IN5480 Lecture notes 1 September 2021

Updated: 30 August 2021

Note: This document will be revised for legibility and clarity. Please do not hesitate to send feedback about the document, info about errors, suggestions for improvement and so forth. I hope to revise and improve this text iteratively (of course without extending it too much - or adding new sections), in the same way as you will do with your individual assignment and the group assignment.

1 Introduction

The 2020 annual conference by LDO (Likestillingsombudet - Equality and Anti-Discrimination Ombud) was “Kunstig intelligens: Fare eller fremgang for likestillingen?”. The theme AI and universal design is on the political agenda – where the stakes are high. Will AI make our society more or less fair? This is an example of AI and Universal Design that are discussed on the societal level, macro level. What challenges and threats does AI imply for equality, inclusion and anti-discrimination? In what ways can we meet the challenges that arise? What opportunities arise? These and other questions are discussed today.

Inclusion and Universal Design are important concerns when working with any systems, also for AI systems. In what way does the system we are working with include users - and in what ways does it exclude?

This lecture will attempt to discuss these two questions:

1. What *challenges* are there when it comes to AI and universal design/equality and discrimination?
2. What is our challenge to each other in this are (students and teachers in HCI/Informatics), given our central role in human-centred/user centred/usability/participatory design? What can we do?

These two questions are broad, and we can address them in different ways. We start off with some background on Universal Design – and move towards some challenges – and how to meet the challenges.

2 Universal Design

Universal Design is concerned with questions concerning inclusion, and hence it is about ethics. There are many ways to talk about ethics when developing systems, testing systems and deploying systems. Both on the individual, micro-level, organizational or meso-level and the macro level - the society. Inclusion, discrimination and fairness are often high on the agenda when discussing ethics; such as in this FATE framework (this is from the book Interaction Design; beyond human computer interaction by by Sharp Rogers and Preece (5th edition)):

Fairness: Is it fair, data handled fair, without discrimination?
Accountability: Is it correct, valid?
Transparency: Is it visible to the user what is going on, how decisions are made?
Explainability: Explanations, is it possible to understand by us?

Another framework is proposed by Jutta Treviranus; presented in her thesis: *"The three dimensions of inclusive design: A design framework for a digitally transformed and complexly connected society."* The three dimensions of the framework are:

1. Recognise, respect, and design for human uniqueness and variability.
2. Use inclusive, open & transparent processes, and co-design with people who have a diversity of perspectives, including people that can't use or have difficulty using the current designs.
3. Realise that you are designing in a complex adaptive system.

Here in Norway, and in Europe, we have a long history of working with fairness and inclusion; both of the built environment, of products and services. As more and more environments, products and services are digital - or in part digital, ICT and inclusion become increasingly important.

The umbrella term here in Norway is "Universell Utforming" or in short UU. Universal Design (UD) is the English term for this. In other parts of the world terms like Inclusive Design (ID), barrier free design, design for all - and accessible design are applied. Universal Design is both discussed on the macro level (politics, policies and UN for example), meso level (organisations and domains), micro level (the individual use of ICT).

UD is about inclusion of all. An aim when making any new systems or functionality is that the "new" systems is not creating new barriers for users, and the barriers that already exist shall be reduced. It is operationalized for certain technologies today, such as web technologies and app technologies. We all know the WCAG 2.1, (Web Content Accessibility Guidelines) standard by WAI (Web Accessibility Initiative) at the W3C (World Wide Web consortium), that are used as guidelines for ensuring accessibility of services on the net.

The four principles within WCAG are:

1. Perceivable: Information and user interface components must be presentable to users in ways they can perceive.
2. Operable. User interface components and navigation must be operable.

3. Understandable: Information and the operation of user interface must be understandable.
4. Robust: Content must be robust enough that it can be interpreted by a wide variety of user agents, including assistive technologies.

These guidelines are used during the design and evaluation of systems to ensure that systems are available to as many users as possible, without special adaptation or aid (hjelpemidler), in as many contexts as possible.

The term “Universal Design” was coined by Ronald Mace. It is a term used for the design of buildings, environments and products, so that these are accessible for all. The seven principles are:

1. Equitable use.
2. Flexibility in use.
3. Simple and intuitive.
4. Perceptible information.
5. Tolerance for error.
6. Low physical effort.
7. Size and space for approach and use.

Ron Mace was mostly concerned with the built environment, but these principles has also been applied and used in the area of digital technologies.

WCAG is specific guidelines for web content; and therefore, more specific than the seven principles above.

UCD (user-centred Design), HCI (Human Computer Interaction), and PD (Participatory Design) are concerned with the human side of computing, and work for inclusion of the user group or population in question. Within these research fields, an attempt is to foreground the “user”, the “human” and the “participants” – and work from there. How to make the technologies “fit” the user? How to make the technology usable and useful for this or that user or user group? We may say that Universal Design is in a way inherent to these research fields.

A new EU directive about inclusion is the Web Accessibility Directive (WAD). You can read more about this at Digitaliseringsdirektorated/DIFI, and learn about “new” initiatives to get more universal designed services and solutions; such as declarations and options to report problems/feedback on accessibility questions – in addition to the requirement of texting video material.

