TInnGO

Transport
Innovation
Gender
Observatory

# Deliverable 4.6

**TInnGO digital analysis:**

**Introducing digital analysis to transport and mobility studies**
**Gender smart mobility and employment.**

TInnGO
Transport Innovation Gender Observatory

## Deliverable fact sheet

| | |
|---|---|
| **Document name** | TInnGO digital analysis report |
| **Responsible partner** | UCPH |
| **Deliverable number** | 4.6 |
| **Due date of deliverable** | June 2020 [M19] |
| **Actual submission date** | July 31, 2020 |
| **Version** | v.2 |
| **Authors** | Hilda Rømer Christensen (UCPH)<br>Michala Hvidt Breengaard (UCPH)<br>Stine Pedersen (UCPH) |
| **Reviewers** | Eileen O'Connell,Interactions, TInnGO partner , Samantha Dawn Breslin, Sodas, UCHP ( external) |
| **Work package no.** | WP4 |
| **Work package title** | Conceptual coordination: Translating theory into practice |
| **Work package leader** | UCPH |
| **Deliverable participants** | UCPH, CU, SOCTR, VTM, INTECO, VTI, |

## Dissemination level (please select one)

| | | |
|---|---|---|
| **CO** | Confidential, only for members of the consortium including the Commission Services | |
| **PU** | Public | X |
| **PP** | Restricted to other programme participants, including the Commission Services | |
| **RE** | Restricted to a group specified by the consortium including the Commission Services | |

## Nature of the deliverable (please select one)

| | | |
|---|---|---|
| **R** | Report | X |
| **D** | Demonstrator | |
| **W** | Websites, patents filing, etc | |
| **O** | Other | |

TInnGO
Transport
Innovation
Gender
Observatory

# Contents

## Abstract

The aim of this report is to present the TINNGO digital analysis as a possible research method for studies of transport, mobility and diversity. The aim has been to explore digital analysis as a new mode of evidence based knowledge production which enables comparative analysis of various gender diversity and transport discourses throughout Europe. More precisely the digital queries throw lights on affiliated networks related to terms such as smart mobility and passenger as well as well as gendered patterns of transport and employment. Moreover it is demonstrated that digital analysis is not an isolated tool, but is in need of interpretation and alignments with conceptual and contextualized analysis.

The digital analysis tools and findings have been developed in an educational process in workshops, individual partner data collections and exercises. Both among TInnGO partners and with a consultation team from University of Aalborg, Denmark. As for the educational and development process of the digital tools it contained a workshop on gender and diversity in digital analysis convened at University of Copenhagen November 19th, 2019 followed by consultations and a workshop for selected partners in Brussels. March 5th, 2020.

## The TInnGO project [1]

The TInnGO project meets the need to address smart mobility in innovative ways in order to improve the access of mobility for all. In so doing, the TInnGO project addresses transport related areas, such as production, consumption and governance as integrated in the notion of *Gender Smart Mobility.* TInnGO aims to:

- Sharpen the critical analysis of gender-blind approaches in the planning, production and policy-making of Smart Transport as well as in the usage and consumption of smart transport services.
- Develop proactive Gender Smart approaches to mobility governances.
- Feed into new and visionary examples and approaches of Gender Smart Mobility.

## Summary

This report reflects the ambition of the TInnGO project of introducing new methods in transport and mobility research. This report presents trends and challenges in the field and demonstrates cases of how digital materials can be applied in studies of mobility and transport. Using visual media archives as data material this report presents a new mode of topical and comparative analysis of various visual patterns of gender, diversity, and transport and employment.

**Aims of the TInnGO digital analysis:**

- To widen the methodological expertise of TInnGO partners and provide a presentation of the targeted digital analysis for a wider public.

[1] https://www.TInnGO.eu/

- To give Smart Transport special attention by exemplifying representations of gender in images of self-driving cars.

- To look at visual representations of transport as a gendered labour market.

- To provide methods of digital analysis as a supplement to various other methods applied in the TInnGO project.

**Findings and Lessons learned:**

- *Analytical:* The TInnGO partners have in their searches provided evidence of the digital tools and methods as useful for the assessments of various intersections of gender and technology, e.g. the overrepresentation of technical devises and masculine images of employment across national borders and sectors.

- *Technical:* In order to refine the comparative dimensions the tool 'google Trends' was applied to find relevant national and comparative search queries. Combined with the tools 'google image scarper', 'Clarifai recognition' tagging data service, and the visualizing 'Gephi' tool the applied method provide the technical innovation of the analysis.

**Digital analysis – potentials and reflections :**

The TInnGO Digital analysis presents a new mode of knowledge-based comparative analysis of patterns of gender, diversity and transport throughout Europe, using visual media archives as data material. In the following we list potentials and reflections in a range of questions and preliminary answers, which came up in the process:

- The TInnGO digital analysis focuses on the visual representations, which average European citizens in and across countries encounter when searching for web-based information on smart transport, i.e. the ways in which google represents the field of smart transport.

- The aim has been to provide knowledge of patterns of connections and visual clusters of representations, which might inform our perceptions, behaviours and practices.

- The visual digital method conducted in the TInnGO project allows for a special version of "mixed methods," i.e. a combination of quantitative (distant) and qualitative (close) readings of a given visual data. The method allows for data visualization by means of topic modelling in word clouds, network diagrams and visualizations of various background variables.

- Applying this tool of digital analysis to the research field of transport is innovative as it focus on visual representations and patterns in large data-sets which could not be analysed otherwise

- This method discloses current and future scenarios of how gendered discourses are embedded in existing transport employment and emerging smart mobilities.

## Workshops and consultations:

The digital analysis tools and examples in the report have been developed in a stepwise educational process within the TInnGO project consisting in workshop, consultations and individual data collections by TInnGO partners. The process has been supervised and assisted by a consultation team, led by Anders Munk and Matheu Jeremy, from University of Aalborg, Denmark. A TInnGO workshop on gender and diversity in digital analysis was convened at University of Copenhagen, November 19th, 2019, The aim of the workshop was to educate all TInnGO project partners in how to apply a gender and diversity focus in digital methods The overall objectives of the workshop were to widen the methodological scope of all TInnGO partners, using visual media archives as data material. The digital workshop also provided a kick-off event for the targeted digital analysis, which were conducted in spring 2020. A follow up consultation and workshop was held in Brussels, March 5th, 2020. The workshop included introduction of extended guidelines, sensitive to the geographical context and language among the TInnGO partners in Romania, Spain, Portugal, UK, Sweden and Denmark. Furthermore, the workshop offered consultations with the Aalborg team on the set up of networks and other types of visual disseminations.

## Structure of the report:

In order to contextualize the use of digital analysis with the overall concepts of the TInnGO project, this report contains of the following sections:

1. Gender and Diversity in digital analysis: Brief introduction to the fields of quantitative analysis and big data in gender studies.
2. TInnGO digital analysis: Core concepts into digital analysis – challenges and potentials.
3. TInnGO digital analysis: Main steps and lessons learned
4. TInnGO digital analysis: Ways forward.

# Gender and diversity in digital analysis: Brief introduction to the fields of quantitative analysis and big data in gender studies.

Introduction

*"Data has become a currency of power. The most successful Internet businesses make their money by aggregating data. Decisions of public import, ranging from which products to market, to which prisoners to parole, to which city buildings to inspect, are increasingly being made by automated systems sifting through large amounts of data."*

(Data visualization from a feminist perspective: Interview with Catherine D'Ignazio
by Nanna Thylstrup and Kristin Veel, in *Women, Gender and Research* 1. 2017)

At a comprehensive and general level this report addresses the emerging interest in quantitative methodologies and big data in women's and gender studies as well as the emerging interest in introducing gender and diversity in transport and mobility studies. In the report, we use new digital methodologies to connect gender and diversity to European transport and mobility. The report somehow reflects the growing hegemony of evidence-based policy and strategies in 21s century policymaking, which has turned statistics and quantitative methodologies into key data with wide-ranging effects at both the institutional and individual levels. As a result, knowing how to collect, find, analyse, and communicate data is of increasing importance not only in transport analysis, but in general in present-day society. This development has made ownership of data pivotal, along with access to IT equipment, resources, and expertise. Data is today mainly collected and stored by big corporations and governments, who have the resources to do so and who often control access. People today, as argued by Catherine D'Ignazio, are far more likely to be discriminated against with data or surveilled by data than they are to use data for their own civic ends.(Thylstrup & Veel: 2017). In the TInnGO project for example, it has been difficult to reach a comprehensive and comparative level of transport data, which included gender and diversity such as age and ethnicity

This report, and the TInnGO digital analysis itself, aims to explore critical aspects of power and inequality in prevailing big data provisions and the current "google world" which meets the average EU citizens  also when it comes to transport and mobility.  The aim has been to  to explore the

potentials of online data for alternative or even subversive use in transport and mobility studies and how to apply new technologies of data collection and visualization. How does the provision of new data feed into existing practices and politics of diversity and gender equality? Is it possible to collect and organize data collections in ways that support new forms of equality and democratic governance? And what are the potentials and pitfalls of emerging methodologies? How can bodies be made visible without creating new essentializing categories? It turns out that the TInnGo focus on the field of gender smart mobility at one at the same time made up an appropriate field for explorative analysis of potentials and limitations: on the one hand the following analysis can feed into both transport and mobility studies, as well as into the larger field of digital methodologies. And on the other hand the limitations in current visual data collections methods came to the fore, because of the poor reading abilities of personal attributes and diversity.

## Overcoming the paradigm war [2]

While addressing cutting edge issues of data collection and methodology, this report also seeks to nuance or even overcome the old 'paradigm war' in feminist scholarship between quantitative and qualitative methodologies. It is well known that gender research took off in the 1970s along with political currents such as the student revolt and the women's liberation movements. This also implied a break from established and dominant scientific paradigms, and not least from quantitative methodologies in the social sciences. In parallel with other oppositional trends, feminist research defined itself as opposed to 'positivist, quantitative research methodology', which became regarded as the bias of masculine knowledge and women's invisibility. Feminist criticism of quantitative research culminated over the first decades of women's and gender studies in the 1970s, being concerned with the power implications of research methodologies. It was argued that choosing topics in quantitative research implicitly supported sexist values on a broad scale. It was further argued that female subjects were excluded and marginalized and furthermore that the relationship between researcher and research subjects was intrinsically exploitative. Moreover, the resulting data were regarded as superficial and over-generalized, and it was argued that quantitative research was not

[2] This chapter is inspired by the volume Quantitative methods, big data and gender. Special issue of *Women, Gender and Research* 1. 2017 pp 1-90..
https://tidsskrift.dk/KKF/issue/view/6751

being used to overcome social problems (Oakley 1998: 709). In contrast, feminist research was branded as research with, for and about women. And appropriate methods included participant observation, semi-structured interviewing, life-histories and focus groups. Such methodologies came to be seen as epistemologically distinct from the quantitative methods of surveys, experiments, statistical records, structured observations and content analysis. All in all, feminist scholarship contributed to the paradigm debate by introducing new themes which confronted the gender-blind and sexist core of much research (Oakley 1998: 708). On the other hand, it seems as if mainstream transport research and data collection have been heavily influenced by the post-world war paradigm of mobility, growth and the idea of universal and "neutral" individual users, passengers etc. If and when gender is included it is often in a simplified matter, and in a binary male/female variable, while mobility is carefully measured along multiple dimensions / travel mode, distance, trip linking etc. (Breengaard, M., Christensen HR, Levin, L. 2019).

In the twenty-first century, several scholars have entered the field of quantitative methodologies and seem to be in favour of ending the war and bridging the debate (McCall 2005, Hughes and Cohen 2012). Two alternatives seem to have come to prominence. One is seen in the efforts to develop research practices where quantitative and qualitative methodologies are applied in a mixed-methods approach. Another alternative is to accept the academic division of labour and specialization and the fact that particular methods can have particular assets and limitations, issues that are related to research questions and research interests (McCall 2005: 1791). McCall in particular argued for the potentials of advancing both quantitative analysis in combination with intersectional analysis. The potential of using big data resources, the potential of combining various sociocultural categories at big scales, and the possibility of creating localized /or regional specific analysis, from where appropriate gender and diversity strategies could be established were demonstrated in relation to socio economic variations in the US. We have been inspired by McCall's main arguments, while taking some of her points into a slightly other direction in the mapping of visual representations and networks (of association) in comprehensive image scrapings.

In a further note McCall also addressed the various ideas of intersectionality, which she connected to a broader discussion between epistemology and ontology. Do intersections take place between already fixed categories, such as gender, ethnicity and class, or do they produce new categories? McCall suggested a systematic division of the complexity of intersectionality into anti-categorical,

intra-categorical and inter-categorical complexities. Here inter-categorical complexity is used by researchers who adopt existing analytical categories to document relationships of inequality among social groups along multiple and conflicting dimensions, while intra- and anti-categorical complexity is seen as being connected to qualitative micro- and meso-level studies. McCall's arguments and methodological differentiations assist in the clarification of how to work with multiple categories – also in the analysis of mobility and transport (Breengaard, M., Christensen HR, Levin, L. 2019).

