NAACL-Short '06). Association for Computational Linguistics, Stroudsburg, PA, USA, 89-92.

[Download paper](#)

Abstract: *We present a method for summarizing speech documents without using any type of transcript/text in a Hidden Markov Model framework. The hidden variables or states in the model repre-*

sent whether a sentence is to be included in a summary or not, and the acoustic/prosodic features are the observation vectors. The model predicts the optimal sequence of segments that best summarize the document. We evaluate our method by comparing the predicted summary with one generated by a human summarizer. Our results indicate that we can generate 'good' summaries even when using only acoustic/prosodic information, which points toward the possibility of text-independent summarization for spoken documents.

- ▷ Paper: Levitan, S. I., An, G., Ma, M., Levitan, R., Rosenberg, A., & Hirschberg, J. (2016). Combining Acoustic-Prosodic, Lexical, and Phonotactic Features for Automatic Deception Detection. In INTERSPEECH (pp. 2006-2010).

[Download paper](#)

Abstract: Improving methods of automatic deception detection is an important goal of many researchers from a variety of disciplines, including psychology, computational linguistics, and criminology. We present a system to automatically identify deceptive utterances using acoustic-prosodic, lexical, syntactic, and phonotactic features. We train and test our system on the Interspeech 2016 ComParE challenge corpus, and find that our combined features result in performance well above the challenge baseline on the development data. We also perform feature ranking experiments to evaluate the usefulness of each of our feature sets. Finally, we conduct a cross-corpus evaluation by training on another deception corpus and testing on the ComParE corpus.

- ▷ Paper: Mendels, G., Levitan, S. I., Lee, K. Z., & Hirschberg, J. (2017). Hybrid Acoustic-Lexical Deep Learning Approach for Deception Detection. Proc. Interspeech 2017, 1472-1476.

[Download paper](#)

Abstract: Automatic deception detection is an important problem with far-reaching implications for many disciplines. We present a series of experiments aimed at automatically detecting deception from speech. We use the Columbia X-Cultural Deception (CXD) Corpus, a large-scale corpus of within-subject deceptive and non-deceptive speech, for training and evaluating our models. We compare the use of spectral, acoustic-prosodic, and lexical feature sets, using different machine learning models. Finally, we design a single hybrid deep model with both acoustic and lexical features trained jointly that achieves state-of-the-art results on the CXD corpus.

4. Understanding & Enhancing Video

4.1. Visual Lecture Transcripts

- **Topic not assigned**

- ▷ Paper: V Shin, F Berthouzoz, W Li, F Durand: Visual transcripts: lecture notes from blackboard-style lecture videos. In: ACM Transactions on Graphics (TOG), Volume 34 Issue 6, November 2015, pages 1-10. ACM

[Download paper](#)

Abstract: Blackboard-style lecture videos are popular, but learning using existing video player interfaces can be challenging. Viewers cannot consume the lecture material at their own pace, and the content is also difficult to search or skim. For these reasons, some people prefer lecture notes to videos. To address these limitations, we present Visual Transcripts, a readable representation of lecture videos that combines visual information with transcript text. To generate a Visual Transcript, we first segment the visual content of a lecture into discrete visual entities that correspond to equations, figures, or lines of text. Then, we analyze the temporal correspondence between the transcript and visuals to determine how sentences relate to visual entities. Finally, we arrange the text and visuals in a linear layout based on these relationships. We compare our result with a standard video player, and a state-of-the-art interface designed specifically for blackboard-style lecture videos. User evaluation suggests that users prefer our interface for learning and that our interface is effective in helping them browse or search through lecture videos.

- ▷ Paper: Monserrat, T. J. K. P., Zhao, S., McGee, K., & Pandey, A. V. (2013, April). NoteVideo: facilitating navigation of blackboard-style lecture videos. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1139-1148). ACM.

[Download paper](#)

Abstract: Khan Academy's pre-recorded blackboard-style lecture videos attract millions of online users every month. However, current video navigation tools do not adequately support the kinds of goals that students typically have, like quickly finding a particular concept in a blackboard-style lecture video. This paper reports on the development and evaluation of the new NoteVideo and its improved version, NoteVideo+, systems for identifying the conceptual 'objects' of a blackboard-based video – and then creating a summarized image of the video and using it as an in-scene navigation interface that allows users to directly jump to the video frame where that object first appeared instead of navigating it linearly through time. The research consisted of iteratively implementing the system and then having users perform four different navigation tasks using three different interfaces: Scrubbing, Transcript, and NoteVideo. Re-

sults of the study show that participants perform significantly better on all four tasks while using the NoteVideo and its improved version - NoteVideo+ - as compared to others

4.2. Enhancing Educational Videos

□ **Topic not assigned**

- ▷ Paper: Pavel, A., Reed, C., Hartmann, B., & Agrawala, M. (2014, October). Video digests: a browsable, skimmable format for informational lecture videos. In UIST (pp. 573-582).