The work within UD is often about making sure that people and groups with physical, cognitive or motor disabilities (blind, deaf, wheelchair users) etc are included and not excluded. However, the idea - and ideals behind UD is broader. For example, situational disability (tired users, power outage/pitch dark) etc happens for “all”. And of course, old age, where body functions and cognitive capacities get reduced. There is a social model of disability and ability at the core of these discussion, where disability is something that is created in the meeting between citizen and the society. Stairs make you disabled in a wheelchair for example, or you are disabled by a combination of cold weather, gloves and

the mobile telephone screen. Diana Saplacan at IfI has worked with this in her PhD project, with the final title: *Situated Abilities: Understanding Everyday Use of ICTs.*

Interestingly, the development of AI technologies - current and previous research projects - have often focused on inclusion and universal design. The MECS project here at IfI is one such example, where we investigated the possible use of robots at home for the elderly living independently at home. A group that potentially can stay longer at home with some help from various (AI) technologies such as “robots” moving around in the home. The RHYME project is another example, where interfaces and novel interaction mechanism of playing music for all was addressed.

Other examples of technologies that “initially” was designed and developed for people that had challenges of using “standard” solutions are:

- Speech synthesis and speech recognition. Audio books for visually impaired.
- Remote control: In order to help out with motor impairments.
- Video conference: Used by visually impaired for assistance.
- Eye tracking/detection: Used as input for motor impaired users.

So one part of “AI” and “inclusion” is the positive side for inclusion - where AI technologies are developed, applied and used to include people and users who are else excluded.

What is peculiar with AI and UD? The are opportunities for including more people, even all! On the other side, there are opportunities to exclude people when statistics and numbers are crunched about people. AI and UD is like a double edged sword, or a dilemma that we will look into one here.

2.1 Challenge - Way of talking/thinking about machines

We have discussed ways of talking about machines that is similar to the way of talking about us humans. For example, that both humans and machines “understand” and both humans and machines “learn”. When this is the case, challenges arise in determining whose understanding is “correct”, “true” or “right”? We have stated previously, with assistance from Searle, that computers do not “really” understand in the same way as humans understand. Nevertheless, the rhetoric’s of AI – and the language used at specific interfaces is often using the concept understanding also for the computer. Output from speech recognition system may for example state that: “I do not understand what that is?” or “Help me to learn more about this by providing examples”.

In what way does this matter? Is it only a manner of speaking?

When things go “wrong”, when errors occur, something is missing or in the way for proceeding – whose understanding is the correct one? The user´s understanding or the system´s understanding? Or are they perhaps both correct, but from different perspectives? Or both wrong?

One problem with using the same word “understanding” and “learning” for both the computer and the human is that it is means different things to “understand” and to “learn” in the two instances. Machine learning and understanding is about numbers, input and processing –

and not about meaning and making sense. Human learning and understanding is about meaning and making sense. It is in a way like comparing cats and automobiles; both are names of “things” in the world, but they are of very different kind.

A traditional model of a computer system is made up of three parts:

- Input: Something is sent “in” to the system by the user.
- Processing: Something is done with the input, calculation/manipulation.
- Output: Something is sent “out” to the user from the system.

With “traditional” systems, the user has overview of what the input to the system is. What happens when the input to the system is “hidden” from the user? Do you have any example of “smart systems” or “AI” systems that operate in that way? Is it then possible for the user to understand that part of the system, if there is no information about what the input is?

When using a traditional text editor, or using a traditional telephone service, there is no need for the user to understand the inner workings of the application, service, CPU, base station, wireless protocols and so forth. If you do not get the correct output when pressing the letter, “K” on the keyboard, or you do not hear the person at the other end of the telephone line, you understand that something is faulty or not working with the technology. What is the difference with AI services that “make decisions” and “predictions” for you?

One difference is about ambiguity and predictability. In traditional “non AI” system, there is usually a direct coupling between input and output. You press the letter “K” and this is displayed on the screen. However, when interacting with AI system, there is not necessarily such direct coupling between input and output. Then the question arise: In what way is it possible for the user to understand this unpredictability? In what way is it possible for the user to determine when something goes wrong?

Everyone who is affected by the introduction of an 'AI service' - should have a real opportunity to understand, criticize and influence the design and use of technology so that it does not compromise with principles of inclusion and universal design. This is a thesis that need some arguments. What are they? If this thesis is not the case, who should understand and criticize and influence? Some “expert” or authorities? Legal experts? Engineers?

To say that “the machine understands” and “the machine learns” may trivialize what is meant by human understanding and human learning. What can we learn from the statement “the machine learns” and “the machine understands me”?

2.1.1 Understanding, principles and guidelines

Understanding how to use a computer has traditionally been important in HCI. For example, in the WCAG 2.1, one of the four principles are «Understandable», and this is described as:

Information and the operation of user interface must be understandable.

In the Microsoft guidelines for Human AI Interaction, the second guideline is:

Make clear how well the system can do what it can do. Help the user understand how often the AI system may make mistakes.

