IRIE
International Review of Information Ethics
ISSN 1614-1687

Vol. 5 (09/2006)

Ethics of Information Technology in Medicine and Health Care

Vol. 5 (09/2006)

Content

Ethics of Information Technology in Medicine and Health Care

Miscellaneous

Editorial: On IRIE Vol. 5

Have you seen a doctor lately? We hope not (at least not in an unpleasant matter). Anyway, if so you could have seen for yourself that modern medicine has become almost inconceivable without the use of information technology: from getting an appointment in the first place up to the making of a diagnosis, the treatment and not to forget all the accounting and necessary archiving. In fact it radically changed the delivery of health care in various aspects. And even more fundamentally – at least that is what some authors of this issue argue – information technology has transformed the medical construction of the human body and the scientific understanding of disease itself.

No one can deny the great improvements that have been made possible by these developments. Being aware of this massive further research and development is undertaken in the field. Not less important it is then to be aware of the ethical issues raised: the benefits and risks for the patient, the practitioner and the society thus developing guide lines for an appropriate use of information technology in medicine and health care.

Any normative analysis in this field has to be based on a thorough factual understanding of the technological developments, their medical applications and qualified philosophical interpretations. Therefore, we are very thankful that this issue is once again co-edited by two very experienced experts in the field: Georg Marckmann, Dept. of Medical Ethics at the University of Tübingen, Germany, and Kenneth W. Goodman, Director of the Bioethics Program and Co-Director, Ethics Program, at Miami University, USA. Due to their excellent piece of work, we can publish six very profound contributions to the subject in this issue ranging from an analysis of the role of software in the patient care process up to questions on property and availability of genetic information. For an overview of the different contributions see the very compound and well informed introduction Georg Marckmann and Kenneth Goodman wrote at the beginning of this volume.

We therefore are convinced that this issue of IRIE provides a most welcome forum to analyse and discuss the ethical and social issues raised by the various applications of information and communication technology in medicine and health care. May it lead to deeper insights, open controversies and new perspectives for scholars and practitioners all over the world.

Having no book reviews included in this issue we take the advantage to encourage you to contact us actively if you intend to contribute one or would like to have a specific book reviewed. IRIE is open to reviews of all kind of books in the field of Information Ethics - not only those dealing with the current subject of an issue.

Once again an article not belonging to the major subject of this issue complements this volume. It is James D. Caufiled's 'Myth of automated meaning' unmasking uncredited human processes in the generation of search engine results. We didn't want to withhold this interesting point of view supplementing our issue No. 3 explicitly dealing with the subject of search engines. That issue gained not only a great deal of attention but also the matters it dealt with keep on to be debated vividly in the scientific community. That is finally what we intended and what we hope for this compelling issue on the Ethics of Information Technology in Medicine and Health Care as well.

Yours,

Rafael Capurro, Thomas Hausmanninger, Karsten Weber and Felix Weil

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

Georg Marckmann and Kenneth W. Goodman

Introduction: Ethics of Information Technology in Health Care

Abstract:

Computer-based information and communication technologies continue to transform the delivery of health care and the conception and scientific understanding of the human body and the diseases that afflict it. While information technology has the potential to improve the quality and efficiency of patient care, it also raises important ethical and social issues. This IRIE theme issue seeks to provide a forum to identify, analyse and discuss the ethical and social issues raised by various applications of information and communication technology in medicine and health care. The contributions give a flavour of the extraordinarily broad landscape shaped by the intersection of medicine, computing and ethics. In fact, their diversity suggests that much more work is needed to clarify issues and approaches, and to provide practical tools for clinicians.

Authors:

Georg Marckmann, M.D., M.P.H.

University of Tübingen, Department of Medical Ethics, Schleichstr. 8, D-72076 Tübingen, Germany

☎ + 49 - 7071 - 29 78032, ✉ georg.marckmann@uni-tuebingen.de, 💻 www.uni-tuebingen.de/medizinethik

Relevant publications:

- Marckmann G. Recommendations for the Ethical Development and Use of Medical Decision Support Systems. Medscape General Medicine June 20, 2001. Available at: http://www.medscape.com/viewarticle/408143
- Marckmann G, Diagnose per Computer? Eine ethische Bewertung medizinischer Expertensysteme. Köln: Deutscher Ärzteverlag 2003

Kenneth W. Goodman, Ph.D.

Director, Bioethics Program, Co-Director, Ethics Programs, University of Miami, P.O. Box 016960 (M-825), Miami, FL 33101, USA

☎ +1 305 243 5723, ✉ kgoodman@miami.edu, 💻 http://www.miami.edu/ethics/

Relevant publications:

- Goodman KW (Hg.) Ethics, computing, and medicine. Informatics and the transformation of health care. Cambridge: Cambridge University Press 1998.
- Anderson JG, Goodman KW. Ethics and Information Technology. A Case-Based Approach to a Health Care System in Transition. New York: Springer 2002.

Computer-based information and communication technologies continue to transform the delivery of health care and the conception and scientific understanding of the human body and the diseases that afflict it. In fact, modern medicine has become almost inconceivable without the use of computers. While information technology has the potential to improve the quality and efficiency of patient care, it also raises important ethical and social issues. Work on ethical issues at the intersection of medicine and computing has so far generated an ensemble of interesting and important questions:

> What are appropriate uses of health information systems?
>
> Who should use these systems?
>
> What benefits and risks do these technologies have for patients?
>
> How does information technology change the physician-patient relationship?
>
> How does (and will) medical decision making change?

Perhaps most fundamentally: How does (and will) information technology transform the medical construction of the human body and disease?

Medical informatics has evolved as rapidly as any science in history, paralleling and relying on extraordinary advances in information collection, storage, analysis and transmission. Indeed, like sciences and technologies that precede it – biochemistry, microbiology, genetics and cell biology, pharmacology and so on – medical informatics is changing the standard of care. It is no longer fanciful to consider whether it might be blameworthy for a physician or an allied health professional to *fail* to use intelligent machines or their accoutrements in clinical practice and research.

Indeed and as ever, one should be forgiven for thinking that the science is progressing faster than the ethics. We are too soon able to do things – before we know whether we ought to or not.

This is not a problem. It is rather a stage-setting in which we have the opportunity to use the tools of applied and theoretical ethics to track scientific change and, eventually, to guide it. If ethics lags too far behind, however, a science without moral underpinning risks losing touch with broadly shared human values. So we worry – we must worry – when use of a new machine poses risks to patients; and we worry when failure to use the machine also may mean that a patient's care might be suboptimal, or that a patient might come to grief.

Ethical analysis of these concerns and tensions thus emerges as a moral imperative itself. Put differently, failure to scrutinize the expanding application of an evolving technology is itself blameworthy.

It follows that ongoing developments in the fields of information and (tele-)communication technology require continuous monitoring of ethical and social implications. This IRIE theme issue, "Ethics and Information Technology in Health Care," seeks to provide a forum to identify, analyse and discuss the ethical and social issues raised by various applications of information and communication technology in medicine and health care.

We have sought to construe the forum as broadly as possible, better to give a flavour of the extraordinarily broad landscape shaped by the intersection of medicine, computing and ethics. The contributions to this theme issue succeed nicely in making plain such breadth. In fact, their diversity suggests that much more work is needed to clarify issues and approaches, and to provide practical tools for clinicians. A patient-centered approach to medical care requires nothing less.

In the first contribution of this issue, *Dirk Hagemeister* investigates the effects of modern information technology (IT) on the working environment in hospitals and outpatient care and on the physician's decision-making process. On the one hand, IT certainly provides benefits for patient care by providing new computer-based diagnostic devices (e.g. the CAT scan), by improving the efficient storage, retrieval and exchange of patient data and as part of a basic infrastructure for integrated care networks. On the other hand, IT in medicine may also have negative consequences for patient care. In many cases, these effects are not a problem of IT per se but rather result from the complex interaction of IT with other structural determinants of a health care system, e.g., physician reimbursement systems. This makes clear that IT is primarily an instrument that can have both intended and unintended consequences, which in turn are heavily influenced by the specific policy framework. An electronic billing system, for example, might lead a physician to up- or downgrade the diagnostic code in order to increase his personal income. Some less visible, but nevertheless ethically important implications might occur if the electronically implemented diagnostic framework restricts the information input and the process of diagnostic reasoning. Users of IT in medicine therefore should be familiar with the specific strengths and limitations of the systems they are using.

Apart from storing and processing patient data, Information and computer technology in medicine can also be used to create virtual realities. *Janne Lahtiranta* and *Kai Kimppa* discuss in their contribution the use of anthropomorphized, human-like artefacts for teaching and training in medicine. Especially in intensive care and emergency medicine, students and residents learn clinical skills with manikins that show a realistic full-body anatomy.[1] According to the Institute of Medicine (IOM) report "To err is human: Building a safer health system", between 44,000 and 98,000 people die in hospital each year as a result of medical errors that could have been prevented.[2] As one measure to improve patient safety, the IOM suggests the use of simulation whenever possible:

> *"Health care organizations and teaching institutions should participate in the development and use of simulation for training novice practitioners, problem solving, and crisis management, especially when new and potentially hazardous procedures and equipment are introduced."*[3]

As Lahtiranta and Kimppa point out, training with anthropomorphized artefacts may prevent harm and provides benefits for the patients. Consequently, we have an ethical obligation to promote training of clinical skills with manikins and dummies. However, these simulated realities will always fall short of the actualities of patient care. Therefore, medical students should be aware of these limitations and enjoy sufficient opportunity to interact with the complexity of real-world patients. Computer-based simulation can be a valuable supplement, but should never be a substitute of conventional bedside teaching.

Britta Schinzel is also concerned with the moral problems created by the use of constructive technologies. Modern imaging technologies like magnetic resonance imaging (MRI), positron emission tomography (PET) or functional MRI provide new and – supposedly – realistic insights into the human body. However, these images are highly constructive artefacts that result from an extremely complicated and in many respects contingent combination of model driven algorithms, computations and visualizations. These epistemic observations become ethically relevant, if the images are used without reflection of their production process. Instead of depicting the world as it is (which in fact might be impossible), the images are heavily loaded with interpretations and create new meanings for "health," "disease," "normality," and "gender." Developing standardized atlases of the brain also elicits ethical issues given the plasticity and inter- and intra-individual variability of the cerebral structure and functioning. Due to their realistic appearance, these computed images create the illusory impression that certain differences between groups and populations are biologically fixed within the human brain. This can promote stereotypes and false dichotomies that are embedded in the seemingly "objective" results of scientific imaging techniques.

Information and communication technologies (ICT) in medicine are not only used to store, process, model, and transmit patient data. Innovative approaches try to integrate computer technology with the human body in order to obtain insights into the functioning of the body and to develop new diagnostic and therapeutic interventions. With these ICT implants, the use of computers in medicine seems to enter a new dimension, raising not only intriguing ethical issues but also anthropological issues. With the following two articles, we stay within the field of neuroscience research, but move from imaging to intervention.

Jens Clausen focuses on the ethical aspects of neuronal motor prostheses. Advances in the neurosciences and in micro system technology provide the potential to connect computer-systems with the human brain via brain-computer interfaces. This might offer new therapeutic perspectives especially for paralysed patients (e.g., hemiplegia). The goal is to bridge the interrupted nerve fibres with microtechnical devices and connect the cortex to an artificial limb or – even better – with the paralysed limb of the patient. On the one hand, brain-computer interfaces raise general ethical issues related to the protection of human subjects and the limits of a man-machine-integration. On the other, neuronal motor prostheses raise ethical issues that can be attributed to technological components themselves. Will the implanted electrodes of the input component that registers the cortical field potentials alter the patient's personality traits in an unacceptable way? Indeed, what should count as an "acceptable" alteration? Who bears responsibility for

---

[1] Spence 1997

[2] Kohn et al. 2000

[3] Kohn et al. 2000, 179

actions of the artificial limb that result from an indissoluble interaction between the patient's brain and the decoding algorithm? Will the (probably) wireless output component allow unwanted external control or interference? Risks for patients due to malfunctioning of the system must be balanced against the benefits of restored limb function. It certainly will require many year of development and clinical testing before neuronal motor prostheses will become available for patients. Nevertheless, ethical analysis should accompany these technological developments right from the start as part of the timely anticipation of possible adverse effects on the patients.

Elisabeth Hildt discusses another application of ICT-implants in the neurosciences: With electrodes implanted in the brain, deep brain stimulation provides the possibility to influence the functioning of certain brain regions, e.g. for tremor control in patients with Parkinson's disease. Other potential applications of neurostimulation include the treatment of severe neuropsychiatric disorders like obsessive-compulsive disorders. Like many other invasive medical technologies, deep brain stimulations involve benefits (e.g. improvement of tremor) and risks (e.g. infections, intracranial haemorrhage, adverse neuropsychological effects). However, the effects of deep brain stimulation deserve special attention because they interact with brain structures that might influence the patient's personality and character traits: How far will the patient be able to control the effects of neurostimulation? Will these changes be reversible? This calls for a thorough assessment of clinical benefits and adverse effects, especially in the long term. With regard to other possible applications – e.g. enhancement of cognitive functions – Hildt concludes that deep brain stimulation should be restricted to severe disorders with a well-known pathophysiological basis for which there are no other less invasive treatments with comparable effectiveness.

Michael Nagenborg and *Mahha El-Faddagh* conclude this issue with a paper on the availability, appropriate use and ownership of genetic information. In this case, ethical issues arise from the application of computer technology for processing a specific type of medical information. While genetic information plays an important role in the diagnosis of hereditary diseases, it also provides opportunities to select effective therapeutic interventions according to the specific genotype of the patient. However, other players like employers or insurance companies are interested in genetic information. This gives rise to the question: Who has legitimate access to the genetic information? Who owns or controls the genetic information? Nagenborg and El-Faddagh apply the concept of different information spheres, and argue that the usage of genetic information should be confined to the medical sphere. Based on a Kantian concept of ownership the individual is the "owner" of information about his genes in the sense that he has the right to exclude other people form using this very personal information. Hence, information infrastructures in health care should be organized in a way that preserves the boundaries of the medical sphere and precludes unauthorized usage of genetic information by non-medical institutions. As the authors point out, this will be a difficult task given increasing economic pressures on the world's health care systems.

The contributions in this IRIE issue present a colourful picture of the rapidly expanding field of information and communication technology in medicine and health care. New ethical issues arise depending on the specific context in which computer technology is applied. The articles here provide several examples to illustrate the point that it is not ICT per se that raises ethical issues, but rather the connection of ICT to certain concepts, types of medical information and other technologies. Especially in the field of neuroscience research, the increasing integration of human and machine with ICT-implants will pose intriguing epistemological, ethical and anthropological questions. Depending on the specific context, ethical analysis of ICT applications in medicine therefore will increasingly have to combine insights and approaches from several different disciplines.

References

Kohn LT, Corrigan JM, Donaldson MS (eds.). To err is human. Building a safer health system. Washington: National Academy Press; 2000.

Spence AA. The expanding role of simulators in risk management. Br J Anaesth 1997;78(6):633-4.

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

Dirk Thomas Hagemeister

"Software must not manipulate the physicians:" The IT Challenge to Patient Care

Abstract:

Information technology plays an increasingly important role in the medical working environment. Besides facilitating improvements in the quality of health care, it might also bear some unwished effects. Examining the 'making' of a diagnosis and the role it plays in modern medicine leads to the question how far this process of 'diagnosing' might be affected by the 'technical surroundings'. A number of examples from clinical medicine in the hospital and the ambulatory sector illustrate the way IT is being utilised in modern medicine. A twofold negative effect could result from this 'computerisation': Firstly, the technical requirements for the use of IT might force the process of diagnosing to be adapted with subsequent wrong or altered diagnoses. Secondly, constraints like cost control might be facilitated by IT and thus its application might cause the doctors trying to avoid such pressures by modifying the diagnosis and potentially worsening treatment and outcome.

Agenda

Author:

Dr. med. Dirk Thomas Hagemeister, B.A. M.P.H.:
Rustenburg Provincial Hospital, POBox 6299, Rustenburg 0300, South Africa
☎ + 27 - 14 – 533 0 916 , ✉ Dirk@Hagemeister.net, 💻 www.Hagemeister.net

"Programme dürfen Ärzte nicht manipulieren" (Software must not manipulate the physicians) reads a recent headline of an article in the German medical journal[1], discussing the effects of drug advertisements by sponsoring drug companies hidden in software designed for the administration of GP-practises.

The use of modern IT equipment is changing the working environment at medical practices and in hospitals alike. From patient administration to the documentation of the clinical course of events, from delivering and storing the results of special examinations to writing discharge letters and the billing of the care, the health care professional nowadays is facing computers everywhere.

Has this 'computerisation' of the medical workplace negative consequences for the quality of medical care? Does it affect the health carer-patient relationship? To answer these questions, the 'function' of a diagnosis needs to be understood and the influence that IT might have on the making of diagnoses and on the choice of treatment to be analysed.

'Making' of a diagnosis

"How is the appendicitis in room 12 doing?" - A patient's diagnosis is one of the key concepts in modern medicine and medical practitioners have been criticised for a tendency to treat diagnoses rather than patients. But despite of its central role, even in times of 'evidence based medicine' the 'making of a diagnosis' is a rather weakly defined process. A sequential approach is being taught at medical schools and usually, is followed by the medical practitioner (i.e. general history, history of complaints, physical examination, specific 'diagnostic' procedures)[2]. Yet 'diagnosing' lacks clear epistemological rules as they might be expected in science. The diagnosis rather acts as a working hypothesis for the further care of the patient, based on an initial assumption that is generated by using a combination of the patient's complaints, clinical symptoms and the doctor's experience. Thus a diagnosis is not a definite and ultimate entity but rather 'the best guess under what is currently known', and such a diagnosis is continuously reassessed and modified by additional data gained either from further investigations or from the clinical course and the response to specific therapy[3]. Therefore a diagnosis contains the influence of interpretation by the medical professional and is determined positively (by supporting findings) and negatively (by the exclusion of other diagnoses through contradictory findings), in extreme, a diagnosis can be made entirely by exclusion.

The role of a diagnosis

First and foremost, the classical function of the diagnosis is to predict the natural course and (if necessary) to choose an adequate therapy for the patient's condition[4]. Adequate in this case means a therapy that is likely to succeed. In times of 'evidence based medicine' (EBM) this necessitates proof of effectiveness and efficiency in random controlled trials, but the choice of the right therapy as well involves the physician's experience (e.g. "Is the patient likely to adhere to such a therapy?").

In a broader sense, 'adequate therapy' can imply the inclusion into comprehensive treatment programs, e.g. the 'disease management programs' (DMP) offered by health insurances[5] or treatment programs run by governmental institutions[6].

---

[1] Rabbata, Samir: Praxissoftware: Programme dürfen Ärzte nicht manipulieren. 1346

[2] rooting in traditions as old as the Hippocratic era, the introduction of the 'Clinical Method' has been reported for the end of the 19th century: McWhinney, Ian: A Textbook of Family Medicine. 130-4

[3] some of these epistemological features the 'diagnosis' shares with 'scientific theories' in the way Karl Popper defines it, cf. K.P.: Alles Leben ist Problemlösen. 26ff

[4] McWhinney, Ian: A Textbook of Family Medicine.152

[5] as one example the diagnosis of Diabetes mellitus Type II ‚qualifies' a patient for the inclusion into the respective 'disease management program' of the Bismarckean health insurances in Germany

[6] following WHO recommendation, the tuberculosis treatment in many countries is offered free of charge, and in South Africa a confirmed positive HIV test allows the patient access to the 'wellness program', including prophylaxis and treatment of opportunistic infections, regular blood investigations and the provision of antiretroviral medication according to the stage of the disease

In a similar way, the diagnosis can justify social benefits as well, e.g. if the patient qualifies for social grants[7], sick-leave or even just attention and support by others.

On the other hand, diagnoses can exclude the affected person from intended activities such as practising a certain profession (e.g. becoming a pilot) or even from social 'privileges' such as immigration[8] or the granting of a life insurance policy.

In extreme cases, a diagnosis even can cause specific medical treatment not to be given to the patient, such as dialysis (if infectious diseases such as Hepatitis or HIV are present) or transplantation (if a malignant disease is diminishing the potential recipient's prognosis).

In certain diseases, the diagnoses have an 'alarming effect' for the community as well. Public-health-practitioners get alerted by 'Tuberculosis', 'meningococcal meningitis' or 'Ebola' and consecutively infection control mechanisms are implemented. Clusters of occupational illnesses may lead to a thorough investigation of the underlying cause and to the introduction of better protection for the worker.

Sadly, a diagnosis can qualify a deceased person as potential organ donor ('brain-dead' but otherwise 'healthy'), thus giving new hope to others.

And last but not least, the diagnosis satisfies the doctor's curiosity and enables the doctor to communicate his observations. Whether this happens in form of the description of a single case (case report) or in a large multi-centre randomised study, the diagnosis serves as an essential tool for the medical scientist to identify what he is talking about.

The benefits of IT in medicine

IT has given rise to completely new diagnostic tools such computer tomography and has revolutionised the effectiveness of other techniques like ultrasound. In addition to this, many results of examinations are being stored digitally today (e.g. x-rays), with a number of beneficial effects like the reduced need for archives, the easier, faster and more reliable accessibility and the opportunity to process such images. In hospitals and medical centres with a dedicated IT-support, IT-networks can fasten the transfer of requests and make results that have been generated elsewhere (X-ray, ultrasound, laboratory, ECG etc.) immediately available to the doctor who ordered the investigation[9]. This is probably one of the applications of IT that is most beneficial for the patient since both, making the diagnosis and treating the suffering are accelerated.

In other situations 'telemedicine' is used to offer specialised care that would not be available otherwise, e.g. a neurosurgical opinion for patients in small hospitals without such specialists[10].

The aim to deliver better care or to achieve better efficiency and continuity of care has given birth to other kinds of networks for integrated care, too. In the British NHS clinics, hospitals and 'mobile doctors' are linked by one computer-network and findings during an out-of-hours home visit to a patient are electronically forwarded to the patient's GP[11]. A similar approach is planned in Germany with the 'elektronische Gesundheitskarte'[12], where the digitally served data in the card are intended to prevent unnecessary double investigations or accidental application of drugs when a known allergy exists[13].

An IT-application of less direct benefit for the patient is the digitalised claiming, where the provider (doctor) uses a computer and software to capture the relevant data on the patients he sees (name, insurance number, diagnosis etc.) and forwards it in an electronic format defined by the medical aid to

---

[7] in South Africa a social grant is given to HIV-patients solely dependent on the CD4-count (<200)

[8] e.g. when countries ask for medical certificates

[9] Krüger-Brand, Heike: Picture Archiving and Communication System – Abschied von der „Bildertüte". 1949-50

[10] e.g. within a network in the north-eastern German state of Mecklenburg-Vorpommern, connecting smaller hospital with the university hospitals

[11] though the EMIS© system used in the NHS still does not allow immediate access to e.g. specialist reports - personal experience in the UK

[12] Krüger-Brand, Heike: Orientierung am Anwender. 2001

[13] cf. on the advantages in family practice: McWhinney, Ian: A Textbook of Family Medicine. 379-80

claim the payment for the treatment[14]. Similarly hospitals use special software not only to administrate the patients but to capture the patients' diagnoses and procedures, something essential in times when electronic 'grouper' generate a 'diagnose related group' (DRG) from these data and these DRGs e.g. in Germany are defining the financial remuneration for the respective case.

In other applications as well, the computer is used to make data capturing easier and more efficient, when public health authorities or research institutions utilise these aides to collect and process large amounts of data, hoping to identify relevant threats, risk factors and the same.

Potential negative influence of IT on doctors' behaviour and on the quality of medical care

The possible adverse effects of the use of IT in medicine can be grouped into two classes, i.e. those effects that are intrinsic to the use of IT (e.g. the conversion of thoughts or verbal information into digital data) and those effects where IT 'facilitates' negative influences from other sources.

The quest for an early diagnosis

Converting a diagnosis into a computerised code first necessitates certain skills, the medical practitioner (or somebody else who forwards and processes the information) will have to be able to enter the diagnosis and supporting evidence into the computer system. The change of 'medium' this information is undergoing, restricts its range of content, the differentiated thoughts and doubts about such diagnoses in a doctor's mind or even as shared knowledge in a medical firm in the ward is reduced to a definite entity. In principle, this phenomenon is not a new one, as diagnoses had been fixed in written medical works (books, discharge letters etc.) for centuries. Yet, the difference might be that the computerised data may be easier accessible to others, without any additional comment by those who 'created' the diagnosis, and thus might be more prone to misinterpretation[15]. Due to the fact that the actual diagnosis is e.g. in Germany's DRG system relevant for the financial outcome, administration and the senior clinical staff pay a high degree of attention to an early and 'appropriate' encoding of the admission diagnosis. In this manner, a potentially 'premature' diagnosis-making is motivated for. Despite of the possibility to alter the diagnosis during (and even after[16]) the hospitalisation of the patient, an initially made wrong diagnosis has a strong potential to affect the further treatment ('being on the wrong track').

Similarly, even in the ambulatory sector in Germany and in South Africa the medical practitioners in private practise are forced to justify their payment claims for treatment and medications by (ICD-encoded) diagnoses. Since in German family medicine, the intentional delay of making the ultimate diagnosis ('abwartendes Offenlassen der Diagnose')[17] is a well accepted procedure in many minor complaints, this 'demand for a diagnosis' contravenes the recommendable approach for the case management. At least the German system offers an additional code to express whether a diagnosis is suspected (**V**erdacht), confirmed (**G**esichert) or excluded (**A**usgeschlossen) with the respective letter to be added to the ICD-code. In this way, a bit of the considerations and doubts can be mirrored in the digital system - an opportunity the system in the hospital sector, unfortunately does not allow for.

To adjust the diagnosis or the patient?