[Download paper](#)

Abstract: *Increasingly, authors are publishing long informational talks, lectures, and distance-learning videos online. However, it is difficult to browse and skim the content of such videos using current timeline-based video players. Video digests are a new format for informational videos that afford browsing and skimming by segmenting videos into a chapter/section structure and providing short text summaries and thumbnails for each section. Viewers can navigate by reading the summaries and clicking on sections to access the corresponding point in the video. We present a set of tools to help authors create such digests using transcript-based interactions. With our tools, authors can manually create a video digest from scratch, or they can automatically generate a digest by applying a combination of algorithmic and crowdsourcing techniques and then manually refine it as needed. Feedback from first-time users suggests that our transcript-based authoring tools and automated techniques greatly facilitate video digest creation. In an evaluative crowdsourced study we find that given a short viewing time, video digests support browsing and skimming better than timeline-based or transcript-based video players.*

- ▷ Paper: Kim, J., Guo, P. J., Cai, C. J., Li, S. W. D., Gajos, K. Z., & Miller, R. C. (2014, October). Data-driven interaction techniques for improving navigation of educational videos. In Proceedings of the 27th annual ACM symposium on User interface software and technology (pp. 563-572). ACM.

[Download paper](#)

Abstract: *With an unprecedented scale of learners watching educational videos on online platforms such as MOOCs and YouTube, there is an opportunity to incorporate data generated from their interactions into the design of novel video interaction techniques. Interaction data has the potential to help not only instructors to improve their videos, but also to enrich the learning experience of educational video watchers. This paper explores the design space of data-driven interaction techniques for educational video navigation. We introduce a set of techniques that augment existing video interface widgets, including: a 2D video timeline with an embedded visualization of collective navigation traces; dynamic and non-linear timeline scrubbing; data-enhanced transcript search and keyword summary; automatic display of relevant still frames next to the video; and a visual summary representing points with high learner activity. To evaluate the feasibility of the techniques, we ran a laboratory user study with simulated learning tasks. Participants rated watching lecture videos with interaction data to be efficient and useful in completing the tasks. However, no significant differences were found in task performance, suggesting that interaction data may not always align with moment-by-moment information needs during the tasks*

4.3. Survey: Human Activity Recognition from Video

□ **Topic not assigned**

- ▷ Paper: L Onofri, P Soda, M Pechenizkiy, G Iannello: A survey on using domain and contextual knowledge for human activity recognition in video streams. Expert Systems with Applications, Volume 63, 30 November 2016, Pages 97-111. Elsevier

[Download paper](#)

Abstract: *Human activity recognition has gained an increasing relevance in computer vision and it can be tackled with either non-hierarchical or hierarchical approaches. The former, also known as single-layered approaches, are those that represent and recognize human activities directly from the extracted descriptors, building a model that distinguishes among the activities contained in the training data. The latter represent and recognize human activities in terms of subevents, which are usually recognized by means of single-layered approaches. Alongside of non-hierarchical and hierarchical approaches, we observe that methods incorporating a priori knowledge and context information on the activity are getting growing interest within the community. In this work we refer to this emerging trend in computer vision as knowledge-based human activity recognition with the objective to cover the lack of a summary of these methodologies. More specifically, we survey methods and techniques used in the literature to represent and integrate knowledge and reasoning into the recognition process. We categorize them as statistical approaches, syntactic approaches and description-based approaches. In addition, we further discuss public and private datasets used in this field to promote their use and to enable the interest readers in finding useful resources. This review ends proposing main future research directions in this field.*

4.4. Event Recognition in Unconstrained Videos

□ **Topic not assigned**

- ▷ Paper: Jiang, YG., Bhattacharya, S., Chang, SF. et al.: High-level event recognition in unconstrained videos. *Int J Multimed Info Retr* (2013) 2: 73.

