



universität  
wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

## Can Computers be Non-Binary?

Studying the role of gender in the four research areas of VRVis Vienna

verfasst von / submitted by

Cyntha Wieringa

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of  
Master of Arts (MA)

Wien, 2021/ Vienna, 2021

Studienkennzahl lt. Studienblatt /  
degree programme code as it appears on  
the student record sheet:

A 066 906

Studienrichtung lt. Studienblatt /  
degree programme as it appears on  
the student record sheet:

Masterstudium Science-Technology-Society

Betreut von / Supervisor:

Assoc.-Prof. Dr. Maximilian Fochler



# Acknowledgements

At the moment of writing these acknowledgements, 11<sup>th</sup> of March 2021, it is almost exactly a year ago that Austria went into its first lockdown in response to the developing COVID-19 infections in the country. At the time, I viewed the six-week lockdown that followed as a quick fix to a situation that would return to normal shortly after. Little did I know that I would not have another physical class or see my student peers in real life for the rest of my masters. The pandemic made many aspects in life more challenging, and writing a master thesis was no exception to this. Therefore, in this time where many people had to adapt to quickly changing situations, a sincere thank you to those who helped me throughout the thesis process is in its place.

First, a big thank you to my supervisor Maximilian Fochler. With your many roles as program director, researcher, teacher, and father I cannot imagine how challenging it must have been to keep all those balls in the air during these times of home confinement. Regardless of those challenges, I never felt like they took away from the quality of your feedback and patience. I always felt your commitment to my thesis project, and was therefore never shy to ask for your help or advice. More you cannot ask from a supervisor.

Thank you to Brigitte Rutzer from the TU Wien for being involved with this project from the start. I highly appreciate you being there during the presentation day at VRVis to serve as a gatekeeper for the boundaries of this thesis, and offering your expertise when I needed it.

Another big thank you to the people at VRVis Vienna and in particular to Gerd Hesina, Werner Purgathofer, and Franziska Steyer-Beerman. Thanks to Gerd and Werner for giving me the opportunity to be an actual researcher in the field, and being patient when the COVID situation put certain delays on this project, and thanks to Franziska for always being my first point of contact within VRVis whenever I needed something. Of course, also a huge thank you to all the people at VRVis who showed interest in my research and who were always more than willing to participate.

Lastly, I am very grateful for all the support the department of Science and Technology Studies has offered during the whole thesis process. I am grateful for my peers

who gave me helpful feedback in the Coping seminars, and engaged in helpful discussions about our shared struggles and how to overcome them. Also, not unimportant to mention, is my gratitude for the Teaching Assistant position I had at the department. It allowed me the luxury of having a workspace outside of my home, which hugely improved my workflow and motivation. I know it was a luxury that many people in these times did not have and so I did not want to take it for granted.

# Table of Contents

<b>Acknowledgements .....</b>	<b>i</b>
<b>1. Introduction .....</b>	<b>3</b>
Structure of thesis.....	4
<b>2. State of the Art .....</b>	<b>6</b>
<i>Defining sex and gender</i> .....	6
This rise of the term gender.....	6
Definitions .....	8
<i>Gender Aspects in Science and Technology</i> .....	10
Gender and science .....	11
Gender and technology.....	15
<i>Responsible Research and Innovation</i> .....	21
<b>3. Sensitizing Concepts: methodologies for analyzing gender .....</b>	<b>26</b>
<i>Sex and gender analysis</i> .....	26
Approaches for reducing bias in AI .....	27
Pathways for studying sex and gender .....	29
<i>Gendered Innovations</i> .....	30
Relevant case studies .....	31
Gendered Innovations suggested methodology for gender analysis .....	33
<i>Four computer scientist ‘traps’</i> .....	36
<b>4. Materials and Methods.....</b>	<b>38</b>
<i>Case Description</i> .....	38
Visual Computing.....	38
The VRVis company .....	40
The four different project areas.....	41
VRVis as initiator for this thesis project .....	43
<i>Research Questions</i> .....	43
<i>Qualitative Interviews</i> .....	45
<i>Analysis</i> .....	49
<i>Ethics</i> .....	50
<b>5. Empirical analysis .....</b>	<b>52</b>
<i>Technology design as practice: an overview of VRVis</i> .....	53
Visual Analytics.....	54
Smart Worlds .....	57
Immersive Analytics .....	61
Complex Systems.....	65

<i>Detecting the social: finding the human in the binary</i> .....	67
Talking about gender .....	68
Influence of the designer .....	73
Black-boxing.....	76
Preventing bias in machine learning.....	77
Prototyping and Responsibility.....	83
<b>6. Discussion.....</b>	<b>87</b>
<i>Visual Analytics</i> .....	87
<i>Smart Worlds</i> .....	89
<i>Immersive Analytics</i> .....	91
<i>Complex Systems</i> .....	93
<b>7. Conclusion.....</b>	<b>97</b>
<i>Main findings</i> .....	97
Conceptualizations of gender .....	98
Virtual Reality and Machine Learning.....	100
Prototyping and Responsibility.....	102
<i>Further research</i> .....	103
<i>Experience as a researcher</i> .....	105
<b>References.....</b>	<b>107</b>
<b>Abstracts.....</b>	<b>115</b>
<i>English version</i> .....	115
<i>German version</i> .....	116

# 1. Introduction

Contemporary society leaves hardly any aspect of human lives untouched by technology. Visibly, we constantly carry around tech in our pockets and spend most of our working days behind a computer screen. On the other hand, invisibly, through datafication and algorithms, decisions are constantly made for and about us. The COVID-19 pandemic has further made visible the reliance of society on technology. Examples include the medical interventions that were quickly developed during the pandemic, and allowing education at all levels, different types of industry, conferences, and social happenings to function whilst people were confined in their homes during the lockdowns. Although technology tends to offer a sense of neutrality and objectivity because it was developed through computer programming and raw calculation (Bath, 2014), this definitely is not always the case. A big topic within the field of science and technology studies (STS) is showing how technology can both be shaped by and reinforce negative social structures such as harmful racial and gender stereotypes.

When it comes to developing technology, even though men and women have differing needs, end-products are often skewed towards the male-default. For example, female drivers have a 47% higher chance to sustain severe injuries during car crashes because crash dummies are based on average male proportions (Bose, Segui-Gomez & Crandall, 2011). This also shows in other technologies such smart phones on average being too large for female hands and Virtual Reality (VR) headsets being too large for female heads (Criado Perez, 2019). An explanation that is given for the development of this male-skewed technology is because people active within the Science, Technology, Engineering and Mathematics (STEM) fields are predominantly male. “When it comes down to the tech that end up in our pockets, it all comes down to who is making the decisions. And like the world of venture capitalists, the tech industry is dominated by men” (ibid, p. 180).

Because awareness about such gendered factors in science and technology is rising, a growing number of initiatives vouch for including sex and gender analysis in innovation processes. Research shows that including sex and gender analysis in experimental and technological design “[...] foster[s] scientific discovery, improve[s]

experimental efficiency and enable[s] social equality” (Tannenbaum et al., 2019, p. 137). Therefore, frameworks such as Responsible Research and Innovation (RRI) are of increasing relevance. Furthermore, gender is also becoming an important topic in other layers of science such as funding agencies. These agencies put external pressures on scientific and research institutes to incorporate sex and gender analysis into their research design. Since these criteria are at times handled quite rigorously, ignoring them can therefore have negative consequences such as the loss of funding.

In line with these initiatives and types of research, VRVis Zentrum für Virtual Reality und Visualisierung forschungs -gmbh, a visual computing company located in Vienna, was looking for a master student who could write their thesis about the role of gender in the technology development at the company. Part of the reason being that this was strongly encouraged by the funding programme VRVis is part of. The funding programme, called COMET, aims to fund projects that carry out high-quality collaborative research between industry and science that is situated more on the industry side than university groups would normally be. It was by request of COMET to encourage research about the possible role of gender within the projects it funds. Within VRVis, the work and projects are divided between four different research areas: Visual Analytics, Smart Worlds, Multiple Senses, and Complex Systems, and COMET funds projects in all four of these areas. Therefore, this thesis aims to uncover the answer to the question: *Does gender play a role within the development of technology in each of the four different research areas at VRVis? And if so, how?* In order to answer this question, separate subquestions will focus on the conceptualizations of gender the VRVis researcher hold, what steps they take in order to prevent or counteract potential gender issues in their work, how much importance is given to gender considerations in projects, and where in the project process they identify possible gender aspects playing a role.

### Structure of thesis

This thesis is structured along seven main chapters. The next chapter sets out different strands of STS literature that serve as a context in which the research in this thesis can be placed. It explores the historical emergence of and defines the concept of gender. Then it

engages in an exploration of specific STS literature that focuses on the role gender plays in the development of science and technology, and how science and technology can have an influence in determining gender relations in society. The chapter finishes with an explanation of the Responsible Research and Innovation framework, which, amongst others, has a focus on developing gender inclusive research and innovation.

Chapter 3 of this thesis elaborates on existing methodologies for analyzing sex and gender in specifically computer science and engineering fields. These methodologies come from three specific pieces of literature: (1) two Gendered Innovations reports written by the European Commission describing specific case studies in which sex and gender analysis led to more socially robust research and innovation, (2) a publication in Nature by Tannenbaum et al. (2019) that describes when and how to engage in sex and gender analysis in research, and (3) a book chapter from Corinna Bath (2014) describing four 'pitfalls' computer scientists could fall in which leads to the development of gendered technology.

Chapter 4 gives a broad elaboration of the materials and methods utilized for the research in this thesis. It starts with a description of the visual computing field and the VRVis company. This is then followed by the research questions asked in this thesis, and by a description of the method used to answer them, namely qualitative interviews.

Chapter 5 features the empirical analysis of the conducted interviews. It starts with a broad description of the specific practices and projects within each of the four VRVis research areas, which is then followed by an elaboration of the findings ranging from the conceptualizations of gender the VRVis researchers have, to specific practices they employ to circumvent any possible gender bias in their work.

Because it was a specific wish of VRVis to look at each of the four different research areas instead of one specific one, chapter 6 will tie the findings to each of the specific areas, and elaborates how they associate with the literature described in chapter 3.

Finally, chapter 7 shortly summarizes the main findings of this thesis, and makes some suggestions for possible future research. Then it reflects on the experiences I have had as a researcher during this thesis process, as the COVID-19 pandemic has had overall influences that I think are relevant to share for anyone who reads this thesis after these challenging times.

## 2. State of the Art

The following chapter includes an exploration on existing literature and research that can serve as the context for the study at the core of this thesis. The exploration includes literature from the wider social sciences but maintains a special focus on the field of science and technology studies (STS). The main aim of this chapter is to get sensitized to important concepts related to the case study of this thesis, and to identify a research gap to which it can add.

### Defining sex and gender

The categorization of 'female' and 'male' is a tale as old as time. The process of figuring out the differences, conceptualizations, and roles in society of the two sexes has differed over time and space but seems to have been present throughout different periods. Think about Adam and Eve, the multiple waves of feminism in the twentieth century, and even the different roles of females and males within hunter-gatherer societies. However, even though this process has been around for a long time, the term 'gender' is a fairly new concept. In an extensive bibliography research of 12.000 titles that were published between 1900 and 1964, conducted by Aldous and Hill (1967), it came forward that the term 'gender' was not used a single time. In 2017, the website of the World Health Organization defined gender as "[...] the socially constructed roles, behaviors, activities, and attributes that a given society considers appropriate for men and women" and sex as "[...] the biological and physiological characteristics that define men and women". The following section wishes to explore how the term 'gender' went from hardly being mentioned in the early twentieth century, to an important concept in our current day society. How can gender and sex be defined? How do these two concepts differ, and how are they similar?

### This rise of the term gender

As stated above, the term gender was, at least within the scientific community, not used much in the early twentieth century (Aldous & Hill, 1967). During this time, the words 'gender' and 'sex' were often used interchangeably and purely referred to someone's

biological sex (male of female). Twentieth century sexologist John Money declared in 1973 that he was one of the first scientists to draw a distinction between sex and gender in the 1950s, and certainly the first person to define it in print. In 1955 Money defined the words 'gender' and 'gender role' as:

"All those things that a person says or does to disclose himself or herself as having the status of boy or man, girl or woman, respectively. It includes but is not restricted to sexuality in the sense of eroticism. Gender role is appraised in relation to the following: general mannerisms, deportment and demeanor; spontaneous topics of talk in unprompted conversation and casual comment; content of dreams, daydreams and fantasies; replies to oblique inquiries and projective tests; evidence of erotic practices and, finally, the person's own replies to direct inquiry." (Money, 1955, p. 397)

As is becoming clear from this quote, Money started to view gender in more of a social context, and started to focus on the societal constructs beneath the expression of one's biological sex. His work introduced terms such as 'gender identity' and 'gender role', which in turn started to pull apart sex and gender. Money (1973) continues to explain that the need to have a term like gender came from the challenges presented by hermaphroditism, which is a type of intersex malformation leaving an individual to be born with both ovarian and testicular tissue. He found that sex fails to describe those individuals born with hermaphroditism on whom sex reassignment is not possible until after early infancy. Furthermore, Money (1973) pointed out that there are a variety of civil and legal roles the term 'sex' has to play that have nothing to do with one's genitalia and found the term therefore not multiple enough. In his own words "So as not to confuse the sex of the genitalia and their activities with non-erotic and non-genital sex roles and activities that are prescribed culturally and historically, I settled on the term, gender role, as the all-inclusive one, as defined above." (ibid, p. 398). So, to conclude, in Money's terms sex is seen more as the biological qualities, such as the reproductive organs and sex chromosomes, of women, men, and intersex whilst gender encompasses more of the socio-cultural associations and attitudes.

Although Money already laid the groundwork of defining gender and pulling the term apart from sex in the 1950s, his work only became popularized in the 70s. This popularization is often connected to the rise of the mainly US-based feminist movement in the late 60s and 70s, and to feminist scholars using the term "[...] as a way of

distinguishing “socially constructed” aspects of male–female differences (gender) from “biologically determined” aspects (sex)” (Haig, 2004, p. 87). Feminists during that time, having acquired their right to vote, started to fight for further emancipation of women’s rights with a main focus on obtaining equal treatment as men (Baxandall & Gordon, 2008). Therefore, the goal of using the term gender in scientific publications by feminist scholars was to make visible the previously invisible or ignored social structures that shaped the roles of males in females in society. Haig (2004) studied the occurrences of the terms gender and sex in scientific publications between 1945 and 2001 and concluded that by the end of this period the uses of gender highly outnumbered the uses of sex specifically in the social sciences, arts, and humanities. He additionally assigned this increase to a rise in publications within the field of psychology concerning individuals who did not conform to sexual stereotypes such as hermaphrodites, transgenders, homosexuals, feminine boys, and masculine girls. All in all, it is possible to conclude that the terms gender and sex were pulled apart and given different meanings both for feminist social needs as described by Haig (2004), and because it became clear that the term sex was not multiple enough to describe individuals with a more complex self-identity than just ‘female’ or ‘male’ (Risman, 2018).

### Definitions

The field of gender studies quickly developed, and additionally other scientific disciplines began to see gender as an interesting topic for potential research. Within STS, big names that have worked on the topic are Donna Haraway (1989) and Londa Schiebinger (1989), who each have done research about certain gender aspects within the development of science and technology. Gender was furthermore picked up by institutions outside of science, including the European Commission. In 2011 they established a research group, of which Schiebinger took part, that focused on the topic ‘Innovation through Gender’ (Stanford University, 2011). The final report of the project ‘Gendered Innovations’ was published in 2013, with a renewed version ‘Gendered Innovations 2’ coming out in 2020. Through the means of researching the role of gender in different case studies in domains such as engineering and technological development, health and medicine, environment,

and food and nutrition, the report aims to show how raising awareness on certain gender aspects improves science as a whole.

“The case studies demonstrate that differences between the needs, behaviors and attitudes of women compared to men really matter, and accounting for them in research makes it relevant to the whole of society. They also show that these differences can vary over time and across different sectors of society and require specific analyses” (European Commission, 2013, p. 5).

Therefore, one of the outputs of the final report was to provide scientists and engineers with practical tools they could utilize to formulate relevant questions and rethink frameworks and concepts for gender analysis. This thesis broadly elaborates on these practical tools in the next chapter, but in this section, it suffices to solely state the definition of gender used in the Gendered Innovations report. Gender is defined as “a socio-cultural process, [that] refers to cultural and social attitudes that together shape and sanction “feminine” and “masculine” behaviors, products, technologies, environments, and knowledge.” (European Commission, 2013, p. 9). Considering the focus of this thesis is on the development technology and not solely on aspects regarding self-identity, this definition of gender seems most suited. Sex, on the other hand is defined as “[...] biological qualities characteristic of women (females) and men (males) in terms of reproductive organs and functions based on chromosomal complement and physiology” (ibid, p. 9).

The 2020 Gendered Innovations report brings up three more concepts that are described as related dimensions of gender. The need for having these dimensions is because the authors of the report view gender as a “[...] set for organizing principles that structure behaviors, attitudes, physical appearance and habits” (European Commission, 2020, p. 14). Having these organizing principles broken down into different dimensions makes the concept of gender and its different facets easier to study. The first dimension is gender norms which are produced through social institutions, social interactions, and wider cultural products. Gender norms assign men, women and gender-diverse people with a set of appropriate behaviors, knowledges, preferences, products, and professions through the means of socially and culturally shaped expectations and attitudes. They can therefore, as described later in this chapter, also influence the development of science and technology. Gender norms draw upon and reinforce gender stereotypes, which is

certainly not always a desirable process. That is because both of these concepts shape and are simultaneously shaped by unequal distribution of resources and discrimination in social settings such as the workplace, families and other institutions. Lastly, gender norms vary and are constantly in flux across different eras, cultures, social settings, and geographical locations. The second dimension is gender identities which “[...] relate to how individuals or groups perceive and present themselves in relation to gender norms” (European Commission, 2020, p. 15). So, whereas gender norms push from a societal context onto the individual, gender identity pushes from the individual back into society in relation to different types of gender norms. Similar to gender norms, gender identities are also highly context-specific and interact with other social dimensions such as ethnicity, socioeconomic background, or cultural heritage. The third dimension is gender relations which “[...] relate to how we interact with people and institutions in the world around us, based on our sex and our gender identity” (ibid, p. 15). Gender relations focus on how our social interactions within workplaces, families, schools, friend groups, and other public settings are shaped by gender. An important aspect of gender relations are the social divisions of labor. For example, it is generally known that women usually take care of the majority of the unpaid care work in their households (Criado Perez, 2019). A consequence when particular type of work has such a gender segregation is that it “[...] becomes marked symbolically with the (presumed) gender category of the larger group” (ibid, p. 15). Think about how construction worker is usually seen as a male profession and nursing as something female. These three dimensions of gender help within research about the topic to conceptualize certain behaviors, power relations, or, more specific to this thesis, developments within science and technology.

### [Gender Aspects in Science and Technology](#)

Questioning the supposed objectivity of science and technology and pointing at the situatedness of knowledge is one of the major topics within the field of STS. Founding works within the field such as the ones from Bruno Latour and Steve Woolgar (1979) and Karin Knorr-Cetina (1981) show that conducting research is often a matter of tinkering with locally available resources, instead of performing it according to pre-planned lines. These works therefore show that even though science and technology present themselves

as objective, they often have an inherent messy side to them that can be affected by situated social circumstances. And it is because of this messy social side that science and technology can at times carry with them potential harmful gender aspects.

Furthermore, the fact that science and technology are affected by the social also works the other way around, as there are many works within STS that focus on this aspect. One of the most famous examples within the field are the bridges of Langdon Winner (1986). Winner describes how designers in Long Island New York in the beginning of the twentieth century made the bridges in that area very low on purpose so that the public busses would not be able to cross under them. This had as a social consequence that these busses, mostly utilized by the poor and black minority, were not able to transport their passengers to locations where, in order to get there, crossing under these bridges was necessary. Since this publication, more research has been done on the implicit and explicit social consequences of technological and scientific development. This includes the topic of certain gender issues. *Feminist technoscience*, with Donna Haraway (1989) as its spearhead, developed with the purpose to create awareness of these issues, to eventually have more inclusive science and technology.

The goal of the following subsections is to provide an overview of existing works within STS on gender aspects within science and technology. How has gender influenced science and technology, and how have they in their turn generated possible gender issues? The first subsection shortly focuses solely on gender and science, but mainly to serve as a step up to elaborate further on *feminist technoscience*. The rest of the subsections focus on the relationship between gender and technology with a special focus on gender and computer science.

### Gender and science

A classic starting point to study the interaction between gender and science is the aforementioned Londa Schiebinger (1989) with her work on the depiction of the female skeleton throughout the eighteenth century. Looking for different ways to show the difference between males and females other than through their reproductive organs, anatomists started depicting the female skeleton in the hope to find them. Schiebinger (1989) showed that societal norms and values played a substantial role in the depiction,

which in turn strengthened certain gender stereotypes. For example, females were depicted with smaller heads than their male counterpart, for their inferior intelligence, but with larger pelvises. For “[t]he purpose of the superior pelvis was ultimately to provide passage for the superior skull. [...] Thus male and female bodies were indeed complementary.” (ibid, p. 209). Such a depiction therefore strengthened the superior position of men over women in society and makes a great case on how scientific knowledge is constructed by societal norms and values and vice versa.

Due to the observation of this relation between science and society, Donna Haraway (1989) urges scientists to be more critical towards these possible oppressive and exclusive effects. One of her case studies is the research conducted by the primatologist couple Harry and Margaret Harlow in the 1950s. The Harlows separated infant monkeys from their mothers and gave them the option between a metal surrogate mother with food and a cloth surrogate mother without food. Because the baby monkeys more often sought comfort with the cloth mother, they concluded that love and affection were crucial for a human child’s upbringing. This conclusion in turn had effects on gender debates going on in mid 20th century America. It namely pointed to possible negative effects of women joining the workforce and thus being less engaged with their children.

A more recent example is current work going on in the field of epigenetics. In a 2014 article in *Nature*, Sarah Richardson et al. warn about the possible harming effects towards women of the conclusions in the field of epigenetics. One of the aims of the field is to research differences in DNA expression through the behavior of pregnant women. Richardson et al. (2014) describe the harmfulness for women of epigenetic findings in the following way:

“Headlines in the press reveal how these findings are often simplified to focus on the maternal impact: ‘Mother’s diet during pregnancy alters baby’s DNA’, ‘Grandma’s Experiences Leave a Mark on Your Genes’, and ‘Pregnant 9/11 survivors transmitted trauma to their children’. Factors such as the paternal contribution, family life and social environment receive less attention.” (p. 131).

The authors urge that these simplified conclusions may lead to irresponsible discussions that would make women into scapegoats, and thus recommend epigenetics to include more interdisciplinary engagement in their research.

Other times, instead of reinforcing societal gender norms, science has to 'fight' them, or change them so they work in their favor. Nelly Oudshoorn (2003) analyzes the attempts to develop a male contraceptive pill. She describes how the trials for the drug had difficulties coming of the ground because taking a contraceptive pill was not considered to be fitting with the existing masculine gender norms. Therefore, researchers who were attempting to make the trials work had to actively 'create' a non-hegemonic male gender norm in order to get participants. For example, they only accepted men into the trials who were in monogamous relationships thus creating the image of "[...] caring, responsible masculinities of various types." (ibid, p. 223). On the other hand, the trials also spoke to already existing male gender norms. They framed the participants of the trials as trailblazers, and men that engage in an adventure in order to find the new. With all these points Oudshoorn (2003) tries to underline that "[...] technologies have thus the capacity to make visible and to destabilize dominant cultural narratives on gender" (p. 227).

### *Feminist technoscience*

The works mentioned in the previous subsection can all fall under a certain stream of STS research called *feminist technoscience*. But before elaborating on how this is exactly conceptualized, it is good to trace the meaning of the latter of those two words. The term technoscience was popularized within the academic language by Bruno Latour in 1987, and has since then had its true meaning contested. With the term, Latour aims to combine four different strands of arguments that are present within the field of STS. First comes the high level of intertwinement between scientific and technological developments. Earlier work from Latour and Woolgar (1979), in which they conducted ethnographic research in laboratories, shows that it is near impossible to come to new scientific findings without the support from technology and vice-versa. The word technoscience therefore points towards this intertwinement. The second argument relates to the power of laboratories and engineering. Latour (1987) writes that "laboratories are now powerful enough to define reality" (p. 93). What he refers to with this claim is that scientific and technological discoveries from laboratories do not merely uncover facts about nature but can travel outside of the labs and deeply affect the societies in which we live. Third is what

is called the 'seamless web' that connects engineers, scientists, and societal actors in the practice of science. This web normally stays relatively invisible or unconsidered, but the term technoscience aims to bring it forward with its focus on the process of practicing science. An example that shows this intertwining between different actors in the practice of science is establishing scientific facts. Latour (1987) states that "[a]lthough it sounds counter-intuitive at first, the more technical and specialized a literature is, the more 'social' it becomes, since the number of associations necessary to drive readers out and force them into accepting a claim as a fact increase" (p. 62). Lastly, technoscience embraces the ability of science to create nature-culture hybrids. Since the beginning of modern science, scientists have functioned under the pretense that they were 'discovering' knowledge about nature independent from cultural influences. However, the examples in the last subsection have showed that science can indeed be touched by social and cultural influences and are thus creating these nature-culture hybrids.