On the other hand, the process of 'digitalising' the diagnosis might influence its content by more fac-

---

[14] in Germany, where the ‚Kassenärztliche Vereinigung' (KV) acts as an intermediary between doctors and medical aids, the KV Hamburg in 2004 coined the expression 'Papierabrechner' ('paper-claimer') for those practitioners who still used the 'archaic' procedure without computer

[15] despite of the principle aim to achieve it, the allocation of diagnoses and ICD 10 codes (tenth revision of the international classification of diseases, a code that consists of a letter and one to five digits) is unfortunately not always one-on-one ('ein-eindeutig' – reciprocal unambiguity), thus a single ICD 10 code can represent different diseases or the same disease might be encoded using different codes

[16] e.g. in case of pending histological results

[17] in cases of minor complaints and after excluding live threatening complications, it is justified to monitor the patient since the natural course of a disease often gives important additional information on the kind of the disease

tors than the mere limitation of information entered, stored and made accessible in the computer, since the fact of working with information technology might have a direct effect on the process of 'diagnosing' itself. In an extreme situation, this could result in the diagnosis being 'adjusted' to the framework defined by the information technology. This could simply be the case if the encoding system does not offer the real diagnosis as a selectable option and therefore, the diagnosis needs to be substituted by an 'available' code, a problem that occurs especially in the initial phase of a new system.

A more 'sinister' variant of this phenomenon occurs when the computer is used to optimise the financial outcome and therefore the patients' diagnosis is altered. The extreme 'Brave new world'[18] version of this kind of reaction is sketched by Samuel Shem in his novel 'Mount Misery': The patients are automatically 're-'diagnosed, leading to those diagnoses the health insurances are currently paying well for[19].

Unrealistic as this scenario might seem, there are tendencies towards this direction. In German hospitals, after a patient has been hospitalised, the software generates not only the DRG according to the entered diagnoses, but gives values of minimum[20] and average duration of the hospitalisation of such cases. If the patient is discharged sooner the remuneration is reduced accordingly, what might obviously affect the motivation to send the patient home. But once the minimum duration has been covered, the hospital administration might 'ask' not to keep a patient too long. Similarly, the DRG gives the hospital already an instant idea of how much a payment can be expected for the case – a situation that might cause pressure on the attending medical practitioner not to opt for too expensive procedures (or in extreme cases to refuse the hospitalisation).

Yet the task of converting a case into a DRG is even more complicated. As mentioned above, the ICD-10 might offer different codes for the same disease, resulting in different DRGs and ultimately different payment. And even the different weighting of the patient's diagnoses might show such a result: It has to be chosen one 'main-diagnosis' with the other ones being 'additional diagnoses' and the right choice of the main-diagnosis can more than double the remuneration. With this 'science' of encoding being so important for the financial wellbeing of the hospital, some journals offer monthly exercises in 'proper' encoding[21]. Obviously these exercises put more emphasise on the best financial outcome within the given regulations, rather than a most conclusive description of the 'real' diagnosis of the patient[22].

'Plausibility-check' and 'End of funds'

The use of IT-technology in processing financial claims for health care facilitates a thorough 'examination' of the providers' claims like, in particular if hospitals or private practitioners hand in digitalised data.

Obviously, the above mentioned claims for hospitalisation can be checked by the health insurance the money is demanded from most easily, with sometimes painstakingly detailed 'requests for explanation' by the insurances. But even the private practitioner is exposed to this kind of 'quality control': Since every consultation or prescription needs a (justifying) diagnose, the health insurance can easily run a test with the data set handed in for claims. This 'plausibility-check' will reveal whether all prescriptions were 'adequate' or whether e.g. a prescribed drug does not match the diagnosis it was prescribed for.

---

[18] in his foreword to the 1946 2nd edition of his novel Aldous Huxley states: "The theme of Brave new world is not the advancement of science as such; it is the advancement of science as it affects human individuals". 10 – the very theme of this paper, too!

[19] Shem, Samuel: Mount Misery: "They spent hours (…) reading (…) whichever DSM diagnosis was now bankable. If suddenly insurance was paying top dollar for, say, 301.13, Cyclothymic Disorder, the big Toshiba would be reprogrammed to reprogram all the little Toshibas, so that for data I typed in (…) now it would spew out, time after time like a run of luck at craps in Vegas, 301.13, 301.13, 301.13 (…)". 213

[20] ‚Untere Grenzverweildauer'

[21] e.g. the journal of the German Association of surgeons ‚Der Chirurg BDC'

[22] it is not the author's intention to claim that the above mentioned journal or others would propagate an untrue encoding of the patients' diagnoses, but since it is known that it is difficult to optimise a system with regard to two parameters, the preference might happen to be 'financial outcome' before 'adequacy'

Between different countries, the constraints for the private practitioners (and their potential responses to the quest) differ. In South Africa e.g. the funds for drugs within a given 'medical aid[23]' often are limited *per patient* – once the patient has exceeded these funds, the medical aid is not going to pay for any more drugs under the specific scheme. Here the practitioner should be aware of the 'limitations' if he dispenses drugs from his practise[24], an adequate software in the practise will help to get aware that the funds are exhausted.

On the contrary, in Germany the funds for drug prescription in the private health insurances usually do not have a ceiling. The limitations (and financial risks) for the private practitioner are different: The Bismarckean insurances (with assistance by the above mentioned KVs) under the current legislation, aiming at cost containment, may implement a limit for prescription of drugs, laboratory investigations or physiotherapy *per doctor* (or practice). The software to be used in such a practice in Germany, usually offers functions to estimate whether the recent prescriptions have exceeded such limits. The doctor faces the risk of having the expenses for these excessive drugs/ therapies deducted from the remuneration, within this so called 'Regress'-system, the medical practitioner in effect pays for the prescriptions. The resulting dilemma isn't too difficult to imagine, since the doctor (with assistance by the IT) can figure out when the limits are reached – and then either decide for an alternative therapeutic approach (something that has not been 'overused' yet) or accept a reduced income for the sake of granting the patient the first choice of treatment[25].

Conclusions

As computers and software assist the practitioner in identifying pending financial threats, it is obvious that the use of IT under these circumstances does not necessarily serve the patients' best interest: Given the subjective influence on the making of a diagnosis and the epistemologically 'fragile' quality of the diagnosis, it is not too difficult to imagine that external pressures that are either exercised onto the practitioner or identified by the practitioner by using IT could significantly affect and alter both the diagnosis and the consecutive treatment decided upon by the doctor.

Certainly, information technology has a great potential to assist in making patient care faster, more efficient and more continuous. But as nothing comes without a price, IT in health care could have detrimental effects on the quality of care by

-causing the doctor to *forget* important information (due to 'early diagnosing' and the reduced content of information in digitally 'encoded' diagnoses)

-forcing the doctor to *falsify* diagnoses (to justify treatment or prescriptions if the 'true' diagnosis doesn't serve this purpose or simply doesn't exist in the system)

-tempting the doctor to *avoid* 'expensive' diagnoses and treatments (those identified as might not be paid for by the health insurance)

-motivating the doctor to *discharge* the patient sooner than appropriate (as further hospitalisation might reduce the revenue)

As often with new inventions, the use of IT in health care might thus bear negative effects as well. To gain a broader idea of the potential effects, to help to prevent the negative outcomes and to contribute to an adequate assessment of its role in health care is the intention of this paper.

References

Der Chirurg BDC, Springer Medizin Verlag, Heidelberg

Huxley, Aldous: Brave new world, Reclam Verlag, Stuttgart 1992

Krüger-Brand, Heike: Orientierung am Anwender – Forderungen der Leistungserbringer zur Gesundheitskarte, Deutsches Ärzteblatt 2006 (103) 30, A2001

Krüger-Brand, Heike: Picture Archiving and Communication System – Abschied von der „Bildertüte". , Deutsches Ärzteblatt 2006 (103) 28-29, A1949-50

---

[23] the expression used in SA for a private health insurance

[24] a usually quite lucrative option if the practitioner has acquired a dispensing licence

[25] the actual trilemma has been demonstrated recently by demonstrating private practitioners in Germany, where after at a certain point in each remuneration term they might rather completely close their practices, as any further activity effectively results in reduced income

McWhinney, Ian: A Textbook of Family Medicine, 2nd ed., New York/ Oxford, Oxford University Press 1997

Popper, Karl: Alles Leben ist Problemlösen, München, Piper Verlag 1996

Rabbata, Samir: Praxissoftware: Programme dürfen Ärzte nicht manipulieren, Deutsches Ärzteblatt 2006 (103) 20, A1346

Shem, Samuel: Mount Misery. New York, Dell Publishing 1997

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

J. Lahtiranta and K. K. Kimppa

The Use of Extremely Anthropomorphized Artefacts in Medicine

Abstract:

Anthropomorphized, or human-like artefacts, have been used in teaching and training of medical skills for a long time. The most famous artefact used today is probably the Resusci ® Anne CPR training manikin, which is used for training of resuscitation skills. However, what has changed over the lifespan of these artefacts is the level of human-like features in them. All around the globe, highly anthropomorphized ICT artefacts are used in training of medical skills. Amongst others, the UNAM University in Mexico City and Royal North Shore Hospital in New South Wales use artefacts, which are closer to human-like robots than traditional manikins in teaching. The purpose of this article is to look deeper into this phenomenon, consider its potential implications for the patient-physician relationship and quality of patient care, and to propose some practical methods for minimizing the possible risks emerging from the use of these extremely anthropomorphized artefacts.

Agenda

Authors:

J. Lahtiranta:

Department of Information Technology, Information Systems, University of Turku, FIN-20014 Turku, Finland.

☎ + 358 - 2 - 333 8674, ✉ janne.lahtiranta@it.utu.fi

K. K. Kimppa:

Department of Information Technology, Information Systems, University of Turku, FIN-20014 Turku, Finland.

☎ +358 - 2 - 333 8665, ✉ kai.kimppa@it.utu.fi

There are various arguments for and against anthropomorphism in the Information and Communication Technology (ICT). Some argue that anthropomorphism, or introducing human-like features, into ICT artefacts ensures better user acceptance and creates a positive user experience (Cassell et al., 1999; Gong, 2002). On the other hand, others are adamant about potential risks and threats of anthropomorphism, claiming that it has a negative impact on the user interaction, creating unpredictability and vagueness (Shneiderman, 1988).

Regardless of the arguments for and against anthropomorphism, it is inherent to humans to anthropomorphize different things. Every day we encounter anthropomorphism in various situations. For example, anthropomorphism is widely used in entertainment and advertising. Movies like "Madagascar" by the DreamWorks Animation SKG (2005) or the famous cartoons by Walt Disney are casebook examples of anthropomorphism. Unintentional, or implicit, anthropomorphism is probably even more common. For example, we have a tendency of rationalizing actions of some (computer) artefact by referring to it as "he" or "she".

Already during the 60's there were clear cases of computer programs being anthropomorphized. One of the most famous case being ELIZA, a computer program by Joseph Weizenbaum (Weizenbaum, 1966), which parodied a Rogerian therapist. However, recently the anthropomorphism has been used in the ICT applications probably more than ever before. For example, in applications and systems of medicine (i.e. in electronic health, telemedicine, consumer health informatics, etc.) anthropomorphism is used not only in graphical user interfaces, but also in patient-physician interaction, teaching and patient care (figure 1).

Figure 1. Examples of the use of anthropomorphism in medicine.

In the first example, a remote controlled robot is used to visually examine a patient. With this solution, physicians are capable of interacting with the patients whether they are in another part of the hospital, or even in the another part of the world (Reuters, 2005a). In the second example, medical students are trained by using robot dummies. The robots are equipped with mechanical organs, synthetic blood and mechanical breathing systems (Reuters, 2005b). In the last example a robot nurse called "Pearl" assists elderly people with their day-to-day activities (Jajeh, 2004).

While these are just examples about anthropomorphism in medicine, they raise a question of what kind of impact, if any, anthropomorphism has on patient-physician relationship or on quality of patient care. Patient-physician relationship has always been one of the cornerstones for comprehensive care. The development from human-to-human to human-and-computer-to-human, and finally to computer-to-human "relationship" creates further space for the critique directed towards the use of applications and systems of ICT in the established medicine. A potential consequence from this is that the human is more easily regarded as an object rather than a real feeling subject.

What is known about anthropomorphism in human-computer interaction is that it has some effect on human behaviour and decision making. For example, in some cases a human-like (or social) interface can be used to invoke trust in the user (Bickmore & Picard, 2004). It is also demonstrated in numerous psychological studies, that people treat computers and new media as real people and places (Reeves & Nass, 1996). Considering the applications and systems of electronic health in widespread use today and the future possibilities created by a general technological advancement, the question about the potential impact is a current one and requires close attention.

Anthropomorphism and Medicine

Susceptibility to anthropomorphism is always individual. Some of the users are more prone to it than others. Similarly the effects of anthropomorphism change from one individual to another. Due to this individuality, defining what could be called as optimal anthropomorphism, or creating a balance between positive user experience and complicated user expectations, is extremely difficult. On the far end of the scale there are interfaces and technologies stripped from all human characteristics, including the used style of writing. On the other end are the highly anthropomorphized human-like artefacts that mimic human looks and behaviour as far as technologically possible.

In medicine, there is a need for the whole scale from "man to machine". For example in the clinical decision making the users need information, such as numerical laboratory results, in a concise and explicit form. In training the users practice their clinical skills with anthropomorphized artefacts, such as the Resusci ® Anne CPR training manikin with realistic full-body anatomy. Between these two examples, there is a wide range of different artefacts with varied degrees of anthropomorphism for different types of users (i.e. consumers, patients, doctors, etc.).

In a situation when the users operate with a single artefact at a time, they are usually more aware of its type, role and significance. For example, in health care an ICT artefact could be a conveyer of information, source of it, or just a tool for practicing and demonstrating clinical skills. Unfortunately, especially in today's operational health care work, the users operate with various types of artefacts at the same time. Furthermore, these artefacts usually form a mosaic-like structure and hierarchy, potentially concealing some of the artefacts (and even actors) from the user. Especially asynchronous Internet-based services where the user and service provider do not interact face-to-face in real time, or services which employ (semi)autonomous agent technology, are prone to this kind of complicating effect.

A potential problem occurs when the user becomes blinded by the layers of technology and the true identity of the service provider becomes obscure. Such situation could occur due to the intermediating channel, type of the service or to the aforementioned mosaic-like structure. The situation is more problematic if the identity of the service provider changes between an ICT artefact and a human actor. This change becomes a liability when it occurs unnoticed, without informing the human actor or the one acted upon. In this kind of situation the user may consider the source of the information to be a human actor, while in reality it could be an ICT artefact. The impact and consequences of this kind of muddling effect is further explored in Lahtiranta & Kimppa (2004).

A more interesting scenario, in the context of this article, occurs when the user is given a single highly anthropomorphized tool to operate with and all potentially muddling effects of different layers of technologies are stripped down to minimum. In this kind of situation the user is prone to the most basic, implicit, anthropomorphism typical to the human nature. What are the potential effects of anthropomorphism to patient care when the user is trained with a highly anthropomorphized artefact, such as the robot dummy used in simulating childbirth (figure 1) and the students at some point must apply what they have learned to a living subject?

Extremely Anthropomorphized Artefacts in Medicine

It is true that medical students must practice their skills using different kinds of resources. Some of the necessary information comes from the literature, some from the teaching and training, and some from adapting their skills into practical patient care. Practical training requires real, live patients but before the students are skilled enough, they are usually trained with dummies, charts, ICT artefacts, etc. For ages, different kinds of models and manikins have been used in training. However, now more than ever, these manikins are given extremely lifelike qualities. In medicine the ICT artefacts, namely manikins and robots, have become subject to unconstrained anthropomorphism.

In robot design, or in design of "a machine that looks like a human being and performs various complex acts (as walking or talking) of a human being" as defined by the Merriam-Webster's Online Dictionary, the robot features can be analyzed in a design space consisting of three dimensions (figure 2).

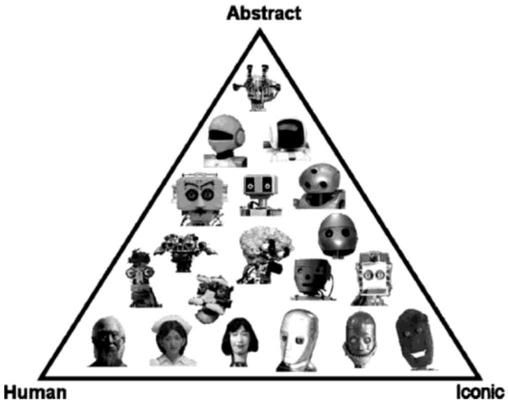

Figure 2. Anthropomorphism design space for robot heads (Duffy, 2003).

The design space defined by Duffy (2003) specifies an illustrative map for designing robot features and defining how closely these correlate to human physiognomy. Features in the iconic dimension employ a very minimum set of human characteristics often found in comics and cartoons. The abstract dimension refers to functional or mechanistic design. Naturally, the human dimension describes features with close resemblance to a human counterpart.

In medicine, the appearance of robots and manikins belong to the human dimension for practical reasons. For example, in CPR training the used manikin must have a close resemblance to the human physiognomy in order to provide the most effective training experience. If the manikin used in training would have iconic or even abstract features, it would probably have a negative impact on the overall medical training. It is not feasible to make abstract patient manikins and robots when the purpose is to prepare the students for a real-life experience.

Mori (1982) defines a region in the robot design when the robot is not quite humanlike as "The Uncanny Valley". Robots, which design falls to this region, are considered a bit peculiar or even absurd and interacting with them can feel somewhat uneasy. Considering that a close resemblance, or even an exact match, to a human counterpart is a necessity in designing certain ICT artefacts used in the training of medical skills, it can be argued that their design is crossing the region. When the design of a robot or manikin crosses the Uncanny Valley, they are potentially becoming more accepted as humans than ever before (Mori, 1982).

In addition to the physical appearance of the artefact used in training, the familiarity of the artefact to the user can be analyzed by using the core set of features defined by Duffy and Joue (2005). These features are:

> Control: the ability to influence the robot, or our environment through it.
>
> Predictability: a human-like, not too simplistic, predictability of the robot.
>
> Dependency: how dependent the robot is of our actions.

When the focus is on training of medical skills using extremely anthropomorphized ICT artefacts, the control feature is highly presented. For example, when medical students practice CPR with a manikin, they strive to stabilize and gain control of the artefact's life-like functions. In this context, the control is realized as medical care and nursing. In training, the used artefact should be as human-like as possible, including appearance and physical responses to the care. Therefore, it can be assumed that the predictability, namely human-like predictability can be high from mechanistic perspective. Furthermore, if the manikin is extremely human-like in appearance, it may create an unfounded illusion that it reacts in all the possible physical responses like a real human would even though it is unlikely to cover them all. Misinterpretations like this with anthropomorphized manikins without any kind of social interface, such as the Resusci ® Anne CPR training manikin, are clearly less likely to happen. The psycho-social dimension of predictability is lower however since with the current level of technology it is hard or nearly impossible to mimic psychological and social behaviour patterns of a human. Since the purpose of the artefact is to imitate physical responses of a human patient in need of a care, the dependency from the care provider is highly presented as well.

From the perspective of the three core features of control, predictability and dependency, it can be argued that the extremely anthropomorphized ICT artefacts used in the training of medical skills have a high potentiality of familiarity to their users. However, we must acknowledge the limitations and restrictions of the used technology and design, which primarily focuses on mimicking human-like appearance and physiological responses to the care.

Optimal Anthropomorphism in Medicine

In the context of medicine, and especially in training of medical skills with extremely anthropomorphized ICT artefacts, it is hard to define an optimal degree of anthropomorphism. From purely didactic perspective, one could be anxious to say that optimal anthropomorphism from design perspective is achieved when the artefact matches to a human counterpart. In addition, one should be able to analyze the performance and effect of the care, which may require dismantlement of the artefact or connecting it to other ICT artefacts, etc.

This mechanistic view however has some drawbacks. A human is a social and psycho-physical entity, an individual. With a mechanistic view to design it is impossible to capture the nuances of working with a living patient. From this perspective optimal anthropomorphism in design should be able to mimic the "vagaries of a human mind" as well as the "fragility of the human body" in all its dimensions. This however, is beyond current technology.

In training of medical skills using extremely anthropomorphized ICT artefacts an optimal anthropomorphism is a delicate balance of three factors. The first two, as suggested by Duffy (2003) are human expectations and capabilities of the artefact. In the context of training, the third factor is formed from the didactic requirements. In addition to the physical form, or completeness of it, the key question is the social interface. The social interface plays a significant role in how the students perceive the artefact to be emotional, sentimental of even intelligent.

A social interface as a supplement to an extremely anthropomorphized form can be viewed as a disadvantage or hindrance, or as a benefit. The interface can be perceived as a hindrance if it has a negative effect on the overall learning experience, such as when the interface diverts the attention of an unpractised medical student away from the physical indications and from the real learning emphasis. On the other hand, a social interface can be used to convey additional information about the success or progress of the administered care (e.g. when the artefact can demonstrate changes in the degrees of pain), or in some cases, about the preconditions of care (such as way of living, use of medications, and so on).

Considering the potential benefits and drawbacks of a social interface it is hard, or nearly impossible, to define what level of anthropomorphism is optimal in this matter. In some cases it could be justified from pedagogical point of view that the artefact shows intentional behaviour, carries out a meaningful conversation and creates an illusion of a living patient. An alternative could be drawing a line to the type of discipline in question. A meaningful restriction for not using extremely anthropomorphized artefacts could be that they should not be employed in teaching and training in the fields of science that deal with mental processes and behaviour of a human being (e.g. psychology), or in other fields that deal with mind and behaviour. This line, however, is drawn in water since it can be argued that social or mental aspect of care is inseparable from medical care.

Conclusions and Recommendations

It is true that it takes more than a human face on a computer monitor to create an illusion of life and intelligence. It may require a use of human-like verbal cues or body language to facilitate social interaction; or just a little bit of user imagination. However, with the examples presented in this article, it may not require a great leap of faith to consider a robot closely resembling a newborn child as its human counterpart. Actually, the situation may be quite the opposite: the users must actively remind themselves that they are practicing with a dummy.

Extremely anthropomorphized artefacts have been used in medicine only such a short period of time that it is difficult to give a definite answer to the question about potential effects to patient care. However, it is clear that the boundaries between a target of care and the tools used for training will become muddled. The history has proven that we will create as anthropomorphic robots and manikins as possible. For example, toy and entertainment industries are full of vivid examples of this.

There is a great potential value in practicing health care with the anthropomorphized ICT applications and systems of medicine. Being able to make the first tries, and the first mistakes, on a lifelike application instead of a real person is an indisputable benefit. It even seems to be our duty to steer the development of the artefacts used in teaching and training of medical skills to this direction. It benefits the patients, both in not having to submit to the fumbling tries of the unskilled trainee, as well as possibly saving lives from the first mistakes. However, we express concern that if the artefacts used in teaching and training of medical skills become human-like enough to fool the medical students and yet the artefacts do not actually react like a human being, the students may consider that this is the way a real human being acts under the treatment. This perception is likely to be false since patients are individuals and they can respond very differently to the treatment, for example due to the individual's pain threshold – or even due to one's imagination.

It is already clear that the current extremely anthropomorphized artefacts such as the robot dummy used in simulating childbirth (figure 1) cannot be fully equivalent with the actual counterpart they simulate – even though the artefacts seem very real. Especially in the near future when the artefacts are probably even more life-like, it is paramount not to mistake the artefact's operation or artificial life-signs it produces as an extensive description of all possible outcomes in real-life. Similarly, the conclusions drawn from the artefact's operation should not always be interpreted in the same way as in a real-world situation.

Until we can rest assured that the used technology is sufficiently human-like to represent an actual correspondence to a real human and the training is repeated long enough to make the differences human beings have in responding to the treatment clear, the artefacts used in teaching and training of medical skills should be clearly marked as tools, however anthropomorphized. Before the technology reaches its peak and the students can practice with artefacts that respond exactly like human beings we must define principles and ensure in practice that the transition from "manikin to man" goes with as few complications as possible. To decrease the potential complications later during the career of the medical students, the instructors should be able to orient the students into a right state of mind prior to actual training.

Both regulators and ethicists are often criticised for being late with their recommendations on how to utilize or apply ICT applications. Therefore it is important to take a proactive approach to the technology development and consider its potential impact on the patient-physician relationship and quality of patient care. Even though the technology today is limited to mimicking human vital functions mechanistically, there are clear sings that the robot functionality is coming closer and closer to imitating human activity. See for example work by Cynthia Breazeal (Breazeal & Scassellati, 2002; Breazeal, 2003; Breazeal et al., 2004) and her expressive robotic creature, Kismet[1].

References

Bickmore, T. W. & Picard, R. W. (2004) Towards Caring Machines. Proceedings of CHI, Vienna, April 2004

Breazeal C. (2003). Emotion and sociable humanoid robots, E. Hudlika (ed), International Journal of Human-Computer Studies, 59, pp.119-155.

Breazeal, C., Brooks, A., Gray, J., Hoffman, G., Kidd, C., Lee, H., Lieberman, J., Lockerd, A., & Mulanda, D. (2004). Humanoid Robots as Cooperative Partners for People. Submitted to IJHR (2004) online at: http://web.media.mit.edu/~cynthiab/Papers/Breazeal-etal-ijhr04.pdf accessed 17.8.2006.

Breazeal, C. & Scassellati, B. (2002) Robots that imitate humans, Trends in Cognitive Science, 6, pp. 481-487.

Cassell, J., Bickmore, T., Billinghurst, M., Campbell, L., Chang, K., Vilhjailmsson, H., & Yan, H. (1999). Embodiment in Conversational Interfaces: Rea Proceedings ACM SIGCHI '99. Pittsburgh, PA, 1999, 520-527.

Duffy, B. R. (2003). Anthropomorphism and the Social Robot. Special Issue on Socially Interactive Robots, Robotics and Autonomous Systems 42 (3-4), 31 March 2003, pp. 170-190 (extended version of IROS02 paper).

Duffy, B. R. & Joue, G. (2005). Why Humanoids, 4th Chapter Conference on Applied Cybernetics 2005 September 7-8, 2005, City University, London, United Kingdom.

Gong, L. (2002). Towards a theory of social intelligence for interface agents. Workshop on Virtual Conversational Characters: Applications, Methods, and Research Challenges Conjunction with the Human Factors. Conference, Melbourne, Australia, November 2002.

Jajeh, P. (2004). Robot nurse escorts and schmoozesthe elderly. August 24, 2004, online at: http://www.intel.com/employee/retiree/circuit/robot.htm accessed 10.10.2006.

Lahtiranta, J. & Kimppa, K. K. (2004). Telemedicine and Responsibility: Why Anthropomorphism and Consent Issues Muddle the Picture. 5th International WE-B (Working for e-Business) Conference.

Mori, M. (1982). The Buddha in the Robot. Charles E. Tuttle Co., 1982.

Reeves, B. & Nass, C. (1996) The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, New York.