[Download paper](#)

Abstract: The goal of high-level event recognition is to automatically detect complex high-level events in a given video sequence. This is a difficult task especially when videos are captured under unconstrained conditions by non-professionals. Such videos depicting complex events have limited quality control, and therefore, may include severe camera motion, poor lighting, heavy background clutter, and occlusion. However, due to the fast growing popularity of such videos, especially on the Web, solutions to this problem are in high demands and have attracted great interest from researchers. In this paper, we review current technologies for complex event recognition in unconstrained videos. While the existing solutions vary, we identify common key modules and provide detailed descriptions along with some insights for each of them, including extraction and representation of low-level features across different modalities, classification strategies, fusion techniques, etc. Publicly available benchmark datasets, performance metrics, and related research forums are also described. Finally, we discuss promising directions for future research.

5. Social and Explanatory Behaviour

5.1. Explanation and Trust in Security and AI

☒ **Presented by: Ardian Musolli**

- ▷ Paper: Pieters, Wolter. "Explanation and trust: what to tell the user in security and AI?." *Ethics and information technology* 13.1 (2011): 53-64.

[Download paper](#)

Abstract: There is a common problem in artificial intelligence (AI) and information security. In AI, an expert system needs to be able to justify and explain a decision to the user. In information security, experts need to be able to explain to the public why a system is secure. In both cases, an important goal of explanation is to acquire or maintain the users' trust. In this paper, I investigate the relation between explanation and trust in the context of computing science. This analysis draws on literature study and concept analysis, using elements from system theory as well as actor-network theory. I apply the conceptual framework to both AI and information security, and show the benefit of the framework for both fields by means of examples. The main focus is on expert systems (AI) and electronic voting systems (security). Finally, I discuss consequences of the analysis for ethics in terms of (un)informed consent and dissent, and the associated division of responsibilities.

5.2. Explanation in AI: Insights from the social sciences

□ **Topic not assigned**

- ▷ Paper: Miller, T. (2017). Explanation in Artificial Intelligence: Insights from the social sciences. [arXiv:1706.07269]

[Download paper](#)

Abstract: There has been a recent resurgence in the area of explainable artificial intelligence as researchers and practitioners seek to make their algorithms more understandable. Much of this research is focused on explicitly explaining decisions or actions to a human observer, and it should not be controversial to say that looking at how humans explain to each other can serve as a useful starting point for explanation in artificial intelligence. However, it is fair to say that most work in explainable artificial intelligence uses only the researchers' intuition of what constitutes a 'good' explanation. There exists vast and valuable bodies of research in philosophy, psychology, and cognitive science of how people define, generate, select, evaluate, and present explanations, which argues that people employ certain cognitive biases and social expectations towards the explanation process. This paper argues that the field of explainable artificial intelligence should build on this existing research, and reviews relevant papers from philosophy, cognitive psychology/science, and social psychology, which study these topics. It draws out some important findings, and discusses ways that these can be infused with work on explainable artificial intelligence.

5.3. Explainable Intelligent Robots

☒ **Presented by: Ghassen Trabelsi**

- ▷ Paper: Sheh, Raymond K. (2017). "Why did you do that?" Explainable intelligent robots. In K. Talamadupula, S. Sohrabi, L. Michael, & B. Srivastava (Eds.) *Human-Aware Artificial Intelligence: Papers from the AAAI Workshop (Technical Report WS-17-11)*. San Francisco, CA: AAAI Press.

[Download paper](#)

Abstract: *As autonomous intelligent systems become more widespread, society is beginning to ask: “What are the machines up to?”. Various forms of artificial intelligence control our latest cars, load balance components of our power grids, dictate much of the movement in our stock markets and help doctors diagnose and treat our ailments. As they become increasingly able to learn and model more complex phenomena, so the ability of human users to understand the reasoning behind their decisions often decreases. It becomes very difficult to ensure that the robot will perform properly and that it is possible to correct errors. In this paper, we outline a variety of techniques for generating the underlying knowledge required for explainable artificial intelligence, ranging from early work in expert systems through to systems based on Behavioural Cloning. These are techniques that may be used to build intelligent robots that explain their decisions and justify their actions. We will then illustrate how decision trees are particularly well suited to generating these kinds of explanations. We will also discuss how additional explanations can be obtained, beyond simply the structure of the tree, based on knowledge of how the training data was generated. Finally, we will illustrate these capabilities in the context of a robot learning to drive over rough terrain in both simulation and in reality*

5.4. Context-Aware Analysis and Annotation of Human-Agent Interaction

□ **Topic not assigned**

- ▷ Paper: Baur, Tobias, Gregor Mehlmann, Ionut Damian, Florian Lingenfelser, Johannes Wagner, Birgit Lugin, Elisabeth André, and Patrick Gebhard. “Context-Aware Automated Analysis and Annotation of Social Human-Agent Interactions.” *ACM Transactions on Interactive Intelligent Systems (TiS)* 5, no. 2 (2015): 11.