Donna Haraway (1997) conceptualized *feminist technoscience* as a critique on Latour's interpretation of the term. Although praising him for the focus on science-in-the-making, she believes Latour's technoscience only describes the creation of these nature-culture hybrids and does not question the possible values that are inscribed in the process, especially when they can carry negative social consequences. She states:

"Either critical scholars in antiracist, feminist cultural studies of science and technology have not been clear enough about racial formation, gender-in-the-making, the forging of class, and the discursive production of *sexuality through the constitutive practices of technoscience production themselves*, or the science studies scholars aren't reading or listening—or both" (ibid, p. 35).

Feminist technoscience, she suggests, can bring these constitutive practices forward and can offer some critique on technoscientific progress in order to increase inclusivity. A good example of Haraway's feminist technoscientific research is the example mentioned in the previous subsection of her work on the primatologist researchers. She showed that their research on infant monkeys eventually led to negative consequences for women, as the findings of the primatology research could support arguments to refrain women with children from joining the workforce (Haraway, 1989). Therefore, the way in which technoscience is defined by Haraway "[...] raises important questions about boundaries and transgressions between implosion of science, technology, politics and society,

humans and non-humans etc. as well as implosion phenomena within the same spheres” (Trojer, 2014, p. 168).

Haraway’s conception of feminist technoscience has since then inspired much research. However, as comes with many things that are labeled feminist, the concept is still provocative because of its call for political and transforming actions regarding academic discourse unwilling to transform (Trojer, 2014). Nevertheless, research on implicit gender and diversity inscriptions in science and technology has become ever more relevant, especially since feminist technoscience can “[...] expand the knowledge frames and practices for technology development in increasingly complex realities” (ibid, p. 179). Trojer (2014) describes that with regards to technology, which was and still is for a large part a generally male dominated industry, feminist technoscientific research has shifted from ‘counting heads’ – how many women are involved – to researching processes of knowledge production. This led to an epistemological shift where the challenge now is that “[...] gender research within technology and engineering does NOT primarily focus on gender – men and women. It focuses on technology” (ibid, p. 166). The following subsections review feminist technoscientific research that focusses on gender aspects in technology in different ways.

### *Gender and technology*

Analyzing the relationship between gender and technology can happen in many different ways and from different angles. To serve as a proper background exploration for the case study at the core of this thesis, namely possible gender issues within visual computing projects, this section breaks down a variety of gender relations with technology in the following subsections: gender and product design, gender and computer science, and women in technology.

### *Gender and product design*

As mentioned in the previous subsections, both science and technology are both shaped by and can reinforce and renegotiate gender norms, and analyzing the implicit gender inscriptions in technology requires an epistemological shift in focus from the human to the technology itself. A rather obvious case that combines these two lines of

argumentation is Ellen van Oost' (2003) work on different shaving devices. She analyzes Philip's Ladyshave for women and Philishave for men in terms of their *gender scripts*, which she defines as "[e]xisting or even stereotypes images of projected gender identities [which] are transformed into design specifications that are in accordance with cultural symbols of masculinity or femininity" (ibid, p. 195). Gender scripts are inscribed into specific technologies by designers and define their imagination of the specific users. In case of the shaving devices van Oost (2003) describes how their gender scripts speak to existing societal stereotypes, and actually inhibit women from seeing themselves as technologically savvy. For example, the Ladyshave is colored bright pastels or pinks, whereas the Philishave is colored black, grey or dark blue. Furthermore, the Ladyshave is marketed and designed more as a beauty product such as a lipstick, and intentionally has all signs of bolts and screws masked. According to van Oost (2003) this latter mentioned aspect could be interpreted as a representation of female users as technophobic, and even inhibits their ability to see themselves as technologically competent. In such a way "Philips not only produces shavers but also gender" (ibid, p. 207).

On the other hand, more neutral seeming products can also come with negative gender inscription. Sometimes seemingly neutral products are designed on a male norm, and thus carry the risk that people not conforming to these norms can be harmed. A striking example is the overrepresentation of male drivers in traffic data, which has led to the design of seatbelts and airbags being mainly based on the male proportions. A 2011 study analyzed automotive crash data from the United States in the period from 1998 and 2008 by sex, and came to the shocking conclusion that in a case of belt-restrained accidents female drivers had a 47% higher chance to sustain severe injuries (Bose, Segui-Gomez & Crandall, 2011). Further research then pointed out that the introduction of a virtual crash dummy that could account for the smaller female proportions highlighted the need to add a 'female' crash test dummy to regulatory safety testing (Linder & Svedberg, 2019).

Another similar example is how the design of cockpits in commercial and military airplanes can eventually block women from making certain developments within their careers. Rachel Weber (1997) describes how leading up the late nineties, guidelines for designing cockpits were based on male dimensions, and the standard was that only ten

percent of men would not be accommodated by a certain design feature. In the case of sitting height “[...] the five percent of men who are very short and the five percent who are very tall will not be accommodated” (ibid, p. 238). However, women often have smaller dimensions than men and thus the chance that they fall under the bottom five percent is higher. This deemed the women who fall under this percentage ineligible to use a variety of technologies within aircrafts, and consequently blocked them from making certain developments within their careers as commercial or military pilots.

So far, the examples mentioned in this subsection have focused on the effects of physical objects on sex and gender properties. However, at the core of this master thesis lays a case study that is about possible gender aspects within computer science projects which do not necessarily involve physical properties. Therefore, the next subsection is devoted to how gender could play a role in the field of computer science.

#### *Gender and computer science*

More relevant for my research are the potential gender issues that arise within Information and Communication Technologies (ICT) and engineering. Nelly Oudshoorn, Els Rommes and Marcelle Stienstra (2004) focus on the barriers design cultures within ICT can create which influence the diversity of users from a gender perspective. Through the cases of two electronic communication networks established in the late nineties the authors show how design processes caused the technologies to fail in their goal of gender inclusivity. Both networks attempted to initially make their technology accessible for everybody, but in both cases two main reasons were found which caused women to be underrepresented in their user groups. First is the path dependency of the hardware. Path dependency describes how “[...] due to previous material, social, and symbolic investments, future technologies will develop very much in line with existing technologies” (Oudshoorn et al., 2004, p. 36). Both electronic communication networks utilized hardware (computers and teletext) that was, at the time, inherently used more by males, and thus underrepresented women from the get-go. Second is the *I-methodology*, which refers to the process in which designers consider themselves as a good representative of the users (Akrich, 1995). In both cases the vast majority of the designers were male, and due to their lack in receiving feedback from users, they designed features

of both technologies more masculine. However, the abovementioned examples refer to a time where indeed access to and work in ICT was mainly male dominated. And although the work in ICT is still like that, which creates potential hurdles for women which will be discussed in a following subsection, access to technology does not pose big problems anymore.

Being a computer scientist writing an algorithm or code means constantly making abstractions and decisions based on logic. It is therefore hard to imagine how gender plays a part in these actions. Corinna Bath (2014) says the following about it:

“Abstraction, formalization, and classification produce the impressions of objectivity and a neutral research subject. Here, it seems problematic that these processes inevitably entangled with computer scientists’ work disguise explicit and implicit decisions made in the process of technology design, while they in effect establish hierarchies of knowledge, gendered classifications or dichotomies” (p. 69).

An example in which abstraction decisions made by the computer scientist has consequences with regards to gender is the work of Kaiser et al. (2008) on brain imaging. Their research is about a process in which an algorithm transforms raw data from a computer into colored images of the brain. Kaiser et al. (2008) show that gender differences present in the brain appear or did not appear based on where the computer scientists laid the thresholds in their algorithm, thus showing that formalization processes are not necessarily innocent.

Another big area within computer science that carries with it potential negative social and gender issues is machine learning. Machine learning is the study and creation of computer algorithms that improve in executing a certain task through experience (Mitchell, 1997). These computer algorithms are trained through large data sets to eventually perform a task without being explicitly programmed to do so. In this process, there are multiple ways in which such an algorithm can pick up certain biases. The rest of this subsection focuses on examples of gender biases that could be present within machine learning algorithms.

The performance of a machine learning algorithm highly depends on the data on which it is trained. So, if there were to be any bias in the data, the chances are high that the algorithm will adopt this bias. A famous example is facial recognition software, in

which a variety of biases are identified. One type of such software produced and trained in the US performs worse when it comes to the identifying women, and especially women of color. Whereas another, produced and trained in China, has more difficulty recognizing Caucasian faces (Garvie & Frankle, 2016). Although subtle and often hard to detect, these types of biases can have negative social and gender implications. A similar example is a 2015 study that used an automated tool called AdFisher to analyze Google's targeted advertising. One of the findings was that setting the gender on the tool to female resulted in significant fewer instances of advertisements related to high paying jobs than setting the gender to male (Datta, Tschantz & Datta, 2015). Therefore, the best way to avoid these kinds of biases is to train machine learning algorithms on evenly distributed and diverse data (Feast, 2019).

Next to using diverse data to train machine learning algorithms it is additionally important to label this diversity. The current common practice within research to pool data from males and females has the potential to mask possible gender or sex differences (Tannenbaum et al., 2019). Ignoring such differences can lead to difficulties generalizing results, research inefficiency, and inaccuracies. Furthermore, "[p]ooling data across sexes not only assumes that there is no difference between males and females, but also subsequently prevents researchers from testing for the dependency of an experimental response on the sex of a study participant" (ibid, p. 139). The aspects up to this point all relate to the possible gender implication caused on the data side of machine learning, however, designers of algorithms also have influence.

### *Gender and the STEM fields*

Although the goal of the project described in this thesis is not to analyze the role of gender on the work floor at VRVis, it is hard to ignore that the company is situated within in a field that is for a large part still male dominated. According to the 2018 report of the European Commission 'She Figures', women are still a minority in science and engineering occupations in most countries in Europe. Therefore, this subsection shortly analyzes existing literature with regards to women working in STEM fields. Wendy Faulkner (2014) analyzes women that work in the field of engineering, and how they constantly have to resist against existing gender norms within the engineering culture.

The consensus in the field seems to be that “[...] women who are really into engineering are not ‘real women’ and conversely ‘real women’ are not ‘real engineers’” (ibid, p. 187). Faulkner stresses that this struggle for gender (in)authenticity is a battle constantly fought and reinforced by tiny behaviors on the work floor and warns that active intervention to go against it is necessary.

Furthermore, recent research points to gender composition in work teams affecting the experience of women working in STEM fields. Niler, Asencio, and DeChurch (2020) researched the degree of team identification and collective efficacy amongst female college students who worked in semester-long project at a STEM university. The results showed that the number of women on each team significantly increased women’s experience of team identification and collective efficacy, which ultimately led to a better team performance. Strikingly, the number of men on each team did not influence these types of sensations for males. Other interesting research on a similar topic provides compelling evidence that women in leading positions in the fields of biomedicine, public health, and clinical research are more likely to use possible sex and gender dimensions in their published research (Nielsen et al., 2017). However, “[...] this dynamic has not yet been replicated in other research fields, such as computer science, engineering or the physical sciences” (Tannenbaum et al., 2019, p. 138).

Although it seems that there is this all-encompassing male dominance within the field of technology and engineering, this is actually mostly in the Western world and not all across the globe. Ulf Mellström (2009) offers a glance into the field of computer science in Malaysia, in which women are predominantly active. This dominance can be traced back to certain ethnic and social developments that occurred in Malaysia since the 1990s. Mellström (2009) underlines that factors such as ethnic and class inequalities are often as important as gender issues, but that these factors are often ignored by current gender and technology studies. This accumulates in a criticism towards current gender and technology studies for having a strong western bias, for not being critical enough on the embeddedness of specific cultural contexts, and for not bringing local gender discourses into the analysis. Mellström (2009) presses that being more aware of these aspects would allow for better cross-cultural comparisons.

This point also comes forward in the aforementioned report *Gendered Innovations*, in which methodologies are offered to scientists who wish to include gender and sex analysis into their research design. One of the recommendations made in this report clearly states that an important part of analyzing sex and gender is to identify other aspects, such as socioeconomic status or environment, that could possibly intersect with sex and gender (European Commission, 2013). This brings us into the transition towards the final subsection in this chapter which focuses on other methodological practices that can be used to analyze gender aspects in science and technology.

### Responsible Research and Innovation

So far, this chapter has reviewed a handful of telling examples as to how development in science and technology can have strong social effects in society. Because of this phenomenon, paired with the fact that it is increasingly more difficult to oversee these potential social consequences, developments in science and technology become harder to govern. About a decade ago, these observations started to spark discussion about possible new ways to govern and conduct science. Criticism was outspoken that innovation processes should not be aimed solely at bringing products and services to markets to pursue economic growth. Instead, “[i]nnovation should be viewed in a broader social context. That is, innovation is beneficial to people only if it meets the needs of society, providing economic, environmental and social sustainability” (Burget, Bardone, & Pedaste, 2017, p. 3). Furthermore, criticism was outspoken that existing funding strategies for research benefitted the natural science and engineering disciplines disproportionately to social sciences and humanities (SSH). Felt (2014) identified that this challenged and hindered interdisciplinary cooperation between the two sides, even though knowledge production and an exchange of knowledge across the borders of the natural sciences and engineering and SSH should be encouraged. Especially since SSH disciplines can play an important role in analyzing the social landscape in which natural science and engineering research is situated, and can provide it input regarding public debate issues, normative conflicts, and policy tensions. Eventually, a call for more responsible research and research funding infrastructures was made. As a response, a

framework dubbed 'Responsible Research and Innovation' (RRI) was introduced by science policy makers and various funding agencies mostly within the European Commission to ensure more socially robust research in a top-down manner (Burget et al., 2017). The idea of responsible research was not new before the introduction of RRI, but when the European Commission introduced RRI into its 7<sup>th</sup> framework Horizon 2020 it was the first time that the value of responsible research was given an official framework. The Horizon 2020 framework "[...] primarily emphasized cooperation between science and society and strengthening public confidence in science" (ibid, p. 2) At the time of its introduction, and for a while after, the exact definition of RRI and what it included were in constant flux. However, its goal was clear: challenging and rethinking linear models of science and innovation policy, and moving them more in the direction of transparent and interactive processes by which societal actors and researchers are both included in order to create more ethical, sustainable, and societal desirable innovation (Owen, Macnaghten, & Stilgoe, 2012).

RRI has four dimensions that make up the main framework. They originate "[...] from a set of questions that have emerged as important within public debates about new areas of science and technology" (Stilgoe, Owen, & Macnaghten, 2013, p. 1570). The questions are posed by public groups towards scientists. Think about, for example, how are risks and benefits distributed, what impact can we foresee, which ones can we not? The four dimensions of RRI provide a framework for raising, discussing and responding to these raised questions. The different dimensions are strongly related to each other, and at times even overlap. The first dimension is anticipation, which is something that needs to occur at the beginning of a research process in order to steer it in a desirable direction. "Anticipation is a conceptual dimension that aims at envisioning the future of research and innovation and understanding how current dynamics help design the future" (Burget et al., 2017, p. 10). Basically, successful anticipation means understanding how your research processes influence technological futures, and steering them in a direction whereby they provide benefit and avoid potential harmful consequences. Rose (2014) elaborates that anticipation serves four purposes: the reflection on motivations and implications of research, being transparent about potential uncertainties or controversies, opening the visions of research to the public, and using the outcomes of

anticipatory processes to shape the research in the direction of a desirable future. A key to correct anticipatory processes is that they need to be well-timed. If they occur too early they risk not being constructive, but if they occur too late they might not be meaningful (Stilgoe et al., 2013).

The second dimension is inclusion, which aims at including a variety of different stakeholders, including public groups, into discussions surrounding the development of science and technology. Burget et al. (2017) claim that “inclusion is the conceptual dimension that characterizes RRI the most” (p. 10), but that as a major characterizer it still needs to be the topic of reflective and critical academic discussion. This is because different criticisms exist about the shape and format of different types of public engagement in research and innovation. At the moment public engagement initiatives exist such as focus groups, citizen juries and panels, and consensus conferences (Chilvers, 2010). However, it is a challenge to achieve efficiency in these formats with regards to time and taxpayer money spent. Furthermore, STS scholar have criticized that within public engagement initiatives, particular ‘framings’ of the public can reinforce certain power relations and unbalances can be reinforced (Stilgoe et al., 2013). Therefore, the dimension of inclusion is still under critical reflection, but it is clear that the inclusion of relevant social groups in research processes is necessary within the RRI framework to move towards more desirable future social worlds.

The third dimension is reflexivity, which “[...] at the level of institutional practice, means holding a mirror up to one’s own activities, commitments and assumptions, being aware of the limits of knowledge and being mindful that a particular framing of an issue may not be universally held” (ibid, p. 1571). Mechanisms through which reflexive practices can be integrated in scientific practice are codes of conduct, the adoption of standards, or moratoriums. Although Stilgoe et al. (2013) describe that the adoption of such practices has been successful at laboratory level, they argue that wider implementation in other institutions such as research funders and regulators is also necessary since these institutions carry a responsibility “[...] not only to reflect on their own value systems, but also to help build the reflexive capacity within the practice of science and innovation” (p. 1571).

The last dimension of RRI is responsiveness, which is linked to the concepts of ethics, risks, transparency and accessibility (Burget et al., 2017). Responsive RRI entails that research and innovation have the ability to change their shape or direction in response to public and stakeholder values, or changing social circumstances (Stilgoe et al., 2013). Furthermore, social, economic, or environmental risks of new technologies that could possibly emerge on the short or long-term need to be considered. A way in which the European Union is attempting to adhere to responsive standards is to be transparent, and to provide the public full open access to scientific results of publicly funded research (Burget et al., 2017).

When the European Commission officially introduced RRI into its 7<sup>th</sup> Framework Programme Horizon 2020, RRI was divided into six operational 'key areas': gender equality, science literacy and science education, public engagement, open access, ethics, and governance (Bührer & Wroblewski, 2019). Gender equality is defined as:

“[...] a three-dimensional construct which aims to (1) increase female participation in all fields and at all hierarchical levels of science and research, (2) abolish structural career barriers for female researchers and (3) strengthen the gender dimension in research and innovation content as well as in teaching.” (ibid, p. 1).

Bührer and Wroblewski (2019) researched through the means of a survey, distributed amongst researchers who participated in Horizon 2020 projects, how the gender equality area has been picked up amongst RRI researchers. The results show, that when it comes to gender equality, researchers mostly engage in the first dimension: increasing female participation on their teams, whereas the third dimension is not very widespread. Furthermore, the survey showed that female researchers are more prone to integrate all three dimensions of gender equality into their work, further underlining the importance of having gender diverse research teams. As a last interesting result, the survey provides evidence that research and innovation processes change the most if commitment is externally required. This external requirement can come from funding institutions, Gender Equality Plans (GEP) set up by research companies or agencies, or performance agreements such as the ones concluded between Austrian universities and the Austrian Federal Ministry for Education, Science and Research.

This chapter has provided an overview of existing literature that serves as background knowledge for the project described in this thesis and serves as its context. First an elaboration on the historical context and the definition of the terms gender and sex have been given, continued by a discussion on how gender can play a role within scientific and technological development. The findings of this thesis will mainly add to this discussion. Lastly, a description of the RRI framework was given, which is a framework under which I can place the research conducted in this thesis.

What is striking from this literature exploration is that it seems that much of the research that has been done on possible gender issues in technology design focuses on technologies that are intended to be used by the general public. However, in the case of the technologies that VRVis designed, the users are highly specific with a particular knowledge background. Less research has been done about these types of situations. The next chapter describes different methodologies that exist that can be utilized to perform sex and gender analysis.

### 3. Sensitizing Concepts: methodologies for analyzing gender

The following chapter elaborates broadly on three pieces of literature that all have a specific focus on the importance of sex and gender analysis in science and engineering, and additionally all offer methodologies and important questions to ask for such an analysis. These methodologies and questions have a key role in the formation of the research question of this thesis, and an effect on the analysis of the data collected during the research. The three pieces of literature discussed are a 2019 publication in *Nature* from Cara Tannenbaum, Robert P. Ellis, Friederieke Eyssel, James Zou, and Londa Schiebinger, the Gendered Innovation reports brought out by the European Union in 2013 and 2020, and finally a chapter in the edited volume *Gender in Science and Technology* (2014) written by Corinna Bath. This chapter does not synthesize each of the pieces in their entirety, but focuses on specific parts that carry importance for the research conducted in this thesis.

#### Sex and gender analysis

Tannenbaum et al. (2019) argue that “[...] integrating sex and gender analysis into research design has the potential to offer new perspectives, pose new questions and, importantly, enhance social equalities by ensuring that research findings are applicable across the whole of society” (p. 137). In their publication in *Nature* the authors provide a roadmap for sex and gender analysis that can be used across different scientific disciplines. They hope the publication inspires researchers, peer-reviewed journals, funding agencies, and universities to coordinate efforts in order to implement their suggested socially robust methods for sex and gender analysis.

According to the authors, implementing sex and gender analysis not only leads to higher social equality in scientific outcomes, but additionally can improve the reproducibility of research and increase experimental efficiency and accuracy, which are all important aspects for scientific excellence. A reason that reproducibility often lacks in experimentation is that methodological reporting across different disciplines varies

widely (Baker, 2016). Failing to report sex and gender related variables that are tied to the used methodologies make it therefore more difficult to reproduce certain experiments. On the other hand, the authors argue that sex and gender analysis can lead to a higher experimental accuracy and avoid misinterpretation of results by going against the scientific common practice of pooling data from males and females as “[...] pooling the response of females and males or women and men can mask sex differences” (Tannenbaum et al, 2019, p. 138). A telling example the authors use is an experiment from Li, Ju and Reeves (2017) that attempts to research the physiological reactions of people by touching the intimate parts of a humanoid robot. Although an equal number of males and females participated in the experiment, their responses were pooled together and not analyzed separately. However, research has shown that many contextual factors such as age, gender identity and cultural background heavily influence norms for human touch. By pooling the responses, these factors in the context of this experiment are now rendered invisible.

#### Approaches for reducing bias in AI

Potentially useful for the research performed in this master thesis are the three approaches described by Tannenbaum et al. (2019) to reduce gender bias in artificial intelligence and machine learning projects. As described in the State of the Art chapter, alarming examples exist of algorithmic bias such as the systematic misidentification of women of color in facial recognition software or Google algorithms being five times more likely to offer men advertisements of high-paying jobs than women (Datta, Tschantz & Datta, 2015). The first approach suggested by the authors focuses on the data fed to machine learning algorithms and the acquisition of the data sets since this is a huge source of potential bias. For example, the misidentification of women of color is a bias that stems from problems with data because certain social groups or subpopulations are highly underrepresented in datasets. The approach focuses on increasing the quality of datasets by designing ‘nutrition labels’ “[...] to capture metadata about how the dataset was collected and annotated” (Tannenbaum et al., 2019, p. 141). This metadata contains statistics on, for example, sex, age, gender, ethnicity, and the geographical location of the

data in the set. Researchers can then use these nutrition labels to create balanced data sets.

The second approach is a concept developed by Kusner, Loftus, Russell, and Silva (2017) called *counterfactual fairness*. The concept attempts to capture the intuition that “[...] a decision towards an individual is fair if it is the same in (a) the actual world and (b) a counterfactual world where the individual belongs to a different demographic group” (ibid, p. 4066). In the empirical example of the Google algorithm offering job advertisements, the concept translates to that the job advertisements should be shown to individuals with the exact same qualifications and interests regardless of their sex. A simple way to put counterfactual fairness into practice is to feed the developed machine learning algorithm two similar profiles with only a differing demographic detail. If the results come back different it becomes clear that the algorithm has embedded a bias. Important with counterfactual fairness is that it should only be checked on algorithms that are used in cases where gender, sex, or any other demographic context should not matter for the result. In other cases where gender and sex do have an influence on the result, in for example some medical applications, counterfactual fairness is less useful.

The third approach described by Tannenbaum et al. (2019) to reduce potential gender bias in algorithms is called *multi-accuracy auditing*. A desirable goal with any type of machine learning algorithm is to achieve overall accuracy for as many demographic groups in the population as possible. Multi-accuracy auditing is a systematic methodology that aims to achieve exactly this. “[...] The goal is to ensure that the algorithm achieves good performance not only in the aggregate but also for specific subpopulations—for example, ‘elderly Asian man’ or ‘Native American woman’” (ibid, p. 141). With this approach, an auditor systematically uses a machine learning algorithm to work through specific subpopulations and assesses if the algorithm makes more mistakes for any of them. Kim, Ghorbani, and Zou (2019) used this approach in the facial recognition software to exactly pinpoint which artificial neurons in the algorithm were responsible for the systematic misidentification of women of color. Although it is impossible to develop an algorithm that is completely fair across the spectrum, multi-accuracy auditing “[...] improves the transparency of the AI systems by quantifying how its performance varies

across race, age, sex and intersections of these attributes” (Tannenbaum et al., 2019, p. 141).