Reuters (2005a). Medical robots start work at London hospital. May 18, 2005, online at: http://www.msnbc.msn.com/id/7896909 accessed 10.10.2005.

Reuters (2005b). Medical Students Learn on 'Breathing' Robots. September 27, 2005, online at: http://www.findarticles.com/p/articles/mi_zdewk/is_200509/ai_n15615573 accessed 10.10.2005.

Shneiderman, B. (1988). A Nonanthropomorphic Style Guide: Overcoming the Humpty-Dumpty Syndrome. The Computer Teacher. October, 1988. 9-10.

Weizenbaum, J. (1966) ELIZA – a computer program for the study of natural language communication between man and machine, Communications of the ACM, Vol 9, No 1, January, 1966.

---

[1] http://www.ai.mit.edu/projects/kismet

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

Britta Schinzel

Gender and ethically relevant issues of visualizations in the life sciences

Abstract:

Here moral problems created by the use of constructive imaging technologies within the life sciences are discussed. It specifically deals with the creation of dichotomies, such as gender, race and other differences, created and manifested through the contingent use of scientific and computational models and methods, channelling the production process of scientific results and images.

Gender in technology studies has been concerned with destabilizing essentialist and dichotomous co-constructions of gender and technology. In the technological construction process gendered social constructions of stereotypes and inequalities both of the technological models and of the presumptions in life sciences become structural properties of the artefacts, again flowing back into the seemingly objective results and knowledge of the life sciences. Here we will deal with the construction of gender differences via biomedical imaging and the creation of norms in atlases. Additionally, the de-contextualized images, showing idiosyncratic selections and reducing complexity are used to popularize gendered assumptions about biological facts.

Agenda

Author:

Prof. Dr. Britta Schinzel:

Institute for Computer Science and Social Research, University of Freiburg, Friedrichstr. 50, 79098 Freiburg, Germany

☎ + 49 - 761 - 2034953 , ✉ schinzel@modell.iig.uni-freiburg.de, 💻 www.mod.iig.uni-freiburg.de

Relevant publications:

- Schmitz, Sigrid & Schinzel, Britta (2005): Grenzgänge. Genderforschung in Informatik und Naturwissenschaften. Königsdorf: Ulrike Helmer Verlag, 173 S.
- Schinzel, B., Taeger, J., Gorny, P., Dreier, Th., Holznagel, B. (Hrsg.): E-Learning im Hochschulverbund, DUV Wiesbaden, 2004, 272 S.
- Schinzel, B.: Computer Science between Symbolic Representation and Open Construction; in Lenski, W. (ed): Logic versus Approximation: Essays Dedicated to Michael M. Richter on the Occasion of his 65th Birthday Springer Lecture Notes in Computer Science Volume 3075 / 2004 ISBN: 3-540-22562-5, 16 S

Introduction

Computer Science with its technological products has more or less changed all the sciences and their production processes intensively in general. Its symbolic methods and mathematic-technical paradigms penetrate their model building processes and its methodological instruments, since the data memorizing and integrating capabilities and its visualizing potentials are used also in biomedicine and in the life sciences. The uniform computing methodology of formalization, which is usable independent of the subject, leaves much room for different ways of problem solving and their specific realizations in software: specification (which means de-contextualizing a certain part of the world, abstracting from it, such that it can be represented in discrete symbolic items), architecture, algorithmic solutions and the very implementation and coding. The space left open for specification and modelling alternatives is huge and it opens doors for moulding in one-sided selective views, for idiosyncratic and biased design and for contingent re-contextualization.

Technologically mediated textual, imaged and formalized knowledge is currently changing the order of knowledge (Spinner 1994) through new channels of categorization (e.g. building ontologies in semantic web representations, or search machines), the reduction of complexity and context, formalization and standardization. The vast collection of complex data sets, produced with the help of information technologies brings humans' cognitive capacities to their limits. For this reason, visualization technologies are used more and more to display the essence of results instantaneously. Scientific knowledge therefore is increasingly represented in images, graphics, mathematical and biomedical visualizations (called visiotypes by U. Pörksen 1997). But this turn from text to picture is vice versa also forming our knowledge. The imaging of scientific facts is per se ethically relevant, because images do not explicate their semantic content in the same way as text does. Their meanings are much more dependent on culture, pre-knowledge and interpretations one is familiar with than would be the case of text, even from a hermeneutic view of text. In addition images are also stereotyping and contributing to standardization and normalisation. This makes popularization of scientific images even more subject to false interpretations, as e.g. biological determinations. In particular the new medical imaging methodologies, which are opening enormous possibilities for diagnosis and scientific investigation, also are posing new epistemological, ethical and validity problems: E.g., bodily properties that can be visualized on a one-to-one scale are emphasized in favour of those which cannot be locally and distinguishably represented within a picture. Moreover, the abstract and complex character of data extraction and processing produces a very loose referential tie between body and image, but this is hidden by the very realistic appearance of the images. In addition, their use for standardizations and norms are problematic for many reasons: among others, that new definitions of sanity versus sickness arise and new dichotomies are built up.

Computerized Imaging in Biomedicine

The combination of physical and physiological effects with mathematical and information technological methods have brought up many new methods for the introspection of the inner body without dissection nor invasion, such as Computer Tomography (CT), Magnetic Resonance Imaging (MRI), functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET) and others. The production of the final images relies on the electromagnetic exchange between the atomic structures of the body, delivering masses of raw data to be processed, interpreted and visualized in extremely complicated and contingent combinations of model driven algorithms, computations and visualization techniques. For most of the contrast mechanisms established in practice for imaging today there exist plausibility explanations at most, i.e. there is no deep understanding of the microstructure of tissue, which would allow sound interpretation of what is being seen (Hennig 2001). Thus, the impressive images may be misleading as they seem to show realities of the inner body, whereas they show visualizations of interpreted data, i.e. images of bodily properties are derived – in epistemologically problematic ways - from long and complicated chains of interpretations of physiology in models and computational constructions, which always bear the danger of showing medical artefacts that do not correspond to physiological realities within the inner body (see e.g. Schinzel 2003, Schmitz 2004). The naïve use of such images without reflection of their production process is ethically relevant.

These methods have brought a revolution for diagnoses and scientific production, especially within neuroscience. It is obvious that they are very useful and that they deliver insights into the living body that have not been possible before. However, the claim cannot be held that the use of such techniques always leads to "objective" correspondence with the referent body. I.e., the pictures may contain artefacts stemming from technology itself or from the interaction between technical depiction and the living body. Moreover, by the use of contingent physiological and computing models, simulations and image producing technologies they are loaded with added meanings, which may meet the concrete bodily facts or not. The bio-medical images visualize non-pictorial collections of complex data sets, which have been processed through a lot of "cleaning", analyses, transformation and interpretation steps. E.g. the stray and other data collected at the CT wall do not include the body's space coordinates, which makes complex mathematical region reconstruction necessary. And the constructive image giving methods, ruling out supposed fuzzyness and dirt, intrapolating supposed voxel-values between slices for 3D-representation, rendering, i.e. triangulating surface and inner structure of the body for cuts, deformations and transparent views, and sometimes also colouring the images, e.g. the brain's activation areas for differentiation are preparing the look of the pictures for cognitively adequate understanding. These images therefore are stuffed with interpretations of their constructors and they also produce new meanings, e.g. of reading sanity vs. sickness or needs of therapy from pictures instead of from clinical evidence, or of normality though contingent mathematical averaging methods, of life and of sex and gender.

Moreover, the pictures are driven from a moment's situation which might alter within a minute, a day, or more, according to experience. The images fix this moment's appearance as a biological fact, which has emerged due to "embodiment" (Fausto Sterling 2002) through the contingent conditions under which this appearance arose, in particular for the most variable part of the body - the brain, as described below.

Still there are huge projects, like the Human Brain Project HBP, that try to define standards of the human brain. Here standardized anatomical and functional atlases are constructed, through complicated mathematical averaging methods, diversified by age, sex/gender, sicknesses, race/ethnicity, and in all these dimensions at one point in time. However, such dichotomized standard atlases of brain anatomy and function carry with them the danger of localizing sickness, normality, ethnicity and gender within the imaged body and placing other kinds of (non)evidence into the background.

In the medical practice on the other hand, e.g. in neurosurgery, there is an aspiration to refer to norms, like brain atlases, in order to navigate more safely within the brain. Establishing atlases has become a scientific field in itself, between medicine and mathematics. Considering the problems mentioned above, the questionable correspondence between the bodies under inspection and the images constructed, the contingency of the brains' material and functional status, the validity of such standards is problematic as well. Although without such atlases virtual or real navigation in the brain is even more taping in the dark, it is still an ethical question whether to rely on such standard atlases or not, whether to take such pictorial evidence as scientifically sound and to use it as major tool for evidence in medical practice.

The depiction of illness, especially in illness atlases (see e.g. Narr et al. 2001) brings ethical questions as well, such as whether an individual's image that has similarities with an illness atlas shows that he/she really has that illness, or is in danger of contracting it. Making diagnoses and decisions about a therapy in preference of visual evidence instead of on clinical findings could occur as a consequence. Another epistemological question is whether the deviation shown is a cause or an effect of a possible sickness (see also Schinzel 2004). Furthermore, changes in how human beings view themselves, in the body and in "humanness" have been established, such as the assumption that the mind is materially located and pictorially represented in the brain, and that this might be "the whole truth" about human beings' thoughts, mind, feelings and behaviour. The new, momentary neurologically founded debate concerning free will (Geyer 2004, Hochhuth 2005) is one consequence of this new self-image of concretisation of human beings into the neuro-chemical and neural-physiological.

The Plastic Brain

Considering now the interaction of material and experience, brain functions rely on the switching of the brain's nerve cells into an information processing network through the building up of synapses. This network and switching changes with our experiences, both concerning structure and the brain's function, and it needs these sensual inputs and sensori-motoric experiences in order to work at all. The extreme neuronal and synaptic plasticity of the brain is the basis of our potential to learn and memorize: every experience, every action and every thought is physiologically manifested within the neuronal and synaptic switching within the brain, at least temporarily. Individual experience therefore creates not only the vast inter-individual variability of brain structures and functions, but also the high dynamics (through learning) during a lifetime, i.e. the intra-individual variability.

Clearly then, brain imaging will not only show genetically determined structures, but also the organizational material manifestations of different individual lifetime experiences. As life and experiences of different groups and populations, especially of women and men differ in our societies, these differences are to be expected within the brain structure and matter in some way or the other. That is, such differences are not essentially biological ones, but contingent, context dependent and variable within a population and during the lifetime of an individual. As many investigations have shown, alterations of the synaptic and neuronal structure do not hold on, and they may be reversible, unless repeated activation temporarily fixes the structure.

The brain's plasticity is the most evident example of Fausto Sterling's (2000) embodiment theory. This theory states that the interaction of the body with the environment shapes and transforms the whole body, bones, muscles, organs and nerves. In particular, sensori-motoric experiences are imprinted into the human organism, into psyche, behaviour and into the body's material reality. The human brain, both in structure and in function, at any time of our lives, is embedded into and influenced by the relationship with all its endogenous and exogenous conditions. This of course also holds for sex/gender with respect to the brain and its socio-cultural conditioning within our society. Every kind of analysis of sex/gender differences within the central nervous system therefore has to respect these open developmental dynamics of the nervous differentiation with respect to the environment.

On the other hand, brain images, let alone the constructive aspect of their production, present a momentary image that may change within the next moment, during the female cycle, with weight, with pain, with aging and with disease. Therefore, these images may not be considered as representing "the status-quo", but as moments in time during the development of the respective interaction between behaviour and brain structure and vice versa. It is immediately evident with such fluent "material" as the brain's constituents that standardising becomes problematic.

The huge scientific cartographic programs like the mentioned HBP (see e.g. Thompson et al 2000) is held in high regard for its imaging of structures and functions of the brain, for its combination of different ones into standard brains, for its transforming of individual ones into the standards, of building atlases by brain imaging and mapping. The standard atlases are constructed according to the very selections of individuals subject to imaging (large persons do not fit into the tube, more white western persons can afford to be tomographed), according to the different standardizing technologies (e.g. volume based or landmark based averaging, statistical analyses and warping), the mathematical models on which these rely, etc. Standards are always subject to specific ideas, presuppositions, and in case of brain imaging, subject to the contingent brains' state and the contingent conditions under which the images are taken. In particular, with model driven image construction, the normalization process becomes circular. Moreover, normality in our culture selects certain concepts of sanity, of sex/gender and being normal that often also means being male.

Gender and Brain

Since the existence of medical imaging, sex/gender relevant brain areas and cognitive functions have been pointed out, such as the corpus callosum and some of its parts, the splenium and the isthmus, the laterality of the left and right brain halves, or the lateralization of language capabilities. But all these findings can be put into question. Schmitz (2004) und Nikoleyczik (2004) e.g. criticised publications of functional language tests using fMRI: Shaywitz et al. (1995) investigated in rhyme identification and found that 19 test persons showed a strong activation on the left side of the frontal lobe, and that 11 of the 19 test persons showed marked activation on both sides. However, in these tests no parallel differences in performance were found. This much cited study is propounded as evidence for women's strong bi-laterality of general language processing in contrast to men's uni-laterality. In a study by Frost et al. (1999) of 100 test persons, though, no gender differences were found in language performance, nor there was activation-asymmetry identified in the examined brain areas. But this work, in contrast to that of Shaywitz et. al., is seldom cited. In a recent study, Kaiser (2004) was able to show that a small variance in the setting of fMRT-measuring had an influence on the fMRI-imaging of lateralised language performance: at one time gender differences were visible for the same persons, but with other parameter values they were not, and yet with other values even the sides in the mens' pictures became interchanged.

Visualisations of the thinking brain proffer themselves as a result of neutral technical-natural scientific workmanship that is built upon natural scientific objectivity using effects delivered by physics to enlarge human sensory perception. Digital images of the body, its organs and their functions should objectively represent unaffected truths. However, the publications mentioned above serve as examples of the deconstruction of sex/gender differences in scientific publications.

Still, popularized literature on neurology and brain science is keen on showing sex/gender differences, although their complexity is reduced in many respects: in regard to the construction process, the brain's plasticity, and the difference between sex and gender. This is not only problematic, but even dangerous, especially for adolescents without a settled gender identity.

It seems that in societies, and even more in science, there is a desire for categorizing and defining differences. It is well known that in the empirical sciences, which are making use of statistics, there is a severe publication bias, the selection of results which show statistically significant differences (Easterbrook, Berlin 1991). This holds especially true for publications on empirical findings about gender differences in the brain. As a consequence these findings are, oversimplified, often interpreted as (biological) sex differences. In contrast, gender research has shown that there are also contradictory results to any of the research results differentiating sex/gender. Nonetheless, findings not showing gender differences are much less likely to be published (Wacholder 2004). The reason for the unreliability of such findings is the complexitiy of the research question as already mentioned. The great variability inherent within every population would require considering biographic impacts and the contexts of the investigations, larger proband sample sets and more exact interpretations. Moreover the incorrect use of statistics in empirical findings is well known (Joannidis 2005).

Building dichotomies is ethically problematic, because binary relations, such as between women and men, nature and culture, healthy and unhealthy, can be easily put into hierarchical order. Norms standardizing such differing and variable subjects as the human brain introduce another ethically problematic aspect of medical imaging and atlases, as individual brains then are compared to the seemingly healthy standard brain. Through embodiment, that determines the effects of individual experiences and their manifestations within the brain, such "knowledge" contributes to the construction of incorporated and manifested sex/gender differences, both in structure, function and competences. Thereby they are inscribed again into our bodies and then really become scientifically provable facts (Schmitz 2004). This, finally, is the most subtle ethical challenge in regards to the publication of such dichotomies.

References

Easterbrook, Philippa; Berlin, Jesse; Gopalan, Ramana; Matthews, David (1991): "Publication bias in clinical research," Lancet. 1991; pp. 867-872.

Fausto Sterling, Anne: Sexing the Body. Basic Books, NY, 2000

Frost, Julie; Binder, Jeffrey; Springer, Jane; Hammeke, Thomas (1999): Language processing is strongly left lateralized in both sexes: Evidence from functional MRI. Brain 122 (2), pp.199-208.

Geyer, Ch. (Hrsg.): Hirnforschung und Willensfreiheit. Zur Deutung der neuesten Experimente. Frankfurt a.M.: Suhrkamp 2004.

Hennig, Jürgen: Chancen und Probleme bildgebender Verfahren für die Neurologie; in Schinzel (ed.): Interdisziplinäre Informatik: Neue Möglichkeiten und Probleme für die Darstellung komplexer Strukturen am Beispiel neurobiologischen Wissens; Freiburger Universitätsblätter, 3, 2001, Rombach, Freiburg.

Hochhuth, M: Die Bedeutung der neuen Willesfreiheitsdebatte für das Recht, Juristenzeitung JZ, 2005, 745-753.

Kaiser, Anelis; Kuenzli, Esther; Nitsch, Cordula (2004): Does sex/gender influence language processing? Men show bilateralization in a language production task. Poster. Human Brain Mapping Conference 2004, Budapest, Session Language.

Ioannidis, J.P.: Why Most Published Research Findings Are False, PLoS Medicine, http://medicine.plosjournals.org/ August 2005

Narr, K.L., Thompson, P.M., Sharma, T., Moussai, J.; Zoumalan, C.I., Rayman, J., Toga, A.W. (2001): 3D Mapping of Gyral Shape and Cortical Surface Asymmetries in Schizophrenia: Gender Effects. American Journal of Psychiatry 158 (2), pp. 244-255; [http://www.loni.ucla.edu/~thompson/SZ/schizo_atlas.html] (29.06.2004)

Nikoleyczyk, Katrin (2004): NormKörper: Geschlecht und Rasse in biomedizinischen Bildern; in Schmitz, Sigrid; Schinzel, Britta (Hrsg.): Grenzgänge - Genderforschung in Informatik und Naturwissenschaften, Königstein: Ulrike Helmer Verlag, S. 133-148.

Pörksen, U.: Weltmarkt der Bilder. Eine Philosophie der Visiotype; Klett-Cotta, Stuttgart, 1997.

Schinzel, B. (2003): Körperbilder in der Biomedizin; in Frei Gerlach, F., et al (Hrsg.): Körperkonzepte; Münster/New York/München/Berlin: Waxmann; S 245-264.

Schinzel, B. (2004): Digitale Bilder: Körpervisualisierungen durch Bild gebende Verfahren in der Medizin; Coy W. (ed) http://www.inst.at/trans/15Nr/10_1/schinzel15.htm,

Schmitz, Sigrid (2001): Neurowissenschaftliche Informationssysteme. Chancen und Grenzen in Datenmanagement und Wissensrepräsentation, in: Schnitzler (Hrsg.): Interdisziplinäre Informatik. Neue Möglichkeiten und Probleme für die Darstellung komplexer Strukturen am Beispiel neurobiologischen Wissens. Freiburger Universitätsblätter 149 (3), Rombach, Freiburg, S. 51-65.

Schmitz, Sigrid (2003): Neue Körper, neue Normen? Der veränderte Blick durch biomedizinische Körperbilder, in: Weber, Jutta; Bath, Corinna (Hg.): Turbulente Körper und soziale Maschinen. Feministische Studien zur Technowissenschaftskultur. Studien interdisziplinäre Geschlechterforschung 7. Opladen: Leske & Budrich, S. 217-237.

Schmitz, Sigrid (2004): Körperlichkeit in Zeiten der Virtualität; in Schmitz, Schinzel: Grenzgänge - Genderforschung in Informatik und Naturwissenschaften, Königstein: Ulrike Helmer Verlag, S. 118-132.

Shaywitz, Bennet; Shaywitz, Sally; Pugh, Kenneth ; Constable, Todd, Skudlarski, Pawel; Fulbright, Robert et al. (1995): Sex differences in the functional organization of the brain for language. Nature 373, pp. 607-609.

Singer, Hochhut

Spinner, Helmut F.: Die Wissensordnung; VS Verlag für Sozialwissenschaften, 1994.

Thompson PM, Woods RP, Mega MS, Toga AW (2000): Mathematical/Computational Challenges in Creating Deformable and Probabilistic Atlases of the Human Brain. Human Brain Mapping 9, pp. 81-92.

Wacholder, Sh. et. al.(2004): Assessing the Probability That a Positive Report is False: An Approach for Molecular Epidemiology Studies, Journal of the National Cancer Institute.

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

Jens Clausen

Ethische Aspekte von Gehirn-Computer-Schnittstellen in motorischen Neuroprothesen

Abstract:

Title: Ethical Aspects of Brain-Computer Interfacing in Neuronal Motor Prostheses

Brain-Computer interfacing is a highly promising and fast developing field of modern life sciences. Recent advances in neuroscience together with progressing miniaturization in micro systems provide insights in structure and functioning of the human brain and enable connections of technical components to neuronal structures as well. This possibly offers a future therapy for paralysed patients through neuronal motor prostheses. This paper identifies central ethical aspects which have to be considered in further progressing research in this scientific field and the development of neuronal motor prostheses.

Agenda

Author:

Dr. Jens Clausen:

Albert Ludwigs Universität, Lehrstuhl für Bioethik, Stefan-Meier-Straße 26, 79104 Freiburg, Germany
☎ + 49 -761-203 5041, ✉ jens.clausen@uniklinik-freiburg.de,
💻 http://www.natur-des-menschen.uniklinik-freiburg.de/

Relevant publications:

- Biotechnische Innovationen verantworten: Das Beispiel Klonen. Darmstadt. Wissenschaftliche Buchgesellschaft 2006.
- Zur Bedeutung der Natur des Menschen und seiner Verantwortung im Umgang mit Macht und Ohnmacht am Beispiel biotechnischer Innovationen. In: W. Veith & C. Hübenthal (Hrsg.): Macht und Ohnmacht – Konzeptionelle und kontextuelle Erkundungen. Münster: Aschendorff 2005. 204-220.

Gehirn-Computer-Schnittstellen (Brain Computer Interface; BCI) sind ein viel versprechendes und hochinnovatives Forschungsfeld, das einerseits neue Erkenntnisse über Struktur und Funktion des menschlichen Gehirns eröffnet und andererseits die Entwicklung neuer Therapien in Form von motorischen Neuroprothesen möglich erscheinen lässt. In diesem Beitrag sollen die zentralen ethischen Fragen benannt werden, die bei der Erforschung und Entwicklung von BCIs geklärt und berücksichtigt werden müssen.

Einleitung

Motorische Neuroprothesen können als mögliche zukünftige Therapie für Patienten mit Querschnittslähmung angesehen werden. Sind die neuronalen Strukturen für die Übertragung und Weiterleitung der motorischen Signale zwischen dem zentralen Nervensystem, insbesondere dem motorischen Kortex, und den ausführenden Extremitäten durch einen Unfall, eine Verletzung oder eine Krankheit unterbrochen, führt dies häufig zu einer beträchtlichen Einschränkung oder gar dem völligen Verlust der entsprechenden motorischen Fähigkeiten dieses Menschen. Die gegenwärtigen therapeutischen Möglichkeiten zur Behandlung oder gar Heilung von Querschnittslähmungen sind sehr begrenzt. Große Hoffnungen werden daher in die immer weiter voranschreitenden Erkenntnisse der Neurowissenschaften über Struktur und Funktion des menschlichen Gehirns sowie die zunehmende Miniaturisierung technischer, insbesondere mikroelektronischer Bauteile gesetzt. Denn gemeinsam eröffnen sie die Möglichkeit einer Verknüpfung von neuronalen Gehirnstrukturen mit technischen Komponenten über eine Gehirn-Computer-Schnittstelle. Diese Entwicklungen sind für gelähmte Patienten sehr viel versprechend, weil sie eine neue prothetische Therapie möglich erscheinen lassen, die motorischen Neuroprothesen. Die Idee dieses Ansatzes ist, das verletzte neuronale Gewebe, in dem die Übertragung der motorischen Signale unterbrochen ist, durch technische Hilfsmittel zu überbrücken und die verlorenen motorischen Fähigkeiten durch die Ansteuerung einer künstlichen Gliedmaße, wie etwa eines Roboterarms, und als Fernziel möglicherweise sogar der eigenen, natürlichen Extremität wieder herzustellen.

Aufbau einer Gehirn-Computer-Schnittstelle

Um das therapeutische Ziel einer motorischen Neuroprothese realisieren zu können, ist zunächst die Entwicklung eines funktionalen BCI erforderlich, das geeignete neuronale Signale lesen kann, diese in einem zweiten Schritt mittels decodierender Algorithmen interpretiert, um die so gewonnenen Output-Signale zur Ansteuerung einer künstlichen Gliedmaße oder eines anderen Ausgabegerätes (z.B. Computer) zu verwenden.

Die Input-Komponente

Gehirn-Computer Schnittstellen werden seit einiger Zeit eingesetzt, um gelähmten Patienten vor allem solchen mit Locked-in-Syndrom, die Kommunikation zu erleichtern oder überhaupt erst zu ermöglichen. Dazu werden mit einem nicht-invasiven Verfahren mittels EEG Hirnströme abgeleitet und in einen Computer eingespeist. Diese Signale werden dazu verwendet, um einen Computercursor z.B. in einem Buchstabierungsprogramm zu steuern. Die Patienten lernen die zur Steuerung erforderlichen Signale gezielt zu erzeugen und können dann mit Hilfe dieser Gehirn-Computer-Schnittstelle wieder schriftlich in Form von Briefen oder E-Mails kommunizieren. Zwar ist es ein sehr langwieriger Prozess einen solchen Brief zu schreiben, allerdings kann dies gegenüber dem status quo, in dem oft nicht mehr als eine Ja-Nein Kommunikation mittels Augenblinzeln möglich war, dennoch als ein deutlicher Fortschritt angesehen werden[1].

Die Ansteuerung einer motorischen Neuroprothese wird allerdings wahrscheinlich eine Intervention in das menschliche Gehirn erfordern. Denn die mittels EEG zu erhaltenden Informationen sind sehr begrenzt und höchstwahrscheinlich nicht ausreichend für die dreidimensionale Ansteuerung einer artifiziellen Gliedmaße in Echtzeit[2]. Auf der Suche nach dem am besten geeigneten neuronalen Signal für die Kontrolle einer motorischen Neuroprothese werden als Alternative zum EEG stärker invasive Methoden erforscht. Dazu gehören beispielsweise Elektroden, die direkt auf der Kortexoberfläche platziert werden, um neuronale Aktivität in besserer Qualität abzulei-

---

[1] Birbaumer: Nur das Denken bleibt: Neuroethik des Eingeschlossen-Seins.