[Download paper](#)

Abstract: *The outcome of interpersonal interactions depends not only on the contents that we communicate verbally, but also on nonverbal social signals. As a lack of social skills is a common problem for a significant number of people, serious games and other training environments have recently become the focus of research. In this work we present NovA (Nonverbal behavior Analyzer), a system that analyzes and facilitates the interpretation of social signals automatically in a bi-directional interaction with a conversational agent. It records data of interactions, detects relevant social cues, and creates descriptive statistics for the recorded data with respect to the agents behavior and the context of the situation. This enhances the possibilities for researchers to automatically label corpora of human-agent interactions and to give users feedback on strengths and weaknesses of their social behavior*

5.5. Machine Learning for Semi-Automated Annotation of Social Signals

□ **Topic not assigned**

- ▷ Paper: Baur, Tobias, Ionut Damian, Florian Lingenfelser, Johannes Wagner, and Elisabeth André. “Nova: Automated analysis of nonverbal signals in social interactions.” In *International Workshop on Human Behavior Understanding*, pp. 160-171. Springer, Cham, 2013.

[Download paper](#)

Abstract: *Previous studies have shown that the success of interpersonal interaction depends not only on the contents we communicate explicitly, but also on the social signals that are conveyed implicitly. In this paper, we present NovA (NOVerbal behavior Analyzer), a system that analyzes and facilitates the interpretation of social signals conveyed by gestures, facial expressions and others automatically as a basis for computer-enhanced social coaching. NovA records data of human interactions, automatically detects relevant behavioral cues as a measurement for the quality of an interaction and creates descriptive statistics for the recorded data. This enables us to give a user online generated feedback on strengths and weaknesses concerning his social behavior, as well as elaborate tools for online analysis and annotation.*

- ▷ Paper: Wagner, Johannes, Tobias Baur, Dominik Schiller, Yue Zhang, Björn Schuller, Michel Valstar, and Elisabeth André. “Show Me What You’ve Learned: Applying Cooperative Machine Learning for the Semi-Automated Annotation of Social Signals.” *XAI 2018*: 171.

[Download paper](#)

Abstract: *In this paper we suggest the use of Cooperative Machine Learning (CML) to reduce manual labelling efforts while simultaneously generating an intuitive understanding of the learning process of a classification system. To this end, we introduce the open-source tool NOVA, which aims to combine human intelligence and machine learning to an-notate social signals in large multi-modal corpora. NOVA features a semi-automated labelling process in which users are provided with immediate visual feedback on the predictions, which affords insights into the strengths and weaknesses of the underlying classification system. Following an interactive and exploratory workflow, the performance of the model can be improved by manual revision of the predictions, a process that uses confidence values to guide the inspection.*

5.6. Attitudes towards potentially uncanny robots

□ **Topic not assigned**

- ▷ Paper: Rosenthal-von der Pütten, Astrid M., and Nicole C. Krämer. “Individuals’ evaluations of and attitudes towards potentially uncanny robots.” *International Journal of Social Robotics*

7, no. 5 (2015): 799-824.

[Download paper](#)

*Abstract: In the present work we provide an overview and categorization of explanatory approaches for the uncanny valley effect and present an empirical study. Against the background of the uncanny valley hypothesis, the study utilized qualitative interviews in which participants were presented with pictures and videos of potentially uncanny humanoid and android robots to explore participants' evaluations of very human-like robots, their attitudes about these robots, and their emotional reactions towards these robots. In this regard, the influence of the robots' appearance, movement and the context of HRI were examined. Results showed that, contrasting the hypothesis, participants reported not only negative, but also positive emotional reactions towards the possibly uncanny robots. The robots' appearance was of great importance for the participants, because certain characteristics were equalized with certain abilities, merely human appearance without a connected functionality was not appreciated, and human rules of attractiveness were applied to the android robots. The analysis also demonstrated the importance of the robots' movements and the social context they were placed in. The importance of two possible causes and explanations of the uncanny valley, namely uncertainty at category boundaries (cf. Ramey in *Proceedings of views of the uncanny valley workshop: IEEE-RAS international conference on humanoid robots, 2005*; *Proceedings of the ICCS/CogSci-2006 long symposium "Toward Social Mechanisms of Android Science"*, 2006) and subconscious fears of being replaced (cf. MacDorman & Ishiguro in *Interact Stud* 7(3):297–337, 2006) were explored in this work. On this reflective level of evaluation we found some support for the assumptions that participants experienced uncertainty how to categorize android robots (as human or machine) and that some (but not all) participants felt uncomfortable at the thought to be replaced by robots.*