### Pathways for studying sex and gender

Tannenbaum et al. (2019) underline the importance of integrating sex and gender analysis into the design of the research at an early onset. The main reason being that “much of science and engineering research is path-dependent: once research has been designed, it becomes difficult to change” (ibid, p. 142). In order to help researchers to integrate sex and gender analysis into their research the authors offer two decision trees presented in figure 1. The left of these two trees is for analyzing sex, and the one on the

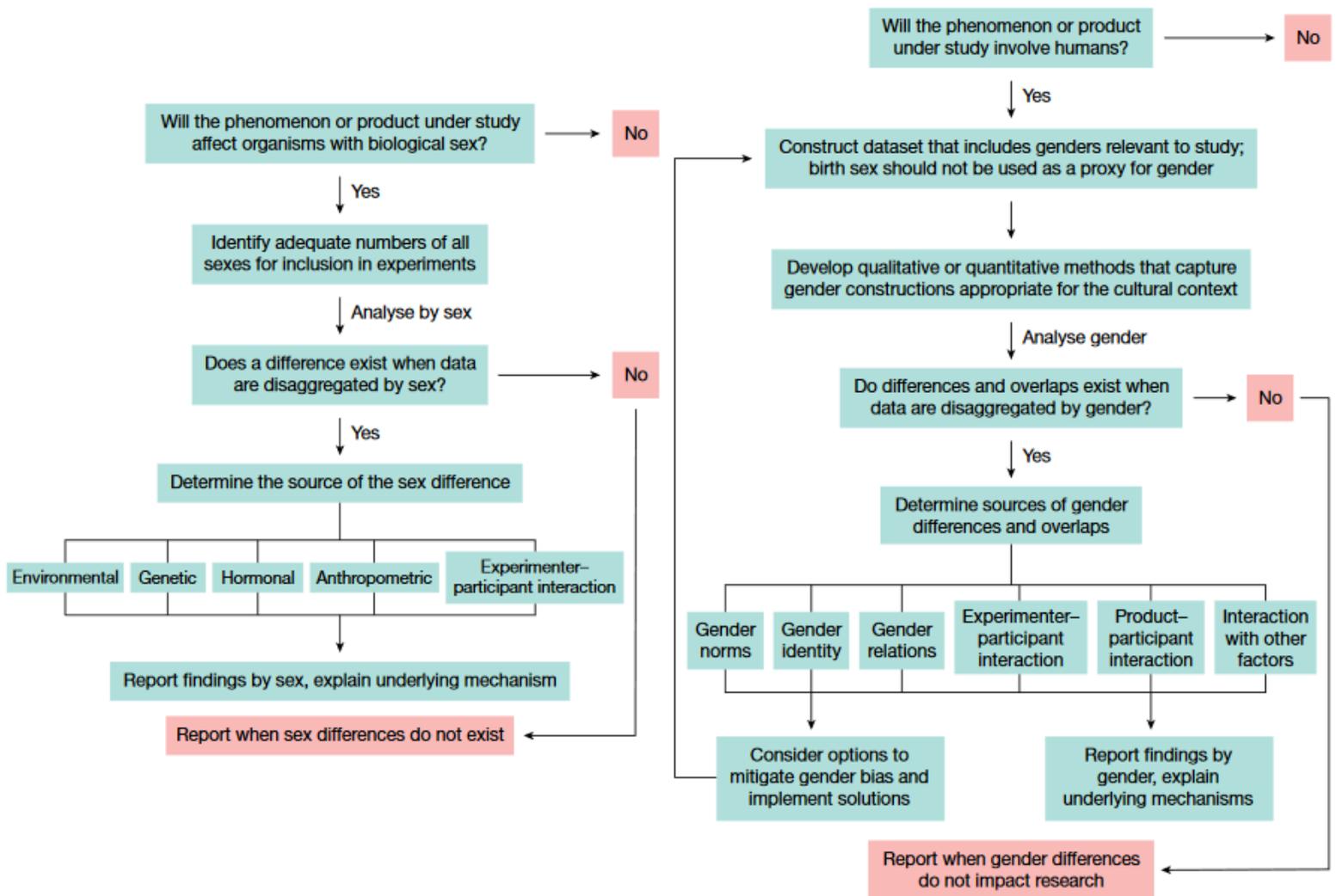


Figure 1: Two decision trees for analyzing sex and gender. Source: Tannenbaum et al. (2019)

right is for analyzing gender. The trees are designed in such a way that a 'yes' suggests the next steps in analysis whereas that is not necessary with a 'no'. In essence what one can conclude from the decision trees is that sex and gender analysis seems relevant in cases where the research under study involves or affects humans. In cases where researchers have considered sex and/or gender but perceive that this analysis is not relevant for their specific hypothesis, Tannenbaum et al. (2019) suggest then that they may rule the analysis out. The last thing the authors underline is the importance of publishing results regardless of whether sex or gender aspects are found. "If researchers expect sex or gender to be important but find no significant differences, this may represent a result worthy of publication. Reporting cases in which sex or gender sameness, overlap or no difference is found may represent an important finding" (ibid, p. 142).

### Gendered Innovations

Policies for conducting science has been one of the main techniques to push for innovations and the integration of sex and gender analysis into the mainstream. For full integration, interlocking policies need to be implemented in the three largest institutions for academic research: peer-reviewed journals, universities, and funding agencies (Tannanbaum et al., 2019). Implementation, however, comes with a large set of challenges. Universities, for example, are for the most part autonomous and thus free to decide how they fill in their curricula. Furthermore, gender analysis is mostly offered within the humanities and social sciences and to a lesser degree in the natural sciences and engineering. A big initiative that attempts to fill this gap is Gendered Innovations<sup>1</sup>. Since its initiation at Stanford University in 2009 it has gained the support of the European Commission and the US National Science Foundation. In 2011, the European Commission established an expert group called 'Innovation through Gender' that focused on developing methodologies for analyzing sex and gender aspects in the field of science and engineering. Their efforts resulted in the Gendered Innovations report published in 2013, in which different case studies are presented where "[...] differences between the

---

<sup>1</sup> <https://genderedinnovations.stanford.edu/>

needs, behaviors, and attitudes of women compares to men really matter, and accounting for them in research makes it relevant to the whole of society” (European Commission, 2013, p. 5). Furthermore, the goal of the report is to offer practical tools for sex and gender analysis to researchers within the natural science and engineering fields to help them rethink concepts and formulate relevant research questions with the appropriate methods. Consequently, the EU Research and Innovation program that started in 2014, Horizon 2020, made the sex and gender dimension an important part of its funding strategy. Under the Horizon 2020 framework, applicants for research funding were asked to integrate sex and gender analysis into their research content when relevant. The efforts resulted in the publication of the second Gendered Innovations report in 2020. The second version builds upon work described in the first report, but also expands it and describes new case studies that were performed under the Horizon 2020 framework. The following section breaks down a couple of specific case studies relevant for the research performed in this thesis, and furthermore elaborates on the specific methodologies suggested by the report for gender and sex analysis.

#### Relevant case studies

Spread over the two Gendered Innovations reports, two case studies carry particular relevance for the research conducted in this thesis. The first one focuses on Virtual Reality technology and the other one on machine learning algorithms. In the 2020 Gendered Innovations report, ‘extended virtual reality’, which can refer to either Augmented Reality (AR) or Virtual Reality (VR) technologies, is described as having potential for “[...] promoting gender equality by reducing bias and increasing empathy” (European Commission, 2020, p. 29). Case studies used in the report describe different scenarios ranging from VR increasing empathy for the other gender by allowing a person to immerse themselves into different environments to VR improving healthcare. However, although these are indeed promising applications of VR technology, more interesting to the research in this thesis are the potential gender aspects that need to be considered whilst developing these applications. Namely, “research shows that women, men and gender-diverse people may differ in how they experience virtual environments. These potential gender differences need to be taken into account in the development and testing

of prototypes” (ibid, p. 29). Since this thesis looks at the development of VR applications, seeing if these aspects are taken into consideration is a potentially interesting point of focus. For example, the level of perceived realism of the virtual environment differs amongst men and women (Felnhofer et al., 2012). Apparently, men report a higher sense of spatial presence due to higher perceived realism, and thus experience a higher sense of actually being present in the virtual environment than women. This is an aspect that can lead to differing levels of interest in VR technology amongst the different sexes. Furthermore, women are more than twice as likely to experience negative virtual reality-induced side effects (VRISE) (Al Zayer et al., 2019). These negative side effects include increased heartrate, disorientation, pallor, sweating, drowsiness, and general discomfort. A potential explanation can be found in the conflict between the body’s visual and inner ear systems which results in a higher sensitivity in women. A method that could potentially overcome this conflict is reducing the field of view in navigation, however, this once again risks undermining women. Women namely tend to base their navigation on landmarks and thus impeding the field of vision highly influences women’s spatial orientation performance. Due to all these gender considerations, the Gendered Innovations report underlines that “it is important to develop virtual settings in a sex-sensitive and gender-sensitive way, applying sex and gender analyses that include both women and men in the development and testing of prototypes” (European Commission, 2020, p. 139).

The case study used for machine learning technologies in the 2020 Gendered Innovations report is one that has been mentioned multiple times already in this thesis: the case of Facial Recognition Software (FRS). Because the case has been previously discussed in this chapter and the previous one, broad elaboration on the gender issues of FRS is not necessary. Other than the call for the use of diverse data sets, the 2020 Gendered Innovations report focuses on the issues surrounding gender and intersectionality in FRSs. “FRSs can perpetuate and even amplify social patterns of injustice by consciously or subconsciously encoding human bias. Understanding underlying intersectional discrimination in society can help researches develop more just and responsible technologies” (European Commission, 2020, p. 146). Intersectionality refers to social categories that could intersect with gender such as age, ethnicity,

socioeconomic status or sexual orientation. The misidentification of women of color by some FRSs is a good example of exactly this. Furthermore, the report brings to light that many of the machine learning studies have a binary view on gender which causes issues for individuals who fall in between the category of either 'male' or 'female'. Scheuerman, Paul, and Brubaker (2019), for example, discovered that commercial FRSs, including those developed by Amazon and Microsoft, performed significantly worse in the recognition of transgender or non-binary genders.

### Gendered Innovations suggested methodology for gender analysis

However, more relevant than the discussion of different potential gender issues in FRSs, are the methodologies consequently suggested by the 2020 Gendered Innovations report to analyze gender and intersectionality in machine learning. The suggested methodologies are developed to go against the reinforcement of negative social structures machine learning algorithms can sometimes play a part in. Three steps, each coming with a list of questions that need to be answered, are suggested for analysis. The questions are all informed by state-of-the-art research in the machine learning field. The first step is an analysis of the social objectives and implications of the work. "Teams need to work together to define the objectives of the system being developed or researched to evaluate its potential social impacts" (European Commission, 2020, p. 220). The questions that need to be answered for this step are concerning the priorities and objectives of the developed technology. Who will it benefit? Who will be left out? The second step is a thorough analysis of the data used for the machine learning technology. Important for this step is available metadata that includes information about gender, sex, ethnicity, age, and geographical location. This metadata can consequently be used for assembling diverse datasets across different variables. Furthermore, careful considerations need to be made on when it is appropriate to include certain variables as gender into the result of the algorithm. For example, advertising for clothing, which is more gender specific, can take gender into consideration whereas advertising for jobs should be equal to all. The third step is an analysis of the model for fairness. Although there is no unified definition of algorithmic fairness "[...] the best approach is to understand the nuances of each

application domain, make transparent how algorithmic decision-making is deployed and appreciate how bias can arise” (Tannenbaum et al., 2019, p. 140). Therefore, the 2020 Gendered Innovations report suggests that developers formulate an appropriate definition and test their algorithm against it. During this analysis attention needs to be paid to intersectionality, meaning isolated tests need to be conducted where the performance of an algorithm is assessed where another characteristic (age, race, etc.) is put in relation to gender. Table 1 below shortly summarizes each of the steps for this methodology.

**Gendered Innovations 2020: analyzing gender and intersectionality in machine learning.**

Step 1: an analysis of the social objectives and implications of the work.

Step 2: an analysis of the data used for the machine learning technology

Step 3: an analysis of the model for fairness

Table 1: Summary of a methodology described in the Gendered Innovation report (European Commission, 2013)

The 2013 Gendered Innovations report offers another useful methodology for incorporating sex and gender analysis into engineering innovation. “Engineering innovation here refers to any product, process, service, or infrastructure in the public or private sector” (European Commission, 2013, p. 116). Integrating such analysis into engineering processes, according to the report, can lead to new products and infrastructures that promote inclusivity across a complex and diverse user group. The process consists of five steps. The first one is evaluating past innovation practices. Where have engineering processes in the past been blind or biased with respect to gender? Something to consider here is the *I-methodology*, “[...] whereby designers create products for users whose interests, abilities, and needs resemble their own” (ibid, p. 117). In male-dominated fields such as engineering or some economic sectors the I-methodology may result in a “male default”, even if certain products are designed for ‘everybody’. An example from the last chapter is the work from Oudshoorn et al. (2004) that described two electronic communication networks failing in their mission to be gender inclusive because of exactly this reason. Another aspect that needs to be considered under this first step is that if differences during men and women are considered during the design process, are they based on stereotypes? “Products or systems based on stereotypes may

reinforce or contribute to gender and other inequalities and not contribute to enhancing social justice or corporate social responsibility” (European Commission, 2013, p.117). An example of products based on stereotypes, that is also mentioned in the previous chapter, is van Oost’s (2003) work on shaving devices. At the end of the first step, developers should have a clearer view on opportunities that have been missed as a result of failing to identify sex and gender factors influencing the design process. The second step focuses on building diverse design teams. This means including women, their experiences, knowledge and networks, into the design teams and to include gender expertise, either developed within a company or coming from outside the project. The third step is very similar to steps previously described in this subsection for developing gender inclusive machine learning algorithms. A thorough analysis needs to be done about the effects of the developed technology on women and men from different social, socioeconomic, and cultural backgrounds. Who does and who does not benefit from the technology? Also important in this step is intersectionality, what other factors could play a part in the differing effects among varying demographic groups? The fourth step is obtaining user input as users are a good potential source for sex and gender intelligence. User input can be obtained through both qualitative methods, being focus groups or interviews, or quantitative. The last step involves evaluation and planning. What are the benefits and problems of the current design process? How was the gender analysis useful and how can it be further developed? Table 2 below provides an overview of the different steps of this methodology.

**Gendered Innovations 2013: incorporating sex and gender analysis into engineering innovation**

Step 1: Evaluating past innovation practices
Step 2: Building diverse design teams
Step 3: Analyzing the effects of the developed technology on women and men from different social, socioeconomic, and cultural backgrounds
Step 4: Obtaining user input
Step 5: Evaluating the current design process and plan for further development

Table 2: Summary of a methodology described in the Gendered Innovation report (European Commission, 2020)

### Four computer scientist 'traps'

The last valuable piece of literature that is elaborated on in this chapter is Corinna Bath's (2014) 'Searching for Methodology. Feminist Technology Design in Computer Science'. The chapter is part of the edited volume 'Gender in Science and Technology'. In the chapter Bath has a specific focus on technology developers active in computer science. She describes four mechanisms that could lead to, what she calls, gendered computational artefacts. Her goal is to not only describe four common 'traps' computer scientists can fall into which can cause gendered technology, she additionally suggests methods which can be used to avoid problematic ways of gendering. The first mechanism, as mentioned earlier in this chapter, is the I-methodology. Technology designers falling into the I-methodology unconsciously assume that the users have the same knowledge, interests, and skills as them, and thus perceive themselves as proper representations of the users. Bath (2014) warns that this process "[...] can inscribe certain gendered backgrounds and knowledge, with concerns and attitudes into the technology" (p. 64), and therefore suggest that involving diverse user groups into the design process could help with its prevention. The second mechanism is the inscription of certain stereotypes into the design of technology. The aforementioned example of the Ladyshave and the Philishave (van Oost, 2003) shows this. Here Bath (2014) presses that such stereotyping should be avoided, and that "a degendering methodology, therefore, has to aim at attributing equal competencies to female and male users" (p. 66). The third mechanism concerns itself with technologies that represent certain human characteristics or behavior. Closely related to the second mechanism, the trap here describes how the characteristics and behavior displayed could portray potentially harmful gender stereotypes. Examples of this phenomenon are Siri, Alexa, or the portrayals of women in videogames. Here Bath (2014) recommends the deconstruction of the binary sex and gender system and allowing users to engender themselves. The last mechanism deals with the abstraction processes present in computer science. According to Bath (2014) abstraction, formalization, and classification produce the impression of objectivity and neutrality. However, these processes can potentially mask explicit or implicit decisions made in the technology design which could lead to possible gender issues. A degendering methodology for these

formal objects “[...] requires questioning assumptions, ontologies and epistemologies of technology design, and a dissolution of dichotomies” (p. 69).

This chapter has reviewed existing methodologies for analyzing sex and gender within fields that are specifically related to engineering and computer science. As is visible, although I discussed literature from three different sources, strong overlaps exist between the different methodologies. Avoiding inscribing stereotypes in technology is something that came up across the board. So was using diverse data sets for machine learning algorithms or being aware of the I-methodology leading to potentially male-skewed technology. This chapter has specifically focused on the parts of the literature that are useful for analyzing the case study at the core of this master thesis: potential gender aspects in the work done at VRVis, a visual computing company in Vienna. The projects at VRVis concern themselves with a range of different application for visual computing, and use varying methods from machine learning to Virtual Reality. Therefore, the methodologies described in this chapter will be serve as good handholds when it comes to guiding my analysis.

## 4. Materials and Methods

The following chapter gives an overview of the case study at the core of this thesis, the research questions asked, and the methods used to answer them. Since some background information is needed in order to understand the research questions asked, the chapter initiates with a description of the field of visual computing, the VRVis company, and the types of projects they engage in. It continues with an elaboration on the main research question and sub questions, to then finish with a description of the main method used for the research, which are qualitative interviews, and how they are analyzed.

### Case Description

#### Visual Computing

The research areas at the center of this master thesis are all situated at VRVis, a company that mainly engages in visual computing. Therefore, before explaining in detail what these projects are and what they aim to achieve, this section will serve as some context and focus on what visual computing is. Especially since this master thesis is situated within a social sciences discipline, it is good to get sensitized to the content and concepts of the visual computing field.

Visual computing is a field that emerged in the early 2000s after it became clear that multiple already established fields within computer science such as computational geometry, computer graphics, and computer vision were increasingly using similar methods (Nielsen, 2005). Although each of these fields has aspects that are not present in the others, the overlapping techniques created a need for a more generic term and as such this became visual computing. Nowadays, visual computing is an umbrella term for an interdisciplinary field that concerns itself with the handling of computer images and 3D models. Virtual Reality (VR), data visualization, and certain types of machine learning can all be considered to fall under the field. Considering these three fields are present within the different project at VRVis, the following paragraphs will elaborate on each of them in further detail.

Virtual Reality can be defined as “[...] graphic environments which both produce in the users the feeling of being physically present in a virtual world and also allows them

interaction in real time” (Foreman, 2010, p. 225). Although this definition allows normal computer monitors to be considered as VR, its more general association nowadays is to immerse a person into a digital environment through a head mounted display (HMD). The interactive quality that VR offers have made it a success within the realm of video games, but both industry and research are increasingly looking into its possibilities.

Data visualization is defined as “the representation and presentation of data to facilitate understanding” (Kirk, 2016, p. 19). Data could be anything ranging from numerical values representing certain types of information, to locations, names or statistics and are typically held in datasets. When datasets get large, it can become difficult to create an understanding of certain patterns and connections that exist amongst the data, and this is where data visualization comes in. Through the creation of images or figures data visualization helps to make large datasets understandable and accessible. However, even if multiple data points were mapped it can be challenging to establish the relationship amongst them, and thus conscious decisions have to be made on how to make understandable representations. This happens through *marks* and *attributes*. Marks consist of points, lines and areas, and attributes are their appearance properties such as the size, color and position. “The recipe of these marks and their attributes, along with other components of apparatus, such as axes and gridlines, form the anatomy of a visualization” (Kirk, 2019, p. 21).

Types of machine learning that fall under visual computing include image classification, speech recognition and signal processing. An example of a suitable application is facial recognition software. Deep learning, which is a type of machine learning, aims to recognize a hierarchy of features in input data through repetitive training, and is therefore especially suitable for such an application (Xie, Zhang & Bai, 2017). Deep learning architectures include multiple layers that are all part of a larger network and are loosely based on mammal brains. Mammal brains namely, similar to deep learning algorithms, also translate information into abstractions in order to recognize certain features, such as edges in visual information.

## The VRVis company

VRVis Zentrum für Virtual Reality und Visualisierung Forschungs -GmbH was founded in 2000 and has by now established itself as an internationally recognized research institution engaged in visual computing. The company has around 70 employees and in 2018 achieved its highest proportion of women in its scientific staff by breaking the 20% barrier (VRVis, 2018). Since 2010 VRVis is part of the COMET programme. COMET stands for Competence Centers for Excellent Technologies and is an Austrian funding programme launched in 2006. The goal of COMET funded projects is to carry out high-quality collaborative research between industry and science that is situated more on the industry side than university groups would normally be. COMET funding allows researchers at VRVis to not only be active within these industry projects, but also to engage in publication activities. So far, employees at VRVis have all together achieved over 500 publications, which helped VRVis to maintain its prestigious reputation as a high-quality research center.

The core of VRVis' mission is to build bridges between science and business. In this age of digitalization companies more often than not acquire large amounts of data, that they sometimes need help with understanding. Here VRVis helps with “[...] present[ing] data, connections and questions in the best possible visual and interactive form” (VRVis, 2020), also because they believe images are a universal means of communication which are accessible and understandable. Figure 1 shows a figure present on the VRVis website that clearly tries to present their mission. If a company were to encounter a problem or issue with their data in the real world, VRVis will be able to digitalize it in the shape of a ‘digital twin’ or a make a different type of simulation/representation. The goal of the simulation will be to provide some sort of visual decision support, so that eventually correct decisions can be made with regards to the real-world problem. The different types of projects at VRVis are divided into four different areas. The next subsection will break down these areas and elaborate on their activities.

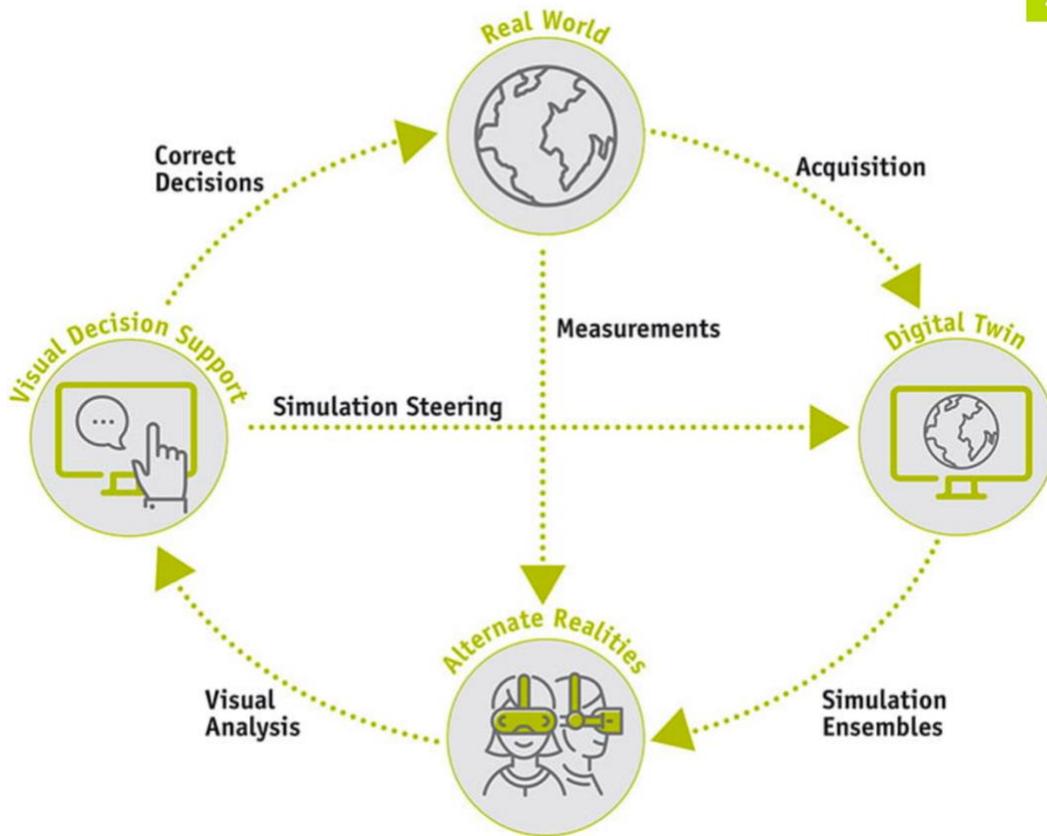


Figure 2: VRVis' goals. Source: <https://www.vrvis.at/en/about>

## The four different project areas

### *Smart Worlds*

Projects that fall under the Smart Worlds research area concern themselves with making digital representations (or twins) of the real world. Making such representations can help with important real-world planning, situation assessment, and decision-making. They are made with geospatial and geometric data and are complemented by additional data channels such as live and recorded sensor data, simulations and abstract data relevant for planning (VRVis, 2020). The end result of the combination of all this data is a complex digital surrogate of a real object, which VRVis denotes as 'smart objects'. An example of a VRVis project that falls under this research area is HILITE. HILITE is a tool that can be utilized by architects to design new lighting concepts. The tool can simulate a specific space and give users a highly realistic and interactively modifiable preview of a certain lighting design.

