Human Collaboration Reshaped: Applications and Perspectives

Martin Bogner, François Bry, Niels Heller, Stephan Leutenmayr, Sebastian Mader, Alexander Pohl, Clemens Schefels, Yingding Wang, and Christoph Wieser
Institute for Informatics, Ludwig-Maximilian University of Munich, Oettingenstrasse 67, 80538 Munich, Germany
bry@lmu.de

Abstract

20th century iconic examples of human collaboration are the assembly line, centralised planning, bureaucracies, vote-based decision making, and school education. These examples, and more generally all forms of human collaboration of the 20th century, are characterized by predefined human roles and little adaptable processes, that is, 20th century collaboration comes at the price of a restricted individual freedom. With the turn of the century, new forms of human collaboration have become widespread that exploit information and communication technologies, data generated by humans, Data Science in general and Machine Learning in particular, and let humans contribute as they like, when they like, and as much as they can, the lack of predefined roles and processes being accounted for by software. The phrase "Human Computation" coined for denoting the new forms of human collaboration stresses a core aspect of the paradigm which can be a downside: With Human Computation, humans become contributors to collaboration-enabling algorithms that can also control and restrict how collaboration takes place. This article introduces to Human Computation and to its role in applications of Machine Learning, presents Human Computation prototype systems developed during the last decade at Ludwig-Maximilian University of Munich, discusses ethical issues of Human Computation and Machine Learning, points to on-going research in the field at Ludwig-Maximilian University of Munich, and concludes with a reflection on the future of Human Computation. The original contribution of this article is a comprehensive presentation of recent research the main part of which has already been published in more detail elsewhere.

1 Introduction

20th century iconic examples of human collaboration are the assembly line, centralised planning, bureaucracies, vote-based decision making, and school education. These examples stress two essential characteristics of 20th century human collaboration: Predefined human roles and little adaptable processes. These characteristics have shaped how our culture views work, education, and more generally all kinds of collective endeavours. Like former cultures could hardly envisage societies but feudally structured, we tend to see in predefined human roles and little adaptable processes unavoidable conditions of collective action. We accept a restricted individual freedom as a necessary price of human collaboration –especially in education and at the workplace. With the turn of the century, however, new forms of human collaboration have become widespread that exploit information and communication technologies, data generated by humans, and Data Science, for letting humans contribute as they like, when they like, and as much as they can or wish. The new forms of human collaboration rely on software for accounting for the lack of predefined roles and processes.

Two businesses among many others owe their considerable success to the new forms of human collaboration: Google and Amazon. The first algorithm used by Google Search for ranking search results, PageRank, exploits the link structure of the Web, that is, data provided by Web document authors. Google Search now relies on many more algorithms, most of them based like PageRank on data humans contribute with –among others, the search results users click on and the rephrasing of queries by users when answers do not satisfy them. Amazon's success builds upon its product search engine and its product recommendations that both exploit data that customers leave when browsing the catalogue, purchasing, making wish lists, and submitting product reports.

The phrase "Human Computation" [1, 2] coined for the new collaboration paradigm stresses a paradigm's core aspect and potential downside: The new forms of human collaboration not only grant humans more freedom than they ever had while participating to collective endeavours, they also turn humans into contributors to algorithms that might control collaboration more than the contributing humans wish.

This article first introduces to Human Computation and to its essential role in many applications of Machine Learning. Then, it presents Human Computation prototype systems developed at Ludwig-Maximilian University of Munich and briefly outlines on-going research at the same university on a Citizen Science project, on algorithmically sustained human participation to Human Computation systems and on ethical issues of Human Computation and Machine Learning. It concludes with reflections on the future of the reshaped human collaboration.

The contribution of this article is an original and comprehensive presentation mostly of recently published research on novel applications of Human Computation. Section 7 on ethical issues of Human Computation and Machine Learning contains material from François Bry's lectures on Human Computation and original, so far unpublished, reflections.

2 Human Computation

The Web, initially conceived as a distributed infrastructure for the exchange of online documents, has made possible outsourcing to wide, often worldwide, workers' pools, what has been called "crowdsourcing" [3]. Online markets like Amazon Mechanical Turk (for micro-tasks), iStock (for photographs), and Innocentive (for technical innovation) have established crowdsourcing as an effective approach to reduce costs by exploiting comparative advantages and by tapping into an expertise outside a business' core competency.

Human Computation [1, 2, 4] builds upon crowdsourcing and goes one step further by relying on "systems that interconnect humans and machines that process information as a system, and [that] serve a purpose" [5] or "systems taking a group of individuals and turning them into a thinking system, a kind of superorganism" [6]. Human Computation systems can also be described as algorithms with "humans in the loop" as conscious or implicit contributors of data or computations. Human Computation thus refers to new, software-based, forms of human collaboration.

Human Computation is ubiquitous: Human Computation is used among others by search engines, for navigation and traffic monitoring (among others by Google maps), for health monitoring (by fitness apps), for recommendations (among others by Amazon), for natural language processing (especially for translation among others by Google and Facebook).

Human Computation mostly takes the form of a "bartered crowdsourcing" on "give and take" Web or mobile apps: A service (like search, navigation, fitness tracking, or translation) is offered for free and the service provider generates a profit from the processing of "footprints" left by users of the service. The data processing of Human Computation systems ranges from relatively trivial computations (like averages) to highly sophisticated Data Science methods (like Latent Semantic Analysis, eigenvector centralities, and Deep Learning). The denomination "Human Computation" stresses that users' footprints often result from non-trivial intellectual activities (like a choice of matching products, face recognition, and natural language translations).

The distinction between Crowdsourcing market and Human Computation might be questioned since, like any other market, a Crowdsourcing market can be seen as a Human Computation system. Indeed, markets interconnect humans and machines, process information as systems, and serve a purpose, namely a resource allocation socially accepted as fair. Admittedly, in the past markets did not rely on computers but instead on the (sometimes sophisticated) computations of market makers, that is, human agents ensuring markets' liquidity. Does a socially accepted resource allocation, the purpose of markets, qualify markets as Human Computation systems? If one takes markets and their functioning for granted, as many "free market" advocates do, then it makes sense to deny markets, among others Crowdsourcing markets, the status of Human Computation systems. If, in contrast, one considers the evolving and mostly complex regulations underlying all markets, then markets can be seen as Human Computation systems. Thus, the distinction between Crowdsourcing market and Human Computation is akin to that between free market and social market economy: It is ideological and depends on whether one disregards or not a market's social underpinning. The distinction between Crowdsourcing market and Human Computation can also be questioned by considering existing systems. In practice, a Crowdsourcing market offers its users services like search, personalisation, and recommendations that are enabled by Human Computation, blurring the distinction between Crowdsourcing and Human Computation.

Users of Human Computation systems can consciously contribute to the systems. This is for example the case when someone "likes" news on Facebook or submits a product report on Amazon. Contributions to Human Computation systems can also be implied. Implied contributions include among others the forwarding by a user to other users of news on Facebook or Twitter and the use of a navigation system. Indeed, forwarding can be interpreted as a strong form of liking and a navigation system knows where its users are and therefore can, among others, "learn" from its users expectable commute times, the location of traffic jams, and traffic jam escape routes. Exploiting implied contributions is not new. Banks and credit card companies have exploited data on their customers' consumptions and insights computed from these data long before the advent of the Internet. The ubiquity of the Internet and of social media have given to the approach unprecedented outreach and power.

An interplay of Human Computation and Machine Learning can often be observed. On the one hand, Human Computation systems often derive new insights from the human contributions they collect by applying Machine Learning methods: Parameters of mathematical models are set using data collected from humans, the quality of the models' outcomes –for example search results or product recommendations– being constantly checked against further data collected from humans, so as to constantly adjust the models. On the other hand, many applications of Machine Learning would not be possible without Human Computation. The interplay of Human Computation and Machine Learning in many current artificial intelligences is addressed in more detail in the next section.

Social media and online commerce platforms all rely on Human Computation and Machine Learning for offering services (like search, recommendation, and personalisation) to their users, sustaining customer loyalty, for targeted advertising and, more generally, in their striving for growth. Many Web sites rely on Human Computation and Machine Learning for collecting data on the sites' usability.

3 Machine Learning, Knowledge Engineering, and Human Computation

In the quest for "artificial intelligences", mostly two approaches have been considered: Machine Learning and Knowledge Engineering.