[2] Nicolelis: Actions from thoughts.

ten[3]. Von diesen auf der Kortexoberfläche applizierten (epicortikalen) Elektroden sind diejenigen Mikroelektroden zu unterscheiden, die direkt in das Gewebe z.B. des motorischen Cortex implantiert werden[4] [5]. Erste Studien, mittels solch intracortikal implantierter Elektrodenarrays beim Menschen motorische Signale abzuleiten, wurden von der Firma Cyberkinetics durchgeführt[6].

Decodierungs-Komponente

Unabhängig davon, welches neurologische Signal mit welcher Ableitungsmethode für die Kontrolle einer motorischen Neuroprothese das am besten geeignete ist, wird es erforderlich sein, die erhaltenen neurologischen Potentiale zu interpretieren. Dazu werden die gewonnenen Informationen an einen als Decodierungs-Komponente fungierenden Computer übertragen. Dort werden aus den Inputsignalen mittels spezieller Algorithmen Prognosen über die antizipierte Bewegung errechnet. Auf der Basis der so berechneten Vorhersage werden dann elektrische Outputsignale generiert, um die prognostizierte Bewegung an der artifiziellen Extremität auszulösen.

Patienten müssen entsprechend lernen, diejenigen neuronalen Signale zu generieren, die tatsächlich die gewünschte Bewegung hervorrufen. Aufgrund der enormen Anpassungsfähigkeit und Plastizität des Gehirns sind Patienten dazu auch wirklich in der Lage, wie die bereits angesprochene Steuerung eines Computercursors oder bei Affen auch die Kontrolle ein Roboterarms belegen.

Output-Komponente

Als mögliche Output-Komponenten der motorischen Neuroprothese zur Behandlung von Gelähmten kommen unterschiedliche technische Geräte in Frage. Am längsten erprobt ist die Ansteuerung eines Computercursors, um beispielsweise ein Buchstabierungsprogramm zu kontrollieren, mit dessen Hilfe die Patienten die Möglichkeit wieder erlangen, in schriftlicher Form zu kommunizieren, Briefe und auch E-Mails zu schreiben. Zu denken ist aber auch an die schon genannten artifiziellen Extremitäten z.B. in Form von Roboterarmen oder die Steuerung des eigenen Rollstuhls. Viele Forschungsprojekte zu Brain-Machine Interfaces sind ursprünglich von der amerikanischen Defense Advanced Research Projects Agency finanziert worden. In diesem Rahmen wurde als mögliches Ziel auch der Einsatz von BCIs zur direkten Ansteuerung militärischer Technik wie beispielsweise eines Kampfjets diskutiert[7].

Ethische Aspekte von Gehirn-Computer-Schnittstellen

Neben der Bewältigung von technischen Herausforderungen sind bei Erforschung und Entwicklung von BCI-basierten motorischen Neuroprothesen eine ganze Reihe ethischer Fragen zu berücksichtigen, die von forschungsethischen Aspekten in Bezug auf die erforderlichen Humanexperimente über neuroethische Fragen bis zur Ethik der Informationstechnologie reichen und damit ein signifikant neues Set ethischer Fragen bilden. Bei der ethischen Beurteilung von BCI-basierten motorischen Neuroprothesen geht es im Sinne einer „innovativen Technikbewertung" (Ropohl) darum, rechtzeitig erwünschte und unerwünschte Entwicklungen und Folgen zu identifizieren, um gegebenenfalls frühzeitig derart korrigierend eingreifen zu können, dass die Entwicklung eine ethisch vertretbare und gesellschaftlich erwünschte Richtung nehmen kann[8]. Die Gliederung der ethischen Fragestellungen orientiert sich an den drei Komponenten einer BCI-basierten motorischen Neuroprothese, der Input-Komponente, der Interpretationskomponente und der Output-Komponente. Abschließend werden dann die generelleren Fragen zur Forschungsethik und den Grenzen einer Technisierung des Menschen behandelt.

---

[3] Ball et al.: Towards an implantable brain-machine interface based on epicortical field potentials.

[4] Donoghue: Connecting cortex to machines: recent advances in brain interfaces; Schwartz: Cortical neural prosthetics.

[5] Schwartz: Cortical neural prosthetics.

[6] Duncan: Hirnimplantate: Fernsteuerung durch Gedanken.

[7] Hoag: Remote control.

[8] Clausen: Biotechnische Innovationen verantworten: Das Beispiel Klonen. 104-129.

Ethische Aspekte bei der Informationsgewinnung

Die invasiven Methoden zur Informationsgewinnung beispielsweise mittels epicortikaler (auf der Cortexoberfläche applizierter) Gitterelektrode oder direkt ins Kortexgewebe implantierter intracortikaler Elektroden bedürfen einer Öffnung des Schädels und letztere zusätzlich noch einer Implantation der Elektroden direkt in das Gewebe des motorischen Cortex oder anderer Hirnareale. Dauerhafte wie vorübergehende Implantationen erfordern ganz offensichtlich eine sorgfältige Nutzen-Risiko-Analyse dieser Interventionen in das menschliche Gehirn. Dabei sind zum einen die medizinischen Risiken des Eingriffs wie beispielsweise Operations- und Infektionsrisiko, mögliche hirnorganische Schädigungen sowie die offene Frage der Verträglichkeit bei Langzeitapplikationen gegen die zu erwartenden therapeutischen Effekte abzuwägen. Insofern allerdings das menschliche Gehirn als biologische Grundlage für zentrale Aspekte unseres menschlichen Selbstverständnisses angesehen werden kann, geht es bei der Nutzen-Risiko-Analyse darüber hinaus aber auch um grundsätzlichere ethische Fragen. In dieser Hinsicht sind mit der Implantation von Elektroden ins menschliche Gehirn als zentraler Punkt mögliche Auswirkungen auf Identität und Individualität des Patienten zu berücksichtigen, da das Gehirn sehr eng mit Bewusstseinsfähigkeit, Persönlichkeit und Identität zusammenhängt.

Medizinische Eingriffe zielen immer auf eine Veränderung des status quo beim Patienten. In diesem Kontext ist daher zum einen zu klären, welche Veränderungen identitätsrelevant sind und zum anderen, welche Veränderungen normatives Gewicht erlangen, und wie diese von den irrelevanten oder zumindest ethisch indifferenten Veränderungen unterschieden werden können. Dieter Birnbacher verwendet dabei die Unterscheidung zwischen zwei Formen von Identität, denen er unterschiedliche ethische Bedeutung zumisst[9]. Die Identität der Person unterscheidet er von der Identität der Persönlichkeit. Dabei spricht er erstere als numerische Identität an, bei der es nur ein Ja oder Nein geben könne, während er die zweite auch als psychologische Identität bezeichnet, bei der Abstufungen und graduelle Veränderungen möglich seien. Die numerische Identität ist eine logische Relation zwischen einem Gegenstand A zum Zeitpunkt p und dem Zeitpunkt p'. Die numerische Identität ist dann gegeben, wenn der Gegenstand A zu den unterschiedlichen Zeitpunkten als ein und derselbe angesprochen werden kann, auch dann wenn er einige Veränderungen erfahren haben mag. Die numerische Identität habe laut Birnbacher besondere ethische Implikationen, da mit einer Veränderung in der numerischen Identität eine Person zu existieren aufhöre. Statt dessen entstehe dafür eine andere Person.

Weniger gravierend seien dagegen Veränderungen in der psychologischen Identität einzuschätzen, da sie als Persönlichkeitsveränderungen derselben Person anzusehen seien. Die psychologische Identität eines Menschen bezieht sich auf konkrete Ausprägungen von Gehirnfunktionen und Charaktereigenschaften, die sich z.B. durch äußere Einflüsse mit der Zeit verändern. Solche Persönlichkeitsveränderungen beeinträchtigen im Regelfall die numerische Identität nicht[10]. Es ist aber durchaus denkbar, dass die Veränderungen bestimmter Ausprägungen der psychologischen Identität so umfangreich sein können, dass damit auch eine Veränderung der numerischen Identität einhergehen würde. Das Verhältnis von psychologischer und numerischer Identität näher zu beleuchten, ist eine philosophisch sehr spannende Aufgabe. Da bei der Verwendung der hier besprochenen motorischen Neuroprothesen allerdings nicht mit einer Veränderung der numerischen Identität zu rechnen ist, haben die Fragen nach möglichen Veränderungen der Persönlichkeit und ihrer ethischen Bedeutung größere praktische Relevanz.

Wenn das Gehirn als materiale Grundlage für die dynamischen Prozesse von Identitätsfindung und -bewahrung im Sinne einer „strukturierten Prozessualität"[11] oder für die Identitätsausbildung im narrativen Vollzug einer Lebenserzählung angesehen wird, scheint es plausibel, dass Eingriffe und Veränderungen in die materiale Grundlage auch die der Identität zugrunde liegenden Prozesse beeinflussen können. Unabhängig davon, welches Identitätsverständnis herangezogen wird, gilt es Kriterien he-

---

[9] Birnbacher: Hirngewebstransplantationen und neurobionische Eingriffe - anthropologische und ethische Fragen. 274f.

[10] Galert: Inwiefern können Eingriffe in das Gehirn die personale Identität bedrohen?.

[11] Beckmann: Menschliche Identität und die Transplantation von Zellen, Geweben und Organen tierischer Herkunft.

rauszuarbeiten, mit deren Hilfe Identitätsveränderungen identifiziert sowie bedenkliche Eingriffe von unbedenklichen unterschieden werden können. Dabei ist insbesondere zu berücksichtigen, dass eine Veränderung der Persönlichkeit zwar ethisch zu berücksichtigen ist, aber kein Ausschlusskriterium darstellt. So können Persönlichkeitsveränderungen akzeptabel erscheinen, wenn sie beispielsweise als Folge eines therapeutischen Eingriffs auftreten, der dem Patienten die Interaktion mit seiner Umwelt wieder ermöglicht[12]. Denn gerade die Interaktion mit anderen kann als ein wesentlicher Bestandteil von Identitätsbildungsprozessen angesehen werden. Für den Einsatz BCI-basierter motorischer Neuroprothesen ist daher zu klären, welche Gehirnfunktionen oder Kombinationen von Funktionen so bedeutsam sind, dass ihre Veränderungen zumindest in bestimmten Ausprägungen zu einer ethisch nicht vertretbaren Identitätsveränderung führen würden. Es ist zu erwarten, dass es sich um relationale Bestimmungen handeln wird, bei denen die ethische Bedeutung von Identitätsveränderungen in Abhängigkeit von den zu erwartenden therapeutischen Ergebnissen beurteilt wird.

Ethische Aspekte bei der Informationsverarbeitung

Die von den Decodierungsalgorithmen der Interpretationseinheit der motorischen Neuroprothese verlangten Signale zur Ansteuerung der Output-Komponente zu generieren, wird die neuronale Aktivität des Patienten verändern. Dies ist einerseits positiv zu bewerten, da es die Voraussetzung dafür ist, dass die Idee zu diesem therapeutischen Ansatz realisierbar ist. Andererseits ist dabei allerdings auch zu fragen, ob die veränderte neuronale Aktivität auch – als unerwünschten Nebeneffekt – spezifische Funktionen des Patientenhirns verändert. Wenn beim Erlernen der gezielten, kontrollierten Steuerung der motorischen Neuroprothese andere Hirnfunktionen verändert werden, ist sorgfältig zu untersuchen, welche Funktionen dies sind, um anschließend zu bestimmen, welche Veränderungen in welchem Ausmaß akzeptabel sind und welche möglicherweise nicht. Die in diesem Zusammenhang zu beantwortenden ethischen Fragen sind denjenigen zur Identitätsproblematik, die im vorherigen Abschnitt behandelt wurden, sehr ähnlich.

Anders gelagert sind dagegen die normativen Implikationen, die mit der Möglichkeit einer falschen Bewegungsprognose durch die Interpretationskomponente verbunden sind. Durch die implizite Prognoseunsicherheit ist es möglich, dass die künstliche Extremität in eine andere als die beabsichtigte Richtung bewegt wird, also z.B. fälschlicher Weise nach rechts statt links. Da sich leicht Szenarien ausmahlen lassen, in denen dies schwerwiegende Konsequenzen auch für Dritte hätte, ist die Frage nach der Verantwortlichkeit für solche Handlungen zu klären. Werden mittels motorischer Neuroprothesen Handlungen ausgeführt, dieauf falschen Interpretationen der abgeleiteten Signale beruhen, ist zunächst fraglich, wer für die resultierenden Folgen verantwortlich ist. Wird die Prothese als integraler Bestandteil des Patienten angesehen, wird man auch bei fehlerhafter Interpretation der Informationen den Patienten für verantwortlich und haftbar halten müssen. Es ist allerdings auch denkbar den Hersteller der Prothese oder den Programmierer des nur bedingt zuverlässigen Algorithmus für verantwortlich für die entsprechende Aktion zu halten. Der technische Imperativ, die Prothese mit allen ihren Komponenten und auch des Decodierungsalgorithmus' so zuverlässig zu gestalten wie möglich, ist selbstverständlich. Daraus resultiert die Aufforderung, den am besten geeigneten Ableitungsort zu identifizieren. Um etwaige kognitive Kontroll- und Vetoinstanzen bei der Signalgenerierung und -modulation durch die Prothese nicht zu umgehen, sondern im Gegenteil gerade so umfangreich wie möglich mit zu berücksichtigen, könnte es erforderlich sein, so „tief" wie möglich im Hirn abzuleiten. Die Elektroden müssten dann an Stellen platziert werden, die hinter diesen Regulierungsmechanismen liegen und nach Möglichkeit nur noch den motorischen Impuls elektrisch kodieren. Allerdings wäre es illusorisch, vollständige Fehlerfreiheit zu fordern. Zu fragen ist daher, wie mit der verbleibenden Unsicherheit umzugehen ist. Mögliche Fehlfunktionen sind allerdings kein Spezifikum von Neuroprothesen, sondern bei technischen Geräten prinzipiell unvermeidbar. Ein etabliertes Verfahren mit möglichen technischen Fehlfunktionen und der damit verbundenen Gefährdung umzugehen, ist die Einführung einer Versicherungspflicht, wie sie bei Kraftfahrzeugen realisiert ist, um drohende negative Auswirkungen soweit wie möglich abfedern zu können. Neben der technischen Optimierung der Geräte, die allerdings Unsicherheit nicht vollständig beseitigen kann, ist zu klären, welche Gefahren vom Einsatz der Prothesen und deren möglichen Fehlfunktionen ausgehen und ob sie als so gravierend eingeschätzt werden, dass eine solche Pflichtversicherung angebracht erscheint. Ist beispielsweise die körperliche

---

[12] Stier: Ethische Probleme in der Neuromedizin: Identität und Autonomie in Forschung, Diagnostik und Therapie.

Integrität oder sogar das Leben Dritter durch den Einsatz BCI-basierter Neuroprothesen bedroht?

Ethische Aspekte bei der Informationsumsetzung

Damit die abgeleiteten und interpretierten bioelektrischen Signale für die Steuerung und Kontrolle einer Output-Komponente verwendet werden können, müssen diese Informationen zum Endgerät übertragen werden. Die wohl komfortablere und auch weniger infektionsanfällige Variante einer kabellosen Übertragung beinhaltet ein gewisses Missbrauchspotential. Mittels einer Fernbedienung mit den entsprechenden Frequenzbereichen könnte es für Fremde möglich sein, die Prothese über externe Impulse ohne Autorisierung anzusteuern, um die resultierende Aktion fälschlicher Weise dem Patienten zuzurechnen. Dass dies dem Respekt vor der Selbstbestimmung des Patienten widerspräche, liegt auf der Hand. Eine technische Lösung, einem solchen möglichen Missbrauch vorzubeugen, könnte sein, bestimmte Frequenzbereiche für Funkverbindungen in medizinischen Anwendungen zu reservieren.

Für die ethische Beurteilung einer militärischen Nutzung von BCIs ist zu klären, in wie weit sich die Wissenschaft zum Komplizen militärischer Zielsetzungen macht, wenn die Forschung aus dem Verteidigungsetat oder von Rüstungsunternehmen finanziert wird. Die Ansteuerung militärischer Geräte verweist allerdings auch auf die im Kontext der Informationsumsetzung zentrale Frage, welche Geräte überhaupt als ausführende Komponente vertretbar sind.

Als ein erstes mögliches Unterscheidungskriterium kann auf unterschiedliche Zielsetzungen rekurriert werden, die mit der artifiziellen Extremität erreicht werden sollen. Die (partielle) Widerherstellung verlorener motorischer Funktionen, erscheint für den Einsatz von BCI-basierten motorischen Neuroprothesen dabei als eine ethisch unstrittige Zielsetzung, die mit dem ärztlichen Auftrag zu heilen und zu helfen abgedeckt wäre. Es ist allerdings zu erwarten, dass der Einsatz der Technik für Ziele, die über die Therapie hinaus weisen und eine Verbesserung der menschlichen Konstitution über ihr Normalmaß hinaus anstreben, sei es durch die Steigerung bereits vorhandener Fähigkeiten, oder durch die Einführung völlig neuer, spezies-untypischer Fähigkeiten, zu einer ähnlich kontroversen Diskussion führen wird, wie dies in anderen Handlungsfeldern des Enhancements bereits der Fall ist. Daher wird auch für die Ausgestaltung von artifiziellen Extremitäten und anderen Ausgabegeräten, die in motorischen Neuroprothesen eingesetzt werden könnten, zu klären sein, wie diese zu konzipieren sind und wo die Grenze zwischen vertretbarem und nicht mehr vertretbarem Einsatz zu ziehen sein wird[13].

Übergeordnete ethische Fragen der motorischen Neuroprothetik

Neben den besprochenen ethischen Aspekten, die sich den einzelnen Komponenten einer BMI-basierten Neuroprothese attribuieren lassen, gilt es zusätzlich auch Fragen zu klären, die sich nicht einer einzelnen Komponente zurechnen lassen, da sie die motorische Neuroprothese als Ganzes betreffen. Dazu gehören insbesondere der Fragenkomplex zur Forschungsethik, der die Voraussetzungen für die Vertretbarkeit einer Erforschung der BMI-basierten motorischen Neuroprothesen behandelt, so wie derjenige zur Grenzziehungsfrage, wie viel Technisierung der Mensch überhaupt verträgt.

Forschungsethik

Zu den zentralen Fragen der Forschungsethik gehört, wer überhaupt als Versuchsperson für die Erforschung dieser medizintechnischen Innovationen in Frage kommen würde. Die Überlegungen zu den komponentenspezifischen Fragestellungen machen sehr deutlich, dass sich dieser innovative Forschungszweig noch mit sehr vielen offenen Fragen konfrontiert sieht. Die in der Grundlagenforschung im Tierexperiment gewonnenen Ergebnisse lassen auf einen möglichen medizinischen Nutzen hoffen, allerdings sind die damit verbundenen Risiken noch nicht hinreichend untersucht. Aufgrund der unklaren Risikolage und den möglichen physiologischen und funktionalen Veränderungen durch den Eingriff ist von einer Erprobung invasiver Verfahren bei gesunden Probanden abzusehen. Eine besondere Problemkonstellation ist bei Patienten mit Locked-in-Syndrom zu berücksichtigen. Dieses Patientenkollektiv gehört zwar zu der Hauptzielgruppe, für die motorische Neuroprothesen überhaupt entwickelt werden, allerdings ist hier die mangelnde Einwilligungsmöglichkeit der ausschlaggebende Aspekt. Durch die vorhandenen kognitiven Fähigkeiten, wird

---

[13] Inwiefern die Natur des Menschen dabei behilflich sein kann vgl. Clausen: Verbesserung der menschlichen Natur zwischen anthropologischer Konstante und Transzendierung des Humanum.

man Locked-in-Patienten kaum als *einwilligungs*unfähig bezeichnen können, da ihnen aber bei entsprechend weit fortgeschrittenem Krankheitsbild die *Kommunikations*fähigkeit fehlt, ist es für diese Patienten unmöglich, eine etwaige Einwilligung oder Ablehnung mitzuteilen. Ein möglicher technischer Ausweg wäre, diesen Patienten mittels nicht-invasiver Ableitungen und der Ansteuerung eines Buchstabierungsprogramms ein gewisses Maß an Kommunikationsfähigkeit zu ermöglichen, um dann auf diesem Wege die Einwilligungsfrage klären zu können. Da auch für diese Maßnahme die Zustimmung des Patienten erforderlich ist, er diese aber nicht selber geben kann, sollten auch nicht-invasive Ableitungen nur dann durchgeführt werden, wenn sie dem mutmaßlichen Willen des Patienten entsprechen. Anschließend ist dann eine direkte Kommunikation mit dem Patienten über bereits erfolgte und etwaige weitere Eingriffe möglich.

Allerdings erscheint die Erprobung auch bei wiederhergestellter Kommunikationsfähigkeit nicht unproblematisch. Wie bei anderen gelähmten Patienten auch ist hier ihre Vulnerabilität besonders zu beachten und bei der Frage nach der Freiwilligkeit einer Entscheidung zu berücksichtigen. Zwar schließen diese Überlegungen eine Erforschung der hier diskutierten motorischen Neuroprothesen nicht grundsätzlich aus, allerdings erscheint ein anderes Patientenkollektiv zum gegenwärtigen Stand der Entwicklung besser geeignet. Zu denken ist hier an Patienten, die aus anderen, beispielsweise diagnostischen Gründen, ohnehin eine Elektrode implantiert bekommen müssen[14]. Diese Patienten scheinen derzeit die am besten geeigneten Versuchspersonen zu sein. Denn ihnen würde durch die Teilnahme an einem Experiment zur Erforschung motorischer Neuroprothesen wohl das geringste zusätzliche Risiko entstehen. Zwar ist die Gruppe dieser Patienten sehr klein, aber für die Klärung der ersten aktuellen Fragen sind diese Patienten – ihre aufgeklärte und freiwillige Zustimmung vorausgesetzt – am besten geeignet. Allerdings ist in diesem Kontext zu klären, ob die in Kapitel „Ethische Aspekte bei der Informationsverarbeitung" angesprochene veränderte neuronale Aktivität und die damit möglicherweise verbundene Veränderung spezifischer Funktionen des Gehirns für diese Patienten ohnehin gilt. Wenn die Veränderung der neuronalen Aktivität bei diagnostischem Einsatz der Elektrode nicht zu beobachten ist sondern nur aus der spezifischen Anwendung in motorischen Prothesen resultiert, ist auch für diese Patienten ein erhöhtes Risiko zu bedenken, über das sie zumindest aufgeklärt werden müssen.

Wie viel Technisierung verträgt der Mensch?

Abschließend sollen zukunftsweisende Überlegungen zu einer möglichen Grenzziehung angestellt werden. Zwar nutzt der Mensch seit jeher technische Hilfsmittel, um sich gegen die aus den Umgebungsbedingungen resultierenden Widrigkeiten zu behaupten[15]. Motorische Neuroprothesen fügen sich in dieser Hinsicht bruchlos in diese Linie technischer (Um-) Weltgestaltung ein. Allerdings stellen sie darüber hinaus auch einen weiteren Schritt hin zu einer unmittelbareren Verbindung zwischen Mensch und Technik dar, in dem die Technik dem Menschen immer näher kommt und sogar im biologischen Gewebe implementiert wird. Die Grenze zwischen Mensch und Technik war bisher scheinbar klar zu ziehen, sie verliert allerdings zunehmend ihre Konturen, wenn Mensch und Technik immer enger ineinander greifen. Die Frage, wie viel Technisierung der Mensch verträgt, thematisiert eine mögliche Grenze für die Technisierung des Menschen, und wie diese gegebenenfalls gefunden werden kann. Für diese Fragestellung gilt es, eine Kriteriologie zu entwickeln, mit deren Hilfe sich die Frage nach einer möglichen Grenzziehung hinsichtlich der Technisierung konkreter adressieren ließe. Dabei ist es unwahrscheinlich, dass die Quantität, die reine Anzahl verwendeter implantierter technischer Geräte dafür ausschlaggebend sein kann. Ob eine Grenzziehung über das Kriterium der Qualität der technisch ersetzten Funktionen aussichtsreicher ist, bedarf weiterer Untersuchungen. Dabei ist ein Rekurs auf unser Konzept, was ein menschliches Wesen ist, und was zur Natur des Menschen gehört ein viel versprechender Zugang für diese Grenzziehungsfrage. Dabei ist insbesondere die zentrale Einsicht der philosophischen Anthropologie des 20. Jahrhunderts (Scheler, Plessner, Gehlen) in den Doppelcharakter[16] des Menschen zu berücksichtigen. Dem Menschen ist seine Natur nicht nur vorgegeben, sondern er ist von ‚Natur' ein Kulturwesen, dem seine Natur gleichzeitig aufgegeben ist. Was diese Einsicht im

---

[14] Ball et al.: Towards an implantable brain-machine interface based on epicortical field potentials. 757.

[15] Clausen: Zur Bedeutung der Natur des Menschen und seiner Verantwortung im Umgang mit Macht und Ohnmacht am Beispiel biotechnischer Innovationen.

[16] Engels: Natur und Menschenbilder in der Bioethik des 20. Jahrhunderts. 19f.

Kontext der Technisierung des Menschen an normativem Gehalt entwickeln kann, ist näher zu untersuchen. Denn die Grenzziehungsfrage scheint ohne den Bezug auf ein normatives Menschenbild und die Konkretisierung derjenigen Aspekte des Menschen, die als werthaft erachtet werden nur schwer beantwortbar zu sein.

Zusammenfassung

Die Entwicklung einer funktionalen Gehirn-Computer-Schnittstelle zur Ansteuerung artifizieller Extremitäten erscheint als ein innovativer Forschungszweig, der große Hoffungen auf die Realisierung einer motorischen Neuroprothese zur Behandlung gelähmter Patienten weckt. Die im Kontext von Erforschung und Entwicklung zu klärenden ethischen Aspekte, verstehen sich keineswegs als grundsätzliche Kritik, sondern sollen vielmehr diejenigen Fragen benennen, die im Zuge des technischen Innovationsprozesses zu berücksichtigen sind, um eine möglichst nachhaltige, ethisch vertretbare Etablierung von BMI-basierten motorischen Neuroprothesen zu ermöglichen. In diesem Sinne erscheint die Entwicklung von Gehirn-Computer-Schnittstellen als ein sehr viel versprechender Ansatz, der mit Blick auf die möglichen Therapieoptionen weiter erforscht werden sollte. Daher sollte die Forschung in diesem Bereich mit Bedacht, unter Berücksichtigung der genannten ethischen Fragen weiter vorangebracht werden.