6. Ethical Computing

6.1. Research Priorities for Robust and Beneficial AI

☒ **Presented by: Thorben Werner**

▷ Paper: Russell, Stuart, Daniel Dewey, and Max Tegmark. "Research priorities for robust and beneficial artificial intelligence." *Ai Magazine* 36, no. 4 (2015): 105-114.

[Download paper](#)

Abstract: Success in the quest for artificial intelligence has the potential to bring unprecedented benefits to humanity, and it is therefore worthwhile to investigate how to maximize these benefits while avoiding potential pitfalls. This article gives numerous examples (which should by no means be construed as an exhaustive list) of such worthwhile research aimed at ensuring that AI remains robust and beneficial.

6.2. Classification of Moral Mediation by Ubiquitous Machines

☐ **Topic not assigned**

▷ Paper: Van de Voort, Marlies, Wolter Pieters, and Luca Consoli. "Refining the ethics of computer-made decisions: a classification of moral mediation by ubiquitous machines." *Ethics and Information Technology* 17, no. 1 (2015): 41-56.

[Download paper](#)

Abstract: In the past decades, computers have become more and more involved in society by the rise of ubiquitous systems, increasing the number of interactions between humans and IT systems. At the same time, the technology itself is getting more complex, enabling devices to act in a way that previously only humans could, based on developments in the fields of both robotics and artificial intelligence. This results in a situation in which many autonomous, intelligent and context-aware systems are involved in decisions that affect their environment. These relations between people, machines, and decisions can take many different forms, but thus far, a systematic account of machine-assisted moral decisions is lacking. This paper investigates the concept of machine-assisted moral decisions from the perspective of technological mediation. It is argued that modern machines do not only have morality in the sense of mediating the actions of humans, but that, by making their own decisions within their relations with humans, mediate morality itself. A classification is proposed to differentiate between four different types of moral relations. The moral aspects within the decisions these systems make are combined into three dimensions that describe the distinct characteristics of different types of moral mediation by machines. Based on this classification, specific guidelines for moral behavior can be provided for these systems.

6.3. Social Dilemma of Autonomous Vehicles

☒ **Presented by: Claudia Sommer**

- ▷ Paper: Bonnefon, Jean-François, Azim Shariff, and Iyad Rahwan. "The social dilemma of autonomous vehicles." *Science* 352, no. 6293 (2016): 1573-1576.

[Download paper](#)

Abstract: *Autonomous vehicles (AVs) should reduce traffic accidents, but they will sometimes have to choose between two evils, such as running over pedestrians or sacrificing themselves and their passenger to save the pedestrians. Defining the algorithms that will help AVs make these moral decisions is a formidable challenge. We found that participants in six Amazon Mechanical Turk studies approved of utilitarian AVs (that is, AVs that sacrifice their passengers for the greater good) and would like others to buy them, but they would themselves prefer to ride in AVs that protect their passengers at all costs. The study participants disapprove of enforcing utilitarian regulations for AVs and would be less willing to buy such an AV. Accordingly, regulating for utilitarian algorithms may paradoxically increase casualties by postponing the adoption of a safer technology.*

6.4. Designing AI for Human Values

□ **Topic not assigned**

- ▷ Paper: Dignum, Virginia. "Responsible Artificial Intelligence: Designing Ai for Human Values." (2017).