Machine Learning comes in two forms, supervised and unsupervised. Supervised Machine Learning consists in the following steps:

1) choosing a mathematical model of relationships between variables of interest,
2) collecting, cleaning, and formatting relevant data,
3) training the model, that is, adjusting its parameters, using some of the data obtained at step 2,
4) inferring relationships for the considered variables by applying the trained model to those data obtained at step 2 that have not been used at step 3,
5) checking the validity of the relationships inferred at step 4,
6) in case the validity check is not conclusive, refining the model and/or adjusting it using new data (that is, repeating the process from step 1 or 2).

Supervision in Supervised Machine Learning refers to training models using data. Unsupervised Machine Learning is similar to Supervised Machine Learning except that it lacks step 3: Unsupervised Machine Learning uses no training data. Unsupervised Machine Learning relies on data only for checking the validity of the relationships inferred from the mathematical model considered. Human Computation is, in practice, the method of choice for collecting the data needed for Machine Learning.

Knowledge Engineering, in contrast to Machine Learning, consists in specifying reasoning systems based on inference rules and logical axioms. Thus, Knowledge Engineering relies like Machine Learning on mathematical models but all models of Knowledge Engineering are "symbolic" in the sense that they are specified in terms of inference rules and logical axioms. Machine Learning, in contrast, relies on models expressed in many more mathematical formalisms including Statistics, Probability Theory and Linear Algebra.

An attractive feature of Knowledge Engineering is that inference rules and axioms are easily understandable for experts of an application field –a physician understands, for example, the inference rules and axioms of a medical expert system. In contrast, most Data Science models –including Supervised and Unsupervised Machine Learning models– are rather arcane and therefore rarely understood by application field experts what shifts the control on applications from application experts to Data Science experts.

Nonetheless, the last fifteen years have seen a considerable rise of Machine Learning and, in comparison, a stagnation of Knowledge Engineering. Knowledge Engineering methods, like ontologies, are still used in practice but mostly ancillary to Machine Learning. An essential reason for today's preeminence of Machine Learning over Knowledge Engineering is that after a few refinements of a model, the aforementioned improvement step 6 can be performed only with data –an easy task that can easily be automated and quickly performed. In contrast, improving a symbolic model of a Knowledge Engineering tool requires, of course, engineering –an expert task which therefore is time-consuming and expensive.

While developing the search engine YASA [7, 8, 9] with Roche Pharmaceuticals we observed the aforementioned superiority of Machine Learning over Knowledge Engineering for the classification of textual documents. YASA classifies textual documents using the Supervised Machine Learning method "Support Vector Machine" [10] with 10-fold cross validation, a state-of-the-art approach to document classification [11, 12]. Knowledge Engineering methods turned out to be useful for expressing the static and limited knowledge needed for specifying YASA's adaptation and guidance. YASA adapts search results and guides a searcher's exploration of search results after her role in the corporation and, to this aim, exploits knowledge available in corporate documents. A role-based instead of user-based adaptation has been retained for YASA for privacy reasons. YASA uses Knowledge Engineering techniques in the form of the following three ontologies –that is, predefined concepts and relations between concepts:

- A "classification ontology" models unstructured textual documents and their properties.
- An "annotation ontology" models named entities referred to by textual documents.
- An "adaptation ontology" models users' roles in the corporation.

YASA's classification and annotation ontologies are built using Machine Learning methods [10, 11, 12, 13].

YASA is a good example of the interplay of Machine Learning, Knowledge Engineering, and Human Computation in many current applications. Machine Learning is used for brute force computations that have to be often updated. Knowledge Engineering is used for specifying more limited and more static information as well as for representing outcomes of Machine Learning. In applications like search that prominently refer to changing information, Machine Learning is prominent. Human Computation is used for collecting usage data –in the case of YASA on searchers' interests or roles– needed for Machine Learning.

In academic circles, Machine Learning often has a restrictive acceptation: An eigenvector centrality index would for example not be seen as a Machine Learning method. Indeed, eigenvector centralities have been developed and investigated in a research field known as Network Analysis. Since this article is about applications of Data Analysis, or Data Science, methods, such historically and socially motivated distinctions are disregarded in the following. The phrase Machine Learning therefore denotes in the following the use of mathematical models of relationships between variables of interest for constructing –with or without training data– algorithmic predictors for these variables. Such a usage is consistent with a widespread practice in the the IT industry.

4 Dual-Purpose Human Computation Systems

Many Human Computation systems are "give and take" systems with two purposes, a first purpose motivating their users to give information to the systems, and a second purpose for which the systems take and process this information. We call such systems "dual-purpose Human Computation systems".

ARTigo: Making Games of Building an Artwork Search Engine. Games serving, in addition to entertainment, another purpose, so-called "Games With A Purpose" short "GWAPs", are a specific kind of dual-purpose Human Computation systems. ARTigo (http://artigo.org) is a Web platform offering for free both, several GWAPs and an artwork search engine. While playing, ARTigo's users leave annotations describing artworks. These annotations are automatically processed by an Unsupervised Machine Learning method, a Higher-Order Latent Semantic Analysis [14, 15] we designed for the purpose which builds, and continuously improves, an artwork search engine.

ARTigo has been conceived in cooperation with art historian Hubertus Kohle within the five-year interdisciplinary research project play4science (http://play4science.org) funded over three years by the German Foundation for Research ("Deutsche Forschungsgemeinschaft").

ARTigo's artwork database includes over 65,000 images from the artemis database (http://artemis.uni-muenchen.de/), the Rijksmuseum (Amsterdam, The Netherland), the Karlsruher Kunsthalle (Karlsruhe, Germany), the Museum of the University of Massachusetts Amherst (Amherst, Massachusetts, USA), and Albertina (Vienna, Austria). Since 2008, over 9 million annotations have been collected from, on average, 150 persons a day playing on ARTigo [20].

ARTigo is an ecosystem consisting of different kinds of games. The annotations collected on the ARTigo platform flow from game to game providing the "seed data" necessary for some games to be playable:

- simple descriptions are collected by "description games",
- more specific descriptions are collected by "diversification games" using simple descriptions collected by description games,
- annotation clusters are collected by "integration games" using annotations collected by games of all kinds.

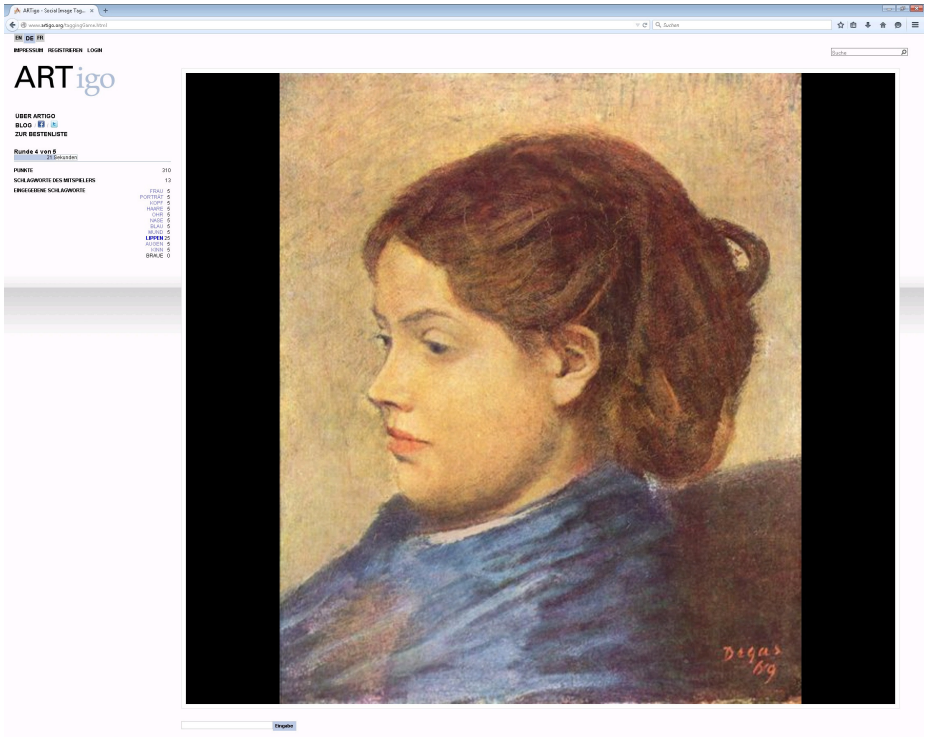

Figure 1: Screenshot of an ARTigo Game session featuring Edgar Degas' "Portrait de Mademoiselle Dobigny" (1869)