### *Visual Analytics*

The Visual Analytics research area deals with data analysis enabled through computer-aided visual representations. One of the main tools developed by this area was Visplore. Towards the end of this thesis project, the software became the center of a new spin-off company due to its large success. Although all dealings with the software currently fall under the new spin-off, at the time of my fieldwork it still fell under VRVis. Visplore mainly concerns itself with data visualization, and its main goal is “[...] to help users acquire new insights and increased decision confidence by enabling an interactive dialogue with big structured data” (VRVis, 2020). An example scenario in which Visplore could be of use is when a certain company notices that the production of one of its machines is going down, but they cannot identify what is causing the decrease in performance. One step they could take in order to solve the problem is to look at the data measured by the sensors in the machine and see if they can identify the cause of the malfunction from there. This is where Visplore could help. The company would talk to the VRVis researchers and explain the information they wish to extract from the data. The researchers would then build a personalized ‘dashboard’ within the Visplore software that would allow the company to easily visualize the machine’s data in different graphs and figures. This process repeats itself for every new customer.

### *Complex Systems*

Projects in this research area engage in practices that include machine learning and deep learning. The majority of the projects within this research area develop applications that can be applied in the medical field. Example projects are the development of an algorithm that, through the means of deep learning, could identify the intervertebral discs in 3D images of a spine, or an algorithm that can identify TBC in chest radiographs.

### *Immersive Analytics*

Research projects within the Immersive Analytics areas aim to push visual computing possibilities further than just the visual element, and therefore focus mostly on Virtual Reality (VR) and Augmented Reality (AR) technologies. AR is similar to VR, but instead of

immersing users in a completely virtual environment, AR technology uses the real-world environment but enhances it with virtual elements. The Immersive Analytics department mainly aims to add the sense of hearing into its projects. An example is an AR project about engine development. Here the Immersive Analytics department developed an app that would allow users an augmented reality view on a combustion engine. The app would enable the user to inspect the engine from each possible angle with the corresponding view, vibrations, and sounds.

### VRVis as initiator for this thesis project

This thesis was initiated by the CEO of VRVis Dr. Gerd Hesina, the company's former Scientific Director Dr. Werner Purgathofer, and Franziska Steyer-Beerman who is part of the Gender and Diversity Management at VRVis. Initially they reached out the head of the Department for Gender Competence at the Vienna University of Technology dr. Brigitte Ratzler. It was at her request that VRVis would reach out to our Department of Science and Technology at the University of Vienna, to which it finally turned into this current master thesis project.

Their specific wish for this project was to analyze possible gender issues in all four project areas, with a special attention on the COMET projects. As explained earlier in this chapter, VRVis is a center funded by research funding (FFG) within the long-term program COMET. Within this program there is a focus on the increase in the proportion of women and the handling of gender aspects in research. Gender criteria are at times handled quite rigorously by the FFG. Simply ignoring them can even lead to the loss of funding. Therefore, there is quite an incentive from the side of VRVis to take an interest in questions regarding gender. Having to look at all four research areas was a specific wish from VRVis since it would be in their benefit to receive feedback on all aspects of their activities.

### Research Questions

As mentioned in the previous subsection, VRVis is the initiator of this project. But I believe the results of this research can also prove useful for the wider field of STS. The previous chapter showed that within STS there is a vast body of research regarding the dual

influence gender and science and technology have on one another. However, when it comes to the topic of possible gender issues in computer science technology, little research is done about technologies that are designed for a highly specific user as is the case at VRVis. The research described in this thesis aims to add to said described research gap. The main question is therefore:

*[MQ] Does gender play a role within the development of technology in each of the four different research areas at VRVis? And if so, how?*

At the end of the research, I wish to create an understanding on if and how gender plays a role within the four different areas at VRVis. Although, within the scope of this thesis, it is impossible to do an analysis of every project within all areas, the hope is that possible findings of specific projects could be applied to a wider scope of projects going on at the company and can create awareness on possible gender aspects in future projects. Because within this study it is only possible to conduct interviews with the researchers at VRVis, as opposed to interviewing the end users of VRVis technology, the analysis of possible gender aspects stays on the side of the researchers and their conceptions. To guide my analysis and help me reach my goal of answering my main question, I pose four further sub questions. These questions are informed by research mentioned in the previous chapter.

*[SQ1] How do the researchers at VRVis conceptualize gender?*

As previously mentioned, the concept of gender is complicated and is at times used interchangeably with the concept sex. Considering VRVis is a company that, next to industry projects, also engages in research and initiated this project that looked into possible gender aspects in their work, it is an interesting starting point to see how researchers at VRVis conceptualize gender. How do they define it, and how do they imagine it having a possible influence on their work?

*[SQ2] What are steps taken by the researchers at VRVis to prevent/counteract any potential gender issues?*

As discussed in the previous chapter, there exists a plethora of different methodologies to counteract, analyze, or prevent gender issues in engineering and computer science projects. This sub-question tries to answer if any of the measures described in the previous chapter, or perhaps other ones, are utilized by the researchers at VRVis to prevent gender issues being inscribed in their technology.

*[SQ3] How much importance is given to gender considerations for a specific project?*

This question mainly aims to uncover when the researchers at VRVis consider it relevant to include gender as an important aspect into their technology design. The question is therefore also linked to the first sub question, as their conception of gender can highly influence whether or not the researcher finds it important to consider.

*[SQ4] Where in the project process do the researchers identify possible gender aspects playing a role?*

At VRVis, running through a project from conception to end goes through multiple steps. There are, for example, meetings with the customers, there is a technology development stage, and data for the project will be provided from somewhere. It will be interesting to see the researchers at VRVis reflect on each of these stages and see if they can identify where possible gender aspects can influence the project. This question can possibly also shed some light on the aspect of responsibility. For example, if the researchers foresee that a possible gender issue is in the data that is provided for their project, but the data is provided by a third party outside of the company, this could potentially lead to them feeling less responsible for the outcoming gender issue.

## Qualitative Interviews

On initial conceptualization of this thesis project, an ethnographic approach seemed to be suitable to uncover the answers to my research questions. I was going to situate myself at the VRVis office for two days a week in order to familiarize myself with the projects, people, and office practices and activities. However, due to the COVID-19 outbreak forcing all universities and businesses to continue their activities from home, an ethnographic

approach became impossible. As a consequence, the best method to study the diverse four research fields has become qualitative interviews conducted remotely.

Gender is a tricky topic to study as people often find it difficult or uncomfortable to discuss about. It could be because the word gender has connotations to other polarizing concepts such as feminism. It was therefore quite intimidating to conduct interviews, as a female social science student, with visual computer scientists who all work in a mostly male-dominated field. Because of this, it is important to reflect on the possible influence this dynamic had on the knowledge production during my interviews. Helpful for this reflection was Silverman's (2006) conception of constructionist semi-structured interviews. According to constructionism "[...] interviewers and interviewees are always actively engaged in constructing meaning. Rather than treat this as standing in the way of accurate depictions of 'facts' or 'experiences', the researcher takes as their topic how meaning is mutually constructed" (ibid, p. 118). This line of thinking is suitable for my research as I did not go into my interviews with a checklist of clear pre-conceptualized ideas on possible gender aspects within the field of visual computing. I wanted to actively engage in a process of knowledge exchange with my interview partners where I could draw their attention to certain aspects that they may not have thought about before, to which they can then come with answers that can enrich my findings. Another big part of the constructionist interviews is the focus on the way in which the interview has been conducted and its effect on how the interview data should be viewed. Silverman (2006) uses a quote from Holstein and Gubrium (1995) that underlines this point.

"Respondents' answers and comments are not viewed as reality reports delivered from a fixed repository. Instead, they are considered for the ways that they construct aspects of reality in collaboration with the interviewer. The focus is as much on the assembly process as on what is assembled" (Holstein & Gubrium, 1995, p. 127).

Being aware of the way the interview is conducted is important in my research for two reasons. First, due to the COVID-19 pandemic I had to conduct all my interviews remotely through Skype. Having online interviews had a large effect on the overall setting and had a negative effect at times when it came to establishing rapport with my interview partners. The online setting allowed for less of a direct communication and made it harder to riff off previous questions or answers. Furthermore, bad internet connection lead to

me having to do some of my interviews over the phone without being able to see my interview partners, which obviously makes it more challenging to establish good rapport. Second, as mentioned before, my interviews were an interaction between me, a female social sciences student, and a highly specialized computer scientist. Considering I lack the technical knowledge to talk in depth about the technologies developed at VRVis, my interview partners often had to explain in simplified language and terms what their practices entailed. This in turn has an effect as well on the type of knowledge that is produced within the interviews.

Semi-structured interviews are the best format to use for my research since my research questions have a focus on individuals' perspectives and conceptions. Furthermore, semi-structured interviews allow researchers the flexibility to explore further interesting threads that come up during an interview, which loans itself well to this thesis where four different research areas are researched with one method. Lastly, according to Jensen & Laurie (2016) the semi-structured format "[...] embraces the collaborative nature of the interview: through the interview process, interviewer and participant work together to develop a shared understanding of the topic under discussion." (p. 173). This latter point is exactly what I underlined earlier in this section: I want my interviews to be a collaborative knowledge production process between me and my partners.

In total I had two rounds of interviews. The first round was between April 2020 and June 2020, during which I conducted six interviews. The second round was in December 2020, during which I conducted two additional ones. Find in the table below an overview of the first round of interviews, all of them lasted between 30-40 minutes. I find it important here to mention the gender of my interview partners, as this is at times a topic of analysis in my empirical analysis chapter.

<b>Research Area</b>	<b>Amount of interviews + gender interview partner</b>
Complex Systems	2, one male & one female
Smart Worlds	1, male
Immersive Analytics	1, female
Visual Analytics	2, one male & one female

Table 3: Overview first round of interviews

The sampling process for the interviews was quite simple. Considering VRVis was the initiator of this thesis project, I did not experience difficulties with gaining field access and all my interview partners were very willing to participate. I would reach out to the head of a certain research area, who would then give me some names of people working within that research area that I could contact personally for an interview. Furthermore, in January 2020 I went to VRVis for a day and received a tour through the different research areas to get briefly acquainted with the variety of projects and people working there.

The main focus for the first round of interviews was to gain an understanding of the practices and activities going on in each of the research areas. What are the types of projects that are going on within each of them? What are the procedures for finishing a project? And what are the personal opinions of my interview partners regarding potential gender issues within their work? Immersing myself in the research areas like this not only deepened my understanding of the activities attached to the areas, but will also familiarized me with the lingo or abbreviations used by the different employees. The number of interviews in each research area is quite low because in order to make the project manageable for a master thesis I wanted to mostly speak to people working on the same project. As these projects are good representations of the wider research area it made analysis a lot less complex. However, since some projects only had a very limited amount of people working on them, the number of interviews I could conduct in that area were limited. Furthermore, as previously mentioned, Tannenbaum et al. (2019) indicate that sex and gender analysis is not necessary for any type of research. With that in mind, the main goal of the first round of interviews was to establish if gender analysis would be useful, and not a high amount of interviews is necessary to determine this. The analysis

of the first round of interviews helped me identify certain ‘areas of interest’ that include potential interesting gender aspects that could be further elaborated on in a second round of interviews.

In the end, I chose to conduct two more interviews with researchers in the Complex Systems and Immersive Analytics research areas in December 2020, as these areas had certain aspects in them correspond with aspects from the literature I described in the previous chapter. Mainly that these areas were working with VR and machine learning technology. The questions in these interviews were much more focused on gender, and went in deeper to specific aspects that came up as interesting after the first round of analysis. Although the main focus of this research are the technologies developed at VRVis, some questions during these follow-up interviews additionally focused on the broader context of VRVis itself.

## Analysis

For data analysis, I transcribed all the interviews with the software ATLAS.ti which is research software specialized in qualitative data analysis. The big advantage of transcribing in ATLAS.ti is that the software also includes tools for coding which is the next step I undertook in my analysis. I transcribed the interviews in a non-verbatim way. Non-verbatim entails that all filler words, stammering, or anything else that takes away from the core meaning of the message is removed by the transcriptionist. I chose this style because the opposite style, being verbatim transcription, where every stumble, non-speech sound, or “mm-mm” (affirmative or negative) is written down would simply not be very useful for my analysis. The focus for my analysis is mainly on what is being said, and not how they are said. Therefore, it was a better decision and certainly a more time-efficient one to choose the non-verbatim style.

After the transcription I engaged in the open coding of my data. For the coding and further analysis I used a combination of deductive coding and grounded theory as described by Charmaz (2006). Deductive coding means that “[...] at least some themes [are] developed before you begin analysis, from previous research or theory or research intuition and experience” (Rivas, 2018, p. 431). The deductive coding for my data is mostly led by Tannenbaum et al. (2019), the recommendations from the Gendered

Innovations report (2020) and the four methodologies from Corinna Bath (2014) as they are described in the previous chapter. A couple of examples of potential themes or focus points I looked out for that were informed by these core pieces of literature are conceptualizations of or assumptions about gender, the handling of data, and mentioning of humans.

Grounded theory, emerging in the 1960s, has been a great tool for qualitative researchers to make their process of analysis more structured and transparent. The theory, in essence, “[...] consists of systematic, yet flexible guidelines for collecting and analyzing qualitative data to construct theories ‘grounded’ in data themselves” (Charmaz, 2006, p. 2). Theories and themes developed over the course of the research can all directly be traced back to what is present in the data. Coding the transcriptions, which is the process of segmenting your data with short names that “[...] simultaneously summarize and account for each piece of data” (ibid, p. 43), is part of grounded theory. After an initial round of coding all my data, the most useful type of coding for my research was focused coding. Focused coding “[...] requires decisions about which initial codes make the most analytic sense to categorize your data incisively and completely” (ibid, p. 57). This is where the deductive themes acquired from Tannenbaum et al. (2019), the 2020 Gendered Innovations report, and Bath (2014) were most helpful as they conceptualized a frame in which I could place codes from my data. Examples of possible focused codes pulled from these pieces of literature are ‘conceptualizations of gender’, ‘mentioning of data’, ‘mentioning of men’, ‘mentioning of women’.

### Ethics

This thesis adheres to the good scientific practice standards as determined by the University of Vienna. However, due to VRVis being the initiator of this thesis project there are some aspects of this thesis that are worth mentioning in this subsection. First, I am writing my thesis in a form of employment for VRVis. I signed a contract with them and will be financially compensated for the final thesis product. I received a work email, access to the online employee platform, and a key card to get into the VRVis office. This latter unfortunately never used due to the changes brought by the COVID-19 pandemic. Furthermore, some of the materials I used for my research contained sensitive

information such as budgets and I was therefore also required to sign a Non-Disclosure Agreement (NDA).

Another challenge that came with this project was attempting to uncover or identify possible gender issues, and to describe them in this thesis without their descriptions having potential harmful effects on the company. A thesis is a public document, and it should be avoided that certain framings of results would have negative effects on VRVis and its employees. Close feedback loops with VRVis ensured that this challenge was overcome in the most efficient way.

Lastly, the fieldwork for this thesis was conducted in a timeframe where Austria's measures for the corona pandemic were strict (around March/April of 2020). All VRVis employees were to work from home, and everyone in Austria was not allowed to meet with other people outside of their household. These circumstances heavily influenced my fieldwork. Instead of conducting interviews face-to-face, they were now all moved to a digital medium. I asked each of my interview partners which software (Skype, GoToMeeting, Zoom etc.) they felt most comfortable with considering some supposedly had privacy issues. Most of the interviews were eventually conducted over Skype. Furthermore, the digital interviews necessitated an adaptation in the informed consent process. For example, me and some of my interview partners did not have access to a printer causing the classical method of signing a consent form unattainable. Therefore, I would send my interview partners the informed consent form a couple of days before the interview with the request that they would read it and get back to me with any possible requests for adaptations. Then, on the day of the interview I would start the recording and ask them to give their verbal request on record. In the informed consent, my interview partner agreed to the interview being recorded and that quotes from the interview would be used in the final thesis. The informed consent furthermore elaborated on the rights of my interview partners to discontinue the interview if they were to want this, and that they had the right to retract their consent at any time.

## 5. Empirical analysis

This chapter provides a discussion and reflection of the different interviews conducted for this thesis, and aims to bring to light aspects that can help answer my research questions. Considering that VRVis engages in projects related to engineering, VR, and machine learning, the methodologies of Tannenbaum et al. (2019), Corinna Bath (2014), and the Gendered Innovation reports (2013 & 2020) are kept in mind during the analysis of these activities. From these pieces of literature, I derived questions and best practice processes and then analyzed how the practices at VRVis relate or do not relate to them. In case something does not relate, this can then potentially be identified as an area where VRVis can improve on their gender awareness.

As a short recap, Tannenbaum et al. (2019) describe three approaches to reduce potential bias in machine learning: capturing metadata from your training data, counterfactual fairness, and multi-accuracy auditing. Furthermore, they provided decision trees that can help with the question about when it makes sense to engage in gender and sex analysis. The Gendered Innovations reports (2013 & 2020) provided some interesting gendered aspects to think about when it comes to developing VR technology, and furthermore underlined that intersectionality is an important factor to take into account whilst developing machine learning algorithms. The report additionally describes two multiple-step programs, one for analyzing sex and gender aspects in machine learning, and the other for analyzing sex and gender in the wider field of engineering. Lastly, Corinna Bath (2014) describes four mechanisms that could lead to gendered technology. First, the I-methodology, which is the process where technology designers see themselves as correct representations of their end-users. Second, the inscription of gender stereotypes within the technology. Third, the inscription of potentially harmful gender stereotypes in technologies that represent human characteristics. Lastly, explicit or implicit decisions that are made in the technology design process which could lead to possible gender issues.

What follows in this chapter is first an elaborate description of the four different project areas within VRVis and the practices that occur in each of them. This is then

followed by a description of specific practices or aspects that are analyzed in the line of the research questions and the mechanisms from the pieces of literature described above. What is important to note before starting the analysis is that any time a quote is taken from an interview, the annotation of the quote mentions which interview it came from, and the gender of the interviewee. Although this is not always usual scientific practice, I believe in this thesis that aims to analyze possible sex and gender aspects, and this is also supported by scientific literature (see Bührer & Wroblewski, 2019, and Tannenbaum et al., 2019), one's gender highly influences how sex and gender aspects are considered and viewed. Therefore, mentioning the gender of my interview partners can further serve as a possible interesting analysis topic.

### Technology design as practice: an overview of VRVis

Studying the developments within science and technology as a practice is one of the cornerstones of the field of STS. Groundbreaking ethnographic works in the early days of the field, such as the ones conducted by Latour and Woolgar (1979) and Knorr-Cetina (1981), show that even with the absence of the deep technical knowledge of a certain discipline, relevant conclusions can be drawn about the possible social influences on science if one has a strong focus on practices. Therefore, I knew that my first step towards immersing myself into the field of visual computing at VRVis in order to identify possible roles gender could play was to gain an understanding of the practices of the four different project areas. What kind of projects do the different areas engage in? What do these projects look like? What are the different phases of a project? How do the researchers at VRVis interact with their customers? Where is the necessary data for the different projects collected? However, as opposed to a classical ethnographic approach where I would be able to situate myself on the work floor at VRVis, something that was made impossible due to the COVID-19 pandemic, I collected the answers to the posed questions through the medium of interviews. Deepening my understanding of the work processes and practices going on within the four different project areas allowed me to identify where these processes could be touched by possible social influences, which could ultimately lead to identifying possible gendered aspects within the technology designed at VRVis.

What follows is a description of the practices and projects within the four different research areas at VRVis as described to me by my interviewees. I asked them to describe the projects they work on and the different steps that exist within their projects, but also about their personal opinions on where they themselves or the end-users have large influences on the technologies they work on. My hope was that this narration of their work would allow my interview partners to become more reflective both about their practices as well where their work could be touched by potential social influences.

Important to note is that each of my interview partners is highly embedded within their own work environment within the company. That is to say, even though the researchers I interviewed are located within a certain research area, they can only talk about their own experiences and the projects they have worked on. Considering I have to stay grounded within the data I have collected, this means I can only draw conclusions about situations and projects that are described to me in the interviews. Important to note is that during these descriptions, possible gender aspects are taken out of account. This is because the purpose of the descriptions is solely to gain a thorough understanding of the activities going on in each of the research areas. Analysis on possible sex and gender issues will be the focus later in this chapter.

### Visual Analytics

From the Visual Analytics research group, I spoke with two people who are involved with the Visplore project. Visplore is a software developed by VRVis over many years that has the goal to help people understand very complex and large data sets. An example of one of those large data sets can be sensor data from a machine production line. Let's say in a particular situation you have one of these production lines that consists out of multiple machines hooked up to each other producing an industrial product. One week the product resulting from this production line is perfectly fine, but the next week the product has failures and can therefore not be sold. In order to find out what is causing the failure, one can turn to the timeseries sensor data that is recorded by the sensors in the machines. These sensors can measure things such as air pressure, vibration, temperature, etc. However, these timeseries data get very large in a small amount of time if you record over

the span of a month or even a couple of days, and thus the data also gets complex. Visplore is a tool where it is possible to upload this data to then explore it and enable users to compare timeseries, and detect patterns in a fast and efficient way. Next to data exploration, Visplore can also be utilized for the validation and selection of forecast models. They have been involved in the validation of energy consumption prediction models with the APG Austrian Power Grid.

The exploration of the data and the detection of patterns is made possible through the visualization of the data in what are called *dashboards*. Dashboards are a collection of different *views*, which are, as my interviewee explained to me, visualizations “[...] where you can visually see something which is in the data” (Q1). Views can take different forms ranging from tables and graphs to histograms and scatterplots. The design of the dashboards with the different views is made specifically for each customer that wishes to use Visplore, considering each of them wishes to engage differently with their data and explore varying facets. Figure 1 shows a variety of different Visplore dashboards. The end product of each Visplore project is a custom designed dashboard that allows customers to analyze their data themselves. So they can, my interviewee explained, “[...] not just work with the data they have currently but they can also load then data from different time periods so they can in two years load another dataset and do the analysis again” (Q1).

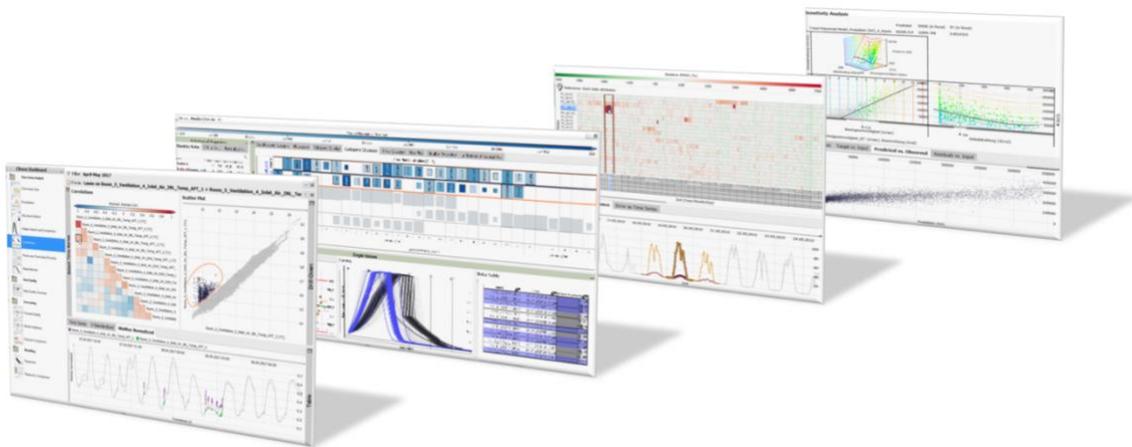


Figure 2: A collection of different Visplore dashboards. Source: <https://www.vrvis.at/en/research/research-topics/visual-analytics>

At the core, the Visplore software does not vary much from project to project. Visplore projects can therefore take significantly less time than some other VRVis projects where new software has to be designed. Most of the time within Visplore projects is

allocated to understanding the customer's data and designing a suitable dashboard. In the following I briefly describe what practices go on within a Visplore project, and how the interaction is with the end-users/customers of the Visplore software.