[Download paper](#)

Abstract: *Artificial intelligence (AI) is increasingly affecting our lives in smaller or greater ways. In order to ensure that systems will uphold human values, design methods are needed that incorporate ethical principles and address societal concerns. In this paper, we explore the impact of AI in the case of the expected effects on the European labor market, and propose the accountability, responsibility and transparency (ART) design principles for the development of AI systems that are sensitive to human values.*

7. Historic Milestones

7.1. Sword of Damocles

□ **Topic not assigned**

- ▷ Paper: Sutherland, Ivan E. A head-mounted three dimensional display. Proceedings of the 1968 Fall Joint Computer Conference, San Francisco, CA, December 1968, Pages 757-764. ACM

[Download paper](#)

Abstract: *The fundamental idea behind the three-dimensional display is to present the user with a perspective image which changes as he moves. The retinal image of the real objects which we see is, after all, only two-dimensional. Thus if we can place suitable two-dimensional images on the observer's retinas, we can create the illusion that he is seeing a three-dimensional object. Although stereo presentation is important to the three-dimensional illusion, it is less important than the change that takes place in the image when the observer moves his head. The image presented by the three-dimensional display must change in exactly the way that the image of a real object would change for similar motions of the user's head. Psychologists have long known that moving perspective images appear strikingly three-dimensional even without stereo presentation; the three-dimensional display described in this paper depends heavily on this "kinetic depth effect."*

7.2. Augmenting Human Intellect

□ **Topic not assigned**

- ▷ Paper: C Engelbart, and WK English: A research center for augmenting human intellect. Proceedings of the 1968 Fall Joint Computer Conference, San Francisco, CA, December 1968, Vol. 33, pp. 395-410. ACM

[Download paper](#)

Abstract: *This paper describes a multisponsor research center at Stanford Research Institute in man-computer interaction.*

7.3. Put That There

☒ **Presented by: Carlo Morgenstern**

- ▷ Paper: RA Bolt: "Put-that-there": Voice and gesture at the graphics interface. SIGGRAPH '80 Proceedings of the 7th annual conference on Computer graphics and interactive techniques, Pages 262-270, 1980.

[Download paper](#)

Abstract: *Recent technological advances in connected-speech recognition and position sensing in space have encouraged the notion that voice and gesture inputs at the graphics interface can converge to provide a concerted, natural user modality. The work described herein involves the user commanding simple shapes about a large-screen graphics display surface. Because voice can be augmented with simultaneous*

pointing, the free usage of pronouns becomes possible, with a corresponding gain in naturalness and economy of expression. Conversely, gesture aided by voice gains precision in its power to reference.

7.4. Tangible Bits

□ **Topic not assigned**

- ▷ Paper: Ishii, Hiroshi, and Brygg Ullmer. "Tangible bits: towards seamless interfaces between people, bits and atoms." Proceedings of the ACM SIGCHI Conference on Human factors in computing systems. ACM, 1997.

[Download paper](#)

Abstract: *This paper presents our vision of Human Computer Interaction (HCI): "Tangible Bits." Tangible Bits allows users to "grasp & manipulate" bits in the center of users' attention by coupling the bits with everyday physical objects and architectural surfaces. Tangible Bits also enables users to be aware of background bits at the periphery of human perception using ambient display media such as light, sound, airflow, and water movement in an augmented space. The goal of Tangible Bits is to bridge the gaps between both cyberspace and the physical environment, as well as the foreground and background of human activities. This paper describes three key concepts of Tangible Bits: interactive surfaces; the coupling of bits with graspable physical objects; and ambient media for background awareness. We illustrate these concepts with three prototype systems – the metaDESK, transBOARD and ambientROOM – to identify underlying research issues.*

8. Transdisciplinary Issues

8.1. Systemic-Functional Theory of Language: Semantic Networks

□ **Topic not assigned**

- ▷ Paper: R Hasan, C Cloran, G Williams and A Lukin: Semantic networks: the description of linguistic meaning in SFL. In R. Hasan, C. MIM. Matthiessen & J. J. Webster (eds), *Continuing Discourse on Language: A Functional Perspective Volume 2*, 1 edn, pp. 697-738, Equinox Publishing Ltd, London, UK and Oakville, USA, 2005.

[Download paper](#)

Abstract: *In SFL, the semantics of a language calls for as much attention as its lexicogrammar: in fact, meaning and wording are two sides of the same coin; the description of both is equally central to understanding 'how language works' (Halliday, McIntosh and Strevens, 1964) – which has been Halliday's agenda since the beginning of his engagement with linguistics. But what actually led SFL into the exploration of semantics as a legitimate domain for description was not these theoretical considerations, per se; rather, like other aspects of the evolution of SFL, interest in semantics too arose in attempts to resolve certain problems in the course of research during the 1960s. This chapter presents one perspective on the course of this development, specifically with respect to semantic networks as a resource for the analysis of meaning.*

8.2. Systemic-Functional Theory of Language: Academic Writing

□ **Topic not assigned**

- ▷ Paper: D Motta-Roth: The Role of Context in Academic Text Production and Writing Pedagogy. In: C Bazerman, A Bonini, D Figueiredo (eds), *Genre in a Changing World*, pp. 317-336, Parlor Press, 2009.