Description Games are games whose players describe an artwork by proposing annotations related to anything referring to the artwork: objects or characters it depicts, its colours, the materials it is made of, etc. ARTigo has a single description game, a variant of the ESP Game [16], called ARTigo Game, with the following gameplay: When two randomly selected players enter a same annotation for a same artwork, they both score and the annotation is validated. Validation is necessary to ensure annotation correctness –among others to exclude malicious annotations. This form of validation has been found sufficient to ensure quality annotations [16]. A description game like ARTigo Game collects simple or "surface semantics" annotations that are needed both as basic descriptions and as "seed data" for games collecting semantically richer, or "deep semantics", annotations.
Diversification games collect annotations that, in general, are more specific than most annotations collected by description games. Diversification games use the simple annotations collected by description games as "seed data". ARTigo has two diversification games: ARTigo Taboo and Karido.
ARTigo Taboo is, like ARTigo Game, a variation of the ESP Game [16]. In contrast to ARTigo Game, the seven most frequently entered annotations for an artwork can no longer be entered by any player of ARTigo Taboo for the same artwork: they are "taboo", hence the game's name. ARTigo Taboo thus forces players to enter novel annotations. As a consequence, ARTigo Taboo yields richer descriptions than ARTigo Game. Taboo-ing annotations can be seen as a form of "scripting" [17], that is, instructing players on the kind of annotations they are expected to enter. Indeed, a list of taboo-ed annotations is an instruction.
Karido [18, 19] is ARTigo's second diversification game we designed specifically for ARTigo. Its gameplay is as follows: Nine similar artworks are randomly selected and displayed to two randomly paired players in 3x3 grids such that the artworks are differently ordered in both players' grids. Artwork similarity is determined from "surface semantics" annotations so far collected for these artworks by description and diversification games. One player is a "describer", the other a "guesser". In the next round, they exchange roles. The describer selects one of the nine artworks on her grid and starts annotating it in such a way that the guesser can recognize and select it. The guesser can ask yes/no questions (like "WATER?" on Figure 2) that are answered by the describer. Since the guesser's grid and the player's grid are differently ordered, locational annotations like "South-East" do not help in recognizing the selected artwork. The sooner the guesser selects the right artwork, the higher the scores for both players. Karido's gameplay incites players to enter annotations distinguishing the selected artwork from others on the players' grids.
Integration Games cluster annotations yielding more precise descriptions than the unstructured sets of annotations collected by description and diversification games. Two integration games have been designed specifically for the ARTigo ecosystem: Combino and TagATag.

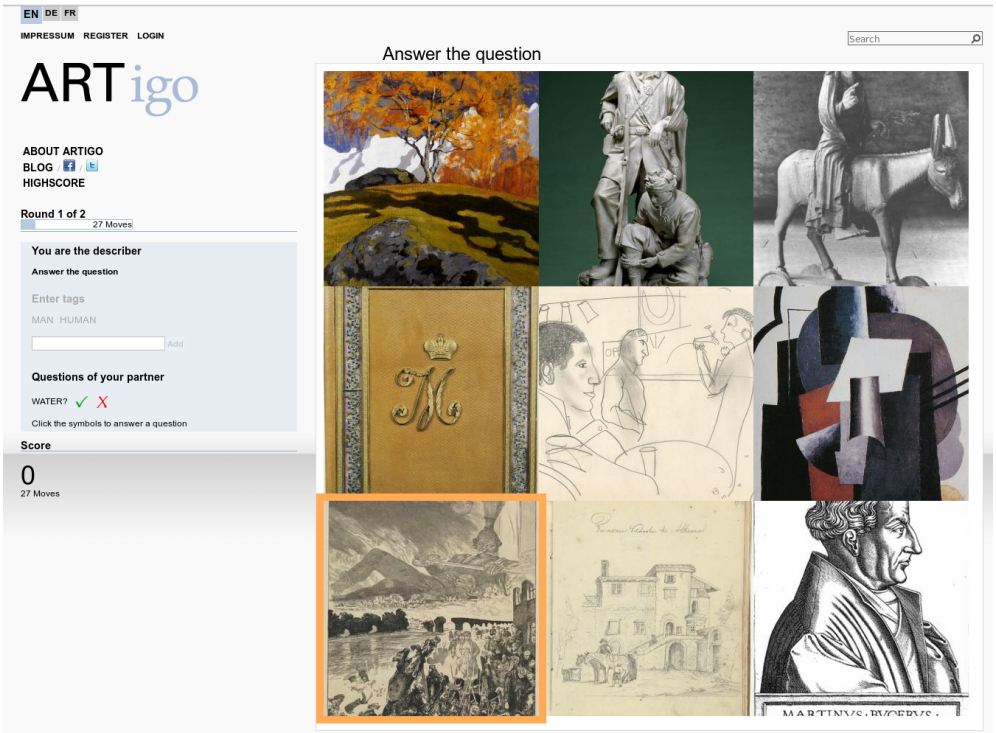

Figure 2: Screenshot of a Karido describer session highlighting the first artwork the player should describe.

Combino makes its players bring formerly collected annotations into relation. An artwork and a set of annotations formerly collected for this artwork are displayed to randomly paired players. Both players score when they

select the same pairs of annotations from the displayed set of annotations. Thus, Combino is, like ARTigo Game and ARTigo Taboo, a variation of the ESP Game [16].
By "squaring" Combino, that is running it with phrases it formerly collected, long phrases like "old man sitting" can be stepwise collected from "old", "man", and "sitting", that is, Combino can collect descriptive phrases.
TagATag is a "squared" [17] and "scripted" [17] ESP Game [16]. TagATag displays to its players an artwork AW, an annotation A formerly collected for this artwork AW and asks to describe annotation A (like for example "man") in artwork AW. Thus, TagATag collects annotations (like "reading" or "old") on an annotation A (like "man") in the context of an artwork AW.
Quantitative and qualitative analyses [20] have shown that the games of various types offered by the ARTigo gaming ecosystem do collect, as intended, different types of annotations and that the ARTigo gaming ecosystem as a whole performs better than each of its games alone.
Conditions for the Success of Games With a Purpose. Over the last eight years ARTigo has been online, it has taught us much on what makes GWAPs successful: Focus, quality data, quality software, and long lasting advertising. Focus means that a GWAP's success depends among others on the GWAP's content being clearly recognizable. One reason for ARTigo's long lasting success is that ARTigo's visitors know what to expect while playing ARTigo games: To see artworks. Quality data is as important as focus. Not only do ARTigo's visitors know what they will see while playing ARTigo games, they also know that ARTigo artwork database is large and contains quality images. Quality software is a further conditions for reaching out to wide player audiences. Indeed, players have no patience with systems not answering timely and are unlikely to come back to a Web platform with an uncertain availability. A reason for ARTigo's success is that it is sustained by a significant amount of Web engineering (including server farms and load balancing). Long lasting advertising is a further condition for the success of GWAPs. Even well-designed and well-working platforms do not gain large audiences simply from their sheer existence. They have to be repeatedly advertised.
BibPad: How Library Users Can Give a Library a Search Engine. Not only games, but all kinds of services can incentivize people to contribute to dual-purpose Human Computation systems.
We devised a Human Computation scheme for the users of an academic library to give the library a search engine. The scheme consists firstly in providing the library users with an online notepad, called BibPad, for their library research using which they can keep track of entities relevant to their research and of how these entities relate to each others. Thus, BibPad is a tool with which library users can build up small ontologies related to their research. BibPad has been designed but not implemented. BibPad's annotation component, Annoto, has been implemented, and tested [21, 22]. Annoto has rich, yet intuitive and easy to use, annotations. With Annoto, texts, images, videos and tunes can be annotated with texts and diagrams (cf. Figure 3).