## The digital turn – new research tools and methods

Existing literature in the field has considered the potentials in aligning political and social research with the digital age, and the question of what newly available online data can bring to research. Among the various critical effects of the so-called "digital turn" is the attention to focus on "pattern finding" rather than interpretations. Another issue is how the study of digital culture informs research. The importance of resisting a monolithic, or unitary, understanding of the changing nature of research in the digital age has been stressed (Lovink 2012). and all along the critical topics and categories involved have been refined. This includes a distinction between the computational and the digital turns[3] and a series of digital research practices have emerged and been labelled as cultural analytics, culturomics, webometrics, altmetrics and digital methods (Priem et al 2010; Rogers 2014). Recent elaborations operate with a distinction between the natively digital and the digitized which may also be applied to methods. We can talk about methods that have been "migrated" to the web, e.g. surveys that are now conducted online. They differ from those natively digital written for IT, e.g. Googles page rank, which privileges one website over another in a ranking or Facebook's edge rank, which privileges friends over others in terms of closeness (Rogers 2014).

All in all, approaches in digital research may be listed in terms of which material are the preferred data (digitized or born digital) and where the methods are situated (emulated or native). Looking at the TInnGO project as a whole, the project applies various methods which also mix up natively digital and digitized, e.g. in the use of "migrated" survey methods in new combinations. The online tools allow for the inclusion and interpretation of more variables in various combinations and for

[3] Currently there exists a variety of approaches with labels such as, e-social sciences and digital metric studies that could be seen as having distinctive ontological and epistemological commitments and positioning.

intersectional analysis. Here the survey method is developed in combination with smart app-based dissemination tools in the TInnGO project count as an example.

What is new in the following digital analysis method is that it uses digitized collections as its material and groups images in relation to key words and specific (digital) software, which assists in creating visual patterns and connections. This method is used as a departure for analysis of representations and patterns and for wider intersectional gender analysis related to smart mobility and employment.

## Absent bodies, powerful data, and visualizations

In Nanna Thylstrup and Kristin Veel's interview with Catherine D'Ignazio, 'Data visualisation from a feminist perspective', D'Ignazio introduces readers to a practical feminist approach on the power and political implications of collecting, analysing and disseminating data visualization. According to D'Ignazio, the main problem with data visualization remains the 'missing body problem', where bodies are extracted, absent, uncounted and rendered invisible from the data presentation. To counter this problem, D'Ignazio suggests a range of design principles for the content, form and process of discovery that can help foster feminist data visualization. Among these are design justice, co-design, and participatory design.

In this section, we will return to the questions related to gender, power, and racism in existing visualizations and visual search tools. Questions which are highly relevant for the TInnGO concepts and project aims. This includes reflections and problematic aspects of power and inequality in prevailing big data archives and provisions. In addition, we want to keep the possible potentials of alternative or even subversive uses of 'big data' and new technologies of collection and visualization open.

In the following, we would like to refer to the interesting study of "fairness" in image tagging algorithms of which findings seem highly relevant for the TInnGO digital analysis. The study, conducted by a team of researchers from the Cyprus Center for Algorithmic Transparency, scrutinizes various search tools , APIs and their potential discriminatory effects on visuals and perceptions (Kyriakou et. al. 2019). [4] Kyriakou et. al. refer to the increasing expectations that algorithms should

[4] See also Buolamwini, J.& Gebru, T.2018.

behave in a matter that is socially just. The study considers such expectations against the case of image tagging APIs and their interpretations of people images notably against the recognition of gender and race.[5] The "fairness" study departs from the fact that image taggers have become indispensable and widely used in our information ecosystem, where new modes of visual communication and sharing have been launched. Recently they have become widely available as cognitive services. The problem is that while tagging APIs offer developers and market agents an inexpensive and convenient mean to add functionality to their creations, most of these tagging APIs are seen as opaque and proprietary and depending on commercial and individual solutions.

Do APIs interpret people regardless of gender and race in a fair manner? This was the guiding question to the comparative analysis of six tagging APIs We will here very briefly summarize the cross-platform comparison of six tagging APIs provided in the by Kyriakou et. al The cross comparison of widely used API image tagging platforms, shows that behaviours differ significantly. While some of the image tagging platforms (APIs) offer more interpretations on images, they may exhibit less fairness toward the depicted persons, by misuse of gender related tags and/or making judgments on a person physical appearance.

As depicted in the chart below the tagging APIs, also called the "tagger API behaviour," is described in manners in which the depicted persons are interpreted by the algorithms. Kyriakou et. Al. make a distinction in this case of gender between:

a. Gender interference: While taggers are general tools and not specifically designed to recognize gender, many gender related tags are used, often inaccurately
b. Judgment tags: Some tags are subjective in nature, and do not logically follow from the content of the input image. In particular many comments on a person's physical appearance. e.g. attractive, cute, pretty.. meaning that they repeat/ and radicalize existing gender stereotypes.

[5] API: **application programming interface** (**API**) is a computing interface which defines interactions between multiple software intermediaries. Clarifai, which has been applied in the TInnGO digital analysis is an image and video recognition tool provided by a New York based company. The API automatically assigns tags to objects and categories taking only the pixels as input, using a wide library of semantic and visual terminologies for artificial intelligence.

c. Abstract interference: Similarly observe tags that describe one's occupation role, e.g. athlete, as well as perceived character traits or emotional states, e.g. using words such **as** serious, friendly, passive etc. (Kyriakou et. al. 2019: 315)

Figure 1: Example from "Fairness in Proprietary Image Tagging Algorithms: A Cross-Platform Audit on People Images" (Kyriakos et. al. 2019: 314 )

Figure 1: Example Chicago Face Database (CFD) images of Asian (AF-204), Black (BF-231) & White (WF-200) women.

| Tagging API | AF-204 | BF-231 | WF-200 |
|---|---|---|---|
| Amazon Rekognition | human, people, person face, head, portrait dimples | human, people, person Afro, hairstyle hair, face | human, people, person face, portrait, head female, woman |
| Clarifai API | one, portrait, cute, child, people facial, wear, man, looking face, isolated, funny, adult joy, casual, happiness, pensive adolescent, eye, serious | people, one, portrait, man wear, adult, side, pensive profile, woman, face, isolated child, facial, Afro, casual fashion, athlete, adolescent | woman, portrait, isolated, one, cute casual, people, fashion, eye young, looking, look, pretty young, wear, face, hair, serious adult, friendly, facial |
| Google Cloud Vision | face, eyebrow, cheek chin, skin, forehead nose, head, jaw, neck | face, forehead, chin, eyebrow cheek, nose, head, jaw, neck, human | face, eyebrow, chin cheek, head, forehead neck, jaw, portrait, ear |
| Imagga Auto-tagger API | beard, man, face, person, male, portrait handsome, child, guy | Afro, man, face portrait, male handsome, head | beard, portrait, face, person man, attractive, model handsome, male, adult |
| Microsoft Vision | person, man, necktie, wearing, indoor, shirt, posing, looking suit, camera, glasses, young, photo dress, black, front, standing, neck white, smiling, male, holding, hair | man, person, wearing, looking necktie, standing, shirt, front face, smiling, white, suit posing, hair, holding, neck, young glasses, black, head, hat, red | person, posing, necktie, wearing shirt, young, man, smiling glasses, photo, holding, camera hair, dress, front, standing, black woman, neck, suit, blue, red |
| Watson Vision | person, skin, light brown color ash grey color | person, woman, female indian red color, coal black color | person, people, face, adult person ash grey color |

Table 1: Output tags for example CFD images produced by six image analysis APIs.

Evaluations of the tabs show that there are shortcomings and misreads in all the selected tagging APIs. There are various misreads of gender, but these become even more pronounced in the case of race. Race is the most misread category, especially in the intersection with gender. Whiteness and recognizable femininity seem to be the most familiar to the machine reads. It is thought provoking that the white woman with feminine hair style is the most recognizable while both the Asian and the black woman is described as man with various judgment tags. Following criticism and critical public

attention some of the API platforms have now omitted sex as a recognizable category in order to pretend neutral– a fact which complicates digital gender analysis or requires different key words.[6] According to D'Ignazio, the main problem with data visualization remains the 'missing body problem', where bodies are extracted, absent, uncounted and rendered invisible from the data presentation. To counter this problem[7], D'Ignazio suggests six design principles for the content, form and process of discovery that can help foster feminist data visualization. Among these are design justice, co-design, and participatory design.

In the study 'The development of gender-responsive indicators: towards a participatory approach' it is argued that, although much attention has been paid to the size and possibilities of big data, including in the field of gender equality, there has been too little concern with the quality of the indicators being measured. The authors therefore suggest a participatory research design where important stakeholders and target groups and their social contexts actively participate in the development of gender-responsive indicators for measuring gender equality. By now examples of digital co-design are scarce in feminist research also when it comes to transport and mobility and the method presented in this TiNNGO digital analysis may have the potential of being elaborated and redesigned in alignments with stakeholders in the various TInnGo hubs. [8]

As for the TInnGO digital analysis , Clarifai was chosen as the Image IPI due to its multiple levels of gender and interferences tags, which was foreseen to be useful for the TInnGO aims. However, even here many misreading of persons were reported from the various partners, as will be demonstrated in following sections.

[6] For a discussion of discriminatory algorithms, fairness and too simplistic ways of solving the challenges, see http://blog.practicalethics.ox.ac.uk/2019/01/cross-post-biased-algorithms-heres-a-more-radical-approach-to-creating-fairness/ -

[7] For examples and reflections see D'Ignazio K & Klein, L F. (2020)

[8] The TInnGO project has set up hubs for developments of Gender and diversity action plans in 10 European cities.

# TInnGO digital analysis: Core concepts and principles assisting digital analysis – challenges and potentials

In order to assist the analysis, we in the following section summarize a range of relevant TInnGO core concepts, which have assisted in the interpretations of the digital analysis.[9] In so doing, we will keep track of the overall aims of the TInnGO digital analysis, which were:

- To introduce analysis of visual digital material with a special focus on gender in smart mobilities and gendered discourses of transport employment.
- To introduce methods where quantitative (distant) and qualitative (close) readings of the visual or textual corpus, e.g. in discourse and cultural analysis was presented together with new topical areas of the transport.

## Combining digital patterns with close readings

At the overall level digital analysis enables the combination of distant and close readings of data and as such new combinations of quantitative and qualitative analysis. While the most common way of combining qualitative and quantitative analysis has been a certain division of work between researchers and experts specializing in either field, the TInnGO digital analysis was intended to work more explorative in integrating the two methods. This has implied the creation of a web corpus consisting of tagged images, guided and interpreted by a range of TInnGO core concepts related to smart transport, gender and employment.

In so doing, we are inspired by Anders Munk, who has spelled out four types of qualitative and quantitative analysis in relation to digital modes and the making of meaning, namely:

1) through *complementarity*, a division of labour in which quantitative and qualitative methods are allowed to unfold relatively undisturbed by one another, the latter performing the job of situating and interpreting the insights gleaned from the former;

[9] For elaborated presentations see Christensen & Breengaard TInnGO Roadmap 2019

2) through a *single level of analysis*, whereby the potential of online traces are seen to reside in their ability to be both qualitatively rich and quantifiable at the same time, enabling an analysis of how apparent macro phenomena are produced on the micro level;

3) through *curation*, a critical practice in which a qualitative understanding of different media environments and their effects on the production of online traces become integral to the way in which such data should be sourced and quantified;

4) through *algorithmic sensemaking*, whereby the relational reasoning typically associated with qualitative fieldwork is emulated quantitatively through techniques like pattern recognition (Munk 2019).

The TInnGO digital analysis has in various ways been inspired by these methodologies while at the same time been, as we will show, a learning process in itself, regarding thematic focus, key words, pattern detection and sense making of the various data sets and network patterns.

## Gender, smart cars and technology – masculine scripts

Technological artefacts such as cars, are commonly perceived as 'neutral objects.' In order to approach the gendered dimension of new technologies, such as smart transport, we might add an understanding of technologies in terms of the social relations, which interact with the technology during design, development, and production. A way to capture the social shaping of technologies is through the concept of 'script'.[10]

A script refers to the processes of making technologies in which designers are defining the imagined users with specific characteristics, such as taste, competences, motives and aspirations, which make them want particular features. As such, the designers are 'inscribing' a certain vision in the new technological object, which attaches particular meanings and lifestyles to the objects. The social shaping of certain technologies, such as cars, includes both the material object, the practices of using it and the identity of the (supposed) user. These are different layers in a script, which are useful for thinking and analysing gendered dimensions of mobility.

[10] We work with the concept of 'script' as discussed by Manderscheid 2018 and elaborated in Christensen and Breengaard: TInnGO Roadmap 2019.

Using this approach can shed light on who might benefit from new technologies and who will not. It also highlights the role of designers and planners and points to how important it is to include these experts in research processes. It serves to show how varied types of transport and mobility practices inscribe women and men in different ways. The question is how new types of inscriptions can be launched in order to achieve more sustainable and just visions of European transport strategies.

Existing studies of the emerging field of smart mobilities point in different directions when it comes to attraction and use by various groups. On one hand, there is a utopian notion of a new beginning with automated cars as an avenue towards a more equal and genderless mobility regime. Here the coming of the driverless car is foreseen, at least in principle, to loosen the strong bonds of men, masculinity, and cars from the automobility era (Dant 2004; Balkmar & Mellström 2018). On the other hand, several studies find serious exclusions of women and potential racism of smart mobility proliferated in the automated car (Manderssheid 2018; Hildebrand & Scheller 2018, Christensen & Breengaard 2019).