References

Ball, T. et al.: Towards an implantable brain-machine interface based on epicortical field potentials. Biomedizinische Technik Vol. 49, 2004, 756-759.

Beckmann, J. P.: Menschliche Identität und die Transplantation von Zellen, Geweben und Organen tierischer Herkunft. Jahrbuch für Wissenschaft und Ethik Vol. 5, 2000, 169-182.

Birbaumer, N.: Nur das Denken bleibt: Neuroethik des Eingeschlossen-Seins. In: E.-M. Engels & E. Hildt (Hrsg.) Neurowissenschaften und Menschenbild. Paderborn: Mentis 2005, 77-94.

Birnbacher, D.: Hirngewebstransplantationen und neurobionische Eingriffe - anthropologische und ethische Fragen. In: D. Birnbacher (Hg.) Bioethik zwischen Natur und Interesse. Frankfurt: Suhrkamp 2006, 273-293.

Clausen, J.: Zur Bedeutung der Natur des Menschen und seiner Verantwortung im Umgang mit Macht und Ohnmacht am Beispiel biotechnischer Innovationen. In: W. Veith & C. Hübenthal (Hrsg.) Macht und Ohnmacht: Konzeptionelle und kontexuelle Erkundungen. Münster: Aschendorff 2005, 119-132.

Clausen, J.: Biotechnische Innovationen verantworten: Das Beispiel Klonen. Darmstadt: Wissenschaftliche Buchgesellschaft 2006.

Clausen, J.: Verbesserung der menschlichen Natur zwischen anthropologischer Konstante und Transzendierung des Humanum. In: T. Runkel et al. (Hrsg.) Jenseits der Therapie. 2007.

Donoghue, J. P.: Connecting cortex to machines: recent advances in brain interfaces. Nature Neuroscience Vol. 5 Suppl, 2002, 1085-1088.

Duncan, D. E.: Hirnimplantate: Fernsteuerung durch Gedanken. Technology Review Vol., 2005, 72-78.

Engels, E.-M.: Natur und Menschenbilder in der Bioethik des 20. Jahrhunderts. In E.-M. Engels (Hg.) Biologie und Ethik. Stuttgart: Reclam 1999, 7-42.

Galert, T.: Inwiefern können Eingriffe in das Gehirn die personale Identität bedrohen? In: A. Bora et al. (Hrsg.) Technik in einer fragilen Welt : Die Rolle der Technikfolgenabschätzung. Berlin: Edition Sigma 2005, 91-99.

Hoag, H.: Remote control. Nature Vol. 423, 2003, 796-798.

Nicolelis, M. A. L.: Actions from thoughts. Nature Vol. 409, 2001, 403-407.

Schwartz, A. B.: Cortical neural prosthetics. Annual Review of Neuroscience Vol. 27, 2004, 487-507.

Stier, M.: Ethische Probleme in der Neuromedizin: Identität und Autonomie in Forschung, Diagnostik und Therapie. Frankfurt: Campus 2006.

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

Elisabeth Hildt

Electrodes in the brain: Some anthropological and ethical aspects of deep brain stimulation

Abstract:

In the following text, medical, anthropological and ethical issues of deep brain stimulation, a medical technology in which electrodes implanted in the human brain electrically influence specified brain regions, will be discussed. After a brief account of the deep brain stimulation procedure and its chances and risks, anthropological and ethical aspects of the approach will be discussed. These relate to the reversibility of the procedure and to the patient's capacity to control the effects it exerts in the brain, to modifications and fluctuations in a person's character traits and individuality brought about by neurostimulation, and to the range of legitimate, adequate uses of the deep brain stimulation approach. The paper concludes that deep brain stimulation should be confined to therapeutic contexts and to severe, otherwise treatment-refractory disorders in which the aim is to normalize brain functioning. Apart from this, it should not be used to modify a person's individual character traits and behaviour or to enhance human traits.

Agenda

Author:

PD. Dr. Elisabeth Hildt:
Chair for Ethics in the Life Sciences, University of Tübingen, Wilhelmstr. 19, 72074 Tübingen, Germany
☎ + 49 – (0)7071 - 2977192 , ✉ elisabeth.hildt@uni-tuebingen.de , 💻 http://www.uni-tuebingen.de/bioethik/index.htm

During the past few years, our knowledge of the structure and function of the human brain has increased considerably, and this has prompted the development and use of new therapeutic strategies, among them technological procedures such as Information and Communication Technology implants (ICT implants) for the human brain.[1] These include direct brain computer interfaces (BCIs) which allow brain signals to be used for communication and control of movement. Apart from non-invasive approaches such as the so-called thought translation device, which is a biofeedback communication system used by locked-in patients, invasive brain computer interfaces are currently being developed which aim at enabling the brain to exert direct motor control. In animal experiments using microelectrode arrays implanted in the motor cortex, it has been possible to control movements merely by "thinking them through." A related research strategy aims at developing implantable brain chips called "artificial hippocampus" that in several years' time might serve to restore memory in patients with Alzheimer's disease or stroke and that may even be used to enhance memory in "normal persons."

Another field which has drawn enormous clinical interest of late is the use of deep brain stimulation techniques in which electrodes implanted in the human brain electrically influence specified brain regions. Until now, this promising technology has mainly been used in the treatment of Parkinson's disease and other movement disorders. Apart from its successful application as a therapeutic approach to treating Parkinson's disease, an intense and controversial discussion is currently being conducted as to whether or not neurostimulation should be used in particular cases of severe and treatment-refractory neuropsychiatric disorders such as obsessive-compulsive disorder (OCD).

The deep brain stimulation (DBS) procedure raises enormous hopes for individual patients suffering from severe diseases. However, the approach also involves considerable risks, and in light of the fact that it exerts a direct technological influence on the human brain, this being the organ which determines, like no other, a person's entire physical and mental existence, it poses complex anthropological and ethical questions. Concerning the individual patient, aspects to be discussed include possible benefits and risks of the procedure, especially issues that relate to claims that it is reversible and that the consequences of DBS can be controlled. Possible modifications in a patient's character traits brought about as a result of the procedure are also an issue of critical inquiry. More general issues subject to anthropological and ethical reflection relate to implications which the direct interplay between man, brain and technical devices might have for our conceptions of human identity and authenticity.

In the following, medical, anthropological and ethical issues of neurostimulation will be discussed, starting with a brief account of the deep brain stimulation procedure.

Deep brain stimulation – the procedure

Until now, deep brain stimulation has mainly been used in patients with idiopathic Parkinson's disease, a slowly progressing neurodegenerative disorder characterized by resting tremor, bradykinesia, muscular rigidity and gait disturbances. Neurostimulation aims at providing symptomatic relief in patients with Parkinson's disease and other movement disorders who are medically refractory to standard treatments.[2]

In this new and promising technology, an electrode lead with four electrode contacts is implanted bilaterally in the subthalamic nucleus under stereotactic conditions. The electrode lead is connected via wires which run subcutaneously to a battery-driven, programmable pulse generator placed below the clavicle. The implantation of the electrode is performed while the patient is awake to allow testing for optimal electrode placement and exclusion of adverse effects using microelectrode stimulation. After the implantation has been completed, the pulse generator is programmed to adjust the electrode system to the individual patient's needs by specifying which electrode contacts are to be used and by optimizing the various stimulation parameters such as pulse width, frequency and amplitude of stimulation. The DBS approach calls for intense post-operative care which involves multiple sessions for device adjustments and medication management.[3]

---

[1] cf. Hildt 2005

[2] Okun & Foote 2004; Rosenow et al. 2004; Benabid et al. 2005; Gharabaghi et al. 2005

[3] Andrews 2003

By effecting chronic bilateral stimulation of the subthalamic nucleus in patients with Parkinson's disease, considerable improvement of motor symptoms and a reduction in medication dosages can be achieved. Best results have been obtained in patients with idiopathic Parkinson's disease who show no cognitive difficulties and who have preserved a good response to L-Dopa but manifest side-effects due to long-term L-Dopa treatment.[4] The scientific community has no good understanding of the mechanism which underlies the positive effect of DBS in Parkinson's disease, however.[5]

The surgery also involves significant risks. These include intracranial hemorrhage in cases where blood vessels are punctured by microelectrodes or the DBS lead, which may cause strokes or stroke-like symptoms, infection, various neurological problems, depression, memory problems and psychiatric and/or behavioural symptoms. Other problems can arise if the lead migrates or an electrode, connecting wire or impulse generator needs to be replaced.[6]

In addition to excellent motor benefits brought about by DBS in patients with Parkinson's disease, various negative neuropsychological consequences of the procedure have been reported. Although DBS usually involves no global cognitive or intellectual deterioration, mild cognitive impairments such as diminished verbal fluency and/or behavioural changes that seem to be comparatively rare and mostly transient have been observed. In individual cases, pronounced modifications in mood and behaviour have been reported, however.[7] In particular, elderly patients with advanced Parkinson's disease are at risk for cognitive impairment. Neuropsychological changes observed in these patients include significant declines in working memory, speed of mental processing, bimanual motor speed and co-ordination, set switching, and declines in verbal learning, visual learning and memory. In addition, changes in mood and personality, especially frontal-lobe-related personality changes, have been observed.[8] At present it is not known whether the neuropsychological effects of DBS are reversible or not.[9]

Several clinical studies are currently investigating whether DBS might have a positive effect on treatment-refractory cases of neuropsychiatric disorders such as obsessive-compulsive disorder.[10] Patients with obsessive-compulsive disorder (OCD) are tormented by obsessions which dominate their lives and impose inflexible rules on them. Although hypotheses have been made concerning dysfunctions in neural circuits connecting the frontal lobe and the basal ganglia (fronto-striatal-pallidal-thalamic-frontal loops), the neurobiological basis of OCD is not well understood.

When used for OCD, DBS entails stimulation of the anterior limbs of the internal capsules or the ventral caudate nucleus. Until now, this strategy has been used in a very limited number of patients with severe, chronic, treatment-refractory OCD. In most of the patients studied so far, a positive effect of DBS on psychopathology and neuropsychological functioning has been reported. In most of the studies, the procedure has evidenced no deleterious impact or any major adverse or harmful consequences on the personality of the patients involved.[11] In one case, a patient is reported to have suffered from DBS-induced panic and anxiety.[12]

Other uses of the neurostimulation technology include hypothalamic deep brain stimulation as a treatment for intractable chronic cluster headache and vagus nerve stimulation in epileptic patients with treatment-refractory partial-onset seizures.[13] In addition, vagus nerve stimulation is currently under investigation in patients with major depressive

---

[4] Saint-Cyr et al. 2000; Okun & Foote 2004; Funkiewiez et al. 2004; Rosenow et al. 2004; Benabid et al. 2005

[5] Garcia et al. 2005

[6] Andrews 2003; Umemura et al. 2003; Okun & Foote 2004; Rosenow et al. 2004

[7] Funkiewiez et al. 2004

[8] Saint-Cyr et al. 2000; Funkiewiez et al. 2004

[9] Saint-Cyr et al. 2000; Bejerot 2003

[10] Gabriëls et al. 2003; Aquizerate et al. 2004; Rauch et al. 2006

[11] Gabriëls et al. 2003; Aquizerate et al. 2004; Cosgrove 2004

[12] Shapira et al. 2006

[13] Andrews 2003; Schoenen et al. 2005

episodes[14] in order to find out whether there is an anti-depressant effect of vagus nerve stimulation.

Direct interaction of man, brain and machine – some anthropological and ethical aspects

Without a doubt, computers and other technical devices are part and parcel of modern everyday life. Technical instruments also constitute a crucial dimension of modern medicine. Recent developments in neurotechnology which involve direct interactions between man and machine clearly have to be discussed against this background.

As opposed to the customary, traditional use of tools and technical instruments such as drills, computers and MP3 players and in contrast to external medical devices such as eyeglasses and hearing aids, neurotechnological instruments used for neurostimulation and other purposes are technical tools which enter the human body and affect it directly. In deep brain stimulation, an electrode is implanted in a specified region of the human brain, with subcutaneous wires connecting the electrode to an impulse generator implanted below the clavicle. The medical use of this technology has been facilitated by the availability of small high-performance computers and other small, biocompatible devices.

Although analogies can undoubtedly be drawn between the normal use of external tools, computers and other technical instruments and the use of "tools" incorporated into the human body, a crucial distinction must be made here: whenever a “tool” is integrated into the human body – especially into the part of the body which is so central to an individual's personality and overall existence, i.e., the brain – the distinction normally drawn between tools and the subject who use them is blurred. What is particularly important in this context is the fact that in contrast to modifications of other parts of the human body, modifications of the brain effected by direct integration of technical tools might subject the persons involved to alterations of individual traits.

In the following, implications of DBS resulting from the direct integration of technical tools into the human brain will be discussed.

What about the capacity to control the effects of neurostimulation?

An aspect of enormous importance relates to the question as to whether and/or to what degree it is possible to control the effects of electrodes on a patient’s brain, i.e. as to what influences electrodes have on a patient’s physical and mental traits. In this context, issues concerning the reversibility or irreversibility of the effects brought about by the treatment play an important role, as does the individual’s ability to exercise control over these influences.

Clearly, DBS is a form of therapy less reversible than pharmacological forms of treatment in which adverse or unforeseen effects of medication can be modified without undue hardship – although even pharmacological treatment modalities may not be totally reversible on a molecular level. Nevertheless, DBS is quite a flexible procedure compared to classical surgical approaches. Unlike traditional psychosurgery which aims at leasioning a specified brain region[15], neurostimulation only aims at exerting an electrical influence on the brain.

In neurostimulation, it is principally possible to adjust the electrode system to the individual patient. By programming the pulse generator, it is possible to modify the stimulation protocol and to adjust stimulation in accordance with the individual needs of the patient, the course of disease or other necessary modifications. One should not ignore the fact, however, that adjusting such devices is a complex, time-consuming procedure with manifold limitations. It may not be possible to achieve optimal therapeutic success, electrode stimulation may cause certain negative effects, etc. To be sure, in cases of severe and long-lasting problems, one does have the option of switching off the electrode or explanting the whole system.

Compared to less invasive approaches, neurostimulation is clearly not a completely reversible procedure, however. DBS involves stereotactic surgery, which entails structural modifications of the brain – albeit limited ones – caused by macroelectrode implantation as well as manifold risks for brain lesions. Currently no one knows how far neurostimulation involves adverse, irreversible, long-term effects resulting from chronic stimulation. In particular, it remains to be investigated whether the

---

[14] Nahas et al. 2005

[15] cf. Bejerot 2003

negative neuropsychological effects observed after DBS are reversible or not.[16]

Adverse effects, especially if they are irreversible, may not only modify a person's individuality, but also significantly reduce her well-being and autonomy. In sum, a person using DBS is strongly dependent on the electrode system. In order to allow the patient to control his overall situation as far as possible, it is necessary to have a specialist team available 24 hours a day and 7 days a week which is able to adjust, repair and (re)programme the system if problems, malfunction or deviations arise. The person's quality of life, her motor performance, her general mood and her overall psychical situation may depend to a high degree on how well the technical device functions. Disturbances, problems and malfunctioning of the electrode system have a direct impact on the individual.

Modifications and fluctuations in personality traits

In evaluating various neurostimulation procedures, the possibility that they might effect modifications in a person's behaviour, mental characteristics and personality plays an important role. On the one hand, psychomodulations or modifications in behaviour may be inherent to the therapeutic procedure; on the other hand, they may develop unintentionally in the form of unwanted or unexpected side-effects. Such psychomodulations pose manifold medical, philosophical and ethical questions as regards the personality, personal identity, individual responsibility, autonomy, authenticity and self-perception of the person involved.[17] They also have complex implications for a person's individual, familial and social life.

Unlike other forms of therapy, deep brain stimulation involves direct modifications of a person's characteristics which can be brought about almost instantaneously. The person's "mode of existence" might be dependent on whether the electrode is activated or not. The following examples serve to illustrate this capacity of DBS to abruptly alter personality traits. In their publication on DBS in patients with OCD, Gabriëls et al.[18] describe the case of a woman who showed positive results after successful DBS. They state that "when stimulation was switched off she almost instantaneously became anxious and in great distress again". Funkiewiez et al. report on a woman with Parkinson's disease characterized preoperatively by fluctuations in mood which manifested themselves in concurrence with her motor symptoms. After deep brain stimulation improved the symptoms significantly, the authors described the patient's situation as follows:[19] "When stimulation was deactivated for motor evaluation, the patient immediately felt overwhelming sadness, dissolved in tears, and experienced a progressive reappearance of severe parkinsonism. When describing the impact of treatment, she said: 'If stimulation is switched off, I am dead; when on, I am alive. With medication on top, it's happiness and well being.' In the case of a woman with significant positive results after DBS for OCD, her husband stated:[20] 'When stimulation is switched ON, she changes and improves suddenly, but we (he and her children) don't have a similar button that can be pressed to adapt at once to this new person.'"

Undoubtedly these cases all show very good therapeutic success; they all involve considerable and highly welcome improvement of symptoms brought about by DBS. On the other hand, they underline the intensive and direct interaction between the patient and the electrode system and the extreme situation the patient finds himself in, being enormously dependent on stimulation. Motor performance, mood and personality may be perceived and interpreted as relying highly on the technical system, i.e. on an external component which exerts direct influence on the person's brain. This may lead to the impression of being – in one's essence – manipulated by the electrode system.

In light of such fluctuations, questions concerning personal identity and the individuality of the patients involved may arise. In which condition is the person really "herself," the authentic person she considers herself to be? When the electrode is switched on or when it is switched off? Manifold difficulties in adapting to the different "modes of existence" may arise. Not only the patient involved but also the whole family has to cope with these abrupt changes in character traits that may occur when the elec-

---

[16] Saint-Cyr et al. 2000, Bejerot 2003

[17] cf. Hildt 1999

[18] Gabriëls et al. 2003, 279

[19] Funkiewiez et al. 2004, 837

[20] Gabriëls et al. 2003, 278

trode is switched on or off or when there are variations in the stimulation protocol.

Beneficence, nonmaleficence, and autonomy

Deep brain stimulation is a promising technology that in principle can be used in manifold contexts. But which uses of the procedure can be considered adequate ones? In order to find an answer to that question, let me argue on the basis of the principle-based approach developed by Tom Beauchamp and James Childress.[21]

With regard to beneficence and nonmaleficence, the risk-benefit-ratio of the procedure is crucial. As discussed above, the surgical procedure bears considerable risks. Only those uses in which a therapeutic benefit might reasonably be expected and in which the (expected) benefits clearly outweigh the risks can be considered acceptable. That's why it should be confined to severe chronic disorders the pathophysiological basis of which is known. Otherwise, the hope for therapeutic benefits might prove to be pure speculation. In general, before DBS is chosen as a treatment by an individual patient, all other less invasive treatment options should have been taken into consideration.

Apart from these medical aspects: Is a person harmed by the incorporation of technical devices into the brain? Manifold concerns, fears and fantasies stem from the option to integrate computers and other technical instruments into the brain and other parts of the human body. These involve aspects such as the technicalization of the human body, the encouragement of a reductionist, technological view on human beings, the fear of losing human identity, and speculations relating to cyborgs (cybernetic organisms).[22]

Undoubtedly, all these concerns – which unequivocally reveal that there are ethical limits to the amount and range of human body parts to be substituted by technical devices – need broad and intense interdisciplinary discussion. However, the decisive criterion in DBS is not the fact that a technical device, i.e. an electrode, is being implanted into the human brain, but the question of whether the device is able to perform more or less the same function as the formerly intact, now degenerated or otherwise dysfunctional neurons do. From the point of view of functionalism, electrode stimulation merely serves to normalize brain function.

This points at another aspect of central relevance: the respect for autonomy of the person involved. In part, concerns and fears related to neurostimulation clearly originate from the perceived power of this technology to modify a person's character traits and individuality – and to do so even against that person's will. Also science-fiction ideas of persons being externally controlled by other persons or by computers may be seen against this background. So, for autonomy reasons neither the implantation of the electrode device nor any modification in the stimulation protocol should be done without having obtained the patient's free and informed consent. DBS should be confined to therapeutic contexts and to severe, otherwise treatment-refractory disorders in which the aim is to normalize brain functioning. Apart from this, it should not be used to modify a person's individual character traits and behaviour or to enhance human traits.

Conclusion

Deep brain stimulation is a new and promising approach. Especially in patients with Parkinson's disease who had been previously treatment-refractory, significant improvements in symptoms have been brought about.

There are several problematic issues that call for further discussion, however. First of all, a detailed assessment of the clinical benefits and the adverse effects of this invasive procedure, especially on a long-term basis, is required. With regard to the individual patient involved, it is absolutely necessary to carefully analyze his situation, to evaluate the chances and risks of the DBS procedure and to seek the patient's free and informed consent. In neuropsychiatric disorders where judgement is impaired, manifold problems arise when it comes to attempting to obtain the patient's free and informed consent to this surgical procedure.[23] Before undergoing DBS, all other less invasive options available should have been taken into consideration. This is especially true as concerns disorders such as OCD. As far as this is concerned, DBS is at present highly investigational and there is currently enormous uncertainty as to possible benefits and harms.

---

[21] Beauchamp & Childress 2001

[22] Hildt 2005

[23] cf. Comité Consultatif National d'Ethique (CCNE) 2003

In view of the considerable risks and the inherent capacity of this technology to modify a person's character traits, the use of DBS should be restricted to severe disorders whose pathophysiological basis is well known and which are otherwise treatment-refractory.

References

Andrews, R.J. (2003): Neuromodulation I. Techniques – Deep Brain Stimulation, Vagus Nerve Stimulation, and Transcranial Magnetic Stimulation, Ann. N.Y. Acad. Sci. 993: 1-13.

Aquizerate, B. et al. (2004): Deep brain stimulation of the ventral caudate nucleus in the treatment of obsessive-compulsive disorder and major depression, J Neurosurg 101, 682-686.

Beauchamp, T.L. & Childress, J.F. (2001): Principles of Biomedical Ethics, Oxford University Press: New York, Fifth Edition.

Bejerot, S. (2003): Editorial: Psychosurgery for obsessive-compulsive disorder – concerns remain, Acta Psychiatr Scand 107, 241-243.

Benabid, A.L. et al. (2005): Deep-brain stimulation in Parkinson's disease: long-term efficacy and safety – What happened this year?, Current Opinion in Neurology 18, 623-630.

Comité Consultatif National d'Ethique (CCNE) (2003): Functional Neurosurgery for Severe Psychiatric Disorders, Jahrbuch für Wissenschaft und Ethik 8, 531-547.

Cosgrove, G.R. (2004): Editorial: Deep brain stimulation and psychosurgery, J Neurosurg 101, 574-576.

Funkiewiez, A. et al. (2004): Long term effects of bilateral subthalamic nucleus stimulation on cognitive function, mood, and behaviour in Parkinson's disease, J Neurol Neurosurg Psychiatry 75, 834-839.

Gabriëls, L. et al. (2003): Deep brain stimulation for treatment-refractory obsessive-compulsive disorder: psychopathological and neuropsychological outcome in three cases, Acta Psychiatr Scand 107, 275-282.

Garcia, L. et al. (2005): High-frequency stimulation in Parkinson's disease: more or less?, Trends in Neurosciences 28(4), 209-216.

Gharabaghi, A. et al. (2005): Wiederherstellung der Funktion: Modulation von Hirnfunktionen durch Neuroprothesen, in: Engels, E.-M. & Hildt, E. (eds): Neurowissenschaften und Menschenbild, mentis: Paderborn, 57-75.

Hildt, E.: Ethical aspects of neural tissue transplantation, Croatian Medical Journal, Vol. 40, No. 3, 1999, 326-333.

Hildt, E.: Computer, Körper und Gehirn: Ethische Aspekte eines Wechselspiels, in Engels, E.-M. & Hildt, E. (eds.): Neurowissenschaften und Menschenbild, mentis: Paderborn, 2005, 121-137.

Nahas, Z. et al. (2005): Two-Year Outcome of Vagus Nerve Stimulation (VNS) for Treatment of Major Depressive Episodes, J Clin Psychiatry 66, 1097-1104.

Okun, M.S. & Foote, K.D. (2004): Deep Brain Stimulation – The 10 Key Issues Every Practitioner, Patient and Family Member should be Familiar With, Business Briefing: North American Pharmacotherapy 2, 1-6, 2204.

Rauch, S.L. et al. (2006): A functional neuroimaging investigation of deep brain stimulation in patients with obsessive-compulsive disorder, J Neurosurg 104, 558-565.

Rosenow, J.M. et al. (2004): Deep brain stimulation for movement disorders, Neurological Research 26, 9-20.

Schoenen, J. et al. (2005): Hypothalamic stimulation in chronic cluster headache: a pilot study of efficacy and mode of action, Brain 128(4), 940-947.

Shapira, N.A. et al. (2006): Panic and fear induced by deep brain stimulation, Journal of Neurology, Neurosurgery and Psychiatry 77, 410-412.

Umemura, A. et al. (2003): Deep brain stimulation for movement disorders: morbidity and mortality in 109 patients, J Neurosurg 98, 779-784.

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

Michael Nagenborg and Mahha El-Faddagh

Genetische Informationen: Eigentumsansprüche und Verfügbarkeit

Abstract:

Titel: Genetic Information: Privacy and Availability

The use of genetic information about a patient may cause serious concern within the discourse on informational privacy. In our article we would like to discuss a positive example of a diagnostic use of genetic information in the field of molecular genetics. With regard to this example we will discuss the question who owns the genetic information to determine who should decide which data is to be stored or deleted. We will use a Kantian concept of property in order to show that the genetic information in the example given is to be considered the property of the patient. We shall argue, that the information should be considered as a part of the medical sphere, which is to be informationally sealed. Although we present hereby a theoretical framework for a design of an appropriate information infrastructure, we will finally point out to the high costs of such an infrastructure.

Agenda:

Authors:

Dr. phil. Michael Nagenborg:

Universität Karlsruhe, Institut für Philosophie, Geb. 20.12, 76128 Karlsruhe, Germany.