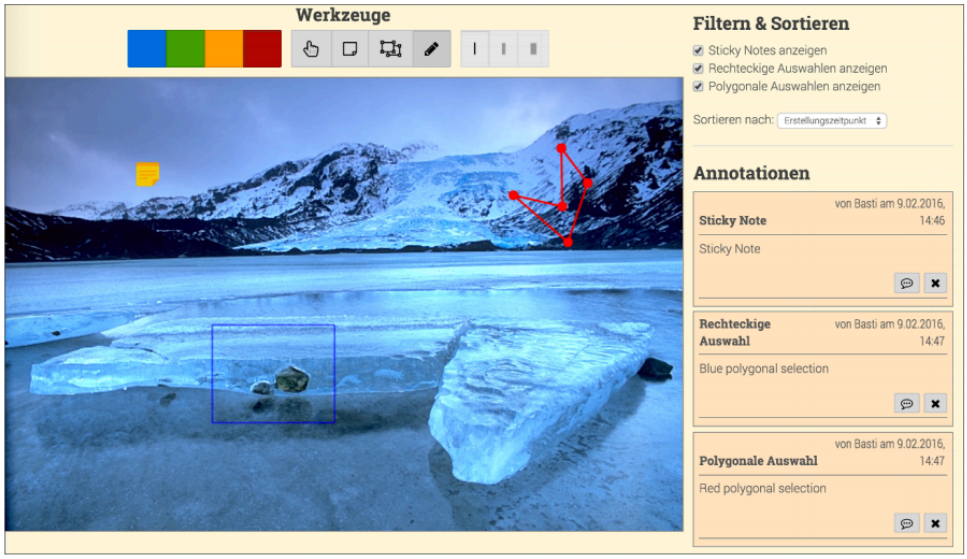

Figure 3 [22]: An Annoto screenshot featuring a "sticky note" (left), a "rectangular selection" (middle), and a "polygonal selection" (right).
(Photograph: Eyjafjallajökull by Andreas Tille published under licence CC BY-SA 3.0 –changed by the annotations– available at: https://en.wikipedia.org/wiki/File:Eyjafjallajökull.jpeg)

The BibPad-based Human Computation scheme consists secondly in exploiting the annotations entered by library users on their BibPad clients (of course after having been granted permission by those users and after anonymization) for automatically building a search engine using convenient Machine Learning techniques (like named entity recognition, clustering, eigenvector centralities, and latent semantic analysis).

5 Backstage: Improving Learning by Giving Students Voices

In higher education, large classes are a salient difficulty for students and lecturers alike. The need for written expression and for guidance while learning in Science, Technology, Engineering, and Mathematics (STEM) suggested using a social medium giving students the possibility to express themselves on lectures' contents by annotating lectures' slides. We hypothesized that a social medium can support large-class teaching by restoring behaviours, especially communication, common in small classes but impossible in large classes and beneficial to learning. Indeed, a salient property of social media is that they enable communication within crowds between people who would hardly communicate without social media. We further hypothesized that new lecture and tutorial formats would positively affect students' participation and hence learning.

We conceived and developed the classroom communication system Backstage [23–30] which is in service since 2012 (http://backstage.pms.ifi.lmu.de:8080). Backstage provides both a backchannel for communication initiated by students or lecturers and an audience response system for communication initiated by lecturers. Lecture-centered communication is incentivised by Backstage constraining to relate every backchannel message to a lecture presentation slide (see Figures 3 and 4). Backstage departs from most social media in a central aspect: Instead of drawing the attention of its users to new contents and instead of fostering new relationships and more communication, Backstage focuses communication on the contents of the lecture and fosters a social regulation of the backchannel communication.

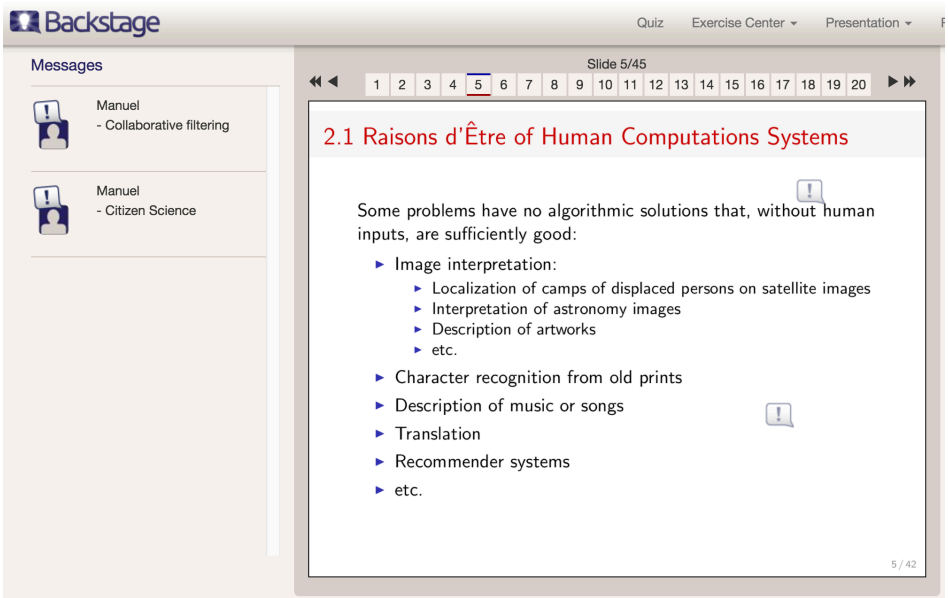

Figure 3: Screenshot of a Backstage-supported course featuring two backchannel posts by a student.

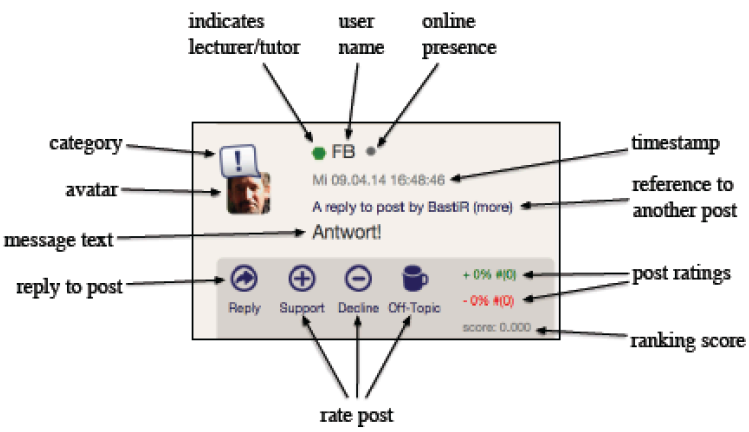

Figure 4 [29]: The components of a post on Backstage's backchannel

The lecture format was redesigned as follows [30]: Strictly serial style, summary of the lecture of the day both upfront and as a conclusion, key elements of a lecture arranged in an eye- and ear-catching manner, information presented in simple, logical, and sequential patterns, 20 to 25 minutes lecturing sessions followed by 5 to 10

minutes quizzes or polls sessions. The tutorial format was redesigned as follows [30]: Peer correction, peer feedback on peer correction, software-based weekly task assignments and work delivery, inverted classrooms.
Extensive evaluations of Backstage in different courses have shown the following [28, 30]:

- Backstage brings back to lectures students distracted by social media.
- Backstage fosters interactivity and awareness in large-class lectures when used in combination with a teaching format providing opportunities for and encouraging lecture-relevant communication.
- Backstage and a teaching format fulfilling the aforementioned conditions are well received by the students.
- When conveniently used by the lecturers, Backstage is useful for students revising lecture contents and the communication on Backstage is lecture-relevant.
- Students appreciate using Backstage during lectures and for revisions.

Computer-Mediated Communication (CMC) can become confusing because it usually varies in both relevance and quality. We investigated how to rely on participants' ratings as indicators of relevance and quality and for message summarisation [29]. We also developed an original eigenvector centrality index for students' reputation on Backstage aiming at incentivizing desired behaviours and regulating communication [29].

6 Human Computation and Markets

Markets, especially contemporary markets that rely on software for pricing, order management and transaction processing, can be seen as Human Computation systems. Indeed, as already observed in Section 2, markets interconnect humans and machines, process information as systems, and serve a purpose, that is, markets fulfill one of the aforementioned definitions of Human Computation systems [5]. This observation suggested to investigate how the market metaphor could be used for designing specific Human Computation systems. This section reports about such systems: Prediction markets, a decision market, and two linguistic field research markets. Finally, a Human Computation scheme is described that by rating credit risk aims at a self-regulation of financial markets preventing market bubbles.