[Download paper](#)

Abstract: *The problem of text production in academic genres has been a challenge for undergraduate and graduate students as well as for writing teachers from different departments. Previous research has provided important results on the structural aspects of academic genres (Swales, 1990) and the discursive construction of identity in academic writing (Ioanic, 1998). However, few studies have concentrated on the contributions of exploration and reflection on context to actual teaching practices. From the perspective of Systemic-Functional Linguistics (SFL), in this paper I would like to focus on the reciprocal relationship between text and context, i.e., the way context can be recreated by analysis of text and vice versa. The aim is to point out some practical implications derived from the use of SFL principles in academic writing teaching and research through context exploration. The focus will be on writing activities that aim at fostering students' awareness about the connections between contextual features (activity, identity, relations as well as the role performed by text in the situation) and their respective linguistic realizations (expression of content, instantiation of relationships between interlocutors, and organization of text). One of the main challenges in language education and research is to teach creative ways to negotiate the norms of the language system (grammar) within the academic culture: the set of meanings, rules, values, power relations and relevant genres that constitute the social practices of a community. Educating students about the uses of language in specific contexts depends on clear descriptions of the connections between text and context. Public discourse on academic publication in Brazil is mainly issued by the Ministry of Education through its two main Research Funding Agencies, CAPES*

and CNPq, which hold quantitative and qualitative expectations about scholars' intellectual production but offer no substantial line of financial support for pedagogic research and course development. Thus full-fledged writing programs are seldom found in Brazilian universities. Very often what we find is some individual or collective teaching initiatives that have survived defying all the odds (e.g., lack of personnel and financial resources) situated in specific institutions.

8.3. Activity Theory: Understanding Acting in Complex Environments

□ **Topic not assigned**

- ▷ Paper: Leena Norros: "Understanding Acting in Complex Environments: Building a Synergy of Cultural-Historical Activity Theory, Peirce, and Ecofunctionalism". In: *Mind, Culture, and Activity*, 25:1, 68-85, 2018

[Download paper](#)

Abstract: *A method of activity analysis is proposed that exploits synergy among cultural-historical activity theory, Peircean sign theory, and ecofunctionalism in response to challenges of modern work. The method comprises an operationalisation of the object of activity via the concept of core task, which enables contextual description of actions. On this basis, a semiotic analysis can be accomplished that reveals generic habitual ways of acting. The variation found in their interpretive strength for comprehending situations predicts differences in the mastery of work and learning in work. The method also allows definition of systems usability as a holistic quality-linked concept for purposes of artefact design.*

8.4. Activity Theory: Plans as Situated Action

□ **Topic not assigned**

- ▷ Paper: Bardram, Jakob E. "Plans as situated action: an activity theory approach to workflow systems." *Proceedings of the Fifth European Conference on Computer Supported Cooperative Work*. Springer Netherlands, 1997.

[Download paper](#)

Abstract: *Within the community of CSCW the notion and nature of workflow systems as prescriptions of human work has been debated and criticised. Based on the work of Suchman (1987) the notion of situated action has often been viewed as opposed to planning work. Plans, however, do play an essential role in realising work. Based on experiences from designing a computer system that supports the collaboration within a hospital, this paper discusses how plans themselves are made out of situated action, and in return are realised in situ. Thus, work can be characterised as situated planning. This understanding is backed up by Activity Theory, which emphasises the connection between plans and the contextual conditions for realising these plans in actual work.*

Source

Texts can be downloaded from the [Learnweb-course](#).

List of the 30 topics which have not been assigned yet: 1.2., 1.3., 1.4., 1.5., 1.6., 2.1., 2.2., 2.3., 2.5., 2.6., 3.1., 3.3., 3.4., 4.1., 4.2., 4.3., 4.4., 5.2., 5.4., 5.5., 5.6., 6.2., 6.4., 7.1., 7.2., 7.4., 8.1., 8.2., 8.3., 8.4., Please see individual papers for a description.