TRIVIIS
Anais de **Conferências** e **Mesas** de **Debates** do
Seminário Internacional de Pesquisas
em **Midiatização** e Processos Sociais

From qualitative classifications to algorithmic distinctions: datafication of the social[1]

Das classificações qualitativas às distinções algorítmicas: datificação do social

Göran Bolin[2]

Abstract: Over time, classification systems have accompanied the human species, and especially so with the rise of modern science. Already during Greek Antiquity, Aristotle made a classification of animals into groups, while Carl von Linné suggested a classification in the form of the botanical system. These classification systems, however, were qualitative, dividing species into groups depending on specified characteristics. With the computerised and algorithmically based classification systems of datafied society, all classification systems become numerically based. Against this background, we need to ask what implications this has for the ways in which humans classify and relate to the surrounding world, and especially social life, since this is at the bottom of the business models that drive the datafication process. How are qualitative evaluations translated into quantitative measures? What are the implications of the quantification of human action? How does quantification affect mindsets of media users in contemporary datafied society? Which are the mechanisms of evaluation involved in contemporary cultures of assessment?
Keywords: datafication; classification; algorithm; social theory.

Resumo: Ao longo do tempo, os sistemas de classificação têm acompanhado a espécie humana, especialmente com o advento da ciência moderna. Já na Antiguidade Grega, Aristóteles realizou uma classificação de animais em grupos, enquanto Carl von Linné

[1] Conferência apresentada no VI Seminário Internacional de Pesquisas em Midiatização e Processos Sociais. POSCOM-UFSM e ECA-USP na "MESA 2 — Classificações: da filosofia-ciência ao ódio classificatório".
[2] Professor de Estudos de Mídia e Comunicação. Sua pesquisa concentra-se em datificação, mercantilização e produção e consumo cultural em mercados digitais.

propôs um sistema de classificação na forma de um sistema botânico. Esses sistemas de classificação, no entanto, eram qualitativos, dividindo espécies em grupos com base em características específicas. Com os sistemas de classificação computadorizados e baseados em algoritmos da sociedade datificada, todos os sistemas de classificação tornam-se baseados em números. Nesse contexto, é necessário questionar quais são as implicações para as formas como os humanos classificam e se relacionam com o mundo ao redor, especialmente a vida social, já que isso fundamenta os modelos de negócios que impulsionam o processo de datificação. Como as avaliações qualitativas são traduzidas em medidas quantitativas? Quais são as implicações da quantificação das ações humanas? Como a quantificação afeta os modos de pensar dos usuários de mídia na sociedade contemporânea datificada? Quais são os mecanismos de avaliação envolvidos nas culturas contemporâneas de mensuração?
Palavras-chave: datificação; classificação; algoritmo; teoria social.

1. Introduction: Histories of classifications

Contemporary datafied society is arguably founded on classifications – the sorting of things, people and practices into categories, possible to enumerate and hence to compute, and ultimately using for the purpose of generating economic value by the production of the audience or media user commodity. Classifications, however, predates datafied society and can arguably be traced back to early forms of social life. Already in Greek antiquity, the philosopher Aristotle (384-322 BC) made systematic classifications of animals into groups based on specific features (ARISTOTLE xxxx). However, it was the Swedish botanist Carl von Linné (1707-1778), or Carolus Linnaeus that he called himself in Latin, that first included humans into the classificatory system in his *Systema Naturae* (LINNÉ 1735), the first edition which Linné published at the age of only 28. In this account, humans were sorted under the main category Anthropomorpha, together with other apes in the animal kingdom. An important quality of Linné's classification system was that it was *qualitative*, making divisions into "kingdoms" (e.g the animal kingdom), "classes", "orders", "families", and "species" depending on specified characteristics.

Linné was of course not alone in his classification task, not even at the time he published his magnum opus. There were indeed others in other parts of the world that engaged in classification of animals and plants, and they also partly build on one another. But Linné's system has become the model for contemporary biological classifications.

Another form of classification arose with structural linguistics in the mid-twentieth century. Within his framework of structural anthropology, Claude Lévi-Strauss in his *Anthropologie structurale* (LÉVI-STRAUSS 1958) and, most importantly, *La pensée sauvage* (LÉVI-STRAUSS 1962), argued that man was the classifying animal. According to Lévi-Strauss, archaic mental structures lie at the bottom of these classifications and to him it is the way in which the human way of thinking is structured that produces classification.

The structuralist approach that Lévi-Struss is representative of, it is often pointed out, is grounded in Saussurean semiology where signs are defined by difference – by what they are not. The letter "A" gets its meaning from what it is not (a "B"), or a "cat" is something that is not a "dog". This is the foundational feature of our ways of perceiving the world around us, and we make distinctions and categorisations out of a continuous social reality by differentiations. As semiotician Roland Barthes once argued: reality is a continuum (i.e. analogue), while analysis is always discreet (BARTHES 1968/1977). The colours of the rainbow follow a continuous colour shifting, while we divide colours into discreet categories by naming them blue, yellow, red, etc.

If Carl von Linné was making distinctions between animals (and other biological entities) by sorting them according to specific traits, semiotics bases distinctions in the descriptions themselves – in language.

From a more constructivist approach, however, we could also emphasise the arbitrariness of signs (that Saussure emphasised) and see them as outcomes of human negotiation in (=inside) communication. What José Luiz Braga referred to in his talk yesterday – that communication is the root to language, that it is the actual development of language; In the drive to understand other human beings we agree on the meaning of signs. But I will not speak about communication today, but about classifications and differentiations.