Prediction Markets. Markets have been used since long for predictions. First and foremost, the prices on financial markets of derivative financial contracts –short "derivatives"– (like options and futures) imply predictions of the estimated likeliness of political or economical events (like the outcome of an election, a future currency rate, a future oil price) derivatives refer to.

Online markets called "prediction markets" have been built for tapping the "wisdom of crowds" [31] for making predictions [32], among others the Iowa Electronic Markets (http://tippie.biz.uiowa.edu/iem/) for political elections run since 1998 by the University of Iowa for research and teaching purposes.

Even though prediction markets have been conceived long before the phrase "Human Computation" was coined, prediction markets are a form of Human Computation systems.

Prediction markets have often, but not always, predicted political elections with better accuracy and at lower costs than polls [33, 34]. One reason for the good performance of prediction market is Keynes' "Beauty Contest Effect" [35]: Keynes described the actions of rational agents on a market in analogy to a fictional contest in which participants have to choose the six most attractive faces from a hundred photographs, those choosing the faces most popular among all participants being rewarded. A strategy for a participant to maximise her winning chances would be not to rely on her own perception of beauty but instead to rely on an estimate of the majority perception. According to Keynes, stock market participants are similarly more likely to act after their perceptions of majority views than after their own views. The "Beauty Contest Effect", for Keynes a reason to distrust stock market-based economies, helps in collecting majority opinions.

Liquid Decision Making: A Decision Market. Inspired from the capability of prediction markets to elicit and gather information from people, we hypothesized that markets could be used not only for predictions (like predictions of the outcome of an election), but also for decisions (like the strategy for a corporation to choose). The difference between the two is that predictions eventually can be checked for their accuracy while decisions shape the future what makes them veracious and their accuracy unverifiable.

Liquid Decision Making is a Human Computation system conceived and tested with Bauhaus Luftfahrt [36, 37, 38] for collective decision making based on the market metaphor. With Liquid Decision Making, decision makers bargain with decision options using play money. System-triggered perturbations ensure that the participants really prefer the choices expressed by market prices [38].

A decision market like Liquid Decision Making overcomes a drawback of voting that has been recently seen in the UK with the Brexit referendum and in the USA with the 2016 presidential election: After voting, the voters cannot react to unforeseen majorities. In contrast, a decision market like Liquid Decision Making allows for a collective decision making over a period of time from a few days to a few months before finally arriving at collective decisions. During this time, emerging majorities can be observed by all decision makers, discussed among them, and, if decision makers deem appropriate, can be reacted to.

Research on Liquid Decision Making includes an investigation of incentives to give to market participants considering two different contexts [36, 37, 38]. The one context consists in striving for a Keynesian Beauty Contest by a pricing amounting to ask the market participants: "What options do you think the majority of the decision makers will choose?" The other context, in contrast, consists in avoiding a Keynesian Beauty Contest by a pric-

ing amounting to ask the market participants: “What option do you favour regardless of what you expect to be the majority choice?” Both contexts are meaningful in practice. Rewards based on market option prices are appropriate in the first context, while rewards based on an estimate of a market participants’ influence on other market participants is the appropriate reward in the second context [36, 37, 38].

MetropolItalia: Markets for Linguistic Field Research. MetropolItalia (http://MetropolItalia.org) is a Web platform running since 2012 two markets, Mercato Linguistico and Poker Parole, both aimed at linguistic field research [39, 40, 41, 42]. MetropolItalia’s market have been conceived in cooperation with linguist Thomas Krefeld within the five-year interdisciplinary research project play4science (http://play4science.org) funded over three years by the German Foundation for Research (“Deutsche Forschungsgemeinschaft”).

The two markets Mercato Linguistico and Poker Parole collect phrases in Italian dialects and metropolises varieties together with data on these phrases such as where they are spoken and their speakers’ ages, genders, and levels of education. Screenshots of the platform are given in Figures 5 and 6.

Figure 5 [42]: Mercato Linguistico during the choice of a region for the displayed sentence.

Figure 6 [42]: Interface Mercato Linguistico and Poker Parole for reviewing one's own assessment's parameters.

On MetropolItalia, one can create so-called "assessments". An assessment consists of a phrase, the region or metropolis where the phrase is spoken, characteristics of the phrase's speakers (like age, gender, and levels of education), as well as an estimate of the proportion of MetropolItalia users are expected to agree with the assessment. An assessment could for example state that "interimme" (North-Calabrian for "in the mean time") is spoken and understood by everyone in Calabria and the expectation that 55% of Italian speakers with high levels of education are likely to agree with this statement. Another assessment could for example state that "frattanto" (Standard Italian for "in the mean time") is used everywhere and by everyone in Italy and the expectation that 95% of Italian speakers would agree with this. The purpose of assessments is not only to collect phrases and where and by whom they are used, but also perceptions among Italian speakers of where and by whom phrases are used. Indeed, such perceptions might significantly depart from reality, North-Italians with high levels of education for example wrongly thinking that certain South-Italian phrases are only spoken by South-Italians with low levels of education.

Assessments are priced as follows: The closer they are to the majority view on the market, the higher their prices. Thus, a Keynesian Beauty Contest [35] is likely to take place on Mercato Linguistico and Poker Parole. This is a highly desired feature of the two markets, for linguistic field research is first and foremost interested in majority opinions.

The markets Mercato Linguistico and Poker Parole differ from each other in the kinds of phrases they collect. While Mercato Linguistic incites through its assessment pricing to enter phrases with widely acknowledged linguistic traits (like where and by whom they are used), Poker Parole incites through its assessment pricing to submit phrases with linguistic traits that most market participants are likely not to properly assess. Phrases of both kinds are important in linguistics.

Figure 6 shows an assessment with referring to the phrase "Quella ragazza è una pittima" ("This girl keeps complaining about nothing") worth 88, a value close to the maximum 100. The user's estimation that 32% of the MetropolItalia's participants would assign this sentence to South Italy is a good guess since 37% of the market participants did. The assessment can be offered for sale for an adjustable price. It can be immediately sold for 83.

MetropolItalia's markets Mercato Linguistico and Poker Parole use play money. The human game drive is a first incentive to speculate on Mercato Linguistico and Poker Parole. A second incentive to speculate on the markets is to confront others, and be confronted, with opinions on a language ones speaks.

Cold Start Problem. MetropolItalia has not been as successful as expected. A first reason is that the platform never had a sufficiently large database of Italian phrases for being sufficiently attractive to Italian speakers. A second reason is that MetropolItalia has not been sufficiently advertised for. A third reason is that all members of the MetropolItalia project but two had no good control of the Italian language. Even though some project mem-

bers began to learn Italian for the project's sake, they never reached a sufficient control of Italian dialects to make it possible for them to contribute on MetropolItalia's markets.
MetropolItalia's limited community of contributors is worth reflecting about. A Human Computation system is almost always confronted with a so-called "cold start" problem: As long as a Human Computation system does not have enough human contributors, it will hardly gain more human contributors. The cold start problem in Human Computation is often overcome, as we did for the gaming platform ARTigo, by having the systems developers being among its first users. It was possible for us to spend five to ten minutes a day over two years playing on ARTigo but, due to our lack of control of Italian dialects, we could not do the same on MetropolItalia.
A Human Computation Scheme for Credit Risk Rating and Self-Regulated Financial Markets. Credit risk rating is essential on financial markets but traditional credit risk rating methods have become unreliable with the advent of derivatives, structured notes, and securitization techniques. The consequence has been a widespread improper credit risk rating that has caused the financial crisis of 2007-2009 which sparked the Great Recession and the European Sovereign Debt Crisis.
We devised a novel credit risk rating method [43] that radically departs from current credit risk rating as follows:

- Relying on a dual-purpose Human Computation approach, it collects credit risk assessments from debtors and not, as usual, from creditors, the incentive for debtors to reveal their estimates of how (un)likely they might fail to honour their debts being an an insurance-like "Grace Period Reward".
- It propagates debtors' risk estimates through the risk dependency graph induced by credit contracts, derivative contracts, and money by aggregating as eigenvector centralities the market agents' contributions to the market's systemic risk.
- It is not based upon stochastic methods and statistical data and therefore keeps its relevance in exceptional situations like rare crises and bubbles.
- It can warn of an increasing credit risk much earlier than current credit risk ranking methods.