A question that arise from this is: Is it possible for the designer to know this in advance – what the various user in various use situations and with various user tasks understands? What role does user involvement and user participation have?

2.2 Challenge about statistics and the human

There are potential threats of AI with respect to «outliers». This topic has been in the press and many public efforts have emerged. Books such as Weapons of Math Destruction (O'Neil, 2017) and Sufiya Ujoma Noble (2018) and Virginia Eubanks (2018) powerfully framed the role of AI in amplifying and demonising poverty and automating racial bias. All three authors highlight the systemic vicious cycles of discrimination experienced by individuals.

“If you are unique (and aren’t we all), numbers are not our friends” (Trevanius) is one statement that needs some explanations. Numbers are potentially very useful for many different purposes within sciences - as well as in everyday life. A meeting starts at 12:15! Or this is 2103 kilogram, whereas this is 2108 kilogram. However, since we are all “unique”, there are some challenges of including all the “uniqueness” of humans into numbers.

Jutta Trevanius explain further this “uniquness”: “.....to recognise the uniqueness of each individual. We are each an irreducible and evolving complex adaptive system of characteristics and needs. This uniqueness and wild, organic diversity have been inconvenient when it comes to designing products, communication, environments, or policies. It defies mass production, mass marketing, mass communication, mass education, as well as simple and straightforward public policies.” Even within one individual, there are idiosyncrasies and differences from situation to situation; after all we are doing a bit different talking and acting here in this lecture for example compared to 3 am Sunday morning at a celebration.

When developing systems in general, very often we tend to design for the “majority”, where limited resources are applied to reach a broad population. The 80/20 principle, or the Pareto Principle, where 20 % of the resources is used to reach 80 % effect is well documented in various disciplines.

This is not an inclusive process; since potentially 20 % (or so) of the population is left out. Another possibility is to aim from the start to reach 100 %, and as a process continuously ask “is anybody left out with this solution?”, and if so, address it. This is at the core of Universal Design. However, this is a challenge when it comes to working with the “average” and modelling probability”, or the “normal” user.

For the argument here, we will only invoke every day, basic statistics about the “data” that is the “input” for machine learning algorithms.

Here we come into some basic statistics:

Distribution. Data about a population has a distribution. Often, we mean probability distribution when talking about distributions. The most common is normal distribution (other names Bell curve (looks like a bell) giving a Gauss Curve), Binomial distribution (two outcomes, true and false) and Uniform distribution (throw one dice many times). In probability theory, the normal distribution is a common continuous probability distribution. Normal distributions are often used in the natural sciences. It is also used in social sciences and humanities.

Representations. Data is about something, and often represented by numbers in computers. It is a lot happening from the “real world” to the representation of the real world, or representing aspect of the real world. How happy are you on a scale from 0 to 10? Let's say you answer 7. This number 7 represents the level of happiness for you (when you answered the question, at that particular place, and at that particular time, in that particular situation). Or, a face recognition system determines by a visual representation of your face that you are 3.27 on this happiness scale.

Average (gjennomsnitt) can be the mean, median, and mode. The "mean" is the "average" we are used to, where we add up all the numbers and then divide by the number of numbers. In the list of numbers, the "median" is the "middle" value. You list the numbers in numerical order from smallest to largest in order to find the median. The value that occurs most often is called the “mode”.

Standard deviation is a number that tell how measurements for a group are spread out from the average (mean). A high standard deviation means that the numbers are more spread out. A low standard deviation means that most of the numbers are close to the average.

Outliers: This are data points that differs significantly from other data points in a data set about a population.

Working with statistics and probability calculations are very valuable for production processes and manufacturing for example. There are some challenges when working to with inclusive design however. Two of the challenge are the notion of the “average” and the notion of the “Normal” in Normal distribution. This may foster comparison of what is above or below the average, as well as what is not-normal or even ab-normal.

"Lies, Damn Lies and Statistics" is the name of episode 21 in the first season of NBC drama *The West Wing.* Many people have used this phrase before, what does it mean? One interpretation is that the use of statistics can be done in many different ways, also for misleading and even lying.

n=1 is another way of saying that each human or user of a system is unique. This is then a challenge when one system is to be developed for a population of n=3.800.000 users!. What are the challenges?

When modelled in statistical models, human traits (exercise, mood, shape, reasoning, navigation activates etc.) get measured and represented in numbers. Outcomes of this is very average, standard deviation - and outliers. How to work with this dilemma?

There is no “final solution” to be found here about inclusion and AI. However, it is about critically asking questions during development of systems, and bringing in a variety of users!! Another way is to get to know, and work with both the WCAG principles and the WAD actively. And of course, to be enthusiastic about participation and users diversity.

3 What can we do – learn more about?

We have addressed two challenges; the way we talk about machines – and statistics and “normal” distribution. In addition, we had the question:

What is our challenge to each other in this are (students and teachers in HCI/Informatics) in this area, given our central role in human-centred/user centred/usability/participatory design?

Continue working with user centred design! And participatory design!

Bring awareness to the language we use when talking about machines and the use of machines.

What is your challenge in this area?

Jo