Telephone, email and personal homepage: ☎ +49 721 3545955, ✉ philosophie@michaelnagenborg.de, 💻 www.michaelnagenborg.de

Relevant publications:

- Privatheit unter den Rahmenbedingungen der IuK-Technologie. Wiesbaden 2005.
- Datenschutz und der Verlust der Bedeutungslosigkeit. S. 61-71 in: Martin Woesler, Hrsg.: Ethik in der Informationsgesellschaft. Bochum 2005.

Dr. med. Mahha El-Faddagh:

Städtisches Klinikum Karlsruhe, Klinik für Kinder- und Jugendpsychiatrie und Psychotherapie, Institutsambulanz, Moltkestr. 90, 76133 Karlsruhe, Germany

☎ +49 721 974-3903, ✉ Mahha.El-Faddagh@klinikum-karlsruhe.com

Relevant publications:

- M. El-Faddagh, M. Laucht, A. Maras, L. Vöhringer, M.H. Schmidt (2004): Association of dopamine D4 receptor (DRD4) gene with attention-deficit/hyperactivity disorder (ADHD) in a high-risk community sample: a longitudinal study from birth to 11 years of age. Journal of neural transmission 111(7): 883-889.
- M. El-Faddagh, P. Schloss, G. Ristow, K. Becker, M. Holtmann, M.H. Schmidt (2003): Untersuchung zur Nonresponse auf Methylphenidat bei Kindern mit hyperkinetischer Störung. In: U. Lehmkuhl (Hg.) „Seelische Krankheit im Kindes- und Jugendalter- Wege zur Heilung", DGKJPP-Kongress 2002, Vandenhoeck & Ruprecht, Göttingen.

Einleitung

Angesichts der zunehmenden Verfügbarkeit von genetischen Informationen, worunter wir im Folgenden Aussagen über die genetische Anlage von Individuen verstehen wollen, wurde bereits das „Ende der Privatheit" diagnostiziert.[1] Dabei wird in vielen Fällen von einer umfangreichen Kenntnis der genetischen Disposition eines Individuums ausgegangen. In der Praxis kann sich das Interesse jedoch auf einen kleinen Ausschnitt beschränken.

In unserem Beispiel geht es um den Genotyp von Rezeptoren an den Nervenzellen im Gehirn des Patienten. Wenn wir im Folgenden auch abkürzend den Ausdruck „Genotyp" verwenden, so ist hiermit stets nur die Bestimmung des Genotyps eines Rezeptors gemeint. Mit „genetischer Information" bezeichnen wir das Ergebnis eines Testverfahrens mit dem dieser Genotyp bestimmt wird,[2] indem z. B. das Vorhandensein einer bestimmten Basenkette überprüft wird.[3]

Wir werden das Beispiel vor dem Hintergrund der Sphärenauffassung der informationellen Gerechtigkeit diskutieren, die von M. J. van den Hoven (1999) entwickelt und in letzter Zeit von M. Nissenbaum (2004) aufgegriffen wurde. Diesen Ansatz werden wir mit einer Kantischen Konzeption von Eigentum verbinden.[4]

Wir wollen zeigen, dass die Informationsethik durchaus Konzepte anzubieten hat, welche es ermöglichen, geeignete Rahmenbedingungen für die Verwendung von genetischen Informationen zu formulieren. Entsprechend gestaltete Informations- und Kommunikationstechnologien könnten sogar dazu beitragen, zur Zeit bestehende Interessenskonflikte zu lösen. Allerdings möchten wir auch darauf hinweisen, dass die angemessene Gestaltung dieser Technologien sowie der notwendigen juristischen Rahmenbedingungen sehr aufwendig ist. Insofern sollten potentielle Kosteneinsparungen, welche durch die Verwendung von genetischen Informationen bei der Behandlung von Patienten möglich sind, nicht der einzige Grund sein, um den hier beschriebenen Verwendungsformen zuzustimmen, da dem Potential zur Kostensenkung im Einzelfall ein hoher Aufwand bei der Gestaltung der (technischen) Infrastruktur gegenübersteht.

Unser Fallbeispiel

Es gibt psychiatrische Erkrankungen, für deren Behandlung verschiedene Medikamente zur Auswahl stehen. Diese wirken jedoch nicht bei jedem Patienten gleich gut, zeigen z. T. sogar keinerlei Wirkung (Problem der „Non-Responder").

Deshalb werden Patienten z. T. ambulant oder auch für eine längere Zeit innerhalb der Klinik mit verschiedenen Präparaten nacheinander behandelt, um ein wirksames Medikament zu finden. Unter der Annahme, dass die Wirksamkeit u. a. von dem Genotyp abhängig ist, könnte der Klinikaufenthalt unter Verwendung genetischer Informationen erheblich verkürzt werden. Studien zur Wirksamkeit von Medikamenten bei verschiedenen Ausprägungen der Genotypen liegen z. B. für Depressionen vor.[5]

Dies erscheint uns sowohl hinsichtlich des Leidens unter solchen Erkrankungen, der Stigmatisierung des Psychiatrieaufenthalts als auch der Kosteneinsparung als wünschenswert.

---

[1] So z. B. der programmatische Titel von Bruen (1999).

[2] Ähnlich: Spinello (2006: 213).

[3] Zwar setzt die Durchführung eines solchen Verfahrens eine Gewebeprobe des Patienten voraus, jedoch liegt das Hauptaugenmerk in diesem Artikel auf dem Testergebnis.

[4] Zur Kritik an dem Schutz von Privatheit durch Eigentumsansprüche im Allgemeinen vgl. Nagenborg (2005: 163ff.). - Mit der Wahl des kantischen Eigentumsbegriffs wollen wir nicht behaupten, dass diese spezielle Konzeption von Eigentum die einzig mögliche und für unseren Gegenstand die beste aller Optionen sei. Da aber eine differenzierte Auseinandersetzung mit den verschiedenen Eigentumsbegriffen den Rahmen dieses Beitrages sprengen würde, haben wir uns dazu entschieden, durch die explizite Bezugnahme auf eine bestimmte Position zumindest unsere Prämissen deutlich zu machen, um nicht den Eindruck zu erwecken, dass ein gemeinsames Verständnis vorausgesetzt werden kann.

[5] Siehe: Rau (2004), Peters (2004).

Die Produktion, Verwendung und Speicherung von genetischen Informationen

Die Bestimmung des Genotyps wird im Folgenden betrachtet als eine Methode der Informationsproduktion unter bestimmten Rahmenbedingungen, die sich wie folgt beschreiben läßt:

(1) Die Informationen werden durch eine medizinisch-therapeutische Institution, zu einem therapeutischen und/oder wissenschaftlichen Zweck

(2) mittels eines technischen Verfahrens erzeugt,

(3) dessen Entwicklung und

(4) Anwendung Kosten verursacht, wobei die

(5) Informationen notwendigerweise zumindest temporär gespeichert werden und

(6) bestimmte Personen Zugriff auf diese Informationen haben müssen.

Zu (1): Die Information wird also innerhalb der medizinischen Sphäre erzeugt, wobei medizinisch-therapeutische Institution hierbei zunächst Krankenhäuser meint, aber auch Arztpraxen umfassen mag. Der Begriff der medizinischen Sphäre wird im folgenden Abschnitt näher erläutert.

Zu (2), (3) und (4): Die Kosten für die Durchführung des Tests sind deswegen zu betonen, da hierdurch die produzierten Informationen einen bestimmbaren, monetären Wert erhalten.[6] Dies bedeutet insbesondere, dass dem Kostenträger ein Schaden entsteht, wenn z. B. die Information gelöscht wird und neu produziert werden muss.

Zu (5): Das Ergebnis des Testverfahrens muss in irgendeiner Form dem behandelnden Arzt mitgeteilt werden, sofern dieser den Test nicht selbst durchführt. Wir gehen hier davon aus, dass die notwendige, temporäre Speicherung elektronisch erfolgt.

Zu (6): Aus dem zu (5) Gesagten folgt dann, dass zumindest der behandelnde Arzt Zugriff auf diese Information haben muß. Wie wir später sehen werden, mag aber auch eine Berechtigung bestehen, auf die Informationen zu Forschungszwecken zuzugreifen.

Der Fluss von Informationen innerhalb der medizinischen Sphäre

M. Walzer hat sein Modell der „Sphären der Gerechtigkeit" (1992, Original 1983) nicht zuletzt als Gegenentwurf zu J. Rawls „Theorie der Gerechtigkeit" (1971) konzipiert, an der er die Vorstellung eines dominanten Gutes kritisiert. Er plädiert statt dessen dafür, zwischen verschiedenen, gesellschaftlichen Sphären zu unterscheiden, wobei Güter ihren jeweiligen Wert innerhalb dieser Sphären erhalten. Eine Ungleichverteilung der Güter innerhalb einer Sphäre wird dabei als unproblematisch empfunden. Allerdings sollen Vorteile innerhalb einer Sphäre nicht zu Vorteilen oder Nachteilen in einer anderen führen. Zudem sollen bestimmte Tauschgeschäfte blockiert werden. M. J. van den Hoven (1999) überträgt dieses Konzept auf Informationen, die innerhalb einer Sphäre ausgetauscht und weitergegeben werden dürfen, wobei H. Nissenbaum (2004) ergänzt, dass auch der Fluss von Informationen innerhalb einer Sphäre Normen unterworfen ist.[7] Die Sphären werden zudem als getrennt (informationell versiegelt) gedacht und der Austausch von bestimmten Informationen zwischen den Sphären ist zu blockieren.[8]

Legitim produzierte genetische Informationen dürfen also innerhalb der medizinischen Sphäre weitergegeben werden, sofern dies aus therapeutischen Gründen oder zu Forschungszwecken geschieht. Der Zugriff auf diese Informationen durch potentielle Arbeitgeber im Rahmen eines Bewerbungsverfahrens also z. B. als illegitim erachtet, da dadurch die Information die Sphäre der Medizin verlässt.[9]

Problematisch an dem Konzept ist die Bestimmung der Sphären, sowohl hinsichtlich ihrer Anzahl als auch ihrer Grenzen. Nissenbaum wählt den pragma-

---

[6] Der Marktwert der Information ist dabei von dem Wert zu unterscheiden, den die Information für das Wohlergehen des Patienten oder für die Forschung hat.

[7] Nissenbaum (2004: 137).

[8] Vgl. Nagenborg (2005: 168ff.) sowie Weber (2005: 228ff.).

[9] Vgl. hierzu z. B. Opinion of the European Group on Ethics in Science and New Technologies to the European Commission, Nr. 18: Ethical Aspects of Genetic Testing in the Workplace (2003).

tischen Ausweg, indem sie betont, dass von der Existenz von unterschiedlichen Sphären mit unterschiedlichen Normen deswegen auszugehen sei, weil wir bestimmte Handlungen in bestimmten Zusammenhängen als Verstöße gegen die jeweiligen Normen empfinden.[10]

Dennoch erscheint uns das Konzept beim Entwurf von Infrastrukturen,[11] welche den Zugriff auf Informationen innerhalb einer bestimmten Sphäre ermöglichen, als regulative Idee sinnvoll.[12] Derartige Infrastrukturen sind so zu entwerfen, dass sie nicht nur den Fluss von Informationen ermöglichen und lenken, sondern auch stoppen können. Hierbei sind insbesondere die Schnittstellen zu anderen Sphären zu beachten, welche so zu gestalten sind, dass der Austausch blockiert werden kann.

Dabei ist zwischen der Binnen- und Außenperspektive zu unterscheiden. Das Modell erscheint uns vor allem für die Gestaltung einer Sphäre aus der Binnenperspektive als angemessen, wenn es also zum Beispiel um den Entwurf eines Krankenhaus- oder Praxisinformationssystems[13] geht. Aus der Außenperspektive können z. B. aber auch Informationsangebote im WWW als Teil der medizinischen Sphäre wahrgenommen werden, was jedoch nicht angemessen ist, da hier andere Normen hinsichtlich des Umgangs mit personenbezogenen Daten gelten können.[14]

Eine Kantische Auffassung von Eigentum

In seiner „Metaphysik der Sitten" (1797ff.) bestimmt Immanuel Kant das Wesen von Eigentum als eine Relation zwischen dem Eigentümer und allen anderen, die von der Nutzung ausgeschlossen werden.[15] Damit tritt er einer Sichtweise entgegen, welche ein unmittelbares Verhältnis zwischen einer Person und einer ihr äußerlichen Sache annimmt.[16] Wir gehen im Folgenden davon aus, dass Informationen analog zu den „äußeren Sachen" behandelt werden können.[17]

Obwohl es in unserem Fall nahe liegend wäre, über ein Selbsteigentum des menschlichen Körpers in Anschluss an Locke zu argumentieren, ist dies im Rahmen einer Kantischen Auffassung nicht möglich. Kant betont nämlich, dass ein Mensch zwar sein eigener Herr, nicht aber als „Eigentümer von sich selbst"[18] sein könne. Ein Eigentümer darf nämlich beliebig mit seinem Eigentum verfahren, was im Falle des Menschen aufgrund der Verantwortlichkeit gegenüber der Menschheit in der eigenen Person nicht gestattet werden kann.

Dementsprechend setzt Kant auch der Verfügbarkeit des eigenen Körpers in der Medizin enge Grenzen.

---

[10] Nissenbaum (2004: 145).

[11] Zum Begriff der Infrastruktur vgl. Star / Bowker (2006).

[12] Umgekehrt läßt sich auch behaupten, dass Infrastrukturen, welche die Weitergabe von Informationen benötigen, die unterschiedlichen Sphären erst konstituieren.

[13] Zu den Begriffen vgl. Dugus / Schmidt (2003: 83f., 90f.).

[14] Nelson (2006: 184) macht dies für den US-amerikanischen Raum anhand der durch die Gesetzgebung geschützten Daten im Gesundheitswesen deutlich, zu denen z. B. auch die IP-Adressen bei Online-Anfragen oder E-Mail-Adressen gehören, die bei privaten Portalen zu medizinischen Themen regelmäßig erhoben werden. Wenn es die Normen sind, die eine bestimmte Sphäre kennzeichnen, so wäre hier von einer anderen Sphäre zu sprechen.

[15] Durch die Erklärung, „dieser äußere Gegenstand ist mein", wird „allen anderen eine Verbindlichkeit auferlegt …, die sie sonst nicht hätten, sich des Gebrauchs desselben zu erhalten" (Die Metaphysik der Sitten, S. 253). – „Die Metaphysik der Sitten" wird im Folgenden als „MdS" abgekürzt. Die Seitenangaben beziehen sich auf die Akademie-Ausgabe.

[16] Vgl. MdS, S. 260.

[17] Wie z. B. Weber (2005: 43) ausführt, kann vor allem dann von Informationen als Gütern gesprochen werden, wenn der Schwerpunkt auf den pragmatischen Aspekt von Informationen gelegt wird, also die Tatsache, dass Informationen Bestandteil von Handlungsabfolgen sein können. Dies trifft auf unser Beispiel zu, da es um die Weitergabe und Nutzung von bestimmten Informationen geht. Dem gegenüber steht eine Sichtweise, welche einen prinzipiellen Unterschied zwischen Informationen und materiellen Gütern einfordert. Im Rahmen dieses Beitrages wollen wir jedoch von einer prinzipiellen Analogiefähigkeit ausgehen.

[18] MdS, S. 270.

Sich selbst eines „Organs zu berauben (verstümmeln), z. B. einen Zahn zu verschenken oder zu verkaufen, um ihn in die Kinnlade eines anderen zu pflanzen … gehört zum partialen Selbstmorde"[19] und ist somit verwerflich. Die Amputation eines abgestorbenen oder absterbenden Organs ist hingegen erlaubt, wenn es die Pflicht zur Selbsterhaltung erfordert. Wir gehen davon aus, dass analog Patienten der Nutzung ihres Körpers zum Zweck der Produktion von genetischen Informationen zustimmen dürfen, sofern dies der Erhaltung ihrer Gesundheit dient.[20]

Nun verfügen Privatpersonen in der Regel nicht über die notwendigen Produktionsmittel, um die Informationen in der benötigten elektronischen Form herzustellen. Allerdings bedarf die Produktion dieser Informationen der Zustimmung der betroffenen Person,[21] so dass aus dem „vereinigten Willen beider"[22] ein Vertrag entsteht. Da es nicht im Interesse des Patienten ist, dass die in Folge des Vertrages entstandene Information ohne jegliche Einschränkung weitergeben wird, sollte die Zustimmung nur dann erfolgen, wenn ihrer Verfügbarkeit Grenzen gesetzt sind.[23] Nach Kant darf er sogar nur unter diesen Bedingungen zustimmen.

Dadurch dass andere von der Nutzung der Information ausgeschlossen werden, wird die Information dann eigentumsähnlich; wenn sich die legitime Nutzung nur auf einen bestimmten Personenkreis beschränkt, erhält sie den Charakter eines gemeinsamen Eigentums. „Gemeinsames Eigentum" deshalb, weil nun eine Gruppe von Personen alle anderen von der Nutzung ausschließen kann. Diese Idee findet sich in der Rechtslehre Kants jedoch nicht. Er kennt zwar den Begriff des „gemeinsamen Eigentums" von zwei Personen, allerdings soll dieser nur dann verwendet werden, wenn einem der beiden Miteigentümer „nur der ganze Besitz ohne Gebrauch, dem anderen aber aller Gebrauch der Sache samt dem Besitz zukommt … ."[24] Da Eigentum an einer Sache aber bedeutet, beliebig mit ihr verfahren zu können, kann in unserem Fall von gemeinsamem Eigentum nicht die Rede sein, denn der Patient kann dem Vertrag nur unter der Bedingung zustimmen, dass die Information nicht beliebig weitergeben wird. Deshalb bleibt der Patient der alleinige Eigentümer, auch wenn er nicht allein in der Lage ist, die Information herzustellen.[25]

Dem entsprechend kann auch der Kostenträger nicht Eigentümer der Information werden, sondern ihm können lediglich gewisse Nutzungsrechte übertragen werden, welche zuvor der medizinischen Institution vom Patienten übertragen wurden.[26]

Der Schutz der kontextuellen Integrität

Da genetische Informationen in der Regel ein Leben lang gültig bleiben[27] und zudem z. T. Aussagen über nachfolgende Generationen zulassen, ist der prinzipielle Anspruch auf Löschung zu betonen. Dieser kommt nach dem bisher Gesagten dem Patienten als alleinigem Eigentümer zu.[28]

---

[19] MdS, S. 423.

[20] MdS, S. 419.

[21] Wie z. B. Savulescu (2002: 649) zu Recht betont, setzt jede Form der ärztlichen Behandlung die Zustimmung des Patienten voraus.

[22] Vgl. MdS, S. 272.

[23] Dies läßt sich mit der Feststellung Kants begründen, dass der Mensch genötigt ist, seine Ansichten in der Regel vor anderen Menschen zu verschließen, damit niemand zu seinem Schaden davon Gebrauch macht. Hieraus ergibt sich der besondere Wert der „moralischen Freundschaft". Vgl. MdS, 471f.

[24] MdS, 270.

[25] Eigentum bedeutet hierbei vor allem, dass andere von der Nutzung der Information zunächst einmal ausgeschlossen werden können, soweit die Information den Besitzer speziell betreffen. Zur Verwendung der Informationen zu Forschungszwecken: s. unten. Zur Kritik an dem Schutz von genetischen Informationen durch Eigentumsregelungen vgl. z. B. Spinello (2006).

[26] Nutzungsrechte sind ein zentraler Bestandteil der deutschen Auffassung des Urheberrechts, nach dem ein Urheber stets im Besitz des Urheberrechts an einem Werk bleibt und lediglich zeitlich, räumlich und inhaltlich begrenzbare Nutzungsrechte veräußern kann (Gesetz über Urheberrecht und verwandte Schutzrechte, § 31ff.).

[27] Ausnahmen, wie die Möglichkeit der sog. „Gentherapie", können hier vernachlässigt werden.

[28] Als rechtliche Grundlage kommen in Deutschland z. B. §20 und §35 des Bundesdatenschutzgesetzes in Betracht.

Dem entsprechend muss nach der Verwendung im Endscheidungsprozess bezüglich des wirksamsten Medikamentes die genetische Information gelöscht werden, sollte der Patient nur der Verwendung zu diesem einen Zweck zugestimmt hat. Denkbar ist jedoch auch, dass eine einschränkende Zustimmung zur weiteren Verwendung der Information gegeben wird, so dass diese z. B. zusammen mit den Informationen von anderen Patienten oder Probanden in anonymisierter Form statistisch ausgewertet werden darf.

Es wäre auch möglich, die Informationen nicht zu anonymisieren, sondern zu pseudonymisieren, so dass eine spätere Zustimmung zu bestimmten Verwertungsformen möglich wird.[29] Forscher könnten dann z. B. Anfragen bei Personen stellen, deren Pseudonym sie kennen, ohne Zugriff auf weitere personenbezogene Daten zu haben. Diese Anfragen könnten z. B. die Einladung zu einer Nachuntersuchung[30] oder die Bitte um weitere Nutzungsrechte sein, wobei vorauszusetzen ist, dass die Patienten über geeignete Kommunikationsmittel verfügen und die Schnittstelle zwischen Patient und medizinischer Institution so ausgestaltet ist, dass sie die Grenzen der medizinischen Sphäre nicht gefährdet.

Die Aufrechterhaltung der medizinischen Sphäre liegt dabei auch im Interesse der medizinischen Institutionen, welche einen geeigneten Rahmen schaffen müssen, innerhalb dessen Patienten und Probanden sicher sein können, dass die Preisgabe bestimmter Informationen angemessen ist und ihre Verarbeitung nur nach den üblichen Normen erfolgt. Dies sind zugleich die beiden Kriterien, die Nissenbaum ihrem Konzept von der „kontextuellen Integrität" zugrunde legt.[31]

Eine geeignete (technische) Infrastruktur könnte dazu beitragen, dass der Anspruch auf kontextuelle Integrität gewahrt wird, wobei der Kontext z. B. in Form von Metadaten mit transportiert werden könnte.[32] In der einfachsten Form könnte die ursprüngliche Sphäre als Metadatum angesehen werden, Metadaten könnten aber den Umfang der Zustimmung des Patienten transportieren, die Zustimmung zu späteren Anfragen bzgl. ergänzender Nutzungsrechte, ein „Verfallsdatum" usw.

Dies wäre auch im Sinne der Forschung, wie die anhaltende Diskussion um die Möglichkeit des „one-time general consent" zeigt.[33] Dabei geht es darum, ob Patienten und/oder Probanden die Möglichkeit haben sollen, einmalig ihre Zustimmung zu allen möglichen, legitimen Verwertungsformen zu geben.[34] Ein entsprechende Infrastruktur, die auch spätere Nachfragen erlaubt, scheint hier eine interessante Lösung. Zudem könnte durch ein solches System die Fähigkeit zur Selbstbestimmung des Patienten gestärkt werden, die z. B. auch über den Fortgang der Forschung in Kenntnis gesetzt werden können.[35]

Die Grenzen der medizinischen Sphäre

Fassen wir das bisher Gesagte zusammen, so können wir zu dem Schluss kommen, dass es Fälle geben mag, bei denen die Verwendung von genetischen Informationen im genannten Umfang als legitim erscheint. Mit der Sphären-Auffassung der informationellen Gerechtigkeit Nissenbaum verfügen wir über ein Konzept, mit denen sich notwendige Rahmenbedingungen formulieren lassen. Eine Kantische Auffassung von Eigentum stärkt dabei die Position des Patienten, weil sie die Möglichkeit begrenzt, Informationen über sich produzieren zu lassen, und dadurch das Konzept des blockierten Austausches zusätzlich stützt.

---

[29] Wir denken hier an ein an System wie es z. B. von Roger-Francis (1998) vorgeschlagen wurde, das wir hier um die Möglichkeit der Kommunikation mit den Patienten ergänzen. Vgl. Nagenborg (2005: 171). – Ähnlich der Vorschlag von DeCew (2006: 128).

[30] Im Falle einer Nachuntersuchung wäre denkbar, dass diese durch einen anderen Forscher oder eine andere medizinische Institution erfolgt, wodurch die Pseudonymität des Patienten bzw. Probanten gewahrt werden würde.

[31] Nissenbaum (2004: 138).

[32] Vgl. Star / Bowker (2006: 238) zur Bedeutung von „metadata standards" in Hinblick auf Infrastrukturen.

[33] Wendler (2006) sowie die Reaktionen auf seinen Beitrag von Hunter (2006) u. a. seien hier als aktuelles Beispiel genannt.

[34] „Legitime Verwertungsformen" könnten z. B. durch Ethikkommissionen gewährleistet werden.

[35] Savulescu (2002: 651) nennt dies als einen Grund, warum das Einholen der Zustimmung von Fall zu Fall von Vorteil wäre.

Allerdings setzt das Modell eine intakte medizinische Sphäre voraus, welche unseres Erachtens durch den zunehmenden ökonomischen Druck in Frage gestellt wird.[36] Denn abgesehen davon, dass die Aufrechterhaltung der Sphäre voraussetzt, dass alle sich nach den Normen richten, trägt auch die zunehmende Privatisierung von medizinischen Institutionen oder die Auslagerung von Teilaufgaben in die Privatwirtschaft dazu bei, dass nicht mehr sichergestellt ist, dass die besonderen Normen der medizinischen Sphäre auch weiterhin ihre Gültigkeit haben. Ein schönes Beispiel findet sich bei Nelson (2006), der von einem Krankenhaus berichtete, das Diktate bei einer Firma schreiben ließ. Diese Firma vermittelte den Auftrag an eine weitere Firma, die schließlich eine Person im Ausland mit dem Abschreiben der Diktate beauftragte und nicht bezahlte, worauf hin die betreffende Person versuchte, das Krankenhaus mit der Veröffentlichung der Patientenakten zu erpressen.[37] Auch jenseits solcher Horrorszenarien macht z. B. ein Blick in den „Health Insurance Portability and Accountability Act" (HIPPA) deutlich, was eine zunehmende Privatisierung der medizinischen Sphäre bedeutet. Dieser regelt nämlich die Weitergabe von Patienteninformationen an "health care organisations, including hospitals, physicians' offices, health care plans, employers, public health authorities, life insurances, clearing houses, billing agencies, information systems and 'any … person or organisation who furnishes, bills or is paid for health care in the normal course of business.'"[38]

Bedenklich erscheint uns deswegen, wenn für die Nutzung von genetischen Informationen allein aufgrund von Kosteneinsparungsgründen argumentiert wird, ohne dabei zu beachten, dass aufgrund der besonderen Natur dieser Informationen auch ein hoher Aufwand notwendig ist, um eine geeignete Infrastruktur zu schaffen, welche den berechtigten Zugriff regelt und den Fluss der Informationen in angemessener Form blockieren kann.

So hilfreich also die Sphärenauffassung für die Beurteilung von Systemen zur Informationsverarbeitung ist, so wenig können wir davon ausgehen, dass die Sphären auf Dauer bestehen, wenn wir hierfür nicht Sorge tragen.

Literatur

Bruen, Robert (1999): Preparing the End of Privacy. The Basis for a New Ethics: Electronic Information and the Human Genome Project. S. 151-164 in: Pourciau (1999).

DeChew, Judith Wagner (2006): Privacy and Policy for Genetic Research. S. 121-135 in: Tavani (2006).

Dugas, Martin,, und Karin Schmidt (2003): Medizinische Informatik und Bioinformatik. Ein Kompendium für Studium und Praxis. Berlin u. a. O.: Springer.

Hunter, David (2006): One-time general consent for research on biological samples: autonomy and majority rules have been misunderstood. In: BMJ, Vol. 332 (18.März 2006), S.665-666.

John C. Moskop, Cathrine A. Marco, Gregory Luke Larkin u. a. (2005): From Hippocrates to HIPAA: Privacy and Confidentiality in Emergency Medicine – Part I: Conceptual, Moral, and Legal Foundations. In: Annals of Emergency Medicine, Vol. 45 (2005), Nr. 1, S. 53-59.

Nagenborg, Michael (2005): Das Private unter den Rahmenbedingungen der IuK-Technologie. Wiesbaden: VS Verlag.

Nelson, Steven (2006): Privacy and Medical Information on the Internet. In: Respiratory Care, Vol. 51, Nr. 2, S. 183-187.

Nissenbaum, Helen (2004): Privacy as Contextual Integrity. In: Washington Law Review, Vol. 79, S. 119-158.

Peters EJ, Slager SL, McGrath PJ, Knowles JA, Hamilton SP: Investigation of serotonin-related genes in antidepressant response. Molecular Psychiatry (2004) 9(9): 879-89.

Pourciau, Lester J. (Hg.) (1999): Ethics and electronic Information in the Twenty-first Centurty. West Lafayette, Indiana: Purdue University Press.

Roger-Francis (1998)

Rau T., G. Wohlleben, H. Wuttke, N. Thuerauf, J. Lunkenheimer, M. Lanczik und T. Eschenhagen (2004): CYP2D6 genotype: impact on adverse effects and nonresponse durcing treatment with antidepressants – a pilot study. In: Clinical Pharmacology and Therapy, 75 (5), S. 386-93.

---

[36] Auch ein Interesse an bestimmten Daten zum Zweck der Kriminalprävention usw. wäre denkbar, soll hier aber nicht diskutiert werden.

[37] Leider war es nicht möglich, die Authentizität des Beispiels von Nelson (2006: 186) zu überprüfen, da die von ihm genannte Quelle selbst keine Nachweise enthält.

[38] Moskop/Marco/Larkin u. a. (2005: 56), vgl. auch Nagenborg (2005: 173) sowie DeCew (2006: 122).

Savulescu, Julian: Against. (For and against: No consent be needed for using leftover body material for scientific purposes.) In: BMJ, Vol. 325 (21.09.2002), S. 649-651.

Spinello, Richard A. (2006): Property Rights in Genetic Information (Original: 2004). S. 213-233 in: Tavani (2006).

Star, Susan Leigh, und Geoffrey C. Bowker (2006): How to Infrastructure. S. 230-245 in: Leah A. Lievrouw und Sonia Livingstone, hg.: The Handbook of New Media. Updated Student Edition. London – Thousand Oaks – New Delhi: Sage.

Tavani, Herman T. (Hg.) (2006): Ethics, Computing, and Genomics. Boston u. a. O.: Jones and Bartlett Publishers.

Van den Hoven, Marcel Jeroen: Privacy or Informational Injustice? S. 140-150 in: Pourciau (1999).

Weber, Karsten (2005): Das Recht auf Informationszugang. Berlin: Frank & Timme.

Wendler, David (2006): One-time general consent for research on biological samples. In: BJM, Vol. 332 (4. März 2006), S. 544-547.

IRIE
International Review of Information Ethics Vol. 5 (09/2006)

James Caufield

The Myth of Automated Meaning

Abstract:

Most discussions of search engines focus on technology or user experience. By contrast, this paper asks about those who produce the recommendations that search engines gather. How are these people and institutions affected when search engines incorporate their work into search results, but no credit is given? The paper argues that the lack of attribution encourages the myth of automated meaning, the false belief that computers and algorithms have created rather than simply gathered these recommendations. It further argues that by concealing the role of these producers, search engines undermine public support for the individuals and institutions that create trustworthy recommendations, especially libraries. Because search engines borrow so extensively from public institutions and the public at large, their ethical obligations are far greater than previously recognized. The paper concludes with some comparisons between the ethical practices of libraries and those of search engines.

Acknowledgements: An earlier version of this paper was delivered at the Symposium *Ethics of Electronic Information in the 21st Century*, 2005. I would like to thank Mardi Mahaffy for commenting on the paper.

Agenda

Author:

James Caufield:

New Mexico State University, P.O. Box 30006, MSC 3475, Las Cruces, New Mexico, 88003-8006.

☎ 505-646-5551, ✉ Caufield@nmsu.edu.

Relevant publications:

- "Where Did Google Get Its Value?" *portal: Libraries and the Academy*, vol. 5, no. 4, October 2005, pp. 555-572.

Introduction

The recent development of search engines has solved one of the most vexing problems of the early Web, namely how to find valuable materials amid the vast number of web pages. *"Until Google, search engines helped a bit, but only a bit. They told you what was popular but not necessarily what you wanted to know. But Google's inventors found a way to put the most useful information at the top of the pile"* (McRae, 2004). Google did this by counting links between web pages and by studying user behavior, thus gathering meaningful recommendations from the Web itself.

Yet the brilliance of this success might blind us to its darker implications. Studies of search engines usually focus on technology or user experience. By contrast, this paper asks about the people and institutions who produce the recommendations search engines gather. While search engines are in the business of borrowing and aggregating trustworthy, relevant, and credible recommendations, they give no credit to the sources that provide them. The thesis of this paper is that by concealing the role of these producers, search engines undermine public support for those who create this meaning, especially libraries.

A few notes about method: private search engine companies are highly secretive about their methods and sources, making direct investigation impossible. This makes the paper in part a speculative effort, and so it pursues a line of thinking as far as possible. It advances what might be called a worst-case scenario, not because I know this to be true in all its details, but rather because I think this question needs to be considered in terms of its furthest possible ramifications.

As for intellectual abbreviations, I will usually call the information that libraries and search engines provide *recommendations*, but sometimes I will use more general terms such as *meaning* or *value*.

Finally, I have treated the positive contributions of Google elsewhere (Caufield, 2005); this paper will have a critical focus.

Review of Some Arguments Concerning the Ethical Obligations of Search Engines

Many reasons have been given for holding search engines to high ethical standards. First, using search engines is a practical necessity, as they are the *de facto* gatekeepers to information on the Web (Reider 2005: 27; Hinman 2005: 21; Neuberger 2005: 5). Thus a search engine such as Google

> *... plays an important role in the information many people rely on for choices in their lives. Cumulatively, the quality of Google's ranking impacts millions of choices made daily* (Wiggins, 2003).

This responsibility is intensified in that these tools are often used to find vital information (e.g., medical; see Fallows, 2005: 1, 6), and users place inordinate trust in search engines (Fallows, 2005: 27) which, as Rieder (2005: 31) and others argue, should not be exploited.

Moreover, the space that is accessed –the Web– is not private property, and it fulfills many of the functions of traditional public spaces such as libraries, museums, educational institutions, and public fora (Nissenbaum and Introna, 2004: 19-22). Finally, as search engines become more central to our information infrastructure, they assume important social and political roles, both in terms of providing the information essential to democratic decision-making, and in assuring the openness necessary for a free society (Letwin, 2004; Hinman, 2005: 21-2, 25). To this list of reasons for holding search engines to high ethical standards, this paper will add one more.

The Myth of Automated Meaning and Search Engines' Ethical Obligations

The most common way to describe Google is in technological terms. Oddly enough, this description is sometimes anthropomorphic, as in the following passage:

> *...the business story of Google is less interesting than the technological one: if information is power, then Google has helped change the world. Knowledge is measura-*

> *bly easier to obtain. Google works. Google knows* (Achenbach, 2004).

Anthropomorphism is one version of the myth that attributes meaning to the machine, but it also serves as a reminder: computers do not know, and assigning relevance is a subjective determination, as is assigning other values (Blanke, 2005: 35-37; Rieder, 2005: 29). Thus Google is a semantic aggregator (Brooks 2004), a machine that only gathers meaning and does not create it. If we forget this, we succumb to the myth of automated meaning.

The implications of this insight have not been sufficiently noted. Most arguments concerning the ethical obligations of search engines point to the vitally important services these companies *provide to* the public. But perhaps the most telling argument concerns what search engines *take from* the public. So long as it is believed that search engine technology creates meaning, it is logical to think that search engines own this information. The recognition that search engines gather meaning from public sources implies a different conclusion. That search engines receive such a massive benefit from the public implies a correspondingly large obligation to the public, an ethical relationship perhaps more akin to stewardship than to ownership.

The remainder of this paper will investigate how and why search engines draw upon the work of public sources, particularly libraries.

Search Engines Need Trustworthy and Unbiased Results

The myth of automated meaning encourages us to think about search engines in terms of technology, so that we tend to forget or de-emphasize their human and ethical aspects. But consider an extreme case, the search service Overture, a competitor to Google, originally founded as Goto.com. In its earliest incarnation all search results were advertisements, with the top slot being given to the top bidder. The private interests of advertisers and the search engine were paramount, and disinterested service to the public was of little concern. Users spurned it: *"Overture soon figured out … that users did not want a search engine that was only advertising"* (Hansell, 2004). While Overture continues to include paid links (advertisements) in its search results, Google has never allowed this practice, so it is hardly surprising that Google results are more relevant and precise than those of Overture (as reported by Goh and Ang, 2003). Thus relevance seems to be a product of ethics as well as of technology.[1]

Although relevance is commonly discussed as if it were the only value by which search engines should be judged, it is not. An exclusive focus on the relevance of results is perhaps appropriate when institutional structures have already assured the trustworthiness of the materials to be searched. This is the case in libraries, where quality is rather rigorously controlled, but it cannot be assumed in the Internet environment, where dubious materials abound. For instance, if a search engine produced recommendations for cancer treatments solely on the basis of advertising revenues and without regard to credibility, this information could indeed be relevant but still worth little. In order for search engine results to be valuable they need to be not only relevant but trustworthy as well.

An all-advertising search engine will not draw users and so, like most other commercial web enterprises, search engines need what the industry calls *'real estate,'* some non-advertising content that will draw *'eyeballs'* to whom it can present advertisements. For some web sites this is bird watching, for others sports news, etc. For search engines it is relevant and trustworthy recommendations. It seems that without this more-disinterested content, search engines will not draw the users they require in order to charge for advertising.

Users' Response to Search Engines that Exploit Trust for Private Benefit

There is evidence that search engine users believe they are receiving disinterested content, though often they are not. Many search engines accept money for including or placing advertising in their results, a fact that many writers on the subject see as problematic (Neuberger, 2005: 10-11; Himma

---

[1] Certainly other factors (including technological ones) play a role, but it stands to reason that when a search engine produces results according to two criteria, relevance and advertising revenue, rather than the criterion of relevance alone, then the relevance of results will be compromised.

2004: 474; Moxley et al., 2004; Scott, 2003; Nissenbaum and Introna, 2004: 14; Larsen, 1999: 22-3). Users are generally unaware of this practice, but once informed they recognize it as an ethical issue. An ethnographic study conducted by Leslie Marable of Consumer WebWatch found that users are frequently unable to distinguish between links that are paid (advertisements) and those that are not (Marable, 2003, p. 15), and this is because searchers have little conception of how search engines work (Marable, 2003: 7, 39-40; see also Neuberger, 2005, p. 5, citing Princeton Survey Research Associates 2002: 17; Fallows, 2005: 27). As Nissenbaum and Introna have observed, *"[g]iven the vastness of the Web, the close guarding of algorithms, and the abstruseness of the technology to most users, it should come as no surprise that seekers are unfamiliar, even unaware, of the systematic mechanisms that drive search engines. Such awareness, we believe, would make a difference"* (Nissenbaum and Introna, 2004: 16).

The Marable study shows that awareness does indeed make a difference. When participants in the study were informed that some search engine results were paid advertisements, *"each participant expressed surprise after learning about the search engine marketing practice. Some had negative emotional reactions"* (Marable, 2003; 39), including anger, a sense of betrayal, being somewhat stunned, disappointment, and a feeling of helplessness. Two people explicitly named this as an ethical issue, questioning *"... the trustworthiness of advertisers that chose to participate in search engine marketing programs"* (Marable, 2003; 18). *"The participants unanimously believed paid search links were too tough to recognize on many sites, and the related disclosure information was clearly written for the advertiser, not the consumer"* (Marable, 2003: 23). A similar example is given by Nissenbaum and Introna: *"[w]hen customers learned that Amazon Books, for example, had been representing as 'friendly' recommendations what were in reality paid advertisements, customers responded with great ire and Amazon hastily retreated"* (Nissenbaum and Introna, 2004: 16). As Fallows summarizes the situation,

> *"[d]espite … the explosion of paid placement sales as a source of revenue for search engines, users remain largely unaware of these issues – even though they still hold strong views of what they consider to be ethical practices by search engines. …[M]any users would object to the practices of search engines they rely on, if they understood more clearly how those engines actually operate"* (Fallows, 2005: 17).

In short, users tend to trust that search engine results will not be biased by private interest, and tend to take offense when they are informed otherwise. As Marable found, the depth of this misplaced trust is indicated by a rather astonishing belief:

> *"Our findings show Web searchers in this ethnographic study chose links from the first page of results nearly 50% of the time because they trusted the search engine to present only the best or most accurate results first. This trust led them to believe it was unnecessary for them to review later results pages"* (Marable, 2003: 38).

Speculations as to Why Users Trust Search Engines

This paper has argued that the quality of information is at least as much a function of ethical decisions as it is of technology. But it is also true that certain technologies have come to be associated with certain ethical practices. There is a long association between the technology of printing and a high quality of information. Of course this is not directly due to the technology, but rather to the culture that has developed in the print environment. Generally, anything published in paper is relatively trustworthy, as the publication process involves many steps of editorial selection and review. Even ephemeral media such as newspapers are held to fairly high standards, and the academic process is especially rigorous, involving a system of critical review on many levels. Thus the printed format has long been associated with a high level of credibility and trustworthiness.

Documents available on the Internet bear a strong resemblance to printed documents. Yet these Internet documents are not as trustworthy and reliable as printed documents because they usually do not have the institutional supports that would vouchsafe their quality. The likely scenario is that because of the similarity of formats users mistakenly extend their trust of printed materials to documents found on the Internet. Since this unwarranted trust of things on the Internet seems to have its origin in the traditions of the print culture, it could be said that digital illiteracy (at least in the sense of an overly-credulous attitude toward information on the Internet) has been caused by … literacy!

The misplaced trust in search engines is a subset of the misplaced trust of things on the Internet. The library online catalog is eminently trustworthy, at least in the sense of being free of commercial bias.[2] Search engines resemble the online catalog, as both accept search terms and then yield results. Users mistakenly assume that a similar appearance implies a similar trustworthiness and credibility. Yet search engines do not have the cultural institutions standing behind them to assure their integrity. Certainly there is confusion: many of my students now refer to the library online catalog as '*a search engine.*' As Nissenbaum and Introna observe, *"[m]ost users tend to treat search engine results the way they treat library catalog search results"* (Nissenbaum and Introna, 2004: 16). Moxley et al. are explicit about the transfer of credibility: *"Web search engines trade on the associations its users make between the libraries and information professionals they have always relied upon for similar needs in the past"* (Moxley et al., 2004: 64-65).

Even if the credibility mistakenly attributed to search engines is not directly due to previous experience with library catalogs, it should still be understood as borrowing from the credibility of the print format and the social institutions it represents, including libraries. Put another way, if libraries and other institutions of the print tradition had not been so ethical in the past, it is unlikely that search engine users would be so trusting today. The users' belief that search engines should be ethical is thus partly a residuum of this legacy. Search engines that tamper with results (e.g., by accepting money for including advertising in the search results) are exploiting and undermining the faith that one will be presented with information that is not biased by direct personal interest, a trust that has been built by libraries and other institutions of the printed medium over many years (for a similar point see Moxley et al., 2004: 62).

A Major Concern: Search Engine Secrecy

Elsewhere I have argued that Google's success should be ascribed mostly to ethical rather than technological factors, specifically that Google has brought many library values (or valuable library practices) to the Internet environment (Caufield, 2005). For instance, Google brought the important library value of disinterested cataloguing to the search environment, as it refused (and continues to refuse) to include advertisements in the search results, placing them instead in a separate column that is clearly labeled as such (a practice that one of Marable's participants found *"honorable."* Marable, 2003: 31).

Google's ethics seem exemplary in comparison to the egregious practices of other search engines, and this is surely one reason that *"Google has the trust"*

---

[2] This is not to say that library cataloging classification is without its flaws. But it is inconceivable that libraries would ever accept money for promoting an item in their catalog, while search engines do this all the time. On the question of agenda-driven bias it is true that libraries are far from perfect, as cataloging classification schemes arguably embody racist, sexist, and classist biases. Yet the bias found in search engine results is of an entirely different degree, one that libraries would never tolerate. There are reasons for this: first, librarians are paid to make professional judgments, and thus stand at a remove from direct interest. The job of librarians is to make this disinterested judgment, however imperfectly they carry it out. Also, the library profession has a code of ethics that proscribes activities motivated by private interest or profit: "We do not advance private interests at the expense of library users, colleagues, or our employing institutions" (American Library Association, 1995, section VI). Finally, librarians' judgments are for the most part open to scrutiny and criticism by professional peers and are at least potentially open to public debate. By contrast, search engines are founded upon private interest. Although the ethics of corporations may vary, the bottom line is that trustworthy and relevant results are a means to draw users to their advertising, not an end in themselves. Search engines rely in part on 'lay indexers' (to be described below), people whose recommendations will often be motivated by immediate private interest. Finally, because of search engine secrecy, oversight is limited to the corporation, which might act to curb bias only if profit is threatened.

So although libraries are not beyond reproach, they set a far higher standard than do search engines. For purposes of this paper it seems justified to make a sharp distinction between the relatively unbiased and disinterested judgments that librarians make and the profit and agenda-driven biases present in search engines. For clarity of exposition the remainder of the paper will treat this distinction as if it were absolute, though I ask that the reader bear the foregoing qualifications in mind.

(Harvey, 2004). Yet despite these positive contributions there is cause for concern, particularly with the secrecy of the ranking processes used by Google and all other commercial search engines.

There are reasons for this secrecy, the most obvious being the need to protect proprietary search methods. If Google were to reveal details of its algorithm, presumably competitors would 'borrow' its techniques and Google would lose its competitive advantage. A second reason for this secrecy is described by Terrence A. Brooks, who points out that traditional information institutions (e.g. libraries)

> *"... were devised, built and managed by information professionals. Professionals were known, publicly accessible and held to high standards of ethics. Information professionals, such as librarians, were considered to be operating a public trust with a view to the best interests of society"* (Brooks, 2004: 7 of 13).

By contrast, search engines rely on a form of collaborative filtering in which large numbers of people contribute to the effort. As Brooks rightly notes, the Google system (or any automated system) gathers information from web authors, who thus act as *"lay indexers"* of web content (Brooks, 2004).

That the people responsible for this are usually not professionals is problematic enough; what is worse is that they are often working toward private goals (promotion of their products and agendas) rather than the public good. Thus any automatic ranking system is susceptible to manipulation or "spamming," the deliberate introduction of recommendations that serve the interests of the lay indexers (Neuberger, 2005: 4; Spencer, 2004; Nissenbaum and Introna, 2004: 13-14). The more that spammers know about the ranking system, the more they will be able to manipulate it; if Google reveals exactly how it gathers recommendations and calculates rankings, spammers will be able to work the system all the better. Brooks understands this danger, and so he stresses *"... the importance of maintaining the ignorance of lay indexers. [....] Google's only possible survival strategy is to keep its parsing and ranking algorithms absolute secrets"* (Brooks, 2004).

This point of view is understandable but, in my view, not sustainable. Certainly Brooks describes the necessity of secrecy for commercial search engines: an end to it would perhaps mean an end to the companies. Yet from a larger social perspective this lack of transparency has enormous implications. What does it mean, when a crucial component of the information infrastructure lacks the transparency that is necessary for accountability, review, or criticism? While Brooks maintains that secrecy is necessary for search engines to function, this secrecy seems to be incompatible with the transparency required of information institutions in a democratic society. [3]

Revisiting the Ethics of Link Analysis: Secrecy Prevents Attribution

When the sources that search engines draw upon are not given credit, what harm is done? The question is difficult enough that I will turn to a philosophical statement of the problem. In Plato's *Republic* (359b-360e), two characters present the argument that it is better to be unjust than to be just. They recite a myth about a ring that renders its wearer invisible and so allows great crimes to be committed with impunity. Reasoning from this example they maintain that any logical person, when granted a complete lack of public accountability, will act unjustly and do what is wrong because it serves their private interest.

In a further refinement of their argument (362d-367e) they ask which is more important, actually doing a good deed or simply having the reputation for doing a good deed? If the action is more important than the reputation, then even if the reputation is transferred to one who does not deserve it there should still be some motivation to do good deeds. But this, they say, is nonsense. No one will continue doing good deeds if the credit goes to others. Hence all that is important is the reputation for doing good. The Socratic response to this argument

---

[3] On this issue there are divergent views. Some, usually citing the necessity of secrecy for private companies, agree with Brooks (e.g., Hinman, 2005: 22). Others, usually pointing toward the public's need for accountability, advocate transparency (e.g., Welp and Machill, 2005: 18; Nissenbaum and Introna, 2004: 23). Rieder (2005: 31) tries to split the difference, suggesting that search engines should be obliged to grant access to their indexes and server farms (but retain secrecy of their algorithms), so that public entities can write and run non-private algorithms.

is, in brief, to consider the good of society as a whole, rather than the good of the individual. From this perspective a different view emerges, a society where political structures ensure that deed and reputation will coincide.

The operations of current search engines involve an analogous form of invisibility. Much of the quality of the current generation of search engines is due to link analysis, a process whereby links from one Web page to another are treated as recommendations. These are then aggregated and, along with other factors, used to rank the results. Generally speaking, the more links pointing to a page, the higher its ranking. What is important here is that these recommendations are borrowed without attribution. [4]

In the early days of the Internet, this practice seemed innocuous, perhaps because there simply was not much information there to borrow and the need to create order was so great. Yet in a more developed web environment the practice has serious implications for traditional information institutions. First, search engines harvest relevance and other values from various sources but do not give credit. So in this sense the source of the good deeds (the meaningful recommendations) is not recognized. Second, the reputation for having produced this good work is transferred from the original source to the search engine. For those who succumb to the myth of automated meaning, the search engine now has the reputation and recognition that the original source deserves. So the argument in this paper is somewhat similar to that advanced by Plato's characters: if there is no recognition of good and bad deeds, if all is hidden behind a cloak of secrecy, then what motivation remains for the good deed of creating meaningful and valuable recommendations?

Where Do Trustworthy Recommendations Come From?

It is a common claim that in this new electronic environment librarians are being "disintermediated," their services replaced by other information services, especially automated ones.

> *"A generation ago, reference librarians -- flesh and blood creatures -- were the most powerful search engines on the planet. But the rise of robotic search engines in the mid-1990's has removed the human mediators between researchers and information. Librarians are not so sure they approve"* (Achenbach, 2004).

While this is the prevalent view, the contention of this paper is that the work of librarians and other information specialists has been incorporated into search engines far more than has been recognized. Their work has not really been removed from the process; rather it has been hidden.

Where do Google's relevant and trustworthy results come from? The short but misleading answer –the myth– is that they come from an algorithm. But mathematical formulae measure and manipulate quantity; they do not evaluate quality. Computers do not decide the relevance and credibility of resources, humans do.

A clue can be found in the sometimes questionable ethical practices of companies that provide 'search engine optimization.' Among other practices used to boost their rankings in Google, spammers set up 'link farms' that create inbound links to their pages. So long as these remain undetected, search engines count these links as legitimate recommendations. Nor is the practice restricted to link farms, as

> *"...bloggers have learned that they can 'game' the system by organizing a campaign in which a large number of blogs mention a URL they wish to boost. This simple technique proves to be an effective method of 'index spamming.' Google has tweaked its algorithm to reduce its effectiveness, but hasn't solved the problem"* (Wiggins, 2003).

---

[4] Reider (2005: 29, 31) points out that complete attribution of the sources recommending any single result might not be possible, due to the complexity of the process. While this is probably an overstatement the point is still valid, for it is difficult to imagine how this highly-complex attribution could ever be represented to the user on the screen of search results. So some form of invisibility, at least in the presentation of results to users, seems difficult to avoid. But the remaining issue is that secrecy prevents any possible examination of who deserves credit for the borrowed recommendations. It is true that library subject headings are also without attribution, but this is far more acceptable, since librarians have agreed to share such things, and since the profession receives recognition for this work. By the same token, even though library subject headings are also opaque, this is mitigated in that we at least know they come from relatively disinterested sources (please see footnote 2).

Of course Google gathers its harvest not only from the spammers' link farms and bloggers' agenda-driven pages, but also from all the rest of the Web, plucking the fruit of meaning, packaging it, and delivering it to the user.

But if link farms produce the spam by which Google results (and hence the information environment) are polluted, whence come the healthy harvests, the valuable recommendations that are not motivated by private profit or agendas? As the extreme case of an all-advertising search engine illustrates, these disinterested and unbiased results are what users desire and, as the Marable study indicates, these are the kind of results they expect.

Evidence that Search Engines Borrow Recommendations from Trustworthy Sources

Because of the manipulation by unreliable 'lay indexers' who try to bias the results, search engines must actively seek out reliable and disinterested sources. In the early days of the World Wide Web most commercial services created their own subject directories. They did this because they were *"…driven by the desire to increase the number of visitors to the site, thereby attracting more advertising dollars"* (Notess, 1997: 72). But information services eventually moved away from investing effort in subject directories. The following discussion is meant to illuminate the economics (or ecology) of information that is responsible for this trend.

Google has given some explanations of how their process works. A famous early paper by Brin and Page describes an important part of the operation. The PageRank algorithm counts links; in general, the more links pointing to a given page the higher its ranking. But PageRank also assigns various weights, so that all inbound links do not count equally:

> *"The reason that PageRank is interesting is that there are many cases where simple citation counting does not correspond to our common sense notion of importance. For example, if the web page has a link off the Yahoo homepage, it may be just one link but it is a very important one. This page should be ranked higher than many pages with more links but from obscure places"* (Page et al., 1999; p. 3).

This is a sensible explanation of a clever process, but we should be clear about what it means: in effect, Google's automatic system 'borrows' the recommendations of Yahoo, and also tries to represent the relevance and credibility of these recommendations by raising them in the rankings. In this fashion Google's rankings will in part draw their value from the work of Yahoo's human editors, but no credit will be given.

What Google's founders proposed in 1999 had, by 2001, become a standard search engine practice:

> *"In order to improve relevancy, search engines have begun to compile potential results into directories that further narrow a user's search to a more specific context. In this way, human intervention can weed out the irrelevant results for a given search…. But for poor, understaffed editing teams, this is a little like sifting sand through a tennis racket. The obvious answer is to staff up the Editing Department, but with the pressure on search companies to deliver actual profits, additional headcount is no way to improve margins"* (Belle, 2001: 49-50).

The solution was this: instead of hiring editors, search engines would mine already-existing directories, using them to increase the value of their search results. The human work that had been invested by others could then be harvested and incorporated into the search results.

Though I have not found any discussion of the topic, it is possible that search engines were harvesting the Yahoo directory to increase the relevance of their results. But another directory, one created by volunteer editors, was frequently discussed as a means of improving relevance. This resource is the Open Directory Project [ODP]. *"Today, ODP's freely available index is used by major search engines on AOL, Google, HotBot, and Lycos, among many others"* (Belle, 2001: p. 51). Thus the ODP editors' judgments about credibility, trustworthiness, and relevance began to be incorporated into the results of these search engines. Indeed, it seems that any search engine not drawing upon this resource would be neglecting a recognized contribution to its competitive edge.

It is important to understand that a direct link from Yahoo or the ODP is more valuable precisely because it is presumably a disinterested judgment, probably more authoritative and less motivated by private interest than the links offered by most web sites. It is also important to understand that this value is mined by search engines and then re-presented in a much more accessible form.

> *"The ODP has very useful information," said Sergey Brin, president and cofounder of Google. "But it's tedious to browse. So we put our technology on top [to make it easier to find relevant results without having to scan through lengthy alphabetized lists of links]"* (Sherman, 2000).

Brin's comment outlines the new ecology of information that has come to dominate the Internet. While human editors provide the essential 'useful information,' search engines will absorb that information and deliver it to the user in a much more convenient form. The automated capture of information practiced by search engines means that potentially all information posted in directories and elsewhere on the Web can be absorbed into search engines. Olsen and Hu (2003) have summed up the situation well: *"Google rides on the back of human screening of web sites."*

Are Libraries an Unacknowledged Source of Trustworthy Recommendations?

As these observations make clear, search engines do not create value but rather harvest it. That is not to say that the people who operate search engines do no work at all. They do not send automated tools out to gather these human judgments from all and sundry; Google does not simply average a mass of decisions to arrive at superior intelligence.[5]

So although Google describes its PageRank system as 'democratic' (Google, 2004b), it is not (Blanke, 2005: 36). This is partly due to the design of the automated system (described above), but there is another reason. Because spammers actively seek to promote their untrustworthy recommendations, search engines will need to actively identify and draw upon trustworthy sources. Google (or any other reasonable search engine company) will begin its search for meaningful recommendations, not by a random canvas, but by targeting institutions known for their quality of judgment. A fairly old web page maintains that *"[Google] has a built-in bias towards EDU and GOV sites that is a refreshing change from the other dot-com spam-laden search engines"* (*Search Engine World*). This undated article was written in approximately 2000, but it seems eminently plausible that Google would especially wish to harvest the unbiased and trustworthy materials of government and educational institutions. It is only logical that these recommendations are less likely to be tainted by private interest and so will, on average, have more trustworthy recommendations than web pages from the dot-com domain.

A crucial step in creating a reliable and trustworthy search engine is the initial crawl to build the index. The search engine sends out a "spider," a program that crawls the web by following links from one web page to another; the index is built by copying these pages. The choice of starting pages is crucial because it establishes the basis for all subsequent rankings. In 2002 Monika Henzinger, a Google spokeswoman, gave an interview to the *Journal of Internet Cataloging*. Speaking about the process of bringing order to the Web, she made clear that Google's relevance and trustworthiness are founded upon web pages selected for their high credibility:

> *"Most people don't realize that Google finds most of its pages when our robots crawl the Web and jump from page to page via hyperlinks. [….] In general, the crawl starts out with a seed set of pages that we know and trust and we think are good. From there we look at what links are going out from that link"* (Monika Henzinger, quoted in De Groat, 2002: 21).

Again, it is only logical that these are not simply web pages, but instead the pages of institutions known to promulgate the most trustworthy and reliable information. The Google representative does not say what those web pages are, though later in her interview she confides *"[t]here is something that*

---

[5] In *The Wisdom of Crowds* James Surowieki has argued that the averaged decisions of many are almost always superior to the judgment of experts, and he attributes Google's success to this phenomenon (Surowieki 2004: 17, 72, 270). Yet Surowieki's hypothesis, when translated to search engines, is a version of the myth of automated meaning, for although it acknowledges the human sources it implies that they are simply averaged. This does not take into account the conflicting biases and agendas on the Internet, which would render all such averages suspect. In such a situation we need the judgment of professionals.

librarians and other expert information professionals could do for us. I would be very curious to get more feedback from librarians. I feel that there is a lot of knowledge and experience that they have and we are not getting enough of that" (Henzinger, quoted in de Groat, 2002: 27).

It is odd that libraries are never mentioned among the resources drawn upon to improve search engine results. After all, a harvest of all library and library-related web sites would likely be larger and of better quality than could be produced from the frequently-mentioned Open Directory Project. Google's stated intentions are consistent with such an activity. In a December 14, 2004 press release Google cofounder Larry Page is quoted as saying *"Even before we started Google, we dreamed of making the incredible breadth of information that librarians so lovingly organize searchable online"* (Page, quoted by Letwin, 2004).

It would be fairly easy to set up systems that would 'borrow' or expropriate the knowledge, expertise, and good judgment of trusted institutions. Would not library web pages be a suitable basis for an algorithm that would produce credible and trustworthy (as well as relevant) results, perhaps even the best set of 'seed pages' from which a search engine can 'start out'? The same operation could be performed on other public institutions, for example schools, universities, public radio, and public television. It could also be done with those private companies that are known to provide reliable information, such as newspapers or medical institutions. To the extent that sources of this kind play an important role in Google's search process, Google is borrowing relevance and trust from libraries and other trusted authorities, what Jose-Marie Griffiths has called *"the knowledge validating institutions"* (Griffiths, 2000: 47).

Is There Surreptitious Gathering of Recommendations from Expert Searchers?

It is well established that search engines monitor user behavior in order to improve performance. Even in the early days of search, *"[t]he Yahoo team quickly realized the value of its users' clickstreams. 'People came to our servers and they'd leave tracks,' Koogle says. 'We could see every day exactly what people thought was important on the Internet.'"* (Battelle 2005: 62, quoting Tim Koogle, Yahoo CEO). Yahoo used this information to build its directories, and by 1999 Brin and Page were apparently using the same kind of data to check the accuracy of the PageRank algorithm (Page et al., 1999: 13).

Other search engines have recently announced that they are gathering user information for similar purposes. For instance, Mooter monitors its users and adjusts search results on the fly (Roush, 2004: 39), and SNAP is *"…a new breed of search engine that ranks sites by factors such as how many times they have been clicked on, among other things"* (Battelle, 2005: 121; see also Walker, 2004). In general, *"[s]earch engines keep voluminous logs of user interactions, mainly to divine patterns to make their engines more efficient and more profitable"* (Battelle, 2005: 196). Google certainly monitors users: *"'the system that has the most users benefits the most,' said Nancy Blachman, a computer scientist and author of an independent guide to using Google (www.googleguide.com). 'Microsoft faces a tremendous challenge because Google fine-tunes its system by watching how users adjust their queries'"* (Markoff, 2004).

Most observers seem to think this monitoring is unobjectionable, worrying only about possible privacy violations. That certainly is a concern. But another implication has gone unnoticed, namely that search engines are harvesting the judgments of individual searchers.

We have already seen that Google does not wish to draw its recommendations from websites at random, but rather seeks out the most trustworthy ones. The same logic would apply to the 'recommendations' made by searchers, some of whom will be more trustworthy and otherwise valuable. Google and other search engines are able to identify IP address and more (Hinman, 2005: 23), and the cookie it deposits on each machine can record search histories (Himma, 2004: 71). It would not be difficult for Google (or any other search engine) to identify people who perform a high number of searches, and this could even be correlated with library IP addresses. Certainly there is a threshold high enough to exclude most casual users, yet keep, for instance, the computer at a busy reference desk. The selections made by these expert users could then be boosted in the rankings, even though librarians and others would not be aware that they were contributing to the authority and relevance of Google's results. Nor does there seem to be anything in Google's current privacy policy that would prevent such 'borrowing':

> *"Google collects limited non-personally identifying information your browser makes available whenever you visit a website. This log information includes your Internet Protocol address, browser type, browser language, the date and time of your query and one or more cookies that may uniquely identify your browser. We use this information to operate, develop and improve our services"* (Google, 2004a).

Due to secrecy the extent of such borrowing is unknown. Still, it is an important ethical issue, one that concerns potentially all searchers, not just information professionals. Because the Google cookie can report search histories for individual machines over time, intensive research would be an especially appealing target. A searcher's choices could be used to improve the results, which would then be delivered to the next person with a similar search, possibly a competitor. This is clearly an issue of intellectual propriety, another problematic aspect of the unannounced 'sharing' that this paper makes thematic.

Mining and Undermining the Public Resources: The New Ecology of Search

By about 2003 we find a general recognition that search engines have once and for all outstripped directories in terms of usefulness:

> *"Yahoo isn't the only directory facing criticism these days. Search engine marketers also point to the Open Directory Project (ODP) as an example of how far directories have fallen behind algorithmic search providers, both in terms of the reach and quality of the results they provide"* (Olsen and Hu, 2003).

Oddly enough, no one other than Olsen and Hu seems to have examined why the semantic aggregators (search engines) are better than the directories. The assumption seems to be that technology has replaced human effort, but this is a version of the myth of automated meaning, as a meaning-aggregating machine will never replace a meaning-creating human.

We can outline the ecological configuration of this new Internet information environment. What does it mean to say that *"Google rides on the back of human screening of web sites"* (Olsen and Hu, 2003)? The contents of any directory are public and are thus available for harvesting by search engines, which can then 'borrow' the relevant and trustworthy recommendations they find there and include them in their results. The reverse process is nowhere near as easy, as it requires human operators to first use the search engine, judge the results, and then place their choices into a directory or web page. Perversely enough, any humanly-produced refinement will once again be 'borrowed' the next time the search engine crawls the web and updates its index. And, as already noted, we can add the corollary that search engines 'ride on the backs' of individual searchers, as the selections made by search engine users can be monitored, turned into recommendations, and then incorporated into the ranking process.

The euphemistic language used to describe these activities is 'harvesting' or 'mining,' but these terms connote productive human activities that extract value from the non-human world. The systematic gathering of recommendations by search engines is, by contrast, an operation that divorces these values from the individual and institutional sources that produce them. The library profession would seem to be especially vulnerable to this unannounced transfer of knowledge on account of the quality and quantity of recommendations it makes publicly available, and because this activity is among the core services of the profession.[6]

Comparison of Library and Search Engine Cultures

How much the information environment has changed! It is frequently noted that libraries are not subject to competitive pressures, and often this is spoken of as a negative thing. As the sole purveyors of a certain kind of information, libraries have occupied a quasi-monopolistic position and thus did not need to pay attention to users' convenience. The recent ascendancy of search engines demonstrates the folly of this shortsightedness.

---

[6] Of course other entities provide the public with recommendations—for instance critics, reviewers, and bookstores. This paper focuses on libraries because search engines are not liable to draw extensively from these other sources, nor are search engines conceivably in competition with them.

But there are values other than ease of access, and in regard to these libraries are generous. The lack of competition has positive aspects, allowing traditional library culture to be collaborative and tolerant of considerable borrowing between its members. One reason for this is that librarians share a common goal –service to the users, often the public in general– which is more completely achieved by cooperation. The tendency has been to treat this kind of information as a public good. Save for limited indexing and abstracting services, most of this information about information was never subject to market competition, and libraries had no need to guard the values they produced. But while the work of librarians is usually made freely and publicly available, search engines are secretive. The contrast illustrates a fundamental change occurring in our information culture.

Because the search engine is more comprehensive (it harvests many directories and web pages), and because it delivers more convenient results, users naturally flock to this new point of access. Since there is no attribution, users come to believe in the myth of automated meaning, failing to recognize that other sources produce this meaning. So it is somewhat ironic to hear that directories

> *"... are being displaced by algorithmic search tools and commercial services that many people... now believe do a better job satisfying Web surfers and advertisers. The transformation is bringing to an end an altruistic era of human editors, who once wielded significant clout in driving traffic to web sites through recommendations made without regard for commercial considerations"* (Olsen and Hu, 2003).

In the same way it is somewhat ironic that the founders of Google would cite a human-powered directory as a source of useful information: ironic, because the links are valuable precisely because they have been judged by professional editors. The problem for the Yahoo directory (as with the Open Directory Project and all library web pages) is that it is open to be mined by all search engines, and thus equally useful to Yahoo's competitors. Despite this, Yahoo is in no position to complain, as it practices the same kind of mining.[7]

Whether it is due to users' preference for ranked results, or because directories can be mined by competitors, or simply because public sources can replace much of the work done by paid employees, it seems that Yahoo has begun to limit the work it invests in creating recommendations: *"At least one search engine marketer has said that Yahoo has scaled back on its directory editors slowly over recent months, giving people new duties or emphasizing paid search listings"* (Olsen and Hu, 2003). A related observation is made by Joseph Janes, who asks *"Why Is the Internet Public Library Broke?"* His answer is that *"[t]he IPL is a really great idea, but the bottom line is that it's not anybody's responsibility to fund a 'public library' for the entire Internet"* (Janes, 2003). This is especially true if any quality recommendations published on the Internet will immediately be siphoned off and repackaged into search engine results. The users (and administrators) must think, if search engine results are more copious and convenient, then why do we need a directory?

It is also ironic to hear the concern, common since the advent of search engines, that librarians might be 'losing their relevance.' This usually refers to the fear that librarians will lose their reputation as the best source for meaningful recommendations, so the public (or students) will no longer seek them out. Yet if the thesis of this paper is even partially correct, librarians have been 'losing their relevance' (and their credibility and trustworthiness) in a much more direct and literal fashion, as their web pages and even their individual searches are harvested by search engines.

Discussion

The distinction used throughout this paper is overly-simplistic but nevertheless useful.[8] Google and other

---

[7] The question as to how much can be 'borrowed' from commercial sites needs fuller investigation. It is possible to block search engine spiders from a website by using a robot.txt file, so commercial sites probably can, for the most part, prevent this form of harvesting. But even if private commercial sites can block spiders and other forms of borrowing, that does not weaken the thesis of this paper, as public sites (including libraries') remain open.

[8] As with all simplifications it is also a falsification and tends toward extremes. Library classification and other services are not so disinterested as herein portrayed (see ftnt. 2). Likewise I simply point to

search engines gather recommendations from two kinds of sources: one has a private interest at stake and is looking to influence search results, making those recommendations biased and untrustworthy; the other is performing a public service by making disinterested judgments about what is most relevant, credible, trustworthy, educational, useful, etc. Search engines remove from these actors all public recognition of the ethical quality of their actions, making them invisible. The reward for the unscrupulous spammer is substantial. Conversely, the reward for those whose disinterested evaluations serve the public good is small to none. This is especially true when the credit for their good deeds has now, under the myth of automated meeting, been transferred to the search engine.

When search engines draw upon this public good and make it appear as a free gift from their private company they are obscuring the source of value. Why would the public support costly institutions such as libraries, when a private entity delivers many of the same things for free, with much greater ease and abundance? Why would a government official or public administrator support public organizations that create order on the Internet? Hiding the sources that create public knowledge encourages disinvestment in them.

Consider an alternative information ecology: if it were not for the semantic aggregators, the Internet Public Library would most likely be flourishing. Does Google do a better job than the directories do? Well, in a way yes, but only because Google draws on all the resources that these directories make public.

Indeed, this seems to be a one-way information economy, where all good is transferred from the creators of meaning to the harvesters of meaning, with no recognition of human labor. This disincentive to invest in the creation of meaningful recommendations, and the corresponding emphasis on harvesting meaning from existing resources, is, I think, pernicious. It is, unfortunately, a major force shaping the information ecology today.

This is not to say that there should not be search engines. Certainly there is value in the work that Google or any competent search engine company does. Librarians know that selecting and organizing resources requires work, and this remains true even when much of the process is automated. Nor is this to say that the semantic aggregation (collaborative filtering) that Google and other search engines perform is inherently bad. On the contrary, it is a very good thing indeed. The point here is simply that some forms of collaborative filtering unfairly distribute the labor and the rewards for that good work. This is particularly true when the library culture of openness and collaboration meets the search engine culture of secrecy and competition. To the extent that the meaning produced by libraries is siphoned off and repackaged by search engines (which only the search engines know), to that extent libraries are collaborating with their competitors.[9]

Conclusion

Many reasons have been given for holding search engines to high ethical standards. Most of these have to do with the essential services search engines provide to the public. This paper adds that search engines are ethically obligated because they draw so much from public sources. To some large extent the intelligence they display is not their own, but is rather borrowed. In a less direct fashion, the trust that users extend to search engines is also borrowed, for it has been built by the ethical practices of other information institutions. For these reasons search engines' ethical obligations are far greater than previously recognized.

I have argued that libraries are especially vulnerable to this borrowing, not because I know that search engines draw extensively upon their work (secrecy prevents that), but because it is logical that they would.

It seems we have entered a new ecology of information, one where search engines capture and recycle all meaningful recommendations. After this content has been repackaged by search engines, the

---

the implications of search engines' borrowing from libraries. These simplistic claims are meant to delineate the basic issues at stake, while the reality is more a matter of degree.

[9] Obviously much could be said regarding the Google Print project and other collaborative efforts. Suffice it for now to quote Rory Letwin, who notes that it "... is not a victory for libraries if the real meaning of this development is simply the transfer of all this information out of our humanistic institution and into the marketplace" (Letwin, 2004).

original sources of meaning appear irrelevant and unnecessary. So even in the unlikely scenario where Google or another search engine creates an ideal aggregation system, one that gives appropriate weight to the disinterested judgment of public institutions and avoids all the corruptions that lay indexers can introduce, even this ideal system will still contribute to the corruption of the information infrastructure. It will do so by mining, exploiting, and ultimately undermining support for all publicly-motivated efforts. Because search engines provide no acknowledgement of these sources they encourage the myth of automated meaning, the belief that these things are produced by technology alone and could be provided without the efforts of people and institutions. If this myth comes to be universally believed, the institutions that create these valuable recommendations will decline. Then much of the trustworthy information that users expect will no longer be there for search engines to harvest, and the world will be an intellectually poorer place.

References

Achenbach, Joel: "Search for Tomorrow," Washington Post, Sunday, February 15, 2004, p. D01.

American Library Association, "Code of Ethics" http://www.ala.org/ala/oif/statementspols/codeofethics/codeethics.htm. (1995) Last accessed March 6, 2006.

Battelle, John: The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture. New York, Portfolio (Penguin Group) 2005.

Belle, Jeff: "ODP Shakes Up Search Directory Content Revenues". EContent. February/March 2001, 49-52.

Blanke, Tobias: "Ethical Subjectification and Search Engines: Ethics Reconsidered," International Review of Information Ethics, vol. 3 (June 2005), pp. 34-38.

Brooks, T. A.: "The Nature of Meaning in the Age of Google" Information Research, 9 (3) paper 180 (2004) http://InformationR.net/ir/9-3/paper180.html

Caufield, James: "Where Did Google Get Its Value?" portal: Libraries and the Academy, vol. 5, no. 4 (October 2005), pp. 555-572.

De Groat, Greta: "Perspectives on the Web and Google: Monika Henzinger, Director of Research, Google," Journal of Internet Cataloging, vol. 5, no. 1, 2002, pp. 17-28.

Fallows, Deborah: "Search Engine Users: Internet Searchers Are Confident, Satisfied and Trusting -- but They Are Also Unaware and Naive," (Pew Internet & American Life Project), January 23, 2005. Available at http://www.pewinternet.org/PPF/r/146/report_display.asp. Last accessed January 12, 2006.

Goh, Dion H., and Rebecca P. Ang: "Relevancy Rankings: Pay for Performance Search Engines in the Hot Seat," Online Information Review, vol. 27, no. 2, 2003, pp. 87-93.

Google (2004a): "Google Privacy Policy" http://www.google.com/privacy.html. Last accessed September 5, 2005.

Google (2004b): "Our Search: Google Technology," http://www.google.com/technology/. Last accessed September 5, 2005.

Griffiths, Jose-Marie: "Deconstructing Earth's Largest Library," Library Journal, August 2000, pp. 44-47.

Hansell, Saul: "Google and Yahoo Settle Dispute over Search Patent", New York Times, August 10, 2004, p. C6.

Harvey, Fiona: "New Media," Financial Times Limited (London) March 30, 2004, p. 8.

Himma, Kenneth Einar: "Legal, Social, and Ethical Issues," The Internet Encyclopedia, Hossein Bidgoli, Editor-in Chief. Hoboken, New Jersey, John Wiley & Sons Inc. 2004, pp. 464-476.

Hinman, Lawrence M.: "Esse est Indicato in Google: Ethical and Political Issues in Search Engines," International Review of Information Ethics, vol. 3 (June 2005), 20-25.

Janes, Joseph: "Internet Librarian: Why Is the Internet Public Library Broke?" American Libraries, March 2003, p. 86.

Larsen, Elizabeth: "Finders Keepers: How Search Engines Control Our Access To Information," Utne Reader, May/June 1999, pp. 22-3.

Letwin, Rory: "On Google's Monetization of Libraries," Library Juice, Vol. 7 No. 26 (December 17, 2004). Available online at http://libr.org/Juice/issues/vol7/LJ_7.26.html. Last accessed May 25, 2005.

McRae, Hamish: "How Google Helped to Change the World," The Independent (London), August 18, 2004, p. 27.

Marable, Leslie: (Consumer Reports WebWatch), "False Oracles: Consumer Reaction to Learning the Truth about How Search Engines Work. Results of an Ethnographic Study." Consumer WebWatch, June 30, 2003. Available online at http://www.consumerwebwatch.org/dynamic/se *arch-report-false-oracles-abstract.cfm. Last accessed May 15, 2005.*

Markoff, John: "The Coming Search Wars," The New York Times, February 1, 2004.

Moxley, David, Joni Blake, and Susan Maze: "Web Search Engine Advertising Practices and Their Effect on Library Service," The Bottom Line: Managing Library Finances, vol. 17, no. 2 (2004), pp. 61-65.

Neuberger, Christoph: "Funktionene, Probleme und Regulierung von Suchmachschinen im Internet," International Review of Information Ethics, vol. 3 (June 2005), pp. 3-17.

Nissenbaum, Helen and Lucas D. Introna: "Shaping the Web: Why the Politics of Search Engines Matter," in The Internet in Public Life (ed. Verna V. Gehring). Oxford, Rowman and Littlefield Publishers, Inc, 2004, pp. 7-28.

Notess, Greg R: "New Databases from the Internet Search Services," Database Magazine, vol. 20, no. 2 (April/May 1997), pp. 72-75.

Olsen, Stephanie, and Jim Hu: "The Changing Face of Search Engines", CNET News.com, March 24, 2003. Available at http://news.com.com/2100-1032-993677.html?tag=fd_lede1_hed. Last accessed May 9, 2005.

Page, Larry, et. al.: "The PageRank Citation Ranking: Bringing Order to the Web", November 11, 1999 http://dbpubs.stanford.edu:8090/pub/1999-66. Last accessed October 2, 2005.

Plato: The Republic, (Paul Shorey, translator), Cambridge, Massachusetts, Harvard University Press 1982.

Rieder Bernhard: "Networked Control: Search Engines and the Symmetry of Confidence," International Review of Information Ethics, vol. 3 (June 2005), pp. 26-32.

Roush, Wade: "Search Beyond Google," Technology Review, vol. 107, no. 2 (March 2004), http://www.technologyreview.com/articles/04/03/roush0304.2asp. Last accessed July 17, 2005.

Scott, David M.: "Paying Your Way to the Top," Econtent, May 2003, pp. 32-38.

Search Engine World: "Search Engine History Part 2" (undated). http://www.searchengineworld.com/engine/players2.htm. Last accessed January 6, 2005.

Sherman, Chris: "Google Introduces Web Directory Using Netscape's Open Directory Project Data." Information Today, March 27, 2000. Available at http://www.infotoday.com/newsbreaks/nb000327-1.htm . Last accessed May 14, 2005.

Spencer, Stephan: "Casing the Competition," Catalog Age, April 1, 2004, pp. 35-7.

Surowieki, James: The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations. New York, Doubleday 2004.

Welp, Carsten, and Marcel Machill: "Code of Conduct: Transparency and the Net: Search Engines," International Review of Information Ethics, vol. 3 (June 2005) p. 18.

Walker, Leslie: "Online Search Universe Is Expanding," The Washington Post, October 21 2004, p. E01.

Wiggins, Richard: "The Privilege of Ranking: Google Plays Ball," Searcher, vol. 11, no. 7 (July/August 2003), p. 23.