As it is described to me, all Visplore projects consist out of many collaborative processes. Each new project starts off with a long deliberation process between the (potential) customer and the people at VRVis. These meetings are described as follows:

“we prefer personal meetings. Especially at the beginning because it is a very difficult and long running process for us to understand the needs of the customers. Because they are deeply rooted within the technical field in which they are working and we have to understand the terms they use and the language that they use and that takes quite some time. Up to several days I would say”  
(Q1, female)

These meetings are therefore all about creating an understanding between the Visplore researchers and their customers. What are the needs of the customer? What type of data do they have? What is the problem they are trying to solve? What are potential views and dashboards that could help them understand their data? During these meetings the customer provides sample data and can also underline how the data should be handled within VRVis. One of my interviewees explained “[...] often they have confidential data that should not get out into the public because of competition. Like when you have a production line that has a certain combination of machines and the parameters and all that stuff is kind of secret, so they don't want the data to get out.” (Q2). After an understanding is reached amongst both parties during these meetings, the data of the customer is sent to the Visplore developer who then starts the design of the dashboard. One of these developers that I interviewed describes the process as follows:

“[...] They think about how it would be best that these dashboards look like for the customer and to develop this prototype just on paper with the customer. As soon as that is put on paper, then one of the developers starts to make a functional prototype out of that, so I do that part. So they give me like two or three pages of a prototype where it was written down in text and also with pictures, sort of like a mock-up, where they explained what they want to do with the dashboard. Then I do the Python dashboard, [...] then they ship that out to the customer, they look at it and give feedback again, and it goes so on and so on. So it's always forward and back.” (Q2, male)

As can be understood from the quote above, the process of designing the dashboards is highly collaborative. This is most likely the case because the Visplore customers are also immediately the end-user, because as stated before, it is the goal that they ultimately do

the analysis of their data and any potential future data themselves. Therefore, although it is the developers at VRVis who design the dashboard, the customer ultimately has the largest influence on its functioning and can sometimes even request completely new functionalities from the Visplore software. The developer that I interviewed further commented on this:

“[...] there is always a collaboration between the designers and the customer. I can remember that one customer wanted [the dashboard] to function in a way that we hadn't had yet. So we had to implement a new way to interact with the software to make that happen. So the customer decided how the dashboard would work.” (Q2, male)

The back and forth process between the VRVis developers and the customers continues until both parties are happy with the dashboard. Meaning that the customer can perform the desired exploration and analysis on their data. From the descriptions I have received during my interviews I conclude that the processes within Visplore projects are so collaborative because the end-user of the Visplore software is highly specific, and there is a high level of codependence between the customer and the Visplore researchers. Both sides are, as stated in a previous quote, deeply rooted within their respective technical field but need to clearly communicate their desires to the other. The customers need to make clear what their data shows and what type of analysis needs to be performed, and the Visplore researchers need to make clear what the possibilities are of the Visplore software, and design it in such a way that the customer can perform the desired analysis.

### Smart Worlds

From the Smart Worlds research group, I spoke with a rendering engine developer of the VISdom project. VISdom is a software that serves multiple purposes but mainly as a decision support system for domain experts in the field of flooding. The program enables an accurate flooding simulation of inhabited areas. As explained to me by my interview partner, a variety of domain experts such as hydrologists, engineers, and flood manager provide VISdom researchers with a mathematical model on which they then provide a visual flood simulation. The domain expert then looks at the important results and the analysis of the visualization and this provides her/him with some kind of decision-making

support. The output is not designed to decide something, but to provide the user, who carries the final authority, with hints that make it easier to reach a conclusion.

Another use of the VISdom software is public communication. According to my interview partner “[...] regarding floods, the public communication is also very important. Because almost everyone is involved in some way because building damage or streets are not reachable, so you would have to communicate this to the public and we also provide visualizations that can be used for public communication.” (Q4, male). The biggest project VISdom is involved in regarding this use is the Hochwasser Risiko Analyse 3 (HORA 3) project, which makes flood risk analysis maps for the entire country of Austria. HORA 3 is a federal project where VISdom simulation results are turned into maps and are published on a government website. Here they can be found by every citizen and they can get informed about whether they live in a risk area or not. Another smaller application is a joint project with a small start-up that develops an app that informs registered users about imminent floods. The app also provides flood forecasting services. Figure 4 gives an impression on what the VISdom software looks like.

The VISdom project has been developed for over 10 years, with its first publication in 2009. About its origins my interview partner explained:

“It started as a toy project to also consider flood simulation which at first was particle based, which is a word that is quite often used in computer games. It means that you are just toying around but it is not physically accurate. We used this to show that it is possible to simulate flooding like this with the system” (Q4, male)

After showing that flood simulation in such a manner can deliver viable results, the research group established a collaboration with a flood protection center in Cologne.



Each simulation starts with a simulation domain, which is a spatial region on which the simulation needs to be run. Usually this is based on a laser scan from aerial photography, which is then the topology on which the water flows. After that, buildings are digitally added because of course the water cannot flow through the buildings. Then extra geo-data, such as existing flood walls a city might have and land-use data is added. Land-use data refers to the surface of a certain area. For example, water propagates different over the paved surface of a parking lot than in a forest. All these approximations and these data are usually provided by the domain expert for which the simulation is done, but in part also comes from open data such as street maps. At the end of the day, simulation results are based on probabilities. My interview partner explains:

“You cannot exactly predict a flooding scenario, because we have the discrete data and we do not know the correct weather conditions and so on. [...] So what we do, we slightly vary the simulation parameters and make different simulation runs. Let's say we make 50 simulation runs with slightly different parameters for the same flooding and then we see okay, this area is always flooded, this area is only flooded in one scenario of the fifty, so it is highly unlikely that it will flood. So we can calculate probabilities like this, and this you also have to communicate to the domain expert.” (Q4, male)

When it comes to the interaction and collaboration with the end-users, my interview partner described that it differs per application of the VISdom software. With the domain experts who use the VISdom software as decision making support there seems to be a really close collaboration, whereas there is little communication with the general public who has access to the flood maps that are made in the HORA 3 project. With the domain experts there are continuous evaluations on the visualization techniques of the VISdom software. There is usually a pre-evaluation where mock-ups are discussed, and then there are later evaluations to discuss if the simulations fit the goals of the domain expert.

“Some of [the domain experts] are using our software all the time and basically when they report something to us, we immediately make a feature for that. So they say “I have this workflow, I have to do this every time and it would be so nice if I could just click on the building itself and then drag something over there”, then we just implement that. So sometimes it is oriented on their workflow, and sometimes we have an idea and we ask them if that makes sense.” (Q4, male)

When it comes to the general public, this tight collaboration is harder to achieve, and thus the feedback loops are very limited. For a paper published by the VISdom group in 2019 a larger user study was performed during which a questionnaire asking for feedback on

the different visualizations techniques was made. This seems to be the only moment when true user-feedback came from the general public to the VRVis research group.

### Immersive Analytics

From the Immersive Analytics research group, I spoke with somebody who works on the project Virtual Reality for Mission Planning. This project was initiated in 2018 for the Austrian Federal Ministry of Defense (BMLV) and the Austrian Institute for Military Geology as an attempt to see if Virtual Reality (VR) technology can improve current mission planning methods. My interview partner explained to me that traditional mission planning happens on 2D maps with a transparent foil on top on which certain paths, routes, and tactical symbols are annotated. Now the Austrian Military wants to see if a 3D immersive environment can improve this existing technique. Although, according to my interview partner, VR applications are not beneficial in every setting and for every problem “[...] in this case you have better spatial understanding of the environment than working with a 2D piece of paper [because] in VR you have a better understanding of the landscape.” (Q3).

In the Virtual Reality for Mission Planning project urban scenes are reconstructed from satellite images on a scale of 1:1, that thus allows a user to walk through a particular landscape as if they were actually on-site (VRVis, n.d.). This allows them to engage in a range of mission-oriented applications such as training their local knowledge, assessing visibility from different points in the area, carrying out certain maneuvers, and planning rescue missions. The technology is envisioned to be used by multiple users: one person who is inside the virtual environment and those outside but who can follow activity through computer monitors. My interview partner describes that this was a desire from the Austrian Military:

“They told us is that they wanted to use it in a setting where the mission planner or the head of the mission basically goes into VR and tells the person outside what he wants to do. And [while] he is in VR, he is interacting with the map and putting in these annotations and then the person outside is also asking question or the person in VR is asking questions. So it is not a stand-alone VR application” (Q3, female)

The person in and outside the virtual environment can interact with the landscape in four different views: a birds-eye view over the whole map, a first-person view, a third-person view, and a different version of the birds-eye view but where you can scale and interact with the map. Figures 4 to 7 give different impressions and descriptions of what the VR application looks like.



Figure 5: A person immersed in the virtual environment, whilst the monitors allow somebody on the outside to participate. Source: <https://www.vrvis.at/en/research/research-projects/virtual-reality-for-mission-planning>

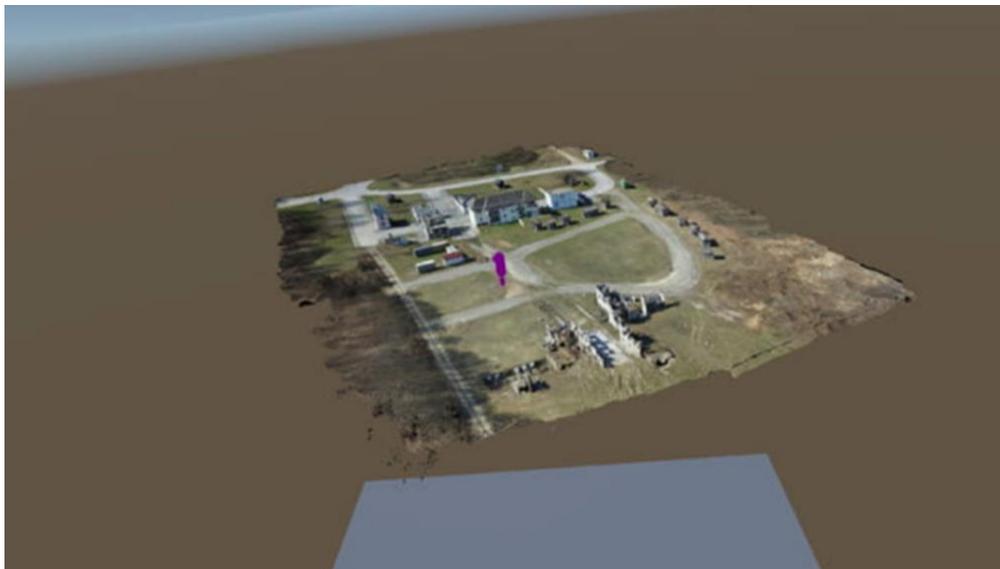


Figure 6: Birds-eye view of the virtual environment. Source: <https://www.vrvis.at/en/research/research-projects/virtual-reality-for-mission-planning>



Figure 7: First-person view in the virtual environment. Source: <https://www.vrvis.at/en/research/research-projects/virtual-reality-for-mission-planning>

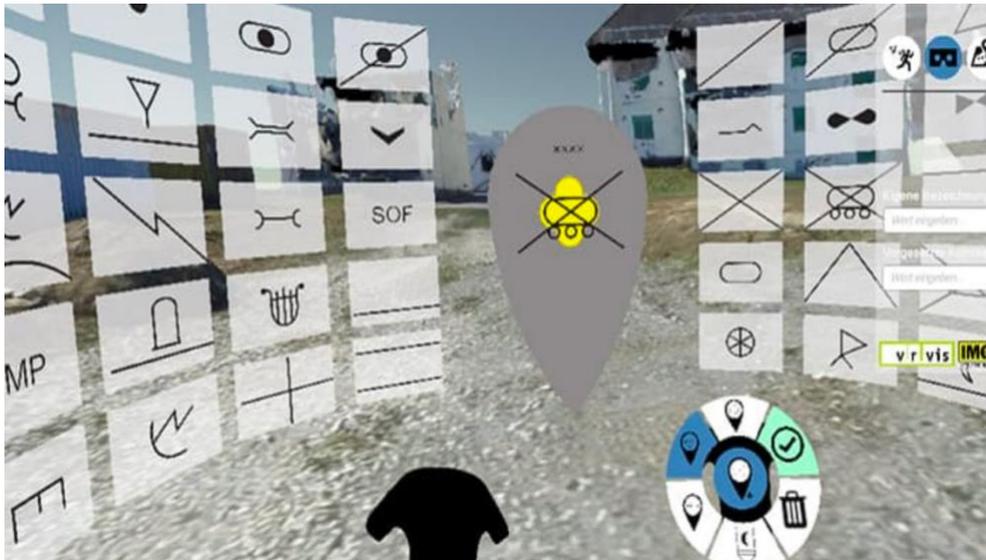


Figure 8: List of symbols which can be placed inside the virtual environment. Source: <https://www.vrvis.at/en/research/research-projects/virtual-reality-for-mission-planning>

The development of the Virtual Reality for Mission Planning started two years ago with a thorough deliberation process with the Austrian Institute of Military Geology. The deliberations were about what could be developed in VR, what would be beneficial to add onto the existing military processes, what kind of functions would the application need, and what is the nature of the data that is provided by the military. According to my

interview partner, this process took the Immersive Analytics research team about three or four months, after which they entered a straight development phase from July 2018 to about February 2019.

During this development phase it seems that the interaction between the researchers at VRVis and the Austrian Military was of a lower intensity than for example in the Visplore projects. For example, the Austrian Military gave VRVis a list of symbols and functionalities that needed to be incorporated into the application, but the VRVis researchers did not know the exact meaning of each of the symbols. About the process my interview partner said the following:

“I remember sitting down and talking about how can we incorporate so many features in the VR application. Because you have to imagine that you have two hands and two controllers. [...] You basically have two controllers with four buttons and two buttons where you have more input space. And now the problem was how can we map so many functionalities on a limited amount of input signals. That was definitely a challenge, so that is where we had the opportunity be a little bit creative and think about how we could design it. [...] We had a lot of freedom in deciding how we do the interaction. Because from the point of the client, the geospatial department, they just gave us a list of functionalities they wanted, or we just want to be able to do this and that. But they didn't say exactly how. So we based on our experience, research and knowledge we then designed and mapped the interaction system into the VR.” (Q3, female)

My interview partner continues to describe that during the development phase there was not a lot of back and forth regarding feedback between VRVis and the Military and Geospatial department. “[Sometimes] we came up with maybe two different options on how we can design that menu, and then we ask them do you maybe prefer this or that. But other than that, I would say we did not have a lot of input.” (Q3, female). The reason for this lesser extent of interaction than in for example the Visplore projects likely can be explained in the nature of the project, but also the expertise of the customer.

A last important aspect to note about the Virtual Reality for Mission Planning Project is that it is still rather young. The application developed is still far from being utilized in the actual military field. This brings along challenges such as a difficulty in defining the exact end-user of the technology. My interview partner explained:

“It is hard to define the end-user in our case. Because the project is a very early project I would say, it is basically our first version of what we would do in the next years even. And right now they are maybe using it internally and using it with their mission command. But it is not very

widely used I would say. It is just for now internally used and they are evaluating it or maybe showing it to other people or doing demos. So it is hard to define the end-user here, I would say the end-users are the people we are working with right now.” (Q3, female)

## Complex Systems

From the Complex System research area, I spoke with two researchers who are active within multiple projects within this department. Furthermore, it is important to note that speaking about the specifications of projects here is more difficult due to the fact that VRVis was careful with specific information that could potentially be disclosed in a publicly available master thesis. Therefore, the following section will describe in more broad terms what type of projects the Complex Systems area engages in, and what kind of practices are involved. Although this research area has some collaboration with non-medical industrial partners, it seems that the majority of the projects are engaged with imaging data from the medical field.

In hospitals nowadays, several types of medical images are recorded with different modalities (think of X-rays, CT, MRI, etc.). Some of these devices can display the image they record immediately, but others require a lot of pre-processing from the moment an image is recorded until the moment the image is showed. VRVis collaborates with companies that concern themselves with this process. My interview partner explains “[...] the company that we are working with is dealing with the pre-processing of the image from the raw acquisition to the displaying for medical doctors, who has no idea about the intrinsic of the machine” (Q5). Another example of a project that my interview partner worked on is a tuberculosis (TBC) classifier. Here a computer is shown a chest radiograph of a patient, to which the computer then analyzes if the patient exhibits TBC markers. The outcome of the analysis is a number between zero and one, one being there is TBC on the image and zero there is none. The result of this algorithm is supposed to provide decision making support for radiologists. Additional to this result comes further visual explanation that elaborates on why the computer says what it does.

Another important aspect to note is that VRVis is never engaged with the actual implementation of these medical imaging technologies into the hospital. My interview partners explain:

"We are doing research about improving several things and problems that [our customers] are dealing with. So they are coming to us and saying "we are having this type of problem" or "we want to improve certain visuals" and to this end we investigate this further. They want to invest more into a subproblem of their problem. So I myself am not dealing with the machines itself, I am also not writing code that is directly applied in these machines. I am only researching this topic and looking into how the problems that arise in this field which can be manifold, can be solved." (Q5, female).

"We are developing prototypes and we are not making the final products. If the prototypes that we make eventually will be developed into a final product then these things are done by the project partners. It's because the prototypes and the full implementation are sometimes also are written in completely different softwares. So then there are other people and departments involved." (Q6, male)

It seems that the work done by the VRVis researchers within the Complex Systems area can be better described as 'performing feasibility studies'.

The reason the Complex Systems research area has its name is due to the nature of the projects it engages in, which are mostly concerned with machine learning (ML) and deep learning (DL). ML and DL are the studies of algorithms that improve themselves automatically through training on large data sets (Mitchell, 1997). Translated to a VRVis project this means that the TBC classifier is an algorithm that increasingly has learned to better recognize TBC by being trained on a large dataset of chest radiographs. My interview partner explained to me that some of the datasets they use for their algorithms are over one million photos. This is of course an incredible amount on data that needs to be collected and, according to my interview partners, this acquisition happens in a multitude of ways.

"Most training is done on public collections that are governed by open collaboration by multiple institutions in order to avoid a single unit to have to come up with this huge amount of data. It's a multicenter approach. So we use this public data which is also important, because even when you use public data and then you publish something it can be reproduced. So it's easier or it makes it possible for other people to check your work. [Other times] our partners provide us with data that they themselves also have collected through different sources. And then they check it themselves of course to make sure that every requirement is completed." (Q6, male)

Projects in the Complex Systems research area, especially those with their partners in the medical field, start with an elaborate problem definition and problem

understanding process where both parties sit together and discuss what the problem would be about. Possible ideas about certain directions and prior knowledge about what works and does not work are discussed and communicated. Then, depending on the project, the VRVis researchers turn to the datasets. As explained above, sometimes they receive data from their partners and other times they will have to look if there are public data sets that can be used for their cause. At times, just from this data or the amount of data the researchers are able to assess whether they problem is solvable or not. As my interview partner put it: “[...] in the first phase [we] want to know: if we would collect this amount of data, would this be useful or can we already rule out that this won't work.” (Q5).

In the development phase for the medical projects, a lot of interaction exists between the VRVis' partners and the VRVis researchers. Depending on the partner, regular meetings are planned where it is discussed how the project is progressing and possible issues that could have come up are exchanged. This interaction is not a one-way street from VRVis to the research partners, but a mutual exchange. My interview partner explains:

“[...] they want us to share knowledge, they want us to share code, they want to learn from us. And also on the other side they are researchers directly working on the same problem or on adjacent problems. There is a collaboration going on where I do some investigations and then I report back and then they also have some ideas which is, I think, a little bit uncommon. That they are so involved in the research.” (Q5, female)

This transparency between both parties is necessary because the use of machine learning algorithms is still relatively new and the use of these algorithms in hospitals is highly regulated. Furthermore, as mentioned in my state of the art chapter, machine learning algorithms are capable of adopting biases and there is a big strive amongst both parties involved in the projects to avoid this as much as possible. More elaboration on this topic will follow later in this chapter.

### Detecting the social: finding the human in the binary

Now that a clearer picture has been drawn about the types of projects that VRVis engages in and the practices within them, we can take a closer look at the findings of the gender

analysis performed on the conducted interviews. As explained in my Material and Methods chapter, this thesis project has been initiated by VRVis. They wished for someone to look into possible gender aspects within each of their four research areas. The following subsections break down aspects that came forward in the analysis of the interviews as relevant to answering the overall research questions. They point to how gender is conceptualized by the VRVis researchers, to places where gender considerations might be important in the technology design processes, or to how possible gender issues are actively circumvented. However, none of the findings at this point are tied yet to specific research areas. It is a conscious decision to group the findings together at this point for two reasons. First, as mentioned earlier in this chapter, my interview partners are highly situated within their own projects and research areas, and I think that grouping the findings together can lead to possible broader conclusions useful and valid for the entire VRVis company. Second, to protect to the best of my abilities the identity of my interview partners. Although the quotes I use in this chapter are not harmful, if I were to analyze each research area separately it will be remarkably easier to deduct the identity of my interview partners. The findings will be tied to specific research areas in the following Discussion chapter.

### Talking about gender

A straightforward starting point if one wants to look at possible gender aspects of technology is how gender is talked about by the people working on these technologies, or how the topic is treated within the wider institution of VRVis. The following subsections break down the inclusive hiring processes VRVis engages in, the influence of these processes on the employees, and how the inclusive hiring processes cause a strong 'counting heads' logic within the company.

### *Inclusive hiring at VRVis*

Although it was specifically requested to research the possible sex and gender aspects inscribed into VRVis technology across the four different research areas, a topic that comes up whilst talking about gender in the broad sense is how it is treated by VRVis. I also wanted to discuss this with my interview partners as research shows that the general

attitude of an institution towards a topic logically influences how that topic is treated by its employees. For example, with regards to how RRI activities are encouraged by research institutes, Bühner and Wroblewski (2019) found that “[...] the institutional environment positively influences the degree of RRI activities and the general attitudes towards more responsible research and innovation: researchers working in an institutional environment that systematically supports the practice of RRI are more active in RRI than researchers who do not rely on such structures” (p. 8). Therefore, I asked my interview partners how the topic of gender is treated by VRVis.

“I think it is a topic all over the institute. Starting from for example an inclusive hiring process, which was as I understand reworked before I came here. So hiring is done very well as opposed to the places that I have seen before that just give you a list of skills and that is what you have to have, and of course it is all the skills and more. Then also students, we have a lot of female students, and there is also the femtech internships which is also a great way for female students to get into these research institutes.” (Q7, female)

“specific hirings I would say definitely, not just because VRVis wants to have a diverse group working environment, they also have regulations where the sponsor [tells you] to have at least a certain percentage of women at the office. So especially during the last years VRVis, from my perspective and understanding, was very focused on hiring women. If there were two people, one man and one woman, with equal amount of talent and expertise they would prefer the women.” (Q8, female)

Inclusive hiring processes are very hands-on methods to balance the gender or diversity quotas on any work floor, but have not always been met without controversy. For example, although these diverse hiring programs have been implemented by many Wallstreet firms, women and minorities have not gained substantial ground in management positions over the past 20 years (Dobbin & Kalev, 2016). The explanation given by Dobbin and Kalev (2016) is that rough top-down requirements put on people in management positions is often met with resistance and methods to circumvent these requirements. I asked my interview partner how she experienced the effect of inclusive hiring methods.

“[...] when I talk about the global field of computer science, not just VRVis but the industry as a total, I feel that there is a change. I feel the change on my male colleagues, not just at VRVis but from my studies and friends I have that are in computer science, male colleagues would say

'maybe you get the jobs easier in the next years because you are a women' and men start to feel like they have a disadvantage on the job market because of those regulations. Which I understand a little, but I guess you need some radical changes to permanently change the course of something. I hope they will stop those female preferences as soon as the quota is balanced." (Q8, female)

Even though it seems from this quote as if the inclusive hiring processes are met with some skepticism or mockery from males within the field of computer science, within the company of VRVis inclusive hiring seem to have a very positive effect. Take a look, for example, at this quote from a male VRVis employee.

"I think in general it is interesting now that VRVis got a lot of new women at the company. I think it only started because the grants we get, the government support, the research grants they demanded that we have more female coworkers. So I am not sure if it is still that way, but for a while the official standpoint was we only accept women. It's not that hard of a limit, because I know that we got some new men as well. But the ratio changed a lot in my general feeling. I think we got a lot of new female coworkers and I think that it is really interesting to see what that does to a field that was male-dominated for so long" (Q5, male)

I find this quote interesting because, as mentioned before, VRVis achieved the 20% barrier in 2018, and by 2019 female researchers made up 28% of the scientific staff (VRVis, 2019). Although this is still far away from being a completely gendered balanced number, this small increase has, in the experience of this male VRVis employee, changed his experience of the male/female ratio a lot and has made him excited to see the changes that could possibly come with the increase. Furthermore, research shows time and time again that having diverse research teams helps the development of gender and diversity inclusive science and technologies (European Commission, 2013; European Commission, 2020; Bath, 2014; Tannenbaum et al., 2019). Another aspect that hopefully becomes an inevitable consequence of actively hiring more women, is that more of them end up in leading research positions. At the moment, there are already female head of departments at VRVis, including the Complex Systems department. In a follow-up interview I conducted with a researcher from this department, this was something that was positively reflected upon.

"Also one definitely relevant aspects is that our group has a female boss which I think is relevant in itself because she does understand things and she understands certain topics that you would deem relevant and she gets it. [...] I think it is just in general if there are meetings with customers and there are ten men and it is just us two women, afterwards it is like "okay we know what was

going on". [...] I think one of the main topics is that she just understands these gender issues as opposed to many other companies where you have men who would probably listen to what you have to say but they would not get it in depth. [She] just understands these things because she also has been through the academic world and she knows what is going on." (Q7, female)

From this quote it becomes clear that having a female in a leadership position has a positive effect on the work experience of female employees. This is something that is furthermore supported by scientific research, and also something that could eventually lead to the development of more gender fair technology. Bühner and Wroblewski (2019) write that "[...] women support female colleagues, encourage gender-balanced teams and consider gender aspects in their research design more frequently than men." (p. 8). In sum, all of these aspects taken into account point towards an overall positive effect of the inclusive hiring efforts of VRVis.

### *'Counting heads' at VRVis*

Next to talking about the inclusive hiring processes, I asked my interview partners about their personal opinion on where gender aspects could be present. More specifically, I broke this question down into possible gender aspects in three sides: the technology designer side, the end-user side, and the data that is used to create the technology. When asked about their opinions, a lot of my interview partners seemed to have the same conceptualization: one where gender overlaps with biological sex. Take a look at the following quotes:

"We have both man and women working on our projects. There are also women amongst the project partners that we work with. [...] I am not sure if our technology is easier to understand for men than for women, because one time we had two female experts using our software and they were in complete disagreement about how it should work" (Q4, male)

"We have a good mix of male and female. I am not making the decision by myself, we usually do it together. I might come up with ideas, and they are coming with ideas too. But in the end we talk about what we are going to do. So maybe in the end possible bias is going to get erased by our communication in the decision-making processes." (Q3, female)

“I was quite happy in my last project that I had a female colleague that I could work with. [...] She was as stubborn with some parts as I saw in male coworkers as well. And it was just a joy to work with her, I think in general she was just a really nice person to work with in general.” (Q2, male)

As can be seen from these quotes, once asked about gender my interview partners referred to the literal amount of men and women on their respective teams and would continue to explain that they do not imagine any gender issues in that aspect. The only quote of the three that leans more towards gender in the sense of social stereotypes and attitudes is the last one. Mentioning that a female coworker can be ‘just as’ stubborn as the males at work is reminiscent of Wendy Faulkner’s (2014) work on women in the field of engineering. She describes how women in that field constantly have to resist against existing gender norms, that is to say, women who are good engineers might not be considered ‘real women’ and the other way around ‘real women’ are not ‘real engineers’. Talking about possible gender aspects in their work in the sense of how many women and how many men are involved in the projects can be seen to fall under Trojer’s (2014) approach of ‘counting heads’. However, this is something that Trojer (2014) identifies as an older branch of feminist technoscientific research.

The ‘counting heads’ framing also seemed to be prevalent when my interview partners were talking about possible gender aspects on the end-user of the technologies they develop.

“I am not sure how many of our customers or our contact persons and the people that use the software, how many are male or female. I feel like most of them are men, but maybe that it's just my, how to say that, because it's a technical field and it still feels like the technical fields are dominated by men” (Q2, male)

“I would say that in our case because of the environment where the application should be used we definitely have some gender aspect on the user side. Because in Austria the majority, I do not know the exact percentage, but the majority of the users are going to be men in our case. Yeah because we are working with the military department. Yeah, so definitely there is one gender aspect.” (Q3, female)

A possible reason that the ‘counting heads’ framing of gender amongst the VRVis researchers is prevalent is due to the focus of the wider institution on increasing the number of women in its staff. This focus manifests itself in the inclusive hiring processes VRVis engages in and is proudly put in the company’s yearly reports. When VRVis broke

the 20% barrier in 2018, it was highly celebrated in their annual report of that year, which mentioned that “[...] we are not intending to rest on our laurels here with regard to our previous success in staff development: we want to do more!” (VRVis, 2018, p. 7). The focus of the company on increasing the number of women amongst its staff can explain why the employees, when asked about gender, initially think about the amount of men and women working on the projects.

### Influence of the designer

During the interviews I conducted, specific attention was paid to the influence my interview partners had on their respective projects. This is in line with my second sub question and the suggested methodology from the Gendered Innovations report (2013) that suggests to look at effect of a person’s assumptions on their work. It is furthermore closely related to one of Corinna Bath’s (2014) suggested traps that lead to possible gendered technologies, namely the inscription of potential harmful gender stereotypes. As can be seen from the quotes below, awareness amongst my interview partners exists that they can strongly influence the technologies that they work on.

“As a developer you always have a lot of influence on the software you produce. When it comes to decision making. So maybe the most influence I have as a developer is when it comes to interaction design” (Q2, male)

“If you develop software you cannot take every little detail into account, so along the way you, as a developer, at some point you make a lot of decisions assuming something. Based on the vision that you have as a developer you create this vision by talking a lot about the kind or evaluating what they want or what they need. So of course, like we made a lot of decisions along the way which might influence the end-user.” (Q3, female)

However, the question is when do these influences lead to possible harmful gender aspects? I think the answer to that question lays in what the nature of a specific project is. For example, according to Tannenbaum et al. (2019) a dissection of the role of gender in a particular technological development is not necessary if the product under study (in this case the technology VRVis develops) does not affect humans. The following quote comes from my interview partner who works on the Visplore projects:

“I would say where the developers have the most influence, maybe I think is the algorithm in the software. Because I would say that mostly that is all too technical and I would say that most customers probably would not care as long as it works.” (Q2, male)

Considering most of the Visplore projects concern themselves with data from industrial machines and have as a goal to improve the functioning of such machines, I tend to agree with my interview partner here. An analysis about the inscription of possible gender aspects in such technology makes less sense than thorough analysis of technology that does affect humans. An example is the Virtual Reality for Mission Planning project, which has humans as their end-users and will potentially affect human behavior when it is implemented in the field. One of the features of the VR application is the calculation of visibility from a certain viewpoint. The user inside the virtual environment is able to drop an avatar somewhere on the scene, and the application will then calculate the visibility of this particular avatar. A big part of the calculation of this visibility is the height of the avatar. My interview partner explains:

“One example would be concerning design choices, for that version we don't have a lot of interactions with settings so for example during the development I would assume that this calculation I'm doing has to have a certain height. If you get the height of puppet correctly only then would it be right. So I would for example say, the eyes are on 1.60 meters, and then based on that we do the calculations. So that would maybe one example of design choices that might influence the user in the end. Because the eyes of the user might not be at 1.60.” (Q3, female)

This is a clear example where a human characteristic is simulated, and a calculation is made based on this characteristic. In my state of the art chapter I point to multiple strands of literature, such as the one from Rachel Weber (1997) about the design of military cockpits, that state it is possibly harmful to make calculations or designs based on certain dimensions as it can exclude people without those dimensions from properly utilize the product or technology. Another VRVis research area that affects humans are the medical projects in the Complex Systems area. However, since these projects are involved in machine learning and deep learning, I devote a separate section to this matter at a later point in this chapter.

### *User interfaces*

Another clear influence that the VRVis researchers in at least three of the research areas have on the technology they develop is the design of user interfaces. In the Visplore projects this translates to the design to the dashboards, in the Virtual Reality for Mission Planning project to the placements of the different symbols and other functionalities into the application, and in the Visdom project to the design of the flood visualizations. My interview partner from the Visdom project also explained that this was his main responsibility.

“[...] You want to show in very few images what is the most important information. So what I do is develop visualization techniques together or to identify important information and to visualize it in a way that the user realizes ‘ah this is important and I have to look at this’” (Q4, male)

As mentioned in my state of the art chapter, through mechanisms such as the I-methodology features and user interfaces of technology can become gendered (Oudshoorn et al., 2004). The question however is if this is the case for the VRVis user interfaces. When asked about the matter, my interview partners mostly focused on the color choices.

“I had some type of project which was some deep learning project which is where you output some sort of segmentation results on the data. So you output visually where is the background in this image, where is the foreground, something like that. And the colors that you use for this, I mean, I can choose the colors where I separate, for example, foreground from background. But when I provide these colors to someone else, that person might be color blind. So in user interface design this is always a topic. If you can create images, you can use whatever color you want. You can have a look into designing this like, with regard to colorblindness. And also in this sense this is also a way of countering gender bias. Because I think in a way color blindness has high correlation with gender. So I then made the first version of this, I used some type of colors, and then I thought it might not be a good idea and so I implemented a second type of color scheme that was more 'color blind friendly'.” (Q5, female)

“The one thing we definitely have to take care of is because we are using visual tools, so we rely on human perception. We rely on that users use their perception to interpret the visual things that they see. And here we definitely have to take care to be as diverse as possible. Because there are people who will be colorblind. [...] But the only thing I see here is that I didn't, I am not aware, of any gender issues when it comes to color and similar things. So at least I am not aware of any study that there might be a difference in color perception for men and women, at least I do not

know any. If there would be some we would take care of that, but so far we concentrate on diversity in the sense that we also make sure that the visual representations are visible for color blind people.” (Q1, female)

Although I do not want to overstate or overgeneralize differences between males and females, it was interesting to me that all three females I interviewed for this research all mentioned color schemes in user-interfaces as a possible place where gender needs to be taken into consideration whereas none of the males I interviewed mentioned it. Whether the user-interfaces actually carry any gender aspects is something that is beyond the scope of this thesis. However, if VRVis would be interested in uncovering this it might be a potentially interesting topic for further research.

### Black-boxing

Writing code in order to develop software includes many processes of formalization and abstraction. Corinna Bath (2014) identifies this as one of the ‘traps’ that could lead to gendered technology. The reason these abstraction processes are a possible trap is because possible gender aspects get inscribed in a technology that is consequently black-boxed. This ‘black-boxing’ of the technology developed by VRVis is something that has come up in multiple interviews, especially when it was in the light of end-user/customer interaction. As described in the first half of this chapter, it is highly dependent on the nature of the project and the research area how the interaction and collaboration is with the end-user, and therefore the amount that a technology is black-boxed also differs. Important to underline is that black-boxing is not something that should be avoided in every project, but certainly is something that needs to be considered in some. These are some quotes from the different conducted interviews:

“But I would say where the developers have the most influence, maybe I think is the algorithm in the software. Because I would say that mostly that is all too technical and I would say that most customers probably would not care as long as it works. But that’s probably just software engineering in general as well. As a user you just want to search stuff, and how it searched through all the data you don’t care as long as it’s fast, and then the developer decides which algorithm he uses to accomplish that.” (Q2, male)

“It’s like a collaborative thing. We inform them about what we are doing, they control stuff, we don’t develop it for ourselves and then give them like a black-box solution. It is openly and

collaboratively developed. So we make them aware of what we are doing and what biases we have found, they also make us aware if they have found something that we have not found, so nothing goes unseen, well hopefully.” (Q6, male)

“[...] they want us to share knowledge, they want us to share code, they want to learn from us.”  
(Q5, female)

The first quote comes from my interview with someone working in the Visplore research area. As mentioned in the previous section, inscription of possible gender aspects is less of a concern here since most projects in this area concern themselves with the functionalities of machines. In the cases of the Immersive Analytics and the Smart Worlds research areas, as explained to me in my interviews, it also seems less important to share code. This is because the researchers at VRVis in these research areas deal with specific types of end-users who are not experts in the field of visual computing but in other domains such as people enrolled in the military (in the case of the Virtual Reality for Mission Planning project) or domain experts in the field of flood prevention (in the case of Visdom). These end-users have therefore less to gain from receiving full access to the code. The other two abovementioned quotes come from my interviews with the researchers working in the Complex Systems area, which is on the other side of the spectrum. Here everything, including the code, is shared with the end-user. This can be explained by the fact that a lot of their projects concern themselves with machine learning. Within machine learning the inscription of possible bias is a serious possibility, as described in my State of the Art and Sensitizing Concepts chapters, and therefore the sharing of the code is more important here. The next section concerns itself specifically with how possible bias is avoided in the machine learning projects.

### Preventing bias in machine learning

“[...] you could say preventing bias is like the main thing that we do. If you want to extract information, what you actually have to do is focus on this information despite a plethora of different biases that are present. It is not just gender, it's also age, it's ethnicity, basically everything you can think of will be there and it influences the network in a way you don't actually want it to influence.” (Q6, male)

“We are constantly talking about bias and constantly reminded of bias because in a way all real-world data is biased. And every dataset that you encounter will have some sort of bias and you will have to be aware of that, you will have to investigate what sorts of biases are there.” (Q5, female)

The fact that machine learning algorithms can adapt certain biases is a big topic nowadays, and has been broadly discussed in my State of the Art and Sensitizing Concepts chapters. In my interviews with the researchers in the Complex Systems research area I therefore spent substantial time talking about this topic, and how they treat potential bias in their practices. As becomes apparent from the quotes opening this section, bias is something that is constantly thought about within the projects, and dealing with it is important in every phase of a Complex Systems project.

What was furthermore one of the goals of this analysis was to see if the researchers at VRVis employ, perhaps in parts, the suggested methodologies to develop fair machine learning algorithms as suggested by the 2020 Gendered Innovations report and Tannenbaum et al. (2019). The suggested methodologies are developed to go against the reinforcement of negative social structures machine learning algorithms can sometimes play a part in. The Gendered Innovations report (2020) suggests three steps for analysis, each coming with a list of questions that need to be answered. The questions are all informed by state-of-the-art research in the machine learning field. The first step is an analysis of the social objectives and implications of the work. “Teams need to work together to define the objectives of the system being developed or researched to evaluate its potential social impacts” (European Commission, 2020, p. 220). The questions that need to be answered for this step are concerning the priorities and objectives of the developed technology. Who will it benefit? Who will be left out? When asked if VRVis considers social objectives of the technologies it develops, my interview partner mentioned the following:

“[...] impact of technology is a conversation topic at VRVis. Also in other groups where people are looking into how to convey information properly and I think it is also a topic for when new subjects are approached. How relevant is this, for whom is this relevant?” (Q7, female).

Staying on the topic of conceptualizing the impacts of technologies, the discussion quickly moved to examples of machine learning where this conceptualization was not done correctly and resulted into problematic algorithms with regards to gender or race. The

overall discussion made it clear that there exists an awareness with the VRVis researchers that gender is an important aspect to consider in certain projects, and that the uncaredful handling of it can cause problematic machine learning algorithms. A famous example that was given by my interview partner is the recent controversy surrounding the automatic picture cropping algorithm from Twitter. The algorithm learned how to automatically crop pictures based on eye-tracking, meaning where people look at in a picture, which eventually resulted in the algorithm having a bias towards white people, and pictures of women getting cropped at the upper body. Twitter has since apologized for its faulty algorithm (Hern, 2020). My interview partner continued.

“So there are a lot of things going on in this direction, and a lot of stereotypes and assumptions already built into the design of the problem. It is not just, you have a problem which you have to answer and you have to take gender considerations into account, but the problems themselves are already stated in a fashion where you want to infer gender from for example handwriting. [...] If someone would come with 'we would want to infer gender from facepictures' and things like that I would be the first to say no, we are not doing that. And I think people would understand.”  
(Q7, female)

This makes it clear that there is awareness in the Complex Systems research area about potential problematic machine learning algorithms with regards to gender, and that there is a responsibility on the side of VRVis to choose projects that do not belong to that category. This was also made clear in the other interview I conducted within this research area.

“What one could add it's also, one is responsible for what kind of projects one does, but yeah. But as you said, we do everything transparently and open and we don't do unethical projects.” (Q6, male).

The second step suggested by the 2020 Gendered Innovations report in order to develop gender inclusive machine learning algorithms is a thorough analysis of the data sets. Important for this step is available metadata that includes information about gender, sex, ethnicity, age, and geographical location. This metadata can consequently be used for assembling diverse datasets across different variables. Tannenbaum et al. (2019) additionally underline the importance of using metadata to compile balanced datasets. When asked about strategies the VRVis researchers employ for analyzing the datasets used, they explained that as a first step, especially in any kind of classification tasks, they

look at any potential class imbalances in the data on which they will train their algorithm. This means that they go over the datasets and try to analyze if there are imbalances in the types of images within the dataset.

“This is one thing that you will look at first before you do anything else. So if you have in a way, meta-data which are descriptions of what is in the image then you can use this meta-data to already investigate. Like you have several columns describing this image and if you say okay, in one column you have ten of this category and ninety of the other category then you can't investigate in this way” (Q5, female)

Any type of imbalances that stand out this way are immediately recognized. This refers to any type of imbalances ranging from image qualities, to the machines which were used to take the images, and color schemes. Furthermore, it is important to check if the images that are labelled with a certain annotation actually show this annotation. One of my interview partners gave an apparently famous example in the visual computing field of a publication that came out about a plural diffusion classifier. Plural diffusion is a condition where water is found in the lungs. Apparently, this research used a big public collection of chest photographs that were labelled as ‘plural diffusion’, but it was discovered later that these photographs showed to condition in a state where it was actually already treated. Therefore, the algorithm developed by this researcher would not be helpful to aid doctors in hospitals. Another strategy employed by the VRVis researchers is analyzing the datasets in terms of compilation.

“You have to look at how data sets are compiled, how they are collected, talk to customers about what the limitations of their data sets are, not only with respect to gender but also with respect to gender.” (Q7, female)

However, a huge problem within machine learning projects in general is the annotation of public data sets, especially when it comes to medical images. This is because medical images are often anonymized and so the ‘meta-data’ of individual images would not include information about age, gender, ethnicity etc.

“[...] with medical data you have anonymous data a lot. So either you don't have this meta-data at all because you could derive some sort of information from it. For example, some public data sets just give you the images and that's it. And even if you get this meta-data, then it might be anonymized. So there might be some kind of algorithms that you can't really see, how can I say this in English, you can't find out who was this person or where was this taken or which device was this taken with, and therefore you sometimes get meta-data, but it's not usable.” (Q5, female).

The fact that the meta-data is unavailable already hinders further strategies suggested by Tannenbaum et al. (2019) such as multi-accuracy auditing because this relies on having meta-data. A method that is used to counter this problem within the broader field of visual computing is hiring people to annotate this data. But this again, according to my interview partner, could lead to some potential bias. For example, humans are only capable of labeling a certain number of images in a fixed amount of time and have to therefore make choices about what they label. If a person is paid by the hour or by the number of images they label, they might skew the labels towards data that is easy to label. Furthermore, according to my interview partner multiple initiatives exist within the field of visual computing that aim to overcome such challenges and develop fair machine learning. These initiatives, such as AI for Good and Fairness, Accountability, and Transparency in Machine Learning, develop and argue for multiple methods to achieve fair machine learning algorithms.

“[...] it goes towards maybe guidelines for collecting data. Also, towards analyzing existing publicly available datasets for biases and like putting the information out there. There is one initiative that is writing guidelines on how to collect data in an unbiased, or as much unbiased, fashion as possible.” (Q7, female)

Moving on from the topic of developing datasets, I focused on how gender is considered whilst developing their algorithms. When asked about it, both my interview partners stated that it highly depends on the project, but that at times a gender issue is also inscribed within the problem that they are working on. An example that they both brought up was that of mammography, which is the process of using chest x-rays to check for chest tumors.

“Mammography is kind of a hot topic in machine learning right now, one of the obvious biases is that most training sets will be actual mammographs from females, but chest tumors are not limited to females. But still the training that you get will be more or less only females. [...] So you cannot then use this kind of software that is then only trained on female data to inspect male data. Or if you do, you need to be aware of that there might be some unforeseen side-effects. So these limitations, these biases have to be made clear. So that the end-user is aware of what can be done or not” (Q6, male)

“[...] you for example have mammography where you are looking for breast cancer. Then most of the images that you will have are images of women, and most of the cases that you will want to

maybe investigate with this type of data will also again be women. So some of the biases in a way are inbuilt in the data and in the problem.” (Q5, female)

Another small example that one of my interview partners brought up is that within these public data sets there is a strong bias towards sick people. This is because these datasets are for a large part compiled out of pictures collected in hospitals, and the majority of people in the hospitals do have some sort of ailment. These type of inscribed biases are hard to overcome and seem near impossible to solve for the researchers at VRVis themselves, since they are not part of the data collection processes themselves. When asked how they overcome such biases, my interview partners described two steps: asking for more data or clearly communicating with their customers on what type of data their current algorithms are trained.

“[...] asking for more data is always an important step I would say. Because in a way a model can only learn so much, and it can only learn, I mean up to some technical details, what you present to it. And therefore the data selection is very critical, and asking for more data if you don't have all that you takes. Or maybe if you cannot get more data right now then you would also have to say okay, this is now the current status, and you always have to say this type of data produces this type of results” (Q5, female)

“regardless of gender bias or regardless of bias, you will always have to communicate clearly 'this model was trained on this type of data' because you might also run into other problems. [...] Of course, we try to make everything as public and transparent as possible. Both for ourselves and both for our partners because they also want to, or they don't want to use anything that would be biased in a negative way. We try to work as transparent as possible and we also publish everything.” (Q6, male)

A big question that is raised here is one regarding responsibility. It seems from the interviews that the researchers at VRVis are doing everything within their capabilities to avoid any sort of bias. However, that does not mean that bias is not accumulated elsewhere. In the data collection or the creation of the data sets for example. The next subsection further reflects on the topic of responsibility for potential bias.

## Prototyping and Responsibility

Since VRVis is a research company, many of its projects focus on developing prototypes. Considering prototypes are preliminary versions of something on which then further development is based, an important question to ask is who is responsible for potential gender issues within these prototypes. The Complex Systems is one of the research areas that is developing solely prototypes. Within this research area there exists close contact between the VRVis researchers and their customer, and due to the fact that bias in algorithms is a large topic within the field of machine learning, conversations about the topics of bias in any form are constantly happening. However, as established in the previous section bias can be accumulated in other places which the researchers at VRVis cannot control such as the data collection or the creation of the data sets. In my follow-up interview with a researcher from this area, my interview partner had the following to say about the responsibility of the possible gender issues inscribed into the prototypes.

“I think our responsibility is to analyze the data in the fashion that we can, or even say 'with this data we cannot do what you are asking' and to communicate this as clearly as possible. Maybe also guide in getting more data, maybe also say 'this is what is missing', 'this is what we don't have and what we would need'. But yeah, at some point we have done our analysis, we communicate this, we deliver something. Of course, there also has to be a focus on what is good and what is bad, what does not work, what is working? And then I think at this point, I would see it like this, maybe other people would not agree. At this point, when we have done our best to communicate what was done then the responsibility also falls onto the customer. If they develop things further and we have made them aware of certain issues, then it is also their responsibility to be aware of what we have made them aware of.” (Q7, female)

It seems from the interviews that the researchers at VRVis are doing, regardless of the aspects they cannot control, everything within their capabilities to avoid any sort of bias. The difficulty in general in this process is that the responsibility for who has to deal with the potential bias is initially moved from the data collectors to the VRVis researchers, who will eventually move it on to their customers who are in the end actually responsible for implementing the technology into the hospitals. However, the Complex Systems research area seems to have a good conceptualization of the responsibility, because there is constant open communication about all the expert knowledge the VRVis researchers can provide to their customers. Open communication about expert knowledge regarding

possible sex and gender aspects with customers can stimulate the responsible development of more gender-inclusive technology.

Another research area within VRVis that develops prototypes is the Immersive Analytics area which works with VR technology. The Sensitizing Concepts chapter reviews the 2020 Gendered Innovations report which discusses gender considerations that should be taken into account in order to make VR technology inclusive. Research namely shows that women, men, and gender-diverse people can differ in how they experience virtual environments (European Commission, 2020). For example, the level of perceived realism of the virtual environment is higher in men than in women (Felnhofer et al., 2012), and women are more than twice as likely to experience negative virtual reality-induced side effects (VRISE) such as disorientation, increased heartrate, sweating and nausea (Al Zayer et al., 2019). Due to all these gender considerations, the Gendered Innovations report underlines that “[i]t is important to develop virtual settings in a sex-sensitive and gender-sensitive way, applying sex and gender analyses that include both women and men in the development and testing of prototypes” (European Commission, 2020, p. 139). In my follow-up interview with the VRVis researcher working on the Virtual Reality for Mission Planning project I asked if these kinds of gender considerations are taken into account in the development of the project.

“Well we always talk about VR related ethics and psychological aspects of VR, but we don't really talk about it in terms of the development itself but just to exchange insights and studies. Yeah as I said, we are mainly working, since we are a research company, with prototypes and those questions are not really the focus of prototyping. I feel like those specific question of gender would rise if a product would go beyond prototyping into serious development.” (Q8, female)

“Color perception differences have not really been taken into consideration. Like I said it is a prototype, it is a very early software. And from the functional part it is way more important than the design choices. It has to work first and if it works, then you can think of design choices” (Q8, female).

I find this approach towards prioritizing function over design choices interesting, especially since it goes against the process recommended by the 2020 Gendered Innovations report namely to consider these gender differences also in the prototyping phase. It is also interesting that it seems that the VRVis researchers are aware of different

ethical and psychological aspects of VR technology and exchange these amongst each other, but deliberately choose not to consider these in the prototypes because the focus lays on the function. This raises the question what it means when a technology, product, or prototype 'works'. By focusing on having certain features function before considering gender inclusive design choices, awareness needs to be raised that the prototype as a consequence functions well in one way, namely performing the functions discussed with the Austrian Federal Ministry of Defense (BMLV) and the Austrian Institute for Military Geology, but does not work in a more gender diverse way. Of course, it is logical that the focus is where it is because you always want to adhere to the wishes of your customer, however, the customers in the case of this project are not VR experts and are therefore likely not aware of these possible gender considerations. Whereas it becomes clear from the quote from my interview that the researchers at VRVis do have discussions about psychological and ethical aspects of VR technology. Having these discussions as well with customers can perhaps spark some reflection or discussions about what functionalities are prioritized in the development of prototypes.

Furthermore, the line between what is a prototype and what is a final product is not always completely clear. As far as I understand from my interview with the VRVis researcher working on the Virtual Reality for Mission Planning project, the developed prototype is internally used by the military at the moment in order to evaluate it or to potentially run demonstrations. In a sense, the prototype is already used in the field as a type of product. This then automatically raises the question of responsibility. Who carries the responsibility for the potential sex and gender issues inscribed in these prototypes? When asked this question, my interview partner said the following:

"I cannot really answer that. Those are questions that are going to be solved when the project is going to be further developed. I cannot really say what we are going to do with the project. I mean, I understand that thinking about gender-friendly software or developing software in general for a diverse target group is very important. So I would definitely, personally as a developer, let it influence my work. But as I said, in a prototype stage there are much more other higher priorities." (Q8, female)

What becomes clear from this quote is that the questions regarding responsibility and sex and gender differences in the experience of VR have not yet been discussed or conceptualized, even though the VRVis researchers are aware of them. If these questions

are raised earlier in the process, a better definition of what needs to be accounted for in the prototype can be conceptualized with the customer, and that way responsibility is then also better conceptualized.

## 6. Discussion

The previous chapter elaborated on the analysis of the conducted interviews for this thesis and set out different strands of potential sex and gender aspects that exist within the work done at VRVis. However, the findings of the different sex and gender aspects were all discussed together and not yet tied to the four specific research areas. Considering, as mentioned in the Material and Methods chapter, it was a specific wish of VRVis to analyze potential gender aspects within each of the four different research areas, this chapter will do exactly this. In the following, each of the four areas (Visual Analytics, Immersive Analytics, Smart Worlds, and Complex Systems) are individually discussed, and specific findings are tied to them based on the conducted interviews and the literature discussed in the Sensitizing Concepts chapter. First, a short recap is given about the type of work and projects the different areas engage in, which is then quickly followed by the specific analysis.

### Visual Analytics

Within the Visual Analytics research area, I conducted two interviews with VRVis employees who mostly deal with the Visplore project. On the VRVis website it is stated that visual analytics “[...] refers to an interdisciplinary approach that deals with how data analysis can be enabled by computer-aided visual representations” (VRVis, n.d.). Mainly they aim to translate large amounts of data, coming from a variety of different areas and industries, into images as visuals are more easily interpreted and understood by people than pure numerical numbers. Within the visual analytics area Visplore is a large project that was so successful that the preceding head of the department now started a spin-off company with the software. Visplore, completely in line with the general goal of the Visual Analytics research area, aims to help customers perform specific types of analysis on large data sets that usually come from sensors through the means of data visualization. In a typical Visplore project, VRVis employees engage in personal meetings with customers to get familiar with the nature of the customers’ datasets and the type of analysis that needs to be performed. When an understanding is reached, VRVis continues to design a specific

*dashboard*, which can be compared to a type of user interface that allows for a variety of data visualizations which all help towards the specific type of data analysis the customer wants to perform on their datasets. The process of perfecting the dashboard is a constant back-and-forth between VRVis and the customer, until in the end a version is made that allows for the specific data analysis the customer wants to perform, and any potential future analysis.

In general, this research area has very little identifiable or relevant gender aspects within the frame of this thesis. The main reason for this is the nature of the application of the Visplore software. Following Tannenbaum et al. (2019), sex and gender analysis is not relevant for every type of research design, and they therefore offered two different pathways (see figure 1 on page 29) to help researchers figure out how relevant it is for their own research. At the top of the two pathways, one for analyzing sex and the other for analyzing gender, a similar question is posed: 'will the product under study affect or involve humans?'. For the vast majority of the Visplore projects, the answer to this question is no. Most of the customers in Visplore projects come from the industrial field where the software is utilized to perform data analysis on industrial machines. The majority of the effects of Visplore are therefore on machines, not necessarily on humans. The analysis of the two interviews conducted with the Visplore researchers additionally did not point to any potential sex or gender issues. No inscribed gender issues can be found on the data side of the projects considering the data mostly comes from machine sensors. Each of the four mechanisms described by Corinna Bath (2014) that could lead computer scientists to possibly create gendered technologies also do not carry relevance in the case of Visplore. The first one, the I-methodology, seems unlikely to have large effects within Visplore projects due to the constant back-and-forth process between the VRVis researchers and the customers. The dashboards are specifically designed and altered until a point is reached where the customers can perform their own analysis. And so because the dashboards are designed for specific users, the I-methodology, where the designer sees themselves as an appropriate representation the use is not in effect. The other methodologies described by Bath (2014), the ones that mainly refer to different types of possible inscriptions of harmful gender stereotypes in either the technology itself

or the abstraction processes that to occur in the technology design processes, are also not relevant in the case of Visplore due to the industrial applications of the software.

A spot where potential gender aspects could play a role is in the design of the user-interfaces of the dashboards. However, this awareness is already present amongst the Visplore researchers. The following quote comes from one of the two conducted interviews within this research area:

“The one thing we definitely have to take care of is because we are using visual tools, so we rely on human perception. We rely on that users use their perception to interpret the visual things that they see. And here we definitely have to take care to be as diverse as possible.” (Q1, female)

Furthermore, analyzing the user-interfaces is not something that has been done within the frame of this thesis and is not something that could have been uncovered through the means of qualitative interviews. If VRVis wishes to uncover potential gender aspects within user-interfaces it will have to do so through the means of further research. Lastly, in one of the interviews it came up that Visplore is a tool that is certainly not limited to applications within the industrial field. A given example of a potential further application was the analysis of medical data. Although this application at the moment is just a hypothetical, it can be strongly recommended that at the moment the answer to the question ‘will the product under study affect humans?’ changes into a ‘yes’ to engage in an analysis of the social objectives of the project in lines with one described in the Gendered Innovations (2020) report, in order to not only avoid any potential sex and gender issues, but also other diversity issues such as age, ethnicity, or culture.

## Smart Worlds

Within the Smart Worlds research area, I conducted one interview with a VRVis employee active within the VISdom project. The Smart Worlds area mainly concerns itself with computer simulations and making digital copies to investigate real world problems. Using simulations allows for a safe framework in which one can run through all conceivable scenarios and parameters and apply the results in the real world. Applications of projects within the Smart Worlds research area are mostly situated within the industrial and spatial planning sector. VISdom is a visual flood simulation program that is supposed to provide a variety of domain experts such as hydrologists, engineers, and flood managers

with decision-making support. Although there are some adaptations of the VISdom software such as providing flood maps to the general public to communicate about potential flood risk, the main goal of VISdom is to provide decision-making support to experts. Developed over ten years, VISdom is now one of the leading flood simulation software that can be specifically adapted for different domain experts. Similar to Visplore projects, there exists a close collaboration between the domain experts and the VISdom researchers, and adaptations can be made to the software until the domain expert can perform the exact data analysis they wish.

For the analysis performed in this thesis, the VISdom project does not seem to have any identifiable sex or gender issues for very similar reasons as the Visual Analytics research area. To the question Tannenbaum et al. (2019) pose in their decision trees ‘will the product under study affect or involve humans?’ the answer is not a straightforward ‘yes’. Of course, decisions made by the domain expert using the VISdom software will down the line affect human beings, and humans might be affected due to the quality of the predictions, however, it seems extremely unlikely that gender is a relevant factor here. What could be perceived in this research is the nature of the data and calculations used for the VISdom software, and identify if these possibly carry sex or gender aspects. Asking the question if the software itself involves human results in a negative answer. The data utilized to make the flood simulations involves technical aspects such as water propagation calculations, height and width of building, or particular types of surfaces over which water propagates. Furthermore, at no point during the flood simulation are any human characteristics simulated so here as well no humans are involved. Therefore, automatically two of Corinna Bath’s (2014) four mechanisms, the ones that talk about the (un)intentional inscription of gender stereotypes, do not carry any relevance. The I-methodology also is not relevant in this case considering the end-user of the VISdom software is highly specific, and there is a high level of collaboration with the two sides which makes it hard for the developers to see themselves as a good representation of their end-users. Furthermore, in the conducted interview it came up that if a domain expert that works with VISdom would want a specific or adapted workflow in the software, this could be specially made for that domain expert. The last of Bath’s (2014) mechanisms, the possible gender inscriptions through the means of the abstraction and formalization

processes that occur during computer science projects, is also unlikely to be of relevance here considering none of the calculations that are made for the software involve humans.

Taking a quick glance at the other projects within the Smart Worlds research area such as making 3D representations of the surface of Mars, or making simulations of large geographical areas in order to engage in infrastructure and spatial planning activities, it seems that the assessments about possible gender aspects in the VISdom project could be drawn for the broader research area. None of these other projects seem to simulate or make calculations about humans. If, in the future, the Smart Worlds research area engages in projects where humans or human behavior is simulated, sex and gender become important aspects to take into consideration. In that case, the suggested methodology described in the 2013 Gendered Innovations report for engineering inclusive innovation might be a good starting point.

### Immersive Analytics

In the Immersive Analytics research area, I conducted two interviews with a VRVis employee that is active within the Virtual Reality for Mission Planning project which is a project in collaboration with the Austrian Federal Ministry of Defense (BMLV) and the Austrian Institute for Military Geology. In the Virtual Reality for Mission Planning project urban scenes are reconstructed from satellite images on a scale of 1:1, that thus allows a user to walk through a particular landscape as if they were actually on-site (VRVis, n.d.). This allows them to engage in a range of mission-oriented applications such as training their local knowledge, assessing visibility from different points in the area, carrying out certain maneuvers, and planning rescue missions. The technology is envisioned to be used by two users: by one person who is inside the virtual environment and one who is outside but who can follow activity through computer monitors. All data used to render the digital environment is coming from the Austrian military. At the moment, the VR software is still in an early prototype phase and a long way off from being utilized in the actual field. At the moment it is mostly used internally by the Austrian military for purposes such as demos.

To the question ‘will the product under study affect or involve humans?’ the answer in the case of the Virtual Reality for Mission Planning project is a clear ‘yes’. Therefore, development of the software needs to ensure that it is not only gender inclusive but in general aims to serve a diverse end-user group. Especially technology within the military, which usually consists out of a vast majority of men (in the case of Austria this is even more the case due to the obligatory military or civil service), has a history to develop technology that is skewed more towards men. A telling example is Rachel Weber’s (1997) study of military cockpits, and how they were mostly designed based on male dimensions. This deemed women, who usually have smaller dimensions than men, to use a variety of technologies within aircrafts, and consequently blocked them from making certain developments within their careers as commercial or military pilots. One of the features within the Virtual Reality for Mission Planning project is to drop an avatar within the virtual environment from which then the visibility of said avatar is calculated. Needed for this calculation is an estimation of the height of the eyes. As explained to me by my interview partner, this height was an estimation and is not variable. This means that the calculation is not going to be accurate for people who do not fit this estimation. Perhaps a way to surpass these possible inaccuracies is to make this the height variable, then it is not just more gender inclusive, but also more inclusive for people who fall outside of the length norms.

Another aspect that came forward from the literature research, and has broadly been elaborated on in the Sensitizing Concepts chapter, is that Virtual Reality is experienced quite differently by men, women, and gender diverse people, and that “these potential gender differences need to be taken into account in the development and testing of [VR] prototypes” (European Commission, 2020, p. 29). For example, the level of perceived realism of the virtual environment differs amongst men and women (Felnhofer et al., 2012). Furthermore, women are more than twice as likely to experience negative virtual reality-induced side effects (VRISE) such as nausea, drowsiness, sweating, and disorientation (Al Zayer et al., 2019). Due to all these gender differences, the Gendered Innovations report underlines that “it is important to develop virtual settings in a sex-sensitive and gender-sensitive way, applying sex and gender analyses that include both women and men in the development and testing of prototypes” (European Commission,

2020, p. 139). However, when asked in the interview if VRVis was considering these gender aspects whilst developing the Virtual Reality for Mission Planning project the answer was a straightforward no. The main reason that was given for this was the fact that VRVis was developing a prototype, and that dealing with these gender sensitive issues was the responsibility of the customer if they were to further develop the VR application. Furthermore, whilst developing prototypes my interview partner mentioned “it has to work first and if it works, then you can think of design choices.” (Q8, female). Whilst this thesis is not trying to criticize this kind of workflow, or trying to assume it should be otherwise, coming from a purely gender perspective, ignoring these kinds of gender differences could potentially lead to the differences being ignored further down the road, which could eventually turn into an issue. A potential method that could surpass the gender differences being ignored further down the line of development, is clearly communicating with the customer, in this case the Austrian Military, that the differences exist and that they should be taken into account whilst the developing and testing of the application.

Other projects within the Immersive Analytics research area include the 3D-printing of art to make it more accessible to the visually impaired, the virtual augmentation of engines, and the simulation of buildings in order to help with restoration activities. The potential different VR experiences amongst the genders might be something to either take into account, or discuss openly about with third party customers. However, since this thesis did not study the other projects, the relevance of these differences is most likely something that needs to be assessed per project.

## Complex Systems

Within the Complex Systems area I conducted three interviews, of which two with the same person. Although the research area has some industrial partners, both interview partners have worked mostly within the medical projects the area engages in. Through the means of machine learning and deep learning, these projects mostly concern themselves with the processing of medical images of different kinds. Companies that VRVis works with are dealing with the pre-processing of medical images from the raw

acquisition to the displaying for medical doctors, who have no idea about the intrinsic of the machines. An example of a project is a tuberculosis (TBC) classifier. Here a computer is shown a chest radiograph of a patient, to which the computer then analyzes if the patient exhibits TBC markers. The outcome of the analysis is a number between zero and one, one being there is TBC on the image and zero there is none. Additionally, the computer offers some visual explanation, so why the computer says what it does.

As elaborated on broadly in my State of the Art chapter and my Sensitizing Concepts chapter through the means of multiple examples, machine learning algorithms can sometimes play a part in reinforcing negative social structures. This happens either through problems with datasets or problems with the algorithm itself. Therefore, the Gendered Innovations reports (2013 & 2020) and Tannenbaum et al. (2019) offer multiple methodologies to go against these possible issues, not just with regards to gender, but also with other factors such as ethnicity or race. Furthermore, it is important to analyze how gender plays a part in projects within this research area because the answer to Tannenbaum et al.'s (2019) question 'will the product under study affect or involve humans?', especially within the medical projects, is yes. Luckily, the researchers at VRVis certainly see gender as a relevant topic for their niche of visual computing, and are aware of problematic examples where algorithms have reinforced negative social structures. Therefore, they conceptualize that part of their responsibility is to not engage in potentially problematic projects.

An important step that is suggested by Tannenbaum et al. (2019) and the 2020 Gendered Innovations report in order to develop fair machine learning algorithms is to utilize meta-data, which describes different characteristics about an image in the dataset, in order to create well-balanced datasets in terms of gender, ethnicity, age, etc. However, a big problem within the broader field of machine learning is that often medical data is anonymized and thus this metadata is not available. Therefore, further suggested methodologies from Tannenbaum et al. (2019) that are also relying to an extent on meta-data such as counterfactual fairness and multi-accuracy auditing are already not possible. Furthermore, checking the algorithm for any intersectionality issues is also not possible without meta-data. The VRVis researchers are fully aware of this shortcoming, and are unfortunately not in a position themselves that they can overcome this considering they

are not involved in the actual acquisition of the datasets. As a result, they have to conceptualize what exactly their responsibility is with regards to delivering their prototype to the customer.

From the interviews it became clear that the responsibility of the researchers is conceptualized in the sense that they analyze the data, whether these come from publicly available datasets or the datasets that are offered to them by their customers, in the fashion they can and then clearly communicate about any potential shortcomings. They clearly communicate to the customer if the type of analysis desired from the customer is possible with the data provided, what is potentially missing from the data, and what they would need in order to improve the prototype. Furthermore, they possibly can guide in acquiring more data. This clear communication process can prevent possible sex and gender issues from getting black-boxed and clearly puts the responsibility of dealing with the issues onto the customer who, in the case of the medical projects of this research area, are responsible for the actual implementation of the algorithms into the hospitals. As was so well put by my interview partner: “if they develop things further and we have made them aware of certain issues, then it is also their responsibility to be aware of what we have made them aware of” (Q7, female).

Lastly, I wanted to analyze how relevant the four mechanisms described by Corinna Bath (2014) are in the practices of the Complex Systems research area. Considering the medical projects the area engages in do not simulate any human characteristics or stereotypical features two of the four mechanisms are already not relevant. Furthermore, the I-methodology also does not carry any relevance since the end-user in these projects are clearly defined through the means of communication with the customers. What could be possible is that the fourth mechanism, the one focusing on the inscription of possible sex and gender issues in the abstraction and formalization processes involved in writing code, carries relevance in the case of the Complex Systems area. An example within the field of visual computing where abstraction decisions made by the computer scientist had consequences with regards to gender is the work of Kaiser et al. (2008) on brain imaging. Their research is about a process in which an algorithm transforms raw data from a computer into colored images of the brain. Kaiser et. al (2008) show that gender differences present in the brain appear or did not appear based on

where the computer scientists laid the thresholds in their algorithm, thus showing that formalization processes are not necessarily innocent. Therefore, it is not unrealistic to assume that similar abstractions can occur within the Complex Systems research area. However, this requires a detailed analysis that involved both computer science and social science competences in close collaboration and is therefore a potentially interesting topic for further research.

## 7. Conclusion

The aim of this thesis was to uncover whether and how gender played a role in the development of technology in each of the four different research areas at the visual computing company VRVis Vienna. VRVis is a center funded by research funding (FFG) within the long-term program COMET. Their specific wish for this thesis was to analyze possible gender issues in all four project areas, with a special attention on the COMET projects. That is because within this program there is a focus on the increase in the proportion of women and the handling of gender aspects in research. During the research, through the means of qualitative interviews, focus was put on how the researchers at VRVis (1) conceptualize gender, (2) take potential steps to circumvent possible sex and gender issues in their work, (3) give importance to gender considerations in the projects they work on, and (4) identify where in their project processes gender aspects could play a role. This final chapter of this thesis starts with a summary of the main findings of the above described research, and relate them to research and literature that exists within the broader field of STS and gender studies. Then, I mention some possible interesting further research that VRVis could engage in if they plan to expand the research conducted about the topics of sex and gender. Lastly, I want to spend some time reflecting on my experience as a researcher throughout this thesis process. When I started this master thesis project in November 2019, the world was a completely different place when compared to world now, which is one that is unfortunately against all hopes, still amidst the global COVID-19 outbreak. How this influenced my thesis and research process is something I reflect on in the final subsection of this chapter.

### Main findings

This thesis aimed to uncover possible gender aspects in the development of technology within each of the four different research areas at VRVis Vienna (Visual Analytics, Smart Worlds, Immersive Analytics, and Complex Systems). The Discussion chapter attached specific findings to each of these areas and therefore in this conclusion I focus more on summarizing the findings in a broader sense and going over the most relevant ones.

## Conceptualizations of gender

The research conducted in this thesis can be placed under the STS branch of feminist technoscientific research. The term, as defined by Donna Haraway (1997), wants to uncover “[...] racial formation, gender-in-the-making, the forging of class, and the discursive production of sexuality through the constitutive practice of technoscience production themselves” (p. 35). By studying the practices of science-in-the-making, feminist technoscientific research shows how the possible inscription of social values into the development of science and technology can have harmful social consequences, and thereby presses for more inclusive innovation processes. Trojer (2014) describes that with regards to technology, which still is for a large part a male dominated industry, feminist technoscientific research has shifted from ‘counting heads’ – how many women are involved – to researching processes of knowledge production. This led to an epistemological shift where the challenge now is that “[...] gender research within technology and engineering does NOT primarily focus on gender – men and women. It focuses on technology” (ibid, p. 166).

However, one of the subquestions of this thesis was how the researchers at VRVis conceptualized gender, and the conducted interviews uncovered that this previous metric of ‘counting heads’ is still very much present amongst the researchers. VRVis as a company has around 70 employees and in 2018 broke the 20% barrier of the proportion of women in its scientific staff (VRVis, 2018). By 2019, female researchers made up 28% of the scientific staff (VRVis, 2019). The increase in female staff was something that was actively strived for by VRVis through inclusive hiring processes partially due to external pressures and the funding advantages they would receive. Therefore, it makes sense that the VRVis researchers conceptualize gender in this way of how many women and how many men are working on the teams. Furthermore, visual computing is still a male-dominated field and VRVis is not exception to this. Adding more women onto the scientific staff is the most straightforward and, more importantly, the most noticeable step towards creating a more gender-inclusive working environment. Therefore, even though the scientific staff only consists out of women for 28%, one of my interview partners

described that the “[...] ratio has changed a lot in [his] general feeling” (Q5, male). The inclusive hiring being the most noticeable effect with regards to gender inclusivity can be an additional explanation why the conceptualization of gender as the number of women and men on research teams is so prevalent in the reflections of the employees. I do not want to state that this conceptualization is the only one present within VRVis, nor do I want to infer that it is the one most commonly held, I just want to underline that according to research (Bührer & Wroblewski, 2019) the institutional environment and attitude towards gender-inclusivity is taken up by the researchers working within that environment.

During my interviews I have additionally encountered conceptualizations that are more in line with the broader definition of gender as “a socio-cultural process, [that] refers to cultural and social attitudes that together shape and sanction “feminine” and “masculine” behaviors, products, technologies, environments, and knowledge.” (European Commission, 2013, p. 9). Namely, I found it remarkable that all three of the women I interviewed for this thesis were talking about the color schemes of their user-interfaces as important gender considerations, whereas none of the men mentioned them. Of course, I do not want to overgeneralize this difference and infer that men do not think about these considerations at all, but I do think it points to the fact that “[g]ender affects the kinds of questions we ask” (Criado Perez, 2019, p. 179). Research shows that women in leading positions are more likely to use possible sex and gender dimensions in their published research (Nielsen et al, 2017). Furthermore, Bührer and Wroblewski (2019) show that off the three dimensions of the gender equality key area of RRI research, namely (1) increasing female participation in all the levels of science and research, (2) abolishing structural career barriers for female researchers, and (3) strengthening the gender dimension in research content and teaching, females are more prone to integrate all these dimensions whereas men mostly just focus on the first. Therefore, although it seems that at the moment VRVis focuses on the first and second dimension of gender equal RRI research (as can be seen in their inclusive hiring processes and the company already having women in leading positions), and the third dimension seems to be less present in the reflections of the employees, continuing down that path is one way that will unavoidably lead to more of a focus on the third dimension. This is because research

shows that “[...] women support female colleagues, encourage gender-balanced teams and consider gender aspects in their research design more frequently than men.” (Brührer & Wroblewski, 2019, p. 8). Another way to create a higher focus on the third dimension could be through internal seminars about gender aspects within research and technology, or further cooperations with social scientists such as this one.

### Virtual Reality and Machine Learning

The last chapter broadly elaborated on how the specific findings of this thesis related to the four different research areas. By using analytical lenses such as the one suggested by Tannenbaum et al. (2019) it was uncovered that some research areas likely had more possible gender aspects in their work than others. Mostly because Tannenbaum et al. (2019) suggest to only perform sex and gender analysis on technology that directly involves or influences humans. In the cases of the Visual Analytics and Smart Worlds research areas, I interviewed VRVis employees that were developing technology where this was not the case. However, some of the projects in the Complex Systems and Immersive Analytics department do directly influence humans, and thus further analysis was certainly relevant here.

Throughout this thesis multiple examples were given from existing literature regarding problematic machine learning algorithms that played a part in reinforcing negative social structures. Think about the alarming examples of algorithmic bias such as the systematic misidentification of women of color in facial recognition software or Google algorithms being five times more likely to offer men advertisements of high-paying jobs than women (Datta, Tschantz & Datta, 2015). Furthermore, developing AI for the medical world to aid with diagnoses comes with its own set of challenges, one of the larger ones being what is dubbed as the ‘gender data gap’ (Criado Perez, 2019). This gap refers to the systematic discrepancy that exists in the collection of any kind of data between men and women, and “[t]he introduction of AI to diagnostics seems to be accompanied by little to no acknowledgement of the well-documented and chronic gaps in medical data when it comes to women” (ibid, p.167).

During the interviews I conducted with the VRVis researchers from the Complex Systems area, we discussed the strategies they performed in order to avoid any harmful

bias in their algorithms. Although it became clear that discussions about bias is a constant process within the research group, one large hurdle prevents the group from engaging in some strategies suggested by Tannenbaum et al. (2019) and the Gendered Innovations reports (2013 & 2020), namely the lack of meta-data in medical data sets. Research discussed in the Sensitizing Concepts chapter constantly underlines the importance of having meta-data in order to develop fair machine learning algorithms. However, because medical data is often anonymized it is difficult to engage in suggested practices such as designing gender-equal datasets. Therefore, it would be desirable if thorough analysis occurs in each medical Complex Systems project that researches whether gender is a relevant variable that determines the outcome of the algorithms. Luckily, it seems that clear communication and collaboration exists with the third-party customers of the research area about the limitations of the data sets used for the algorithms. Also beneficial for this research area is that the teams working on the projects seem to be gender diverse.

The main reason why the Immersive Analytics research area has potential gender issues inscribed in its projects is because Virtual Reality technology itself is already a male-skewed technology.

“From oversized headsets, to research showing that VR causes motion sickness in women to a far larger degree than it does in men, to the fact that narrow computer displays favor men in tasks that require spatial awareness, you’re left with another platform that just doesn’t work well for women- and that is therefore likely to have fewer women on it” (Criado Perez, 2019, p. 182)

The fact that a lot of aspects about the technology itself, without even a specific application yet, skews more towards males already starts off women at a disadvantage. When asked if this was something that was taken into account during the development of the Virtual Reality for Mission Planning project the answer was simply no, mainly because the project was still in the early development. However, considering the researchers at VRVis are the experts in the field of VR technology as opposed to the specific customer for that project, I do suggest bringing conversations about these possible issues to the table. Especially since it is important to avoid that down the line women within the Austrian military are, by design, less likely to fully utilize the designed mission planning software. Furthermore, even though the Immersive Analytics team at VRVis is gender diverse, it is

also important to involve both men and women in the development and testing of prototypes on the user side (European Commission, 2020).

### Prototyping and Responsibility

Because VRVis is a company that stands in between research and industry, a part of its projects focus on developing prototypes that can potentially be further developed into products by third party customers. The Complex Systems and Immersive Analytics research areas both engage in projects focused on developing such prototypes. In the developing of prototypes it is impossible to develop everything perfectly immediately, and thus prioritizations in development have to be made. Or, as put by one of my interview partners: “It has to work first and if it works, then you can think of design choices” (Q8, female).

However, I think it is important to be reflective about these prioritization processes that occur whilst developing prototypes. Because what does it mean when a technology, product, or prototype ‘works’? By focusing on having certain features function before considering, for example, gender inclusive design choices, awareness needs to be raised that the prototype as a consequence functions well in one way but does not work in other ways, in this case a more gender inclusive way.

Furthermore, the line between what is a prototype and what is a final product is not always completely clear. Because it could be that, in the end, little is changed about the prototype, or that the prototype is in a way directly used in the field. As far as I understand from my interview with the VRVis researcher working on the Virtual Reality for Mission Planning project, the prototype is internally used by the military at the moment in order to evaluate it or to potentially run demonstrations. In this sense the prototype is already used in the field as a type of product. This then automatically raises the question of responsibility. Who carries the responsibility for the potential sex and gender issues inscribed in these prototypes?

Important for creating a clear conceptualization of responsibility is clear communication and reflection. Within the Complex Systems area, where there is close contact and collaboration with the customer, constant open communication exists about

the expert knowledge the VRVis researchers have to offer, also because the topic of bias inscription in machine learning algorithms is a big topic within the wider field. Therefore, because all the expert knowledge of the VRVis researchers is shared, the final responsibility of incorporating this knowledge into the further development of the prototype falls onto the customer. In the case of the Virtual Reality for Mission Planning project, some expert knowledge, such as how VR has different effects on the different genders, is not discussed with the customer, and thus part of the responsibility still falls on VRVis. Having open discussions with customers about any kind of diversity issues can perhaps spark some reflection or discussions about what functionalities are prioritized in the development of prototypes. Furthermore, it is desirable for VRVis that possible gender inscriptions are not black-boxed down the line when prototypes go into further development and eventually become products.

Lastly, another aspect that causes this 'pipeline'-structure of responsibility with an overarching question of who carries the final responsibility to deal with any possible gender aspects is inscribed within the setup of funding programs such as COMET. These types of programs fund parts of public universities with public money to build prototypes for the further uptake of private companies. This setup causes that the focus on RRI aspects such as gender inclusivity is enforced from the funder to the funded research (in this case from COMET to VRVis). However, there is no guarantee that this rationale is taken up further downstream these research processes, as the third-party companies that further take up these prototypes do not have to fall under the initial RRI requirements. In other words, the focus on responsible research seems to end in the prototype phase. Important to note is that this is not the fault or choice of VRVis, but rather a consequence of how the funding arrangements are set up. It does, however, underline the existence of an open and important question on when and by whom gender should be reflected upon in the entire trajectory of the innovation process.

### Further research

As mentioned multiple times throughout this thesis, it was a specific wish of VRVis to analyze possible gender aspects in all four research areas. However, because the research had to fit within the scope of a master thesis project, it automatically meant that deeper

analysis into each of these areas was not possible. Furthermore, through the redesign of the originally planned ethnographic study into one solely relying on qualitative interviews, certain processes could not be observed. For example, how do the decision-making processes during projects transpire on the office floor and how are the deliberations between VRVis and the customers and amongst the research teams themselves? Therefore, this last section shortly reflects on some possible further research VRVis can engage in when they plan to expand their research on the topics of sex and gender.

First, quite obviously, the current research can be continued in a more detailed manner through an ethnographic study. This will be more in line with the classical way in which Latour and Woolgar (1979) and Haraway (1989) have performed their researches about science in the making. More of a focus can be put on the practices surrounding the development of VRVis technology and decision-making processes surrounding possible gender considerations can be observed as they come about on the work floor. How are the gender considerations discussed amongst the research teams, and how is the topic of gender more broadly handled by VRVis?

Also, because this study remains a social sciences study about a visual computing company, there is only so much that can be observed without the proper technical knowledge. If VRVis wishes to study in depth if their technology is gender-inclusive, further studies need to be conducted by visual computing experts with the corresponding knowledge. For example, experts in the ethics of machine learning can conduct a detailed study on the algorithms for medical applications that are developed in the Complex Systems research area, or user studies can be conducted about the perception of the technologies and its user interfaces.

Lastly, even though the focus was specifically and by request on gender, I do not think gender is more important than other diversity categories such as age, ethnicity, or cultural background, especially in the Complex Systems medical imaging projects. This was also something that was underlined by multiple interview partners throughout the thesis process. Opening up further research towards these other diversity categories will further encourage the development of socially inclusive technology within VRVis.

## Experience as a researcher

The first steps in this thesis project were made in November 2019 with an initial meeting with representatives from VRVis, the head of the gender department at the Technical University Vienna dr. Brigitte Ratzler, and my supervisor dr. Maximilian Fochler. During the meeting we deliberated how the specific wishes of VRVis could fit within the scope of a master thesis project. This was then quickly followed by a day at VRVis where the same group of people got acquainted with the different research areas and the projects therein through a variety of presentations. In close proximity we hovered over computer screens together and shook hands with many of the people we met that day. Who knew that such simple things were going to be so unthinkable a couple of months after? At the end of 2019, I was assigned a working space at VRVis, given a keycard to the office space, and set days were agreed upon during which I was going to be present on the office floor. In sum: everything was set up so my ethnographic work could start! However, 2020 turned the corner and the COVID-19 pandemic forced everybody into their homes.

The pandemic made any kind of ethnographic work impossible, and since we could not postpone the project because VRVis needed preliminary results by June 2020 and I did not want to delay my thesis process too much, a redesign of the research was needed. The only method possible that would get close to the initial research design was qualitative interviews. However, changing the method also meant a change in what the research can uncover, as the method determines how you look at your empirical data. Steering away from an ethnographic study meant I would not be able to closely observe the practices surrounding the VRVis technologies or the dynamics that exist within the research teams. Although the initial wish of VRVis was to specifically research possible sex and gender issues inscribed in the technologies, through the scope of this thesis with the method of qualitative interviews, I could only observe how the researchers at VRVis themselves viewed the technology they worked on and how they conceptualized gender could play a part in their work.

Another big challenge that came with the changes the COVID-19 pandemic brought to this project was that all qualitative interviews needed to be conducted online. At times, I would encounter situations where bad internet connection would cause answers to be

broken up leading to me having to ask my interview partner to repeat their answer, and sometimes it was even so bad that the connection did not allow video chat, meaning that I only heard my interview partners and could not see them. As one can imagine, these issues heavily hindered the rapport I established with my interview partners and influenced the overall atmosphere of the interview. However, my situatedness as a researcher also heavily influenced the rapport. I was a social sciences master student who came into a visual computing company, which is a field that is mostly male-dominated, to talk about gender. I have experienced that gender is quite a challenging topic to discuss due to the fact that people have different ideas about what is important when discussing the topic. For example, it was difficult to convey to my interview partners that I wanted to mostly study the technologies themselves and not the gender balance of the people working at VRVis. Lastly, throughout the process I often felt doubt about whether I, as a social science student, could infer relevant conclusions about technical computer science work. However, I luckily realized that these feelings were partially there because this thesis project was the largest academic venture I engaged in up until that point, and that they disappeared more into the background as I grew confidence in myself as a researcher.

All in all, although the process definitely came with a set of setbacks, challenges, and difficulties, I found the overall thesis writing process rewarding and challenging in the best way possible. It has taught me about commitment and perseverance, how to think on my feet, and how to make adaptations to my research based on changing settings and contexts. I have learned a lot about the relevance and importance of gender discussions, as this was a topic that I was not very involved with prior to this thesis, and would even consider it something to continue with in the future. But most importantly, this thesis has showed me that STS is a discipline that has relevance in the industry and not just in the ivory tower of science. I hope that STS will expand its influence in the development of science and technology in the future, as this thesis has showed me that the involvement of social scientists in innovation processes can lead to the development of more socially responsible technologies.

# References

- Akrich, M. (1995). User representations: Practices, methods and sociology. In A. Rip, T. J. Misa, & J. Schot (Eds.), *Managing technology in society: The approach of constructive technology assessment* (pp. 167-184). Printer Publishers.
- Aldous, J., & Hill, R. (1967). *International Bibliography of Research in Marriage and the Family: 1900-1964*. University of Minnesota Press.
- Al Zayer, M., Adhanom, I. B., MacNeilage, P., & Folmer, E. (2019, May). The effect of field-of-view restriction on sex bias in VR sickness and spatial navigation performance. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-12).
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533, 452-454.
- Bath, C. (2014). Searching for Methodology. Feminist Technology Design in Computer Science. In W. Ernst & I. Horwath (Eds.), *Gender in Science and Technology* (pp. 57-78). Duke University Press.
- Baxandall, R., & Gordon, L. (2008). Second-wave Feminism. In N. A. Hewitt (Eds.), *A companion to American women's history* (pp. 414-432). John & Wiley Sons.
- Bose, D., Segui-Gomez, M., & Crandall, J. R. (2011). Vulnerability of female drivers involved in motor vehicle crashes: an analysis of US population at risk. *American journal of public health*, 101(12), 2368-2373.
- Bührer, S., & Wroblewski, A. (2019). The practice and perceptions of RRI—A gender perspective. *Evaluation and Program Planning*, 77, 1-10.

Burget, M., Bardone, E., & Pedaste, M. (2017). Definitions and conceptual dimensions of responsible research and innovation: A literature review. *Science and engineering ethics, 23*(1), 1-19.

Charmaz, K. (2006). *Constructing Grounded Theory*. Sage.

Chilvers, J. (2010). *Sustainable Participation? Mapping out and reflecting on the field of public dialogue on science and technology* (Working Paper). Sciencewise Expert Resource Centre. <https://ueaeprints.uea.ac.uk/id/eprint/37545/>

Criado Perez, C. (2019). *Invisible Women. Exposing data bias in a world designed for men*. Vintage Publishing.

Datta, A., Tschantz, M. C., & Datta, A. (2015). Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. *Proceedings on privacy enhancing technologies, 2015*(1), 92-112.

Dobbin, F., & Kalev, A. (2016). Why Diversity Programs Fail. *Harvard Business Review*. <https://hbr.org/2016/07/why-diversity-programs-fail>

European Commission (2013). *Gendered Innovations: How Gender Analysis Contributes to Research*. Publications Office of the European Union.

European Commission (2019). *She Figures 2018*. Publications Office of the European Union.

European Commission (2020). *Gendered Innovations 2: How Inclusive Analysis Contributes to Research and Innovation*. Publications Office of the European Union.

Faulkner, W. (2014). Can Women Engineers be 'Real Engineers' and 'Real Women'? In W.

Ernst & I. Horwath (Eds.), *Gender in Science and Technology* (pp. 187-203). Duke University Press.

Feast, J. (2019). 4 Ways to address gender bias in AI. *Harvard Business Review*.

<https://hbr.org/2019/11/4-ways-to-address-gender-bias-in-ai>

Felnhofer, A., Kothgassner, O. D., Beutl, L., Hlavacs, H., & Kryspin-Exner, I. (2012). Is virtual reality made for men only? Exploring gender differences in the sense of presence. *Proceedings of the International Society on presence research*, 103-112.

Felt, U. (2014). Within, across and beyond: Reconsidering the role of social sciences and humanities in Europe. *Science as Culture*, 23(3), 384–396

Foreman, N. (2010). Virtual reality in psychology. *Themes in Science and Technology Education*, 2(1-2), 225-252.

Garvie, C., & Frankle, J. (2016). Facial-Recognition Software Might Have a Racial Bias Problem. *The Atlantic*.

<https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/>

Haig, D. (2004). The inexorable rise of gender and the decline of sex: Social change in academic titles, 1945–2001. *Archives of sexual behavior*, 33(2), 87-96.

Haraway, D. J. (1989). *Primate visions: Gender, race, and nature in the world of modern science*. Routledge.

Haraway, D. J. (1997).

*Modest-Witness@Second-Millennium.FemaleMan-Meets-OncoMouse: feminism and technoscience*. Routledge.

- Hern, A. (2020, September 1st). Twitter apologizes for 'racist' image-cropping algorithm. *The Guardian*. <https://www.theguardian.com/technology/2020/sep/21/twitter-apologises-for-racist-image-cropping-algorithm>
- Holstein, J., & Gubrium, J. (1995). *The Active Interview*. Sage.
- Jensen, E. A., & Laurie, C. (2016). *Doing Real Research: A practical guide to social research*. Sage.
- Kaiser, A., Haller, S., Schmitz, S., & Nitsch, C. (2009). On sex/gender related similarities and differences in fMRI language research. *Brain research reviews*, 61(2), 49-59.
- Kim, M. P., Ghorbani, A., & Zou, J. (2019). Multi-accuracy: Black-box post-processing for fairness in classification. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 247-254).
- Kirk, A. (2016). *Data visualisation: a handbook for data driven design*. Sage.
- Knorr-Cetina, K. (1981). *The manufacture of knowledge : an essay on the constructivist and contextual nature of science*. Pergamon Press.
- Kusner, M., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual fairness. *Advances in neural information processing systems*, 30, 4066-4076.
- Latour, B. (1987). *Science in action: how to follow scientists and engineers through society*. Harvard University Press.
- Latour, B., & Woolgar, S. (1979). *Laboratory life: The construction of scientific facts*. Princeton University Press.
- Li, J. J., Ju, W., & Reeves, B. (2017). Touching a mechanical body: tactile contact with body

parts of a humanoid robot is physiologically arousing. *Journal of Human-Robot Interaction*, 6(3), 118-130.

Linder, A., & Svedberg, W. (2019). Review of average sized male and female occupant models in European regulatory safety assessment tests and European laws: gaps and bridging suggestions. *Accident Analysis & Prevention*, 127, 156-162.

Mellström, U. (2009). The intersection of gender, race and cultural boundaries, or why is computer science in Malaysia dominated by women? *Social studies of science*, 39(6), 885-907.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Education.

Money, J. (1955). Hermaphroditism, gender and precocity in hyperadreno-corticism: Psychologic findings, Bull. *Johns Hopkins Hospital*, 96, 253-264.

Money, J. (1973). Gender role, gender identity, core gender identity: Usage and definition of terms. *Journal of the American Academy of Psychoanalysis*, 1(4), 397-402.

Nielsen, F. (2005). *Visual computing: Geometry, graphics, and vision*. Charles River Media.

Nielsen, M. W., Andersen, J. P., Schiebinger, L., & Schneider, J. W. (2017). One and a half million medical papers reveal a link between author gender and attention to gender and sex analysis. *Nature human behaviour*, 1(11), 791-796.

Niler, A. A., Asencio, R., & DeChurch, L. A. (2020). Solidarity in STEM: How Gender Composition Affects Women's Experience in Work Teams. *Sex Roles*, 82(3-4), 142-154.

van Oost, E. (2003). Materialized Gender: How Shavers Configure the Users' Femininity

and Masculinity. In N. Oudshoorn & T. Pinch (Eds.), *How Users Matter: The Co-Construction of Users and Technology* (pp. 193-209). The MIT Press.

Oudshoorn, N. (2003). Clinical Trials as a cultural niche in which to consider the gender identities of users: The case of male contraceptive development. In N. Oudshoorn & T. Pinch (Eds.), *How users matter. The co-construction of users and technology* (pp. 209-227). The MIT Press.

Oudshoorn, N., Rommes, E., & Stienstra, M. (2004). Configuring the user as everybody: Gender and design cultures in information and communication technologies. *Science, Technology, & Human Values*, 29(1), 30-63.

Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and public policy*, 39(6), 751-760.

Richardson, S. S., Daniels, C. R., Gillman, M. W., Golden, J., Kukla, R., Kuzawa, C., & Rich-Edwards, J. (2014). Society: don't blame the mothers. *Nature News*, 512(7513), 131-132.

Risman, B. J. (2018). Why Gender Matters: Research explains why gender is so much more complicated than just identity. *Psychology Today*.  
<https://www.psychologytoday.com/us/blog/gender-questions/201808/why-gender-matters>

Rivas, C. (2018). Finding themes in qualitative data. In C. Seale (Eds.), *Researching Society and Culture* (pp. 431 – 453). Sage.

Rose, N. (2014). NeuroView: The human brain project: social and ethical challenges. *Neuron*, 82, 1212–1215.

Scheuerman, M. K., Paul, J. M., & Brubaker, J. R. (2019). How computers see gender: An evaluation of gender classification in commercial facial analysis services. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1-33.

Schiebinger, L. (1989). More than Skin Deep: The Scientific Search for Sexual Difference. In *The Mind Has No Sex?* (pp. 189-213). Harvard University Press.

Silverman, D. (2006). *Interpreting Qualitative Data. Methods for Analysing Talk, Text, and Interaction*. Sage.

Stanford University (2011). *What is Gendered Innovations?*

<http://genderedinnovations.stanford.edu/what-is-gendered-innovations.html>

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research policy*, 42(9), 1568-1580.

Tannenbaum, C., Ellis, R. P., Eyssel, F., Zou, J., & Schiebinger, L. (2019). Sex and gender analysis improves science and engineering. *Nature*, 575(7781), 137-146.

Trojer, L. (2014). Gender Research as Knowledge Resource in Technology and Engineering. In W. Ernst & I. Horwath (Eds.), *Gender in Science and Technology* (pp. 165-185). Duke University Press.

VRVis (2018). *Annual Report 2018*.

[https://www.vrvis.at/fileadmin/05\\_Ueber\\_uns/vrvis\\_geschaeftsbericht\\_2018.pdf](https://www.vrvis.at/fileadmin/05_Ueber_uns/vrvis_geschaeftsbericht_2018.pdf)

VRVis (2019). *Annual Report 2019*. [https://www.vrvis.at/fileadmin/04\\_News-Events/00\\_Downloads/VRVis\\_Geschaeftsbericht\\_2019\\_Webversion.pdf](https://www.vrvis.at/fileadmin/04_News-Events/00_Downloads/VRVis_Geschaeftsbericht_2019_Webversion.pdf)

VRVis (2020). <https://www.vrvis.at/en/about>

VRVis (n.d.). *Visual Analytics*. <https://www.vrvis.at/en/research/research-topics/visual-analytics>

Weber, R. N. (1997). Manufacturing gender in commercial and military cockpit design. *Science, technology, & human values*, 22(2), 235-253.

Winner, L. (1986). Do Artifacts Have Politics? In *The Whale and the Reactor. A Search for Limits in an Age of High Technology* (pp. 19-39). Chicago: University of Chicago Press.

World Health Organization (2017). *Gender, Women, and Health*. <https://web.archive.org/web/20170130022356/http://apps.who.int/gender/whatisgender/en/>

Xie, D., Zhang, L., & Bai, L. (2017). *Deep learning in visual computing and signal processing*. *Applied Computational Intelligence and Soft Computing*, 2017, 1-13.

# Abstracts

## English version

Research has shown that incorporating sex and gender analysis into experimental design improves scientific and technological innovation and increases its social robustness. Furthermore, because scientific institutions such as funding agencies and the European Commission are increasingly bringing gender inclusivity to the foreground, the popularity of engaging in such analysis is additionally on the rise. This master thesis in the field of Science and Technology Studies aims to uncover through the means of qualitative interviews if and how the topic of gender plays a role within the four different research areas of the visual computing company VRVis in Vienna. Utilized for the analysis are existing methodologies for uncovering potential gender aspects in computer science research projects. Within the four areas (Visual Analytics, Smart Worlds, Immersive Analytics, and Complex Systems) the importance of considering gender varies substantially related to how directly the work in that area affects humans. Furthermore, compared to the three gender equality dimensions of the Responsible Research and Innovation (RRI) framework, within the reflections of the VRVis researchers a heavier focus existed on the first two dimensions, which include stimulating the increase of the number of women active within computer science research projects, and less on the third dimension which includes a focus on the effects of sociocultural conceptions of male and female on the development of technology. Lastly, because VRVis is a research institution that is often involved in the development of prototypes and is therefore not in a position to develop all its technology to the fullest, an open but important question exists on when and by whom gender should be reflected upon in the entire trajectory of the innovation process.

## German version

Die Forschung hat gezeigt, dass die Einbeziehung von Sex- und Gender-Analysen in die Versuchsplanung wissenschaftliche und technologische Innovationen verbessert und ihre soziale Robustheit erhöht. Da zudem wissenschaftliche Institutionen wie Förderorganisationen und die Europäische Kommission Gender-Inklusivität zunehmend in den Vordergrund rücken, steigt die Popularität der Beschäftigung mit einer solchen Analyse zusätzlich an. Diese Masterarbeit im Bereich Science and Technology Studies hat zum Ziel, mit Hilfe von qualitativen Interviews herauszufinden, ob und wie das Thema Gender in den vier verschiedenen Forschungsbereichen des Visual-Computing-Unternehmens VRVis in Wien eine Rolle spielt. Für die Analyse werden bestehende Methoden zur Aufdeckung möglicher Genderaspekte in Informatik-Forschungsprojekten herangezogen. Innerhalb der vier Bereiche (Visual Analytics, Smart Worlds, Immersive Analytics und Complex Systems) variiert die Bedeutung der Berücksichtigung von Gender wesentlich in Abhängigkeit davon, wie direkt die Arbeit in dem jeweiligen Bereich Menschen betrifft. Im Vergleich zu den drei Dimensionen der Geschlechtergleichstellung des Rahmenwerks für verantwortungsvolle Forschung und Innovation (Responsible Research and Innovation, RRI) lag in den Überlegungen der VRVis-Forscher außerdem ein stärkerer Fokus auf den ersten beiden Dimensionen, die die Stimulierung der Erhöhung der Anzahl der in Informatik-Forschungsprojekten aktiven Frauen beinhalten, und weniger auf der dritten Dimension, die einen Fokus auf die Auswirkungen soziokultureller Vorstellungen von männlich und weiblich auf die Entwicklung von Technologie beinhaltet. Da es sich bei VRVis um eine Forschungseinrichtung handelt, die häufig an der Entwicklung von Prototypen beteiligt ist und daher nicht in der Lage ist, die gesamte Technologie in vollem Umfang zu entwickeln, besteht schließlich eine offene, aber wichtige Frage: Wann und von wem das Geschlecht im gesamten Verlauf des Innovationsprozesses reflektiert werden sollte.