Instead, both words and images of the self driving/smart cars are routinely presented with an appeal to ideals of middle-class men and masculinity. For example, in an online presentation by big companies such as the German company Bosch, smart technology and automated cars are presented as independent, self-determining and reasonable with a clear visual appeal to white middle-class businessmen. The link between driverless cars and businessmen can be seen as a gendered script, which connect a certain identity (businessman) with a certain type of mobility (automatic driving). A script might also work on other layers, connecting some identities to practices of either driving or being passenger. (Christensen & Breengaard 2019) One example of how both implicit and explicit gendered images were carried into the tagged depictions in the TInnGo analysis was the tag" mini skirt", which referred to a new automated minicar, while the bigger and smarter cars were depicted or implicitly addressing men in the shape of smart technology.

# TInnGO digital analysis – Main steps and lessons learned

In the following sections we demonstrate cases of how big data and digital visualization can be applied in the study of mobility and transport. Digital analysis in this report presents a new mode of knowledge based comparative analysis of various ideas and patterns of gender, diversity, and transport and employment across a selection of European countries, using visual media archives as data material.

The digital analysis tools and findings have as said in the introduction been developed in a stepwise educational process within the TInnGO project consisting in workshop, consultations and individual data collections by TInnGO partners. The process has been supervised and assisted by a consultation team, lead by Anders Munk and Matheu Jeremy, from University of Aalborg, Denmark.

The aim is to outline an account of the particular process conducted by the TInnGO partners who participated in the process. Please note that a a general step by step guide do the conduct of a digital media analysis can be consulted in annex II in this report.

## The TInnGo digital analysis – as a learning process.

From the beginning all, TInnGO partners were made familiar with the basic steps in the TInnGO digital analysis. They were introduced to the same guidelines and conduct of first steps analysis. Following these preliminary exercises, the various steps and key words were corrected and made more reliable for the objectives of the analysis. More specifically, these changes included:

1. Key words became explicitly linked to national language and terms. This was intended both to locate the analysis in language specific contexts and to refine the comparative dimensions of the analysis.
2. Key words of explicit gender / male, female, feminine/masculine were omitted from the general search and only used in subsequent categorizations and pattern making. This meant that gender analysis mainly has been derived from implicit bias / findings, rather than based on data located by the explicit use of male/female. All findings have subsequently been

analysed in comparing and contrasting them with the TInnGO core concepts in relation to smart cars and employment.

The partners were asked to conduct two search queries. As for the specific search terms, the following terms were applied in search query one : "autonomous car and passenger." This search was expected to bring about open-ended depictions of smart transport and various types of passengers, which also reflected that drivers in principle become passengers in the smart self-driving cars. The same considerations applied to search query two, where the search query was "transport and employment." Both searches were broad and open-ended. In order to locate possible national specific terms and patterns to the query, the search words were specifically translated and searched along demographic lines, which complied with national borders. During this process, we realized that "words matter", and that within demographic/national boundaries the most frequently located terms were the following:

Table 1: Context specific search queries

| **Partner** | **Search query 1** | **Search query 2** |
|---|---|---|
| CU, SOCTR | - | *Transport AND Work* |
| CU, UK | *autonomous car AND passenger AND site:.uk*[11] | - |
| INTECO, Romania | *"masina inteligenta" AND "pasager"* | *"transport" AND "angajat"* |
| UCPH, Denmark | *"selvkørende bil*" AND passager* | *transport AND beskæftigelse* |
| VTI, Sweden | *"självkörande bil" AND "passagerare"* | *"transport" AND "anställning"* |
| VTM, Portugal | *carros autónomos AND passageiro AND site:.pt*[12] | *transporte AND emprego AND site:.pt* |

[11] As English is spoken in several countries, search queries only in UK has been searched for.
[12] As Portuguese is spoken in several countries, also Brasil, search queries only in Portugal, has been searched for.

As for the country specific terms, the variations in terms might refer to selfdriving/ smart cars and smart transport as an emerging field, which does not have acquired a fixed meaning by now. We chose the term "autonomous cars" as the most exact and also most radical form of smart transport, because the term smart car might indicate a range of more or less automated cars and therefore was less precise.

All TInnGO partners were asked to follow the five steps in the digital guidelines (see annex X). Put simply this implied the use of query key words and searching for images in the Google Image Scraper: https://tools.digitalmethods.net/beta/googleImages/

All participating partner in the digital analysis were to begin with asked to provide 500 images related to the country specific queries as the basis for the next steps.[13] This included the tagging of all the images in the Clarifai recognition service (API) meaning that every image comes up with a description in a spreadsheet. Next the spreadsheet with images, and tagged concepts forms the basis to discover data in many ways.

From this stage of image sampling and tagging the partners were asked to provide the following assessments:

*Acceptable recognition:* We would like to have an idea of the precision and correctness of the concepts that are tagged to describe the images. Partners are therefore asked to (briefly) compare the pictures with the column "General concepts". Please take a few minutes to jot down if there are any major discrepancies between the concepts and the images, as well as what the concept describes, and how common the error appears. This will give us opportunity to comment on the acceptability of the network, once we move on to step six.

*Description of data:* Please write a description of the tagged images based on your work with data in the spreadsheet. The short description shall include statistical frequencies of relevant categories, e.g. gender and ethnicity. Please do also make comments on the immediate themes and patterns, you have come to think of when you worked with *Reliability of the recognition tool.* Prominent actors, i.e.

[13] The images were simply the first 500 retrieved from the Google image scraper.

which websites appeared frequently, arethe reported findings of these assessments are shown in the chart below:

Table 2: Feminine and masculine expressions in material, Smart car and passenger

| | **Denmark** | **Romania** | **Sweden** | **Portugal** | **UK** |
|---|---|---|---|---|---|
| *Count* | Clarifai demographic tags vs. manual count | Clarifai demographic tags | Clarifai demographic tags | Clarifai demographic tags | Manual scroll through |
| *Scraped images* | 500 | 500 | 500 | 500 | 500 |
| *People* | - vs. 24,2 % | 4% | - | 18,4 % | 10 % |
| *Feminine* | 3 % vs. 6 % | 0 | 4,2 % | 4,6 % | - |
| *Masculine* | 3 % vs. 14,8 % | 0 | 5 % | 2,6 % | - |
| *Comments* | A manual count showed that 6,2 % of the images included person(s) where the gender was unclear.<br><br>Difficulties with the demographic tags, but the "general concepts" were more precise in the given images. Identified police squad, snow and cockpit when not present in image. | Low number of people/ technical dominance | No specific gender pattern / middle class/ mainly white/ young or middle aged/few blacks | Half of the images tagged as feminine were missed classified | Demographic tags incorrect up to 50 % in gender.<br><br>General concepts described snow when not present. |

Table 3: Feminine and masculine expressions in material, Transport and employment

| | **Denmark** | **Romania** | **Sweden** | **Portugal** | **UK** |
|---|---|---|---|---|---|
| *Scraped images* | 500 | 500 | 500 | 500 | 500 |
| *Count* | Clarifai demographic tags vs. manual count | Clarifai demographic tags | Clarifai demographic tags | Clarifai demographic tags | Clarifai demographic tags |
| *Persons* | - vs. 22 % | 8 % | - | 19,8 % | - |
| *Feminine* | 3 % vs. 5 % | 5,2 % | 4,4 % | 4% | 3,6 % |
| *Masculine* | 2 % vs. 10,6 % | 2,8 % | 4,8 % | 3.2% | 2,2 % |
| *Comments* | A manual count showed that 8,6 % of the images included person(s) where the gender was unclear. Difficulties with the demographic tags, but the "general concepts" were more precise in the given images. Identified police squad, crash, glass when not present in image, and "no person" when there was a person. | Intersectionality issues represent a marginal interest with less than 1% according to demographic tags. | Several images with no visual connection to transport and/or employment occurred, e.g. News pictures of crowds or an audience taking part in some event. | Trucks described as police squad, "driver" in driverless images | Five duplicate images. Search included two separate irrelevant images of surgical theatre lamps. |

The general findings showed a rather low number of depicted persons in both smart cars and employment queries. The quieries showed that persons were recognized correctly but that gender and other categories were often misread. Such aspects invites to deeper critical and reflexive analsysi of the algorithmic constructions of categories in follow up studies. Yet in the case of the TInnGo digital analysis we primarily looked at the categorizations/ visualizations to explore patterns and stereotypes relating to gender. Here the overall level the low number of depicted persons might give a first indication of the instrumental and technical upper hands of both smart cars and transport employment sectors. This also points towards an implicit male and masculine bias in the digital image archives, since car technology and engineering has been equated with men and masculinity ever since the birth of the fuel driven motorcar. (Oldenziel, Ruth. 1999. Scharff, Virginia 1991 A point which was stressed in the subsequent depicted networks and clusters provided through the Gephi tool.

## About networks and making sense of patterns

Networks and network analysis represent a second phase in the TInnGO digital analysis. The aim was to transpose the web archives of tagged images into networks that illustrate nodes patterns of multiple encounters and connections, which would be hard to achieve in e.g. qualitative analysis

In order to make the network and pattern finding methods and findings accessible we return back to the suggested mode of analysis called *algorithmic sensemaking* in the four outlined approaches to qualitative and quantitative analysis. Algorithmic sensemaking differs, according to Anders Munk (2019), from the other modes of "mixed methods" in that it does not involve any conventionally qualitative work, but rather solicits sensemaking for quantitative community detection and pattern recognition. Here the hyphen between qualitative and quantitative signifies a specific denomination of quantitative methods that exhibit some of the exploratory affordances otherwise attributed to qualitative work. Here we are moving into a sophisticated method motivated by Michael Agar: "Traditional social science is on the lookout for variables; ethnographers are on the lookout for patterns" (Agar, 2006: 109, in Munk 2019: 165).

In the TInnGO digital analysis, these transpositions were enabled by the application of the Gephi software in combination with Force Atlas2 a specific graph lay out altgorithm for network visualisation. ( Jacomy et al 2014). The ForceAtlas 2 is designed to push nodes apart unless connected by edges, which then act as springs to hold the nodes together (Jacomy et al., 2014). The result is a visual landscape in which nodes are placed in relative proximity to other nodes to which they are densely connected and at a relative distance from nodes to which they are not connected.

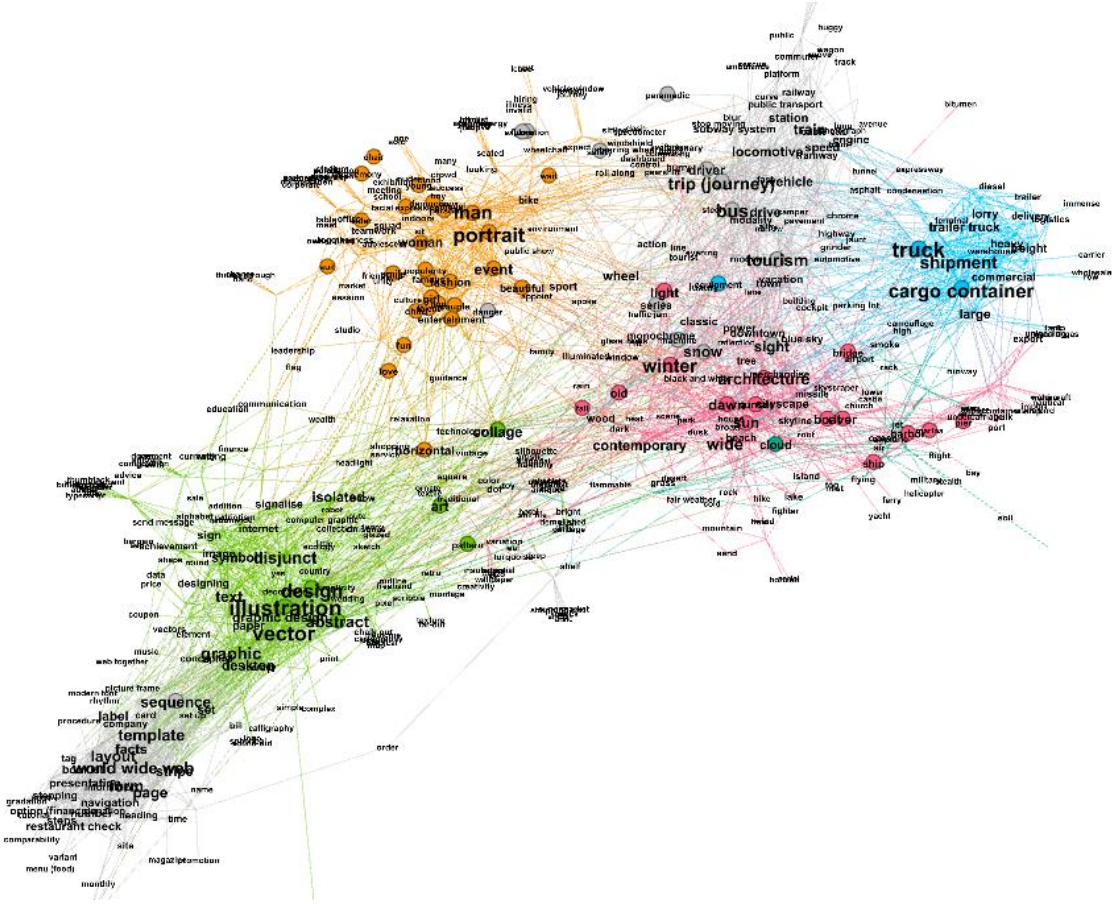
In the following we will present and discuss a sample of the coloured networks that came out of the queries and Gephi translations submittedby by a sample of TInnGO partners,

*Portugal – Smart car and passenger*

The first analysis with the Force Atlas 2 algorithm reveals six poorly defined clusters and a modularity index of 0.37, concordant with the poor definition of clusters. In order to improve the understanding of this set of images, a filter was made based on the number of occurrences of each label, and only the tags with less than 60 occurrences were selected. After filtering the data, eliminating labels like 'transport system', 'panorama' and 'car', we obtained a superior modularity index (0.57), which indicates a better definition of the groups of tags. Five clusters were defined:

- Power/ engine
- Graphics
- Air transport
- Landscape/ tourism
- Road characteristics

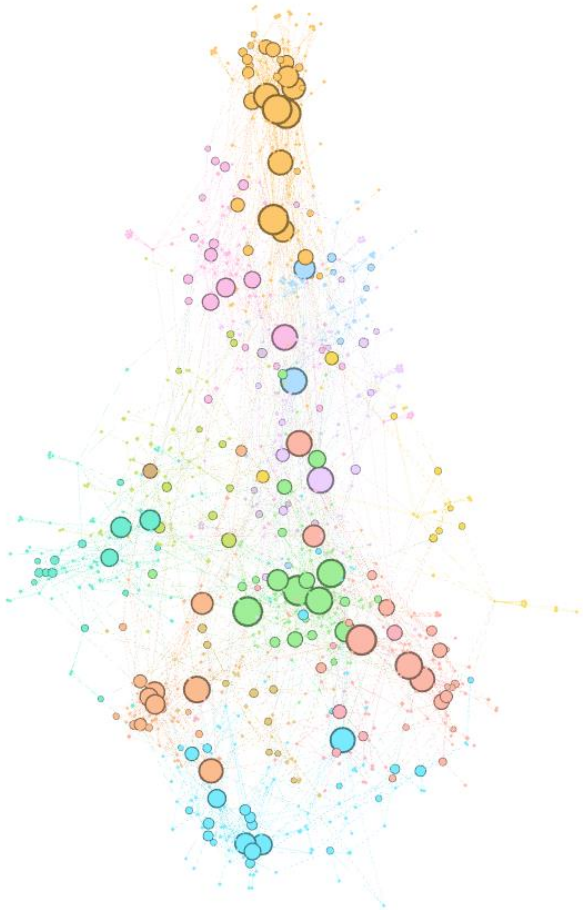
Figure 5: Smart car and passenger, Portugal



The clusters, in general, have different sets of words, which can lead to an inconsistent set of images. For this reason, we have analysed some images of each cluster and found that the Landscape/tourism group is mostly composed of images that are unrelated to the subject. This disconnection of the images to the subject under analysis happens throughout the whole set of data. One possible interpretation for this is that, although Portugal is making investments in the topic 'autonomous cars' and there are more and more tests being done in its territory, it is still not a subject that raises interest to the majority of the population. The algorithm is indeterminate, meaning that it never ends with exactly the same visual result. The same applies to the *Louvain modularity algorithm* used to colour the network (Blondel et al., 2008; Lambiotte et al., 2008). This algorithm attempts to find the most optimal way of cutting the network into smaller components, severing as few edges as possible in the process. Again, there is no exact and reproducible result. It could be said that these types of algorithms work exploratively to find patterns in ways that are in fact rather similar to those that an ethnographer would use (Munk 2019: 165). Apart from ForceAtlas2 layout and the Louvain modularity algorithm, all partners have made use of the function "*Degree Range".* With this function one can exclude tags that appear most or least with other tags. This can be an advantage as we know that the tags that appear most frequent with other tags do not necessarily tell us more about an image. For example will "panoramic" or "transportation system" not give us much more information, if this is occurrent across all images. When filtering and excluding these tags an underlying cluster can instead become more apparent.

*The UK – Transport and work.*

UK made a search on "Transport and Work" as *work* produced more searches according to Google Trends, and furthermore is used in the title of the Government "Department for Work & Pensions" and thus was expected to be a more familiar word in the local context. A network was produced, which can be seen on the right.



Transport and Work, UK.

A high degree of modularity was detected with clear clusters shown. There were two themes detected in the clusters:

- Clusters related to transport modes
- Clusters related to the representation of transport and work

Figure 20:

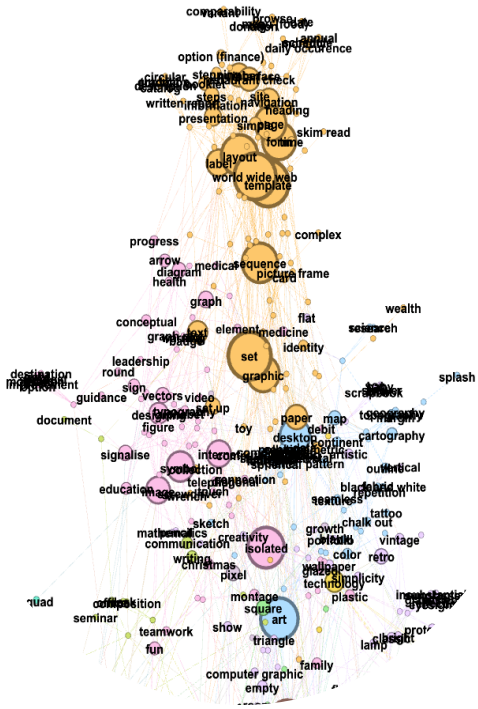



Figure 21: Transport and Work, UK, Enlarged top

**Transport modes**

- Rail (orange)
- Infrastructure (bright green)
- Vehicles – buses, lorries etc (Aqua Blue)
- Marine (Coral red)
- Active Transport (Emerald green)
- Air (Mustard yellow)

**Representations of Transport and work**

- Concepts (yellow)
- Information/education (Pink)
- Art (Sky blue)
- Graphics (lilac)
-

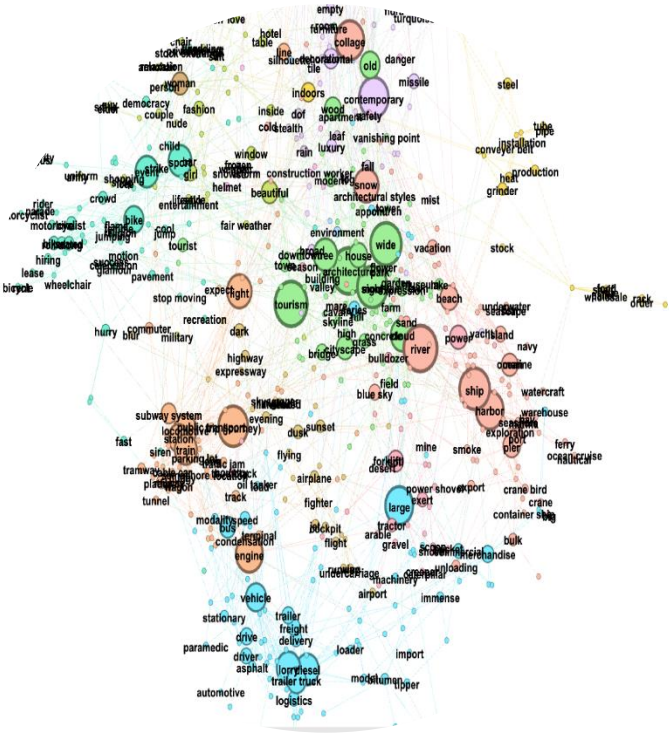



Figure 22: Transport and Work, UK, Enlarged bottom

Two further Gephi analysis plots were produced to explore the male and feminine tags. In figure 23 the nodes which represent a feminine tag are enlarged, and in figure 24 the nodes that reflect a masculine tag are enlarged.

*Feminine Plot*

In the feminine plot women were shown to be significantly clustered in active travel modes (Emerald green). There was a higher degree of descriptive words in this plot, particularly around characteristics of fashion, style, leisure, and travel. When cross referencing the Gephi output against the original images, confirmation was seen that women were mainly shown as passengers on public transport and riding bikes and not as transport workers. Only three images showed explicit reference to women working within the transport sector.

*Masculine Plot*

In the Masculine plot, clusters were observed linking the male images to drivers and cars. Construction workers and the marine sector. This cluster was also show to link with the environment cluster, but this did not appear to be related to work.

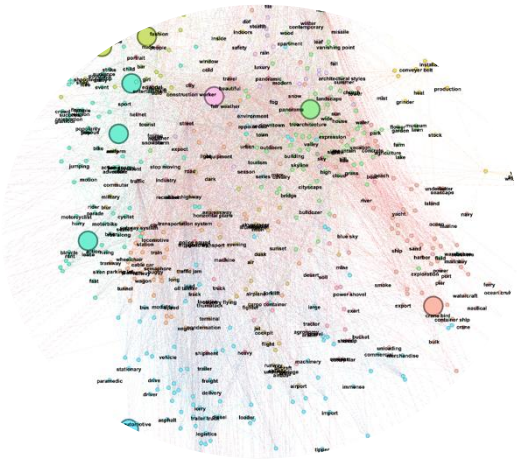


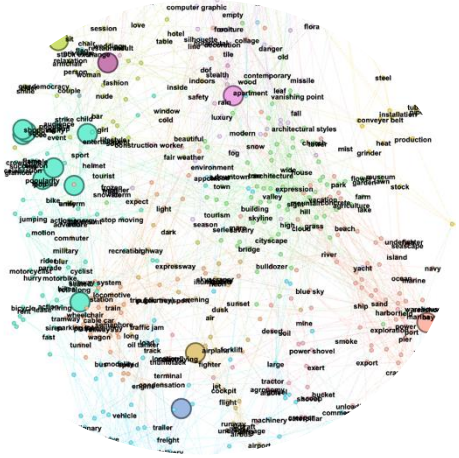
Figure 23:  Feminine plot

Figure 24: Masculine plot

Overall in this google search for the UK it appears more straightforward to establish Transport mode and imagery of transport and work, but not to understand the relationship to work in the transport sector. Gender, race, and age as characteristics were infrequently shown in the images. The images returned in the search were heavily related to transport modes with LGV trucks alone making up over 18% of the total images. None of the truck images featured women as drivers of the trucks yet in the

logistics sector in the UK approximately 25% of the workforce is female. The findings in the image search however would seem to be representative of the very low number of women within the industry who are employed as truck drivers, higher levels of women are employed in a non-driving capacity in roles such as operations management, scheduling data analysis and human resources roles

## Representations of "smart car and passenger"

In the following we look closer at the clustered representations of smarts cars and passengers. This includes examples from Romania and Denmark, and illustrate how data visualizations intersects with qualitative analysis of the patterns found.

### Smart cars and passengers across Europe

Across several of the scrutinized countries, the visualizations show that the clusters are very interwoven, which also results in low modularity scores, showing that the clusters are not clear. That there is no clear clustering also shows that the images representing 'smart cars and passengers' are not divided into clear and deliminated communities. The table below presents an overview of the different clusters in four of visualized networks in this section. Denmark and Romania can be compareddirectly, as the filtering of the networks is similar. ( The findings in representation of smart car and passenger in the UK will be include but cannot be directly compared since they applied a slightly different mode of query)

Table 4: Comparison of clusters in the representation of "smart car and passenger"

| **Denmark** | **Romania** | **Portugal** | **UK** |
|---|---|---|---|
| Everyday transport in city and nature (25 %) | City and nature (25 %) | Power/ engine | Road, asphalt, automotive, bus, and bike |
| Layout and technology (25 %) | Technology and futuristic elements (19 %) | Graphics | City, nature. and tourism |
| Cars (23 %) | Cars (19 %) | Air transport | Vehicle, summer, man, girl, limousine, and technology |
| The users (16 %) | Presentation and documents (13 %) | Landscape/ tourism | Web and computing, graphic design |
| Public transport (2 %) | Interior design (12 %) | Road characteristics | - |

| | | | |
|---|---|---|---|
| Air transport (2 %) | Technology, dashboard, and accessories (8 %) | - | - |
| - | Air transport (4 %) | - | - |

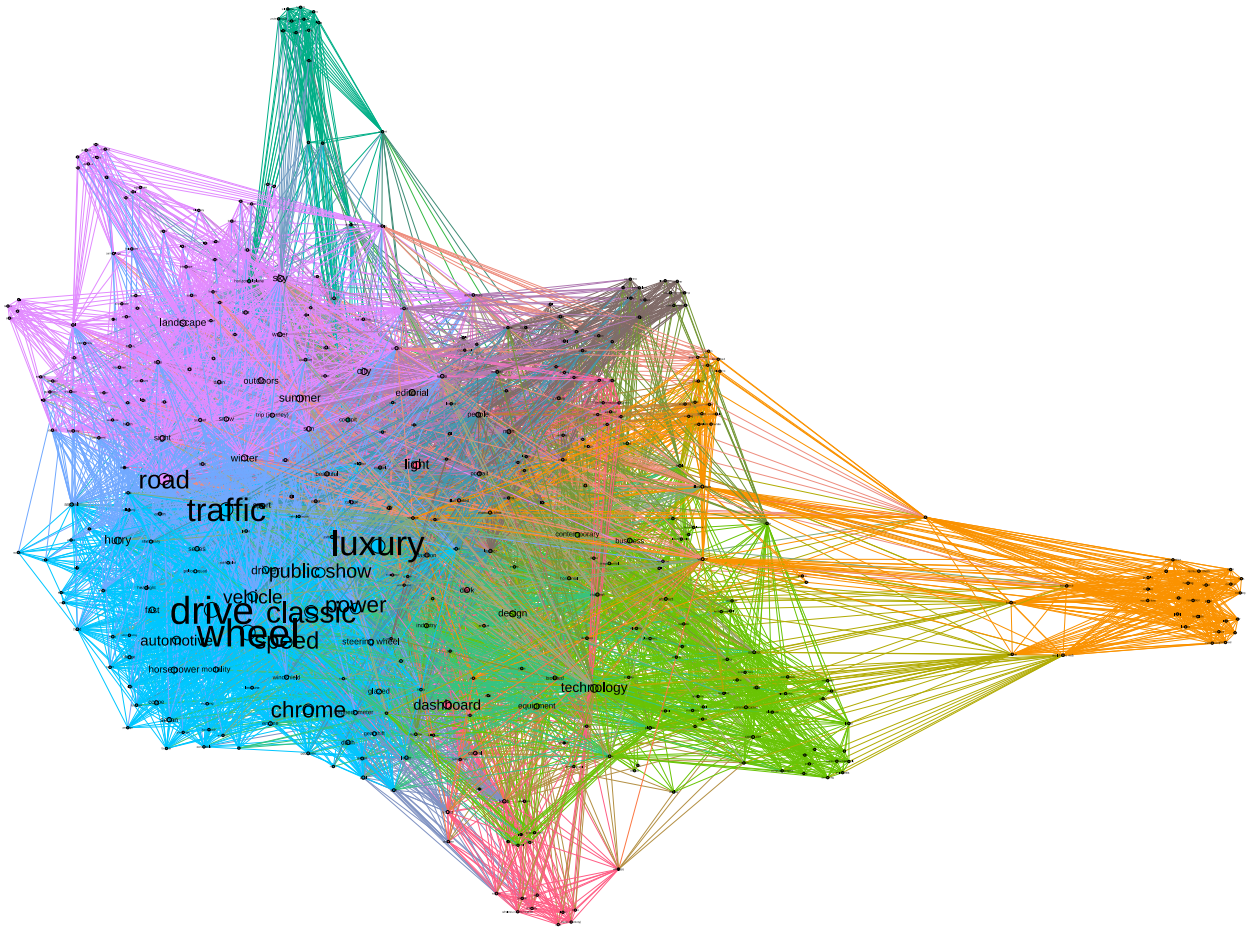
In general, it turns out that majority of the images representing 'smart car and passenger' across the four countries, are not represented with persons (see table 2). Feminine and masculine tags to the images from the UK search showed that both men and women were represented in all clusters, and no immediate gendered patterns were found. A closer inspection is though recommended, as the preliminary qualitative inspection found that "person" was identified well by the recognition tool, but problems in identifying male and female still consisted.

When we look at Denmark and Romania there are similar clusters, though with some differences. The biggest cluster gather in both network 25 % of the total tags, in Romania represented with "city and nature", and in Denmark represented with "everyday transport in city and nature". A qualitative inspection of the Romanian images behind the tags showed that though both male and female car users were present in more images, it was still technology "who was at the wheel". Though it was not the general picture of the cluster, a qualitative inspection from the Danish material showed that there were examples of both men and women as actors and users of technology, without one being subordinated to the other. We need though to have in mind, that a manual count of the total images on 'smart car and passenger' in a Danish context showed that one or more men were present on 14,8 % of the images, where one or more women were only present on 6 % of the images.

### *Romania*

Through a search on *smart car and passenger* in a Romanian context, 500 images were collected and tagged with the Clarifai recognition tool. The figure below illustrates seven clusters, that were identified using the Louvain modularity algorithm, and force atlas 2 layout. Furthermore, the degree range function was used to filter the tags that most often occurred together, to visualize a clearer pattern in data. The clusters are composed of the purple cluster (25 % of the total tags), the green

Figure 3: Smart car and passenger, Romania




cluster (19 % of the total tags), the blue cluster (19 % of total tags), the orange cluster (13 % of the total tags), the grey cluster (12 % of the total tags), the pink cluster (8 % of the total tags), and the darker green cluster (4 % of the total tags).

The biggest cluster is the purple one (25 % of the tags), which contains tags related to **city and nature** e.g. winter, rain, sun, trip (journey), summer, ocean, outdoors, landscape, tourism, bike, mud, pavement, beach, expressway, cityscape, blue sky, and ship. more images were from a qualitative inspection shown indicate vehicles or technologies with persons not present, or the person was in the

background, and the gender could not clearly be identified. Images were also present of both male and female car users, where it was still technology, “who was at the wheel”.

*Image with “trip”-tag*

*Image with “trip”-tag*

At last images of men who drive in a traditional fashion were found in the qualitative inspection of the images behind the tags. Below is shown two seemingly identical images of a man who drives a car. The representation of the men is though very different and indicates different masculinities. The image on the left shows colour, a watch, a suit, relating the car to wealth and perhaps business, and from the window is shown car lights, which could indicate the context is city. The image on the right instead shows a man with a quite neutral t-shirt in a car brand, Seat, that is one of the more affordable brands in the Volkswagen Group, and from the car window one can see sky and something green indicating nature.

*Image with “outdoors”-tag*

*Image with “outdoors”-tag*

The green cluster (19 % of the tags) contains tags related to **technology and futuristic elements**, and contains tags such as technology, device, button, laptop, multimedia, portable, computer, network, display, simplicity, remote control, industry, fun, bright, and empty. In the qualitative inspection the futuristic elements were the images, that distinguished most clearly from the other clusters. See to examples below:

*Image with "technology"-tag*

*Image with "technology"-tag*

The blue cluster (19 %) contain tags related to **cars** e.g. drive, wheel, classic, speed, power, luxury, chrome, traffic, series, vehicle window, exhibition, steering wheel, tractor, engine, tire, slick, super, action, sport, hurry, and accident. The cars represented in the images were not big, and expected to be for private, and not for public use.

The orange cluster (13 %) is quite unattached from the other clusters. It contains tags related to **presentation and documents**, including: facts, steps, donation, heading, form, label, template, interface, password, portfolio, booklet, modern front, tag, presentation, and layout. The grey cluster (12 %) is very interwoven and contains tags related to **interior** design, and contains tags e.g. chair, escalator, hotel, sitting, session, sit, table, lobby, trading floor, apartment, rug, classroom, education, seminar, meeting, desk, and conversation. A qualitatively inspection showed that what distinguishes this cluster is both the visualization of interior design in cars and building

*Images with "chair"-tag*

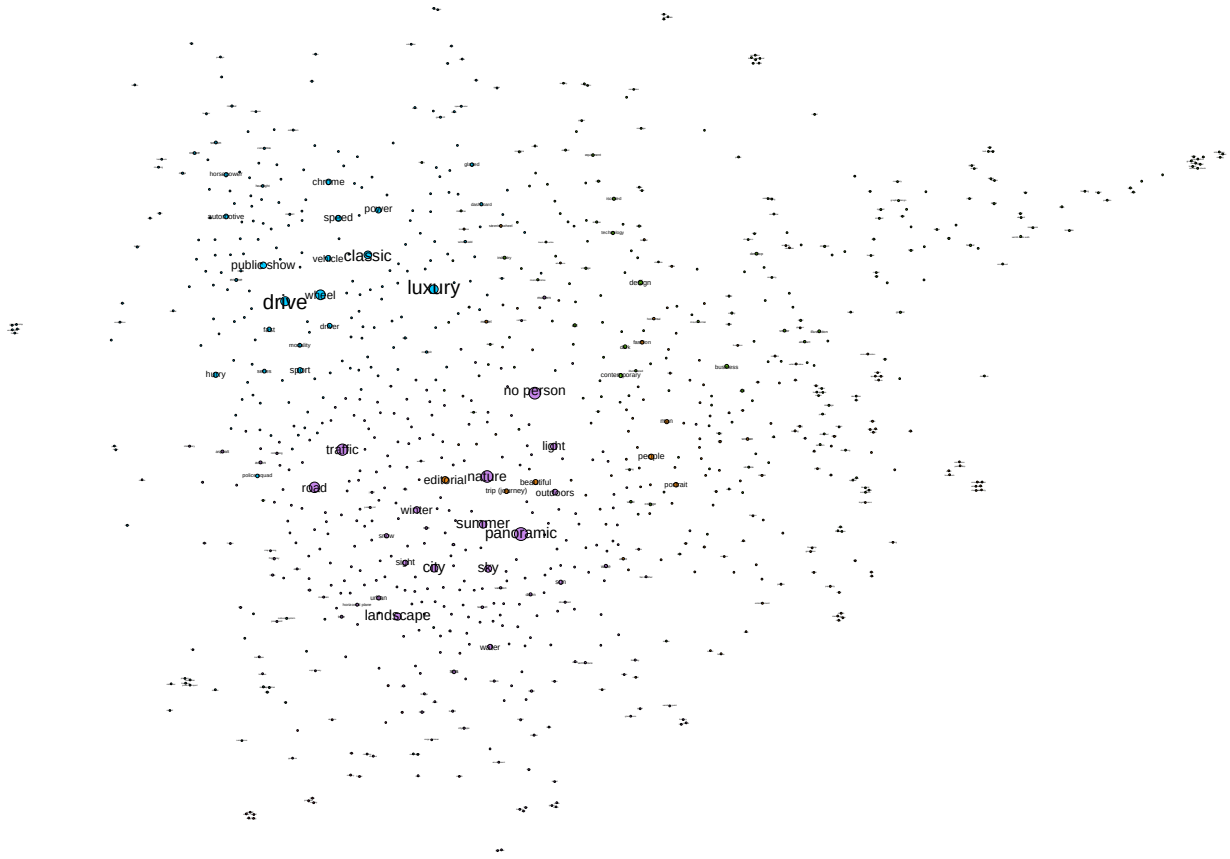
The pink cluster (8 %) contains tags related to **technology, dashboard and accessories** e.g. stereo, wireless communication, speaker, audio, intensity, bass, player, mono, music (auditory), analogue, knob, sound, obsolete, control, temperature, dashboard, time, full moon, light, illuminated, and chrome. The dark green cluster (4 %) contains tags related to **air transport** e.g. undercarriage, silhouette, aircraft, flight, runway, fuselage, airliner, flying, airbus, fighter, and wing. The cluster is quite small, and it is partly reasoned that many tags leads back to the same image (see below), that connects this cluster as something quite distinct.

*Image, that dominates in the dark green cluster*

*Denmark*

Below is found a visualization that reflects the representation of *Smart car and passenger* in a Danish context. The clusters are not clearly separated, which can also be seen from a low modularity score of 0,34. This indicates that the search on 'smart car and passenger' is not divided into different representation communities, which could also be reasoned to it being a specialized category. The tags that most often occurred with other tags were filtered in order to show a clearer pattern. We see in the visualization that there are in all six clusters: The purple (32 %), the green (25 %), the blue (23 %), the orange (16 %) and two small clusters; the dark green (2 %) and the pink cluster (2 %).

Figure 4: Smart car and passenger, Denmark




The purple cluster (32 %) contains tags that reflects **everyday transport in city and nature.** It includes tags such as light, summer, vacation, pavement, landscape, fog, sunset, sky, vacation, sight,

fall, trip (journey), nature, water, and cityscape. A qualitative inspection of the images that the above-mentioned tags represented, showed examples of transport in the city scape, as well as in highway or nature. Many cars were shown, though there also were few images of industry (cargo container) or public transport. The dominant representation of (private) cars can be reasoned to the search query, which were appointed at autonomous car (in Danish *"selvkørende bil").*

*Images with "technology"-tag (1)*



The cars were represented from the outside as full image, in setting as well as presented the interior from the inside. There were more images of cars as technologic wonders with features, but this was not the general theme. Despite the tag "no person" was included in the network, we did find more images where men or women were represented.

*Image with "fall"-tag*

*Image with "fall"-tag*

*Image with "mountain"-tag*

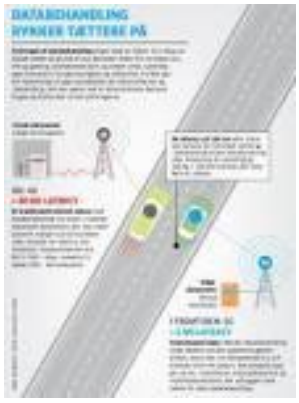
*Image with "fall"-tag*

The immediate images show different settings e.g. highway, mountain (man), farming (man), racing (woman), business (woman). We know from a manual count of the data material that only 24,2 % of the images contain person(s), and that 6,2 % of the total images includes images where the gender is unclear. The main part of the images is not personified. Of the total images 6 % contain at least one woman, and 14,8 % contain at least one man. There are seen a skewing in representation of the total data material, as more men are represented, but when we examine and look closer to the images in this cluster, we find examples on that *the way that women and men are represented in the images* are not subordinated to one another. We find examples where both men and women are actors and users of technology.

If we look back to figure 4, we see the clusters representing 'smart car and passenger'. The green cluster (25 %) have some outliers, that drags it away from the other clusters. It is interwoven with the other clusters, but the least interwoven one. The tags in the green cluster are related to both **layout and technology** and contains tags such as contemporary, business, space, technology, design, number, catalogue, steps, site, presentation, gradation, navigation, booklet, tutorial, label, layout, and illustration. See some examples of the images behind the tags below:

*Images with "technology"-tag (2)*


*Images with "layout"-tag*




If we move forward to the blue cluster (23 %), we see tags related to **cars.** It includes tags such as sport, series, modality, driver, fast, drive, wheel, vehicle, classic, luxury, power, speed, miniskirt, public show, hood, engine, automotive, slick, exhibition, showroom, roadster, sedan, and horsepower. Tags such as "mini skirts", "slick" and "showroom" could indicate a representation of masculinity and sexuality, showing a male dominated discourse within the car industry that links fast cars and women as objects of desires.

*Image with "mini skirt"-tag*

It turned out however, by a qualitative inspection that this is not the case, and miniskirt was not identical with a women in a short skirt. To the right one of the the images, that wrongly tagged as "mini skirt" is shown. The image does not show any lightly dressed woman. As

for gender the minicar associates a driver, where the gender is unclear, but at the same time the miniskirt associates gendered methaphor for a small car.

Such unclear person/gender readings of the driver which as the case for 6,2 % of the total images in the search *smart car and passenger.* Such examples highlight the value of the combined qualitative and quantitative approach to critical and explorative analysis.

*Image with "showroom"-tag*

A qualitative inspection showed that the cars took centre stage in in the majority of the images, though no clear conclusion can be drawn on the masculinity/femininity of the images in general. What can be said is, that this cluster seems to have the main focus on the materiality of the non-human car and devises. If and when persons appear, they seem to be extras and not the centre of focus even though some decictions as shown are implicitly gendered. The cluster contains both images of devices, interior and full picture of the outside of the car.

The orange cluster (16 %) contains tags more related to **the users** than any of the other clusters. It contains the tags, such as people, man, portrait, editorial, beautiful, trip (journey), event, shirt, window, love, girl, woman, couple, one, fun, conversation, seminar, education, desk, university, classroom, and office. The dark green cluster (2 %) represents **public transport** and contains tags, such as subway system, railway, train, line, station, expect, public transport, platform, tramway, commuter. Public transport still finds its way in the search, though the search query were "smart car and passenger". Even though public transport makes up a cluster, the cluster is quite small compared to the blue car cluster, which made up 16 %. At last the pink cluster (2 %) represents **air transport,**

as it contains tags, such as fighter, military, stealth, flight, propeller, runway, airplane, aircraft, airport, wing, flying, air, and jet.

*UK*

In the following is a visualization of the representations of *smart car and passenger* in the UK. The UK team (CU) used Microsoft Visual studio and programming language C# to experiment with filtering and analysing the data, which also included merging tags, for example "woman" + "girl" + "DEM_feminine" in order to simplify the data. The filtering excluded most frequent tags as well as tags that occurred less than five times. This showed clusters: **Green** (most decoupled cluster) relates to web and computing, graphic design, business; **orange** cluster relates to city, nature and tourism; the **purple** cluster relates to road, asphalt, automotive, bus, and bike; and the **blue** cluster relates to vehicle, summer, man, girl, limousine and technology. There were found an unexpected number of productivity aids, few images with negative sides of transport e.g. congestion or pollution, fairly persistent back related to graphic design and a strong set of nodes on nature related themes.

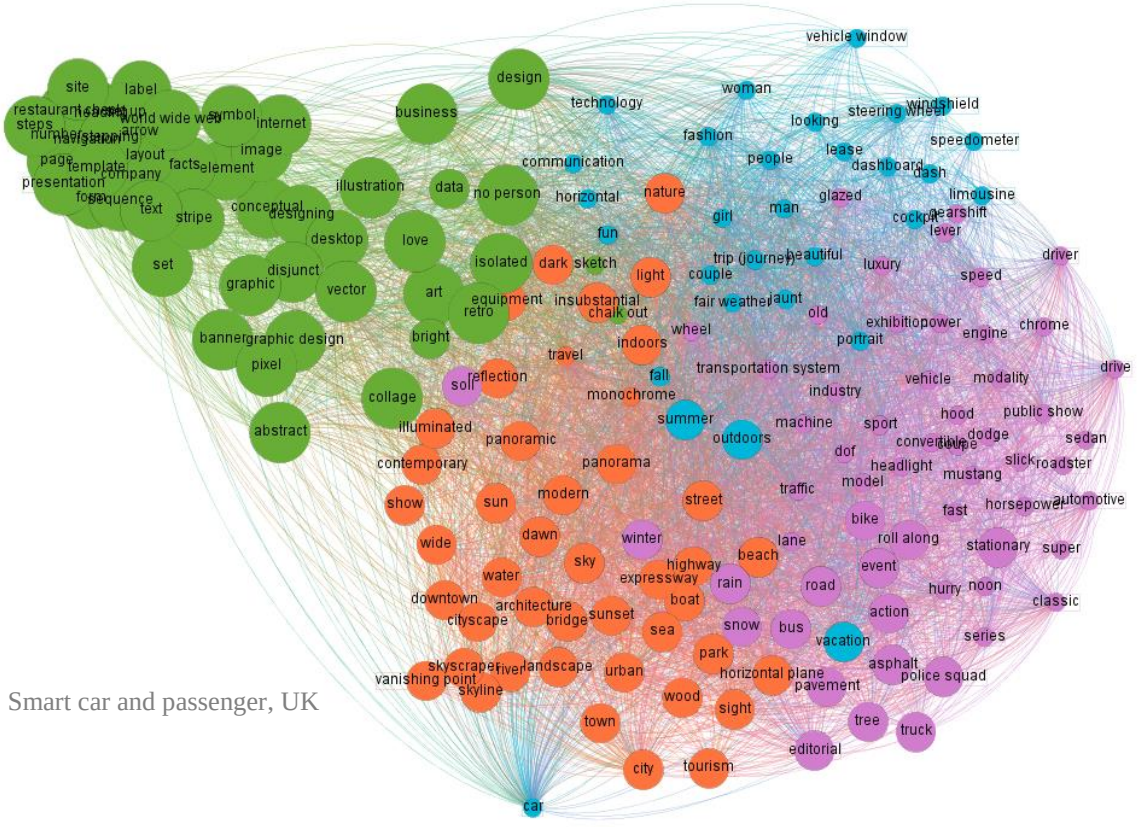



Figure 6: Smart car and passenger, UK

1. Hereafter visualizations were produced to illustrate where “man”, “woman”, and “people” were placed in the network. From these figures, there appears to be no gender bias, as both genders are represented in all clusters.

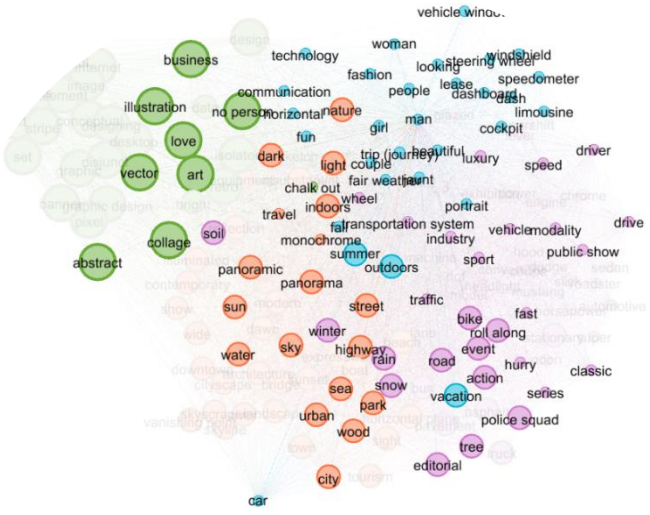



Figure 7: Connections for man

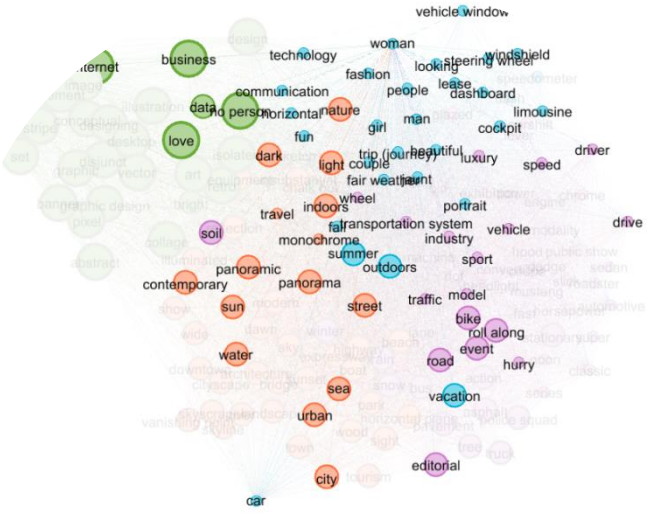



Figure 8: Connections for woman

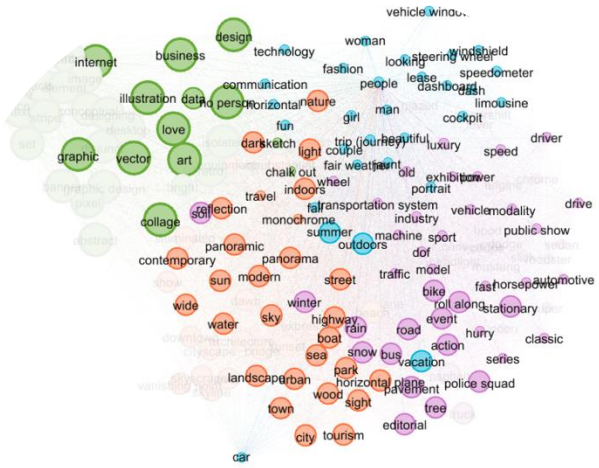



Figure 9: Connections for people

## Summary

In general the digital analysis and pattern findings reflect what has been demonstrated in qualitative analysis of smart cars and gender. In particular the equation of autonomous cars with technology and implicitly/explicitly with men and masculinity. The digital analysis – in addition to the qualitative analysis, using few images – provides a more solid basis of image materials – and also adds to the variations in visual constructions of gendered characteristics and (foreseen) practices. The digital method provides a comprehensive archive of gendered/non gendered representations and a bigger archive from where examples can be drawn and analysed. Also the qualitative inspection of the image challenge the binary gendered depictions and connotations in algorithmic readings. E.g. in the unveiling of various masculinities.

.

## Visual patterns in Transport and Employment

### Gender in the transport labour market – a back drop for digital visualizations

Transport is a highly gender-segregated sector in terms of jobs and employment. . Female staff is mainly found in service jobs and seldom in areas of technology, manufacturing and construction. Further, there is a clear over dominance of men in the sector. The gendered imbalances in the transport sector as labour market are mainly due to issues of:

EDUCATION: The transport related educational system is mainly related to technology and mechanics and is marked by a significant gender-bias, which obviously makes gender-balanced recruitment efforts and policies difficult. Because women make up a small percentage in the transport industry, gender sensitive training tailored is rare and efforts to recuite more women tends to be unsuccessful.

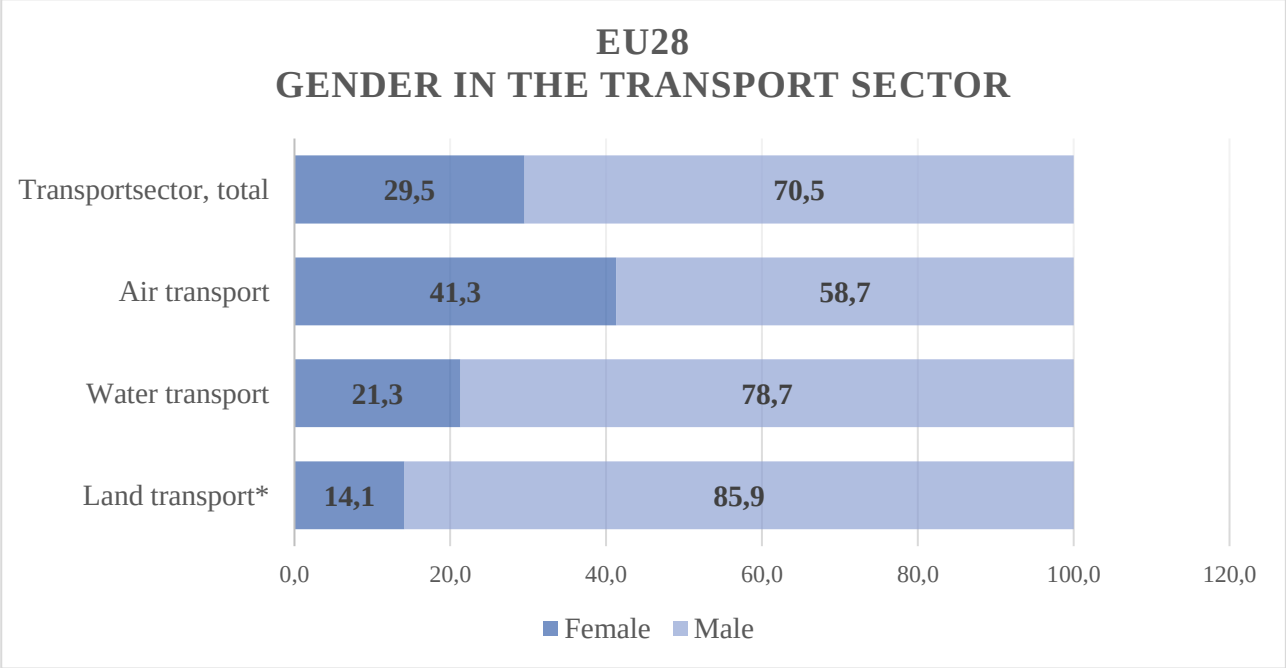
STEREOTYPES: Transport work is imagined as a male basis and scripted masculine, which makes it difficult for women to enter the field. The transport sector has a reputation of consisting of dirty, heavy work – something that is closer connected to masculinity than femininity. Stereotypes of men as technically minded and women as service oriented reproduce their positions in the transport sector.

WORK ENVIRONMENT: That transport is a male-dominated sector might imply little experience, gender, and family frienlye privisions such as issues and rights, i.e. maternity leave, hours of work, part time jobs, shift patterns etc. (Smith & Wigan 2000).

As informal workplace interactions are powerful in the making of inequalities in transport jobs, imbalances in the transport sector can be explained by the concept of unconscious bias and ’homosociality.’ Homosociality means that people prefer persons who are similar to themselves: A recruiting committee consisting exclusively of men will be most likely to also hire a man. Homosociality can influence on the work culture by defining which employees are seen as competent and which are not. As such, not just formal workplace processes but also informal interactions between employees affect women’s position in male-dominated transport domains (Wright 2016).

The figure below shows the gender balance for all EU member states (28) on the same areas of the transport sector. To what extend do the image tagging confirm national traits of segregation and stereotypes?

Figure 2: EU28 - Gender in the Transport sector. Published in Deliverable 9.1. Data from Eurostat.




In the following we will look closer to the clustered representations of Transport and employment. This will include examples from Portugal, Romania, Denmark, Sweden and UK.

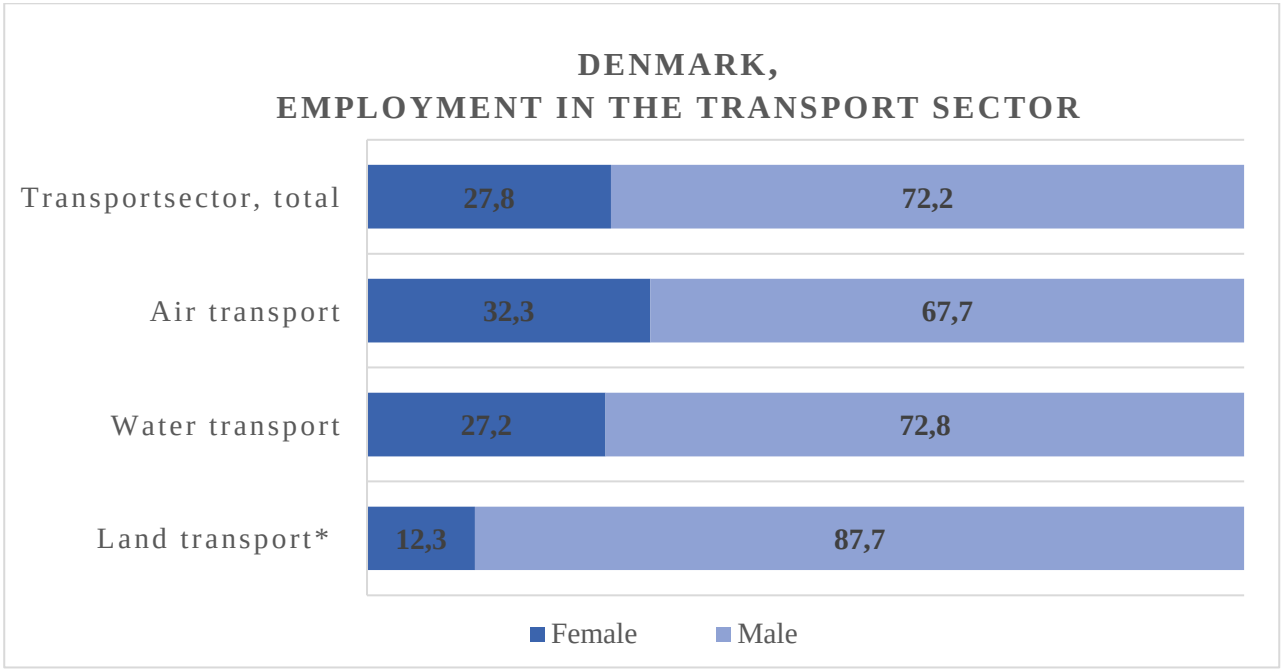
*Denmark*

In the following we will treat the representation of "Transport and employment" in a Danish context. We know from earlier TINNGO deliverable (Deliverable 9.1) that transport is a male dominated work field in Denmark that consist of consisted of 72,2 % men and 27,8 % women [14]. We therefore expect the representation of transport and employment to be masculine, and we will with the visual and

[14] Eurostat Database, Revisited March 14th 2019 (https://ec.europa.eu/eurostat/web/lfs/data/database).

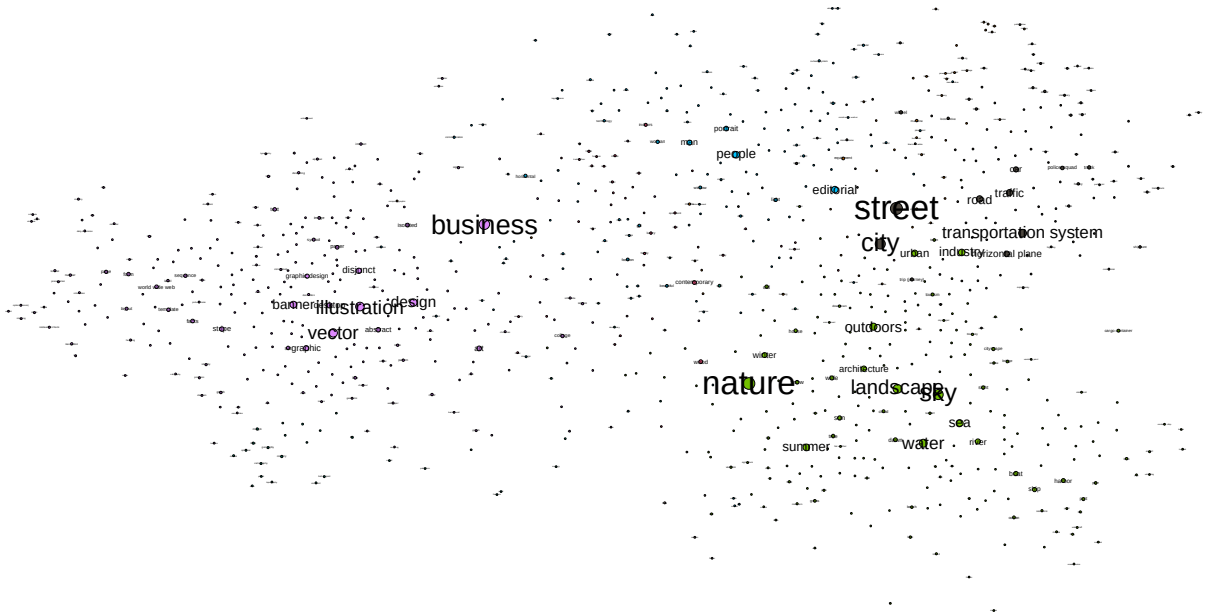
explorative method try to discover contradicting indicators, as well as to describe how the representation is masculine.

Figure 12: Employment in the Transport Sector in Denmark. Published in Deliverable 9.1. Data from Eurostat.




The network illustrated on the following page can be divided into following nine clusters: Purple cluster (30 %), green cluster (23 %), blue cluster (17 %), black cluster (13 %), orange cluster (6 %), pink cluster (6 %), turquoise cluster (3 %), rosa cluster (2 %), and grey cluster (<1 %). The two biggest clusters, purple and green, are placed in each their side, and show more divided clusters, than were found in the search on “smart car and passenger”.

Figure 13: Transport and employment, Denmark




The purple cluster (30 %) contains tags e.g. illustration, design, desktop, wallpaper, layout, calendar, data, business, isolated, money, document. This can overall be considered as business, meeting schedules, and graphic design. This relates to representation in an **office**, which can reflect technical, service or management positions. The green cluster (23 %) contains tags that represent **nature and the construction sector.** This includes tags, such as nature, landscape, sea, cityscape, industry, cargo container, crane bird, crane bird etc. This reflects in a higher degree operating positions. The blue cluster (17 %) contains tags more related to **people**. There is a mixture of context-tags, but include both event, couple, technology, office, smile, helmet, safety, bike, and warehouse. The black cluster (13 %) contains a mixture of tags including: transportation system, car, road, asphalt, logistics, trailer truck, platform, tramway, engine, cable car, and locomotive. This indicates representation of different kinds of both **public, private and industry land transport.** The orange cluster (6 %) contains tags mostly related to **air transport** such as jet, air, airplane, airport, flight, and military. The pink cluster (6 %) contains tags that relate to **private and indoors** including tags as: indoors, table, furniture, hotel, room, lamp, luxury, wood, apartment, family. The turquoise cluster (3 %) distinguishes by having an **international** twist, e.g. map, geography, country, atlas, united, flag. The rosa cluster (2 %) contains tags, that after a qualitative inspection have little to do with representation of transport

and employment. This includes tags such as **flower,** season, flora, and food. The grey cluster (<1 %) tags are not apparent on the network.

We see that the highest occurred tags, which have not been excluded to show a clearer pattern, are “business”, “street”, and “nature” which form a triangle. The biggest cluster (30 %) included the tag ”Business”, which were connected to office-related employment or representations as this included tags as illustration, vector, and design. Tags connected to ”nature” were often related to construction related tags and tags connected to ”street” were related to cars and tramways. Neither of these three clusters showed tags of persons in ”general concepts”. Persons were nested in a separate cluster, and not especially connected to either of the clusters. This could indicate a more impersonal representation.

### *Romania*

Through a search on *transport and employment* in a Romanian context 500 images were collected and tagged with the Clarifai recognition tool. The figure below illustrates eight clusters, that were identified using the Louvain modularity algorithm and force atlas 2 layout. Furthermore, the degree range function was used to filter the tags that most often occurred together to visualize a clearer pattern in data. The clusters are composed of the red cluster (23 % of the total tags), the green cluster (22 % of the total tags), the blue cluster (18 % of the total tags), as well as five clusters that each form 10 % or less of the total tags.

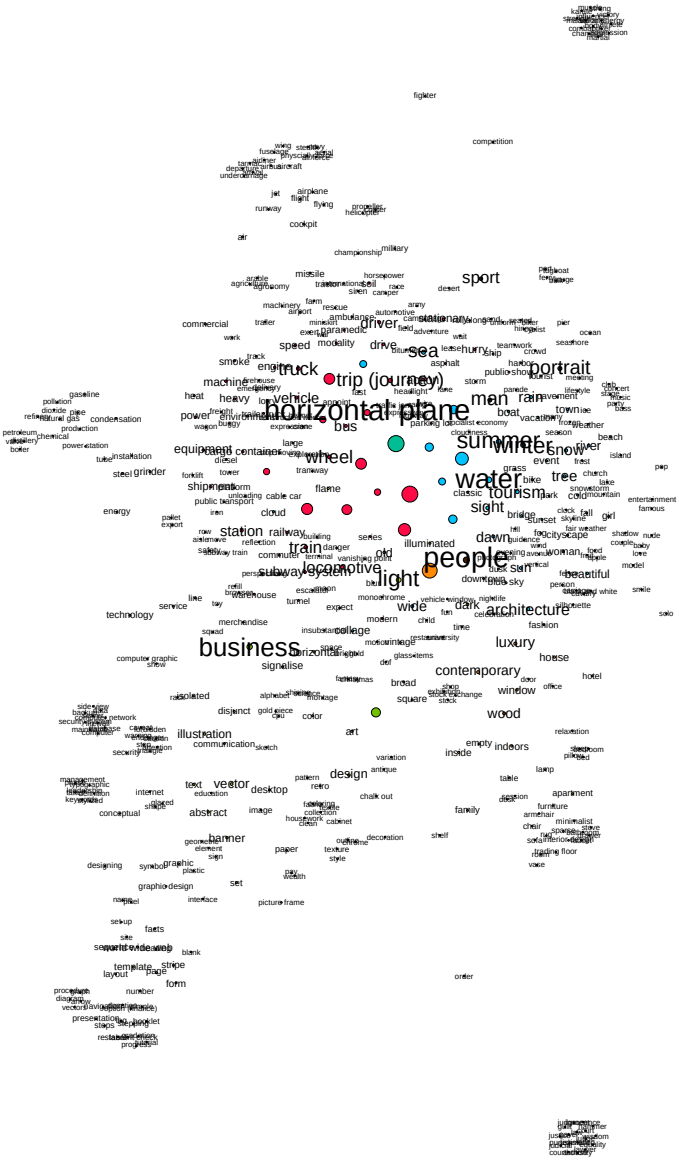



*Figure 10: Transport and employment, Romania*

The red cluster (23 %) contains tags related to **land transport,** and include tags e.g. such as driver, exploration, vehicle, train, action, speed, ambulance, rescue, asphalt, horsepower, safety, tractor, work, and freight. A qualitative inspection of the images that were attached the spoken of tags, showed that many of the tagged images represented different professions and the transport of these rather than public or private transport. These findings included both ambulances, farming, mechanic, police, and trash collector

*Examples of images, that were tagged with "drive"*

The green cluster (22 %) contains tags related to **representation and documents** and includes tags such as graphic, name, facts, set up, world wide web, presentation, booklet, tutorial, form, security, and style. The blue cluster (18 % of the tags) is more interwoven with the red cluster. It contains tags related to **everyday transport**, e.g. tourism, grass, bike, cityscape, wind, and boat. A qualitative inspection shows that there are represented some accidents among the images and much of what is tagged tourism is regular bus routes.

The black cluster (10 %) in the upper right corner is connected by the tag "competition". A qualitative inspection shows, that especially one image dominates the cluster, one of a man in combat. There are many tags attached to this one image that unfortunately do not represent the theme well. On the right you see the image, where many of the following tags are represented: influencer, karate, strong, champion, victory, athlete, combat, body, martial arts, boxer, aggression, muscle, energy, strength. A closer look at some of the other images represented in the cluster shows that a better description from the cluster would **energy production** as images of refinery or power plant.

*Dominating image behind tags, black cluster*

The pink cluster (6 %) contained different tags, but the content was much similar to what was included in the black cluster. Overall, we can call the cluster **tubes and installation,** as it includes tags such as installation, tube, chemical, grinder, equipment, tower, environment, equipment, power, refinery, smoke, pipe, production, pollution, dioxide, natural gas, installation.

*image with "installation"-tag*

*image with "installation"-tag*

The orange cluster (9 %) contains tags related to **office and logistics** such as house, window, door, office, hotel, indoors, inside, empty, table, lamp, relaxation, desk, session, trading floor, sofa, interior design, apartment, office. The turquoise cluster (6 %) contains tags related to **air transport**, e.g. arrival, aircraft, flight, wing, aerial, stealth, flying, physically force, fuselage, runway, airbus. The yellow cluster (3 %) is very interwoven with the other clusters and related to **emergency** as a qualitative inspection showed many images of car lights, fire trucks and ambulances. The cluster

contains tags, such as flame, dark, dusk, nightlife, evening, fun, celebration, insubstantial, space, bright, Christmas, fantasy, motion, science, shining, and moon. The final cluster, the brown cluster in the right bottom corner (3 %), contains tags all related to one image of a hammer in a court house. The many tags such as innocence, judgment, court, law, wisdom, judge, legislation, equality, lawyer, and authority are all related to this images.

*Sweden*

In the Swedish case, the visualization of transport and employment contains tags that divide into nine clusters. The tags that occurred more than 160 times (314 originally) or less than two times were left out in order to see a clearer pattern. The biggest clusters are the green cluster *"Layout and documents"* (25 %), hereafter comes the blue cluster *"City, fast and slow-moving traffic"* (23 %), the orange cluster *"Persons and varied transport"* (14 %), the red cluster *"Emergency, collective transport, trucks and vans"*(12 %), the turquoise cluster *"Agriculture, industry and construction"* (10 %), the purple cluster *"Containers and shipping"* (5 %), the black cluster *"Trains and busses"* (5 %), the pink cluster *"Air transport"* (5 %), and the yellow cluster (>0,1 %). The modularity score was 0,41 before the filter was added and 0,52 after.

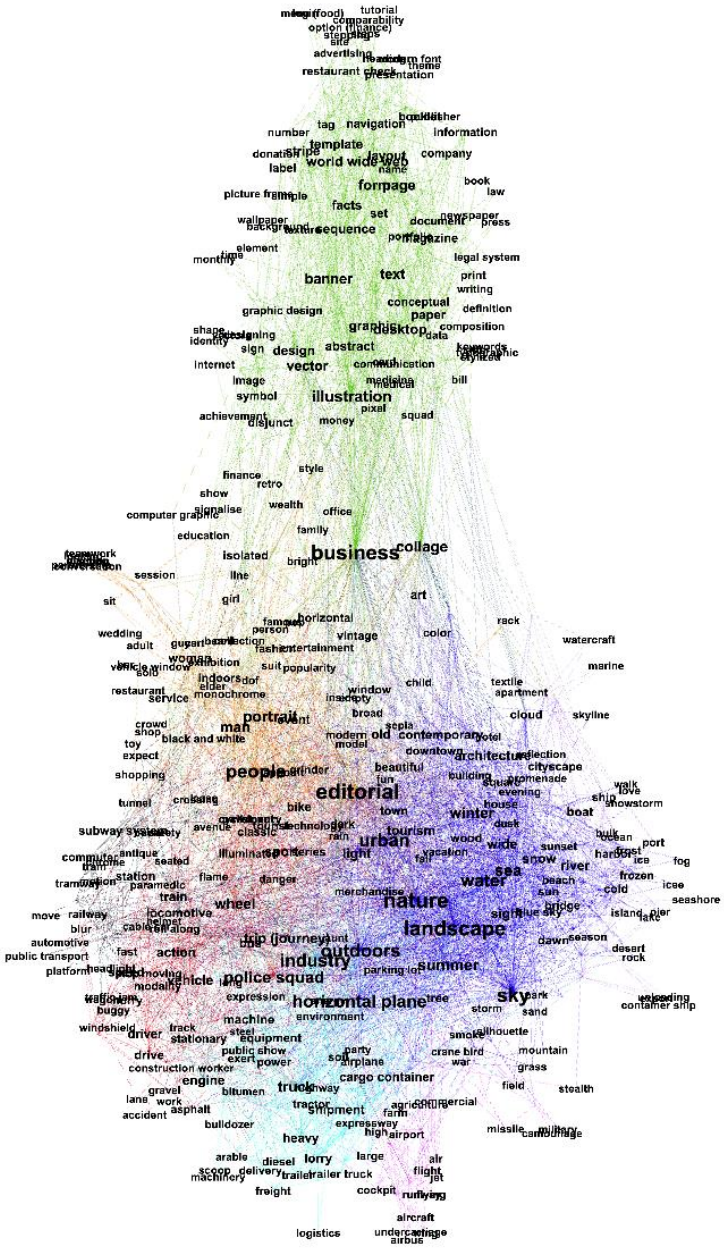



*Figure 14: Transport and employment, Sweden*

We have looked for concepts mirroring female and male employees and/or employment in relation to transport. The concepts ‘women’ and ‘girl’ are found in the outskirts of the **orange cluster**. This means that there are few connections between them and other concepts within the orange cluster and also weak connections to other clusters. Some examples of connections for the concept of ‘women’ are indoors, restaurant, service, solo, vehicle, wedding, fashion, entertainment, vintage, famous person, exhibition, adult and elder. There are some connections with the **green and red clusters**.

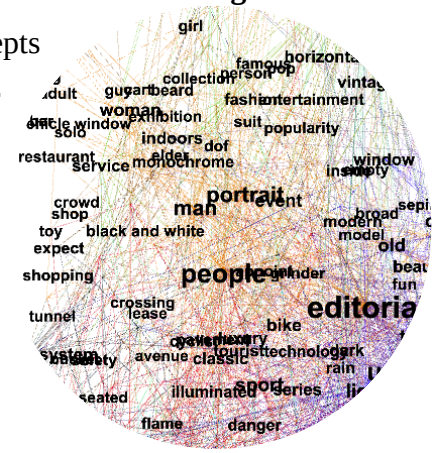



Figure 15: Transport and employment, Sweden; orange cluster

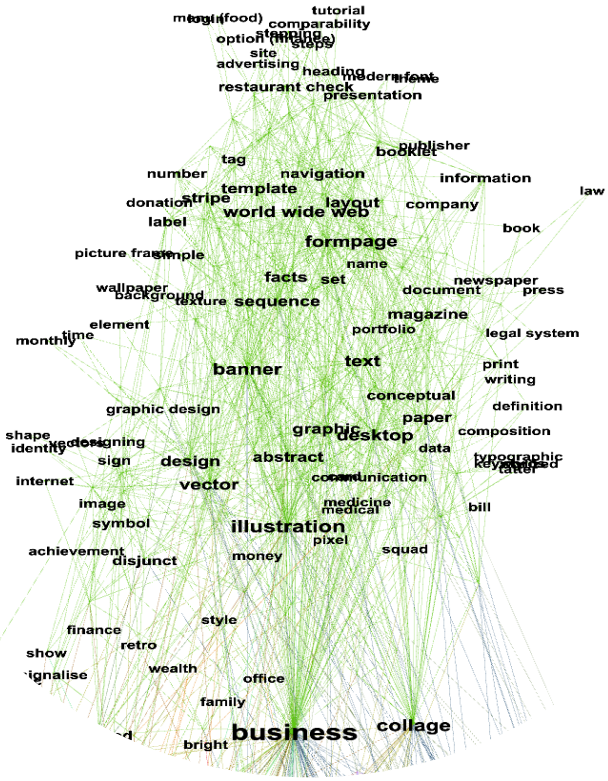



Figure 16: Transport and employment, Sweden; green cluster

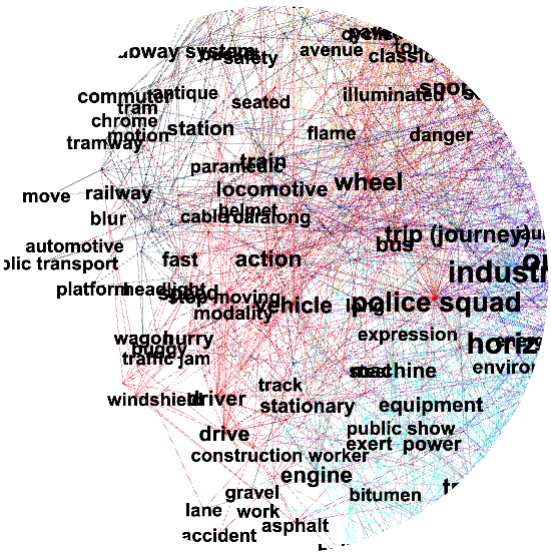



Figure 17: Transport and employment, Sweden; red cluster

For example, education, family, isolated, office, wealth, in the green cluster; and black and white, crowd shop, man, people and shopping in the red cluster. We found some interesting connections

following the concept ‘women’ a bit further into the **red cluster** the line of related concepts. There occurs a line of concepts linked from ‘women’ to ‘public transport’, i.e. concepts on public transport modes and public transport environments:

Tunnel → commuter → subway → tram → tramway → railway → automotive → public transport → platform.

These connections are placed in the left brink of the **orange-red cluster**, rather peripheral in the figure.

The portraying of female employees and/or employment in relation to transportation is to some extent stereotyped, but to some extent it might also be the reflection of working life and women’s travel to work. Women are often referred to public transport for commuting, which is visible in reports published by transport authorities (e.g. Svensk kollektivtrafik, 2019; Transport Analysis 2015a). According to these reports, the highest proportion of travellers in regional public transport are young people (6–17 years) counting 20 percent of the number of journeys. This is followed by women of working age (18–64 years), counting 14 percent. Men have increased their travel by public transport the past ten years and do about 11 percent of the journeys. Thus, the difference between women and men is not that big (Transport Analysis 2015b).

Figures of commuting in Sweden show that men have geographically larger labour markets than women. This may mean that women have less opportunity than men to find jobs that match their skills and experience, which in turn affects their salary development and ultimately the opportunity for financial equality. These gendered differences in the access to labour market are reflected in men’s and women’s travel patterns. Men travel larger distances to work and spend more travel time on matters related to business or education, while women spend more travel time on matters related to service or purchasing. Women in Sweden often work with caring professions, e.g. health care, care for children and older people, and shop selling which is often located locally (SCB 2020). Further, women do trip chaining to larger extent than men, i.e. on their way to/from work doing stops for shopping, accompanying children to school and leisure activities. Women’s travelling reflect (generally spoken) their responsibility for unpaid homework (Transport Analysis 2015a, cf. Dobbs 2007).

Looking into the **red cluster** we see that the concept ‘man’ is closer connected to concepts of danger, modern, technology, wheel. The concept ‘man’ is also more than women connected to the concepts in the **blue and turquoise clusters,** e.g. industry, heavy, horizontal plane, landscape, lorry, machine, nature, outdoors, tractor. Connection lines are also found to airplane, cargo, and container and further into the **pink cluster**. These connections probably reflect the male dominance in transport work and the representation of men in traditional male professions, such as piloting, heavy transport, technology, and the male identity connected to power and machines (cf. Mellström 2017).