So, semiotics builds on differentiation – it classifies objects and things according to a logic of difference. This is also how mathematical calculations are structured, and, hence, also computational principles build on differentiation. It is therefore not strange that one of the precursors of computing, British mathematician Charles Babbage (1791-1871) constructed what he called the Difference Engine already in the early 19th century. The Difference Engine was not a computer because it could only add and subtract, and in that sense, it was more of an early calculating machine, replacing the abacus. It's mechanical calculating capacity, however, made it a tool for more accurate and faster mathematical operations. Eventually, Babbage sought to advance his invention and sketched what he called the Analytical Engine in 1937, a prototype for a mechanical general-purpose computer. Babbage's analytical engine, however, never became realized. His ideas were instead developed by Ada Lovelace (1815-1852), an English mathematician and the only legitimate child of famous British poet Lord Byron and his wife Annabelle Milbank. Ada Lovelace believed that the Analytical Engine could be developed to carry out operations beyond mere calculations, and in a translation of an article on Babbage's analytical engine by the Italian engineer Luigi Federico Menabrea she translated the French original into English, adding an extensive note apparatus where she commented on how the machine could be developed to perform more complex feedback loops, in effect sketching a string of instructions of the kind used in computers (MENABREA 1843). Ada Lovelace unfortunately did not live to see her ideas realized as she died at young age in uterine cancer in 1852, but, although contested by some, she is considered one of the pioneers of computer programming. And the differentiations and categorizations that can be computable are also the basis for later developments, once they become coupled with the business models of late modern capitalism (BOLIN 2011, 2023).

However, while biological or semiotic principles of difference need not be quantitative – they are based in qualitative differences – mathematical calculations build on hierarchical structuring. Numerical value is hence different from semiotic value. In mathematical terms, semiotic value is nominal, while mathematical value is (at least) ordinal.

2. New business models and new forms of classification

As the media became increasingly digitalised since the 1990, it opened for the development of new business models where the telecom industry teamed up with the content-producing media industries to construct the digital consumer who could be individually targeted by advertising and other persuasive messages (BOLIN 2011: 45ff). These business models are built on the monitoring, classification of consumer patterns and the subsequent targeting of ads to media users, ultimately to produce ever more refined consumer segments to target with commercial messages. With the computerised and algorithmically based classification systems of datafied society, most classification systems become subsumed commercial logics.

Against this background, we need to ask what implications this has for the ways in which humans classify and relate to the surrounding world, and especially social life, since this is at the bottom of the business models that drive the datafication process. How are qualitative evaluations translated into quantitative measures? What are the implications of the quantification of human action? How does quantification affect mindsets of media users in contemporary datafied society? Which are the mechanisms of evaluation involved in contemporary cultures of assessment?

Datafied society can be described as an increasingly integrated social sphere where all sectors of society utilize sorting and classification tools for a variety of means. If we look at the agents involved in classifications in contemporary media landscapes, we can conclude that most of them are commercially driven. Since value then become dependent on the datafied classification principles, all kinds of value get dependent on, firstly, what can be measured numerically, and, secondly, what can be monetized. This means that what is valuable – what value actually is – becomes defined by numerical measurement: “What counts – in the sense of what is valued – is that which is counted”, as Alain BADIOU (2008) says. This means that all qualitative value becomes turned into quantity because this is how computers work. It also means that that which cannot be counted, is of no value.

But capitalism also works from a principle of numerical value. The basic principle of capitalism is value accumulation, that is, the production of *more* value. And “more” is

a quantitative category (a qualitative measurement would be "better", for example). Capitalism works from quantity, not quality.

In datafied society, hence, two sorting mechanisms which are both founded on quantitative differentiation and classification coincide: the mathematical principles of the computational sorting mechanism, and the capitalist sorting mechanisms where individual targeting of consumers build on each consumer profile's difference from other consumer profiles.

3. Conclusion

If classification for Aristotle, Linné, and Levi-Strauss were founded on qualitative differentiations, contemporary classifications in datafied society rather build on quantitative assessment. Quality becomes translated into quantity. It can be concluded that classification and differentiation is at the root of human thinking – the way in which humans make order of the surrounding reality. Classification and differentiation is hence not a question of if classification occurs, or not, but rather on how, for what reasons, and by which standards and valuative principles things are classified? This is also the task for contemporary social research.

References

BADIOU, Alain. *Number and Numbers*. Cambridge: Polity, 2008.

BARTHES, Roland. The Death of the Author. In: Roland Barthes: *Image-Music-Text* (pp. 142-148). London: Fontana, 1968/1977.

BOLIN, Göran. *Value and the media: Cultural production and consumption in digital markets*. Farnham: Ashgate, 2011.

BOLIN, Göran. Mediatisation, Digitisation and Datafication: The Role of the Social in Contemporary Data Capitalism. *Central European Journal of Communication* 16(1): 7-18. https://journals.ptks.pl/cejc/article/view/522/pdf

LINNÆI, Caroli. *Systema naturæ, sive regna tria naturæ systematice proposita per classes, ordines, genera, & species*. Lugduni Batavorum, apud Theodorum Haak, MDCCXXXV. Ex typographia Joannis Wilhelmi de Groot. 1735.

LÉVI-STRAUSS, Claude. *Anthropologie structurale.* Paris: Plon, 1958.

LÉVI-STRAUSS, Claude. *La pensée sauvage.* Paris: Plon, 1962.

MENABREA, L.F. 1843. Sketch of the Analytical Engine invented by Charles Babbage, Esq. (including notes by the translator Ada Augusta Lovelace). In: Richard TAYLOR (Ed.) *Scientific memoirs selected from the transaction of foreign academies of science and learned societies, and from foreign journals*. Vol. III. (pp. 666-731). London: Richard and John E. Taylor. https://books.google.com/books/about/Scientific_Memoirs_Selected_from_the_Tra.html?id=qsY-AAAAYAAJ