Independent research [44] has given practical evidence that the approach proposed in [43] is meaningful.

7 Ethical Issues of Human Computation and Machine Learning

Human Computation and Machine Learning open tremendous opportunities but also serious threats. This section, based on François Bry's lectures on Human Computation and on unpublished reflections, discusses these threats and suggests how some of them could be addressed.
Threats of Human Computation. Widely known threats of Human Computation are privacy violations, the difficulty for workers to organize on Human Computation markets like Amazon Mechanical Turk what results in a power asymmetry between workers and requesters [45], and, for GWAPs, game addiction.
Other threats of Human Computation result from the cycles of reflexivity Human Computation systems like markets and social media are subject to:

1) Individual opinions are submitted to the systems.
2) Collective opinions are generated by the systems as an aggregation of individual opinions and delivered to the systems' users what influences their opinions and the process repeats.

Markets are examples of reflexive Human Computation systems: Offer and demand are individual opinions, market prices are collective opinions.
Cycles of reflexivity are virtuous when they contribute to the systems' purposes, vicious when they harm the systems' purposes. Wikipedia articles that are objective in the sense of Wikipedia's Neutral Point of View (NPOV), that is, that remain unchanged, result from virtuous cycles of reflexivity. The viral propagation social media strive for and sustain with recommendations is another example of virtuous cycles of reflexivity. Market bubbles (from the Tulip Mania Bubble in 17th century Holland to the American Stock Bubble in 1920 to the Nifty Fifty Stock Bubble in the 1970s to the Japanese Stock Bubble in the late 1980s to the Dot-com Bubble in 2000–2001 to the Subprime Mortgage Bubble in 2007–2008) are examples of vicious circle of reflexivity: Some prices keep raising more and over longer periods of time than usual, resulting in the "boom" when more and more traders are seduced by the perspective of unexpected gains, lose their sense of risk and buy in the expectation to later sell at higher prices, thus contributing to keep the prices raising; the "bust" eventually takes place when enough traders come to reason and stop buying what makes the prices suddenly and dramatically fall.
Cycles of reflexivity that are virtuous from the standpoint of the operator of a Human Computation system might nonetheless have vicious effects. This has been the case with the viral propagation of "fake news" before the US presidential election of November 2016. The "filter bubbles" that social media create [46] are also the result of cycles of reflexivity that can be both virtuous and vicious. A positive aspect of filter bubbles is that someone interested in art photography soon no longer gets offered pornographic photographs on tumblr as a consequence of her neither liking nor reposting such photographs and not following the blogs where such photographs are posted. A negative aspect of the filter bubble, which is deemed to play nowadays an important role in politics, is that it can insulate from the diversity of news and political opinions and, as a consequence, falsely convey the impression that there might be social agreements on certain positions.
Threats of Supervised Machine Learning With Human Computed Training Sets. Training data might be used in violation of intellectual property rights. This has been the case when members of the Google Brain pro-

ject trained a neural network to generate natural-sounding fluent sentences with recently published novels without the consent of the novels' authors [47, 48] and, at least so far, without Google considering giving them a share of the profit the technology will ensure. The issue here goes beyond a violation of intellectual property rights: The widespread use by businesses of publicly available training data in Machine Learning results in a "New Tragedy of the Commons": Common resources (like languages) are exploited by individual agents acting according to their own self-interest but against interests (like employment) of the society as a whole. Mass unemployment as a threat of Human Computation and Machine Learning is addressed in the following in more detail.

Training data may be –voluntarily or involuntarily– biased what might result in biased predictors. Involuntarily biased training data were the likely reason why Google Photos tagged coloured people as gorillas [49], returned only coloured people for the query "unprofessional hair styles" and only white people for the query "professional hair styles" [49] and answered the query "three black teenagers" with police mug shots while it answered the query "three white teenagers" with photographs of smiling people [51]. Indeed, training data that are selected by humans may reflect these humans' conscious or unconscious biases.

Without the "validity check" mentioned above in Section 3, Supervised Machine Learning may encode human biases. Importantly, the validity check might in some case be hardly possible or even impossible. In other cases, it would be possible but it is not performed so as to save time or reduce costs. A typical case where Supervised Machine Learning methods are used but no validity check is performed is the selection of candidates for jobs, loans, rental properties, and enrolment at universities [52]. Machine Learning methods are trained with résumés of candidates selected in the past who performed well in the job they were given, as creditors, as tenants, or as students. Even though this might at first seem impeccable and even better ensuring objectivity than selections performed by humans, the approach is flawed: Indeed, there is no way to perform a validity check, that is, to estimate how rejected candidates would have performed. A further case where Supervised Machine Learning methods are used but no validity check is performed is warfare predictors. Machine Learning expert Andrew Ng famously said at GTC 2015 that "fearing a rise of killer robots is like worrying about overpopulation on Mars". The catchy phrase was disregarding that robots with a license to kill are already deployed at the Korean Military Demarcation Line and that observations and estimates entered on the battle field by soldiers can be used to train Supervised Machine Learning predictors that in turn can be immediately distributed to all fighting units resulting in an immediate sharing of possibly biased observations and estimates. In this case, like in that of predictors used for candidate selection, validity checks are impossible: After a person is killed, it is no longer possible to establish for sure whether she would have been harmless, would she had lived. Supervised Machine Learning without validity check results in highly unethical "self-fulfilling predictors" that necessarily reproduce prejudices or biases – and incidentally might also harm the businesses or organisations using them.

Even when trained with data that perfectly reflect socially accepted common views, Machine Learning predictors can be threats if they strengthen social biases and social conformism. Admittedly, social biases and social conformism are nothing new. New is however the tremendous power to entrench social biases and social conformism of algorithmic methods that are widely perceived as objective and impartial.

Mass unemployment is the greatest threat of Supervised Machine Learning with human computed training data. This threat has already begun to materialize. We see two main reasons for this "rise of the robots", as Martin Ford calls it [53] using the word "robot" in the sense of "software" or "artificial intelligence", that is, not necessarily referring to hardware.

The first reason is that Machine Learning exploits the capability of computers to process several orders of magnitude more cases than the human brain can. While a simple algorithm can easily be devised for rapidly processing hundred of thousands or even millions of cases, a human brain can process at most a few ten to hundred cases. Humans having to face situations involving much more cases need a long training and a conceptual structuring of the many cases in hierarchies, or conceptual meta-structures, having at each level no more than the few ten to hundred cases the human brain can cope with.

Most administration jobs (from processing insurance claims to personnel management to tax consultancy to assessing student learning) and low qualification jobs (from warehouse workers to delivery workers to drivers to selespersons) require some expertise not because of intellectually challenging tasks but instead because of the many cases that have to be remembered, recognized, and coped with, that is, precisely what Supervised Machine Learning with human computed training data excels at. Indeed, Supervised Machine Learning predictors, from simple Naïve Bayes to more complex Deep Learning methods, are no more than classifiers the stunning effectiveness of which mostly results from the considerable numbers of cases they can rapidly process. We, humans, are proud of the impressive intellectual achievements of our kind. However, most of everyday work, including large parts of the intellectual work of highly qualified experts, does not reach the level of such achievements. Instead, a considerable part of it can easily be automatized as people know very well who, for example, have worked in scientific research before the advent of search engines and recommender systems.

A second reason why Supervised Machine Learning with human computed training data is threatening many jobs is the baffling performances in natural language processing (from classifying documents to understanding casual queries to translating) the technology has already achieved. We submit that it will shortly become clear that natural languages are soon to fall last bastions protecting human work from automatization.

The job erosion Supervised Machine Learning with human computed training data and robotics are already causing will further progress, increasing the payoff imbalance between capital and work, the capital share further growing, the work share further reducing. Recall that this imbalance already threatens developed countries [54]. In [53] Martin Ford stresses that mass unemployment would destroy the developed countries' economies that are based on mass consumption. Indeed, mass consumption is only possible if substantial parts of the population have sufficient incomes. Martin Ford and others therefore advocate firstly reconsidering, and significantly increase, capital taxation, secondly introducing a Guaranteed Basic Income sufficient for all to consume even if they are unemployed. Increasing capital taxation makes sense in case of capital invested in software since a same software can be further deployed at almost no marginal costs, what gives software-based investments an unprecedented economical power. A Guaranteed Basic Income appears more problematic as it would further increase the economical and social gap between developed and under-developed countries: A Guaranteed Basic Income in a developed society would create an irresistible incentive for economic migration from under-developed world regions to that society. A Guaranteed Basic Income could also have within the societies granting it vicious psychological and social effects.

Ethical Imperatives. We submit six imperatives to appropriately address the aforementioned threats.

Firstly, in the spirit of [55], ethics must receive among computer scientists, especially in research and teaching, the importance it deserves in an age in which computing technologies have an unprecedented power to impact on the fabric of society. A prerequisite for this is to strengthen in research and teaching a global view of technologies focused at how distinct technologies (like Human Computation and Machine Learning) can interact.

Secondly, computer scientists must develop and convey a consciousness of the inherent limitation of mathematical models as predictors of human behaviour and for the causes of this limitation:

- Mathematical models of human behaviour, how complex they are, are necessarily based on simplifying hypotheses.
- The aforementioned reflexivity: Human behaviour predictions necessarily modify human behaviour.

Thirdly, an outcome-based social control, or audit, of Machine Learning, and more generally Data Science methods, impacting on the fabric of society (search, recommender systems, social media, etc.) must be striven for. Such a control is the only chance to avoid, paraphrasing Lenin, giving "all the power to the robots". The control must be outcome-based and not technology-based since, although based on simplifying hypotheses, mathematical models are mostly too complex for humans to fully understand and appreciate their practical implications. Furthermore, how desirable it is that the control be informed by experts, paraphrasing Lenin once again, society should not give "all the power to the technocrats" because this would achieve the goal behind Lenin's watchword "all the power to the soviets": The end of democracy.

Fourthly, self-fulfilling predictors should be replaced by software checking the enforcement of publicized criteria. This would achieve the goals that lead to using self-fulfilling predictors in the first place without harming individuals or the society.

Fifthly the aforementioned New Tragedy of the Commons should be addressed by considering new taxation schemes, taxing the exploitation of commons and favouring some human work like social services carried out by humans. Indeed, an essential though often overseen role of taxation is to preserve the commons.

Finally machines should never be the ultima ratio in deciding on human destinies – neither in selecting candidates, nor in selecting war targets, nor in any other situation.

8 On-Going Research

A flavour of Human Computation is Citizen Science, that is, the enrolment of amateurs in scientific projects. We have just completed the set-up of the citizen science platform ARTizen (http://artizen.de) offering laypersons to collaborate with data scientists and art historians in using data science to analyse the European artworks of the long nineteenth century –a term coined by Historian Eric Hobsbawm referring to the period of time spanning from the French Revolution to the First World War.

We are currently improving the functionalities of the Backstage learning and teaching platform for classroom and off-classroom learning and we develop mixed Human Computation–Supervised Machine Learning schemes both for the prediction of mistakes learners can make and for collaboration among learners as well as among teachers preparing lectures.

We are also currently investigating algorithmic non-invasive methods for first detecting and measuring distress – that is, negative stress– and eustress –that is, positive stress– among learners and for sustaining participation of humans to various kinds of Human Computation systems with a focus at the aforementioned learning and Citizen Science platform. We work in particular on distress and eustress detection algorithms relying on data collected by fitness trackers and on algorithms enhancing motivation, ensuring a well regulated participation, tracking participants' reputation, and enabling co-optation, the delegation of tasks by participants to participants with sufficient seniority and reputation.

Finally, we further investigate from a technical perspective ethical issues in Human Computation and Machine Learning, especially how to prevent the mass-unemployment the technologies are threatening to provoke and how society could audit and therefore control the Human Computation and Machine Learning methods it uses.

9 Reshaped Human Collaboration: Hype or Hope?

The human collaboration reshaped by Human Computation and Machine Learning is undoubtedly a hype, and obviously a long-lasting and highly productive one. Whether this hype will lead to a better future is debatable. An essential condition for this hype not to end up in a threat will surely be the society's ability to exercise a sufficient control on the new, software-enabled and software-controlled, human collaboration. We argue that such a control should not rely on expert evaluations of the methods, models and algorithms used, because this would result in empowering a technocracy and, as a consequence, weaken democracy. Instead, a social control based upon audits of the methods, models and algorithms should in our opinion be striven for. A democratic auditing, possibly using Human Computation, of how collaboration software perform must become a social and political objective.

Acknowledgements. The authors are thankful for their contributions to the systems described in this article to: University staff members Dr. Norbert Eisinger, Dr. Vera Gehlen-Baum, Martin Josko, Dr. Fabian Kneißl, Prof. Dr. Hubertus Kohle, Prof. Dr. Thomas Krefeld, Dr. Elena Levushkina, Dr. Stephan Lücke, Dr. Christian Riepl, Dr. Gerhard Schön, Prof. Dr. Klaus U. Schulz, and Prof. Dr. Armin Weinberger; Bauhaus Luftfahrt staff members Dr. Gernot Stenz and Dr. Sven Ziemer; Roche Pharmaceuticals staff members Dr. Alex Kohn and Dr. Alexander Manta; and students Andreas Attenberger, Njomza Avdijaj, Daniel Baumgart, Matthias Becker, Michal Bednar, Alexandre Bérard, Mislav Boras, Fabian Bross, Blandine Bry, Évangéline Bry, Romy Buchschmid, Caterina Campanella, Laura Commare, Silvia Cramerotti, Stefan Fassrainer, Alexander Fischer-Brandies, Daniel Fritsch, Marlene Gottstein, Julia Hadersberger, Patrick Hagen, Diego Havenstein, Marcel Heil, Marco Hoffmann, Werner Hoffmann, Florian Hoidn, Katharina Jakob, Georg Klein, Max Kleucker, Johann Kratzer, Katharina Krug, Richard Lagrange, Philipp Langhans, Stephan Link, Florian Nass, Tien Duc Nguyen, Barry Norman, Julien Oster, Alessandra Puglisi, Anke Regner, Sebastian Rühl, Frederic Sautter, Corina Schemainda, Eva Schmidt, Oliver Schnuck, Jeannette Schwarz, Philipp Shah, Franz Siglmüller, Bartholomäus Steinmayr, Florian Störkle, Sebastian Straub, Daniel Unverricht, and Michael Weisbein.

Literature

1. Luis von Ahn: Human Computation, Ph.D. Dissertation, Carnegie Mellon University, 2007
2. Alexander J. Quinn and Benjamin B. Bederson: Human Computation: A Survey and Taxonomy of a Growing Field. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pages 1403–1412, 2011
3. Jeff Howe: The Rise of Crowdsourcing, Wired, June 2006
4. Edith Law and Luis von Ahn: Human Computation. In R. J. Brachman, W. W Cohen, W. W. and T. Dietterich editors, Synthesis Lectures on Artificial Intelligence and Machine Learning, Morgan & Claypool Publishers, pages 1–121, 2011
5. Pietro Michelucci: Introduction, Handbook of Human Computation, Pietro Michelucci editor, Springer Verlag, 2013
6. Mary Catherine Bateson: Foreword, Handbook of Human Computation, Pietro Michelucci editor, Springer Verlag , 2013
7. Alex Kohn, François Bry, and Alexander Manta: Exploiting a Company's Knowledge: The Adaptive Search Agent YASA, Proceedings of the International Conference on Semantic Systems (I-Semantics), pages 166-169, 2007
8. Alex Kohn, François Bry, and Alexander Manta: Exploiting a Company's Knowledge: The Adaptive Search Agent YASA, Proceedings of the International Conference on Semantics, pp. 166-169, 2008
9. Alex Kohn: Professional Search in Pharmaceutical Research, Doctoral Thesis, Institute for Informatics, Ludwig-Maximilian University of Munich, 2009
10. Corinna Cortes and Vladimir Vapnik: Support-Vector Networks, Machine Learning, Volume 20, Number 3, pp. 273-297, 1995
11. Thorsten Joachims: Text Categorization With Support Vector Machines: Learning With Many Relevant Features, Machine Learning: Proceedings of the 10th European Conference on Machine Learning (ECML), Claire Nedellec and Celine Rouveirol editors, Springer Verlag, pages 137-142, 1998
12. Fabrizio Sebastiani: Machine Learning in Automated Text Categorization, ACM Computing Surveys, Volume 34, Number 1, pages 1-47, 2002
13. David Nadeau, and Satoshi Sekine: A Survey of Named Entity Recognition and Classification, Linguisticae Investigationes, 30, (1) pp. 3-26, 2007
14. Christoph Wieser, François Bry, Alexandre Bérard, and Richard Lagrange: ARTigo: Building an Artwork Search Engine With Games and Higher-Order Latent Semantic Analysis, Proceedings of Disco 2013, Workshop on Human Computation and Machine Learning in Games at HComp, 2013
15. Christoph Wieser: Building a Semantic Search Engine with Games and Crowdsourcing, Doctoral Thesis, Institute for Informatics, Ludwig-Maximilian University of Munich, 2014
16. Luis Von Ahn and Laura Dabbish: Labeling Images With a Computer Game, Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 319-326, 2004

17. François Bry and Christoph Wieser: Squaring and scripting the ESP game, Proceedings of the 4th Human Computation Workshop (HCOMP), AAAI Press, 2012
18. Bartholomäus Steinmayr: Designing Image Labeling Games For More Informative Tags, Diploma Thesis, Institute for Informatics, Ludwig-Maximilian University of Munich, 2010
19. Bartholomäus Steinmayr, Christoph Wieser, Fabian Kneißl and François Bry: Karido: A GWAP for Telling Artworks Apart, Proceeding of the 16th International Conference on Computer Games (CGAMES), pp. 193-200, 2011 (Best Paper Award)
20. François Bry and Clemens Schefels: An Analysis of the ARTigo Gaming Ecosystem With a Purpose, Research Report, Institute for Informatics, Ludwig-Maximilian University of Munich, 2016
21. Sebastian Mader, Christoph Wieser, François Bry, and Clemens Schefels: BibPad as a Library Service or Crowdsourcing a Library Search Engine, Proceedings of the 6th International Conference on Qualitative and Quantitative Methods in Libraries, 2014
22. Sebastian Mader: An Annotation Framework for a Collaborative Learning Platform, Master Thesis, Institute for Informatics, Ludwig-Maximilian University of Munich, 2016
23. François Bry, Vera Gehlen-Baum, and Alexander Pohl: Promoting Awareness and Participation in Large Class Lectures: The Digital Backchannel Backstage, Proceedings of the International Conference e-society, 2011
24. Daniel Baumgart, Alexander Pohl, Vera Gehlen-Baum, and François Bry: Providing Guidance on Backstage, a Novel Digital Backchannel for Large Class Teaching, Education in a Technological World: Communicating Current and Emerging Research and Technological Efforts, 2011
25. Vera Gehlen-Baum, Alexander Pohl and François Bry: Assessing Backstage - A Backchannel for Collaborative Learning in Large Classes, Proceedings of the 14th International Conference on Interactive Collaborative Learning (ICL), 2011
26. Vera Gehlen-Baum, Alexander Pohl, Armin Weinberger, and François Bry: Backstage - Designing a Backchannel for Large Lectures (Demo Paper), Proceedings of the European Conference on Technology Enhanced Learning (EC-TEL), 2012 (Demo Shootout Special Recognition Award)
27. François Bry, Alexander Pohl: Backstage: A Social Medium for Large Classes, Campus Transformation -- Education, Qualification and Digitalization, Frank Keuper, Heinrich Arnold editors, Logos Verlag, Berlin, pp. 255-280, 2014
28. Vera Gehlen-Baum, Armin Weinberger, Alexander Pohl, François Bry: Technology use in lectures to enhance student's attention, Proceedings of the 9th International Conference on Technology Enhanced Learning (EC-TEL), 16-19 September 2014, 2014
29. Alexander Pohl: Fostering Awareness and Collaboration in Large-Class Lectures — Principles and Evaluation of the Backchannel Backstage, Doctoral Thesis, Institute for Informatics, Ludwig-Maximilian University of Munich, 2015
30. François Bry and Alexander Pohl: Large-Class Teaching with Backstage, Journal of Applied Research in Higher Education, Volume 9, Number 1, 2017
31. James Surowiecki: The Wisdom of Crowds, Doubleday, Anchor, 2004
32. Justin Wolfers and Eric Zitzewitz: Prediction Markets, Journal of Economic Perspectives, Volume 18, Number 2, 2004
33. Robert Forsythe, Forrest Nelson, George R. Neumann, and Jack Wright: Anatomy of an Experimental Political Stock Market, The American Economic Review, Volume 82, Number 5, pages 1142-1161, 1992
34. Joyce Berg, Robert Forsythe, Forrest Nelson, and Thomas Rietz: Results from a Dozen Years of Election Futures Markets Research, in Handbook of Experimental Economic Results, Elsevier, 2000
35. John Maynard Keynes: The General Theory of Employment, Interest, and Money. Macmillan Cambridge University Press, 1936
36. Stephan Leutenmayr, Sven Ziemer and François Bry: Decision Markets for Continuously Reflected Collective Decisions, Proceedings of the 3rd International Conference on Social Eco-Informatics, 2013 (Best Paper Award)
37. Stephan Leutenmayr: Liquid Decision Making: Applying the Market Metaphor to Collective Decision Making, Doctoral Thesis, Institute for Informatics, Ludwig-Maximilian University of Munich, 2015
38. Stephan Leutenmayr, Fabian Kneissl, Sven Ziemer and François Bry: Gameful Markets for Collaboration and Learning, Proceedings of Disco, Proceedings of the Workshop on Human Computation and Machine Learning in Games at HComp, 2013
39. François Bry, Fabian Kneißl, and Christoph Wieser: Field Research for Humanities with Social Media: Crowdsourcing and Algorithmic Data Analysis, Proceedings 4. Workshop Digitale Soziale Medien, 2011
40. Fabian Kneissl and François Bry: MetropolItalia: A Crowdsourcing Platform for Linguistic Field Research, Proceedings of the International Conference WWW/Internet, 2012
41. François Bry, Fabian Kneißl, Thomas Krefeld, Stephan Lücke and Christoph Wieser: Crowdsourcing for a Geographical and Social Mapping of Italian Dialects, Proceedings of the 2nd International Workshop on Social Media for Crowdsourcing and Human Computation (SOHUMAN), 2013
42. Fabian Kneißl: Crowdsourcing for Linguistic Field Research and E-Learning, Doctoral Thesis, Institute for Informatics, Ludwig-Maximilian University of Munich, 2014

43. François Bry: Human Computation-Enabled Network Analysis for a Systemic Credit Risk Rating, Handbook of Human Computation, Pietro Michelucci editor, Springer Verlag, pages 215-246, 2013
44. Sebastian Poledna and Stefan Thurner: Elimination of Systemic Risk in Financial Networks by Means of a Systemic Risk Transaction Tax, Research Report, arXiv, arXiv:1401.8026v3 [q-fin.RM], 2016
45. David Martin, Benjamin V. Hanrahan, Jacki O'Neill, and Neha Gupta: Being A Turker, Proceedings of the 17th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW), 2014
46. Eli Pariser: The Filter Bubble: What The Internet Is Hiding From You. Penguin Press Limited, New York 2011
47. Samuel R. Bowman and Luke Vilnis: Generating Sentences from a Continuous Space, Arxiv , Arxiv:1511.06349v4, 1016
48. Richard Lea: Google swallows 11,000 novels to improve AI's conversation, The Guardian, 28 September 2016
49. Jacky Alciné (@jackyalcine): "Google Photos, y'all fucked up. My friend's not a gorilla", Twitter, 29 June 2015 https://twitter.com/jackyalcine/status/615329515909156865
50. Bonnie Kamona (@BonKamona): "I saw a tweet saying 'unprofessional hair style for work'. I did. Then I checked for 'professional' ones", Twitter, 5 April 2016
https://twitter.com/bonkamona/status/717457819864272896?lang=de
51. Kabir Alli (@iBeKabir): YOOOOOO LOOK AT THIS (Video), Twitter, 7 June 2016
https://twitter.com/ibekabir/status/740005897930452992
52. Cathy O'Neil: Weapon of Math Destruction – How Big Data Increases Inequality and Threathens Democracy, Penguin Random House UK, 2016
53. Martin Ford: The Rise of the Robots: Technology and the Threat of Mass Unemployment, Basic Book, 2015
54. Thomas Piketty: Capital in the Twenty-First Century, Belknap Press, 2014
55. Hans Jonas: The Imperative of Responsibility, The University of Chicago Press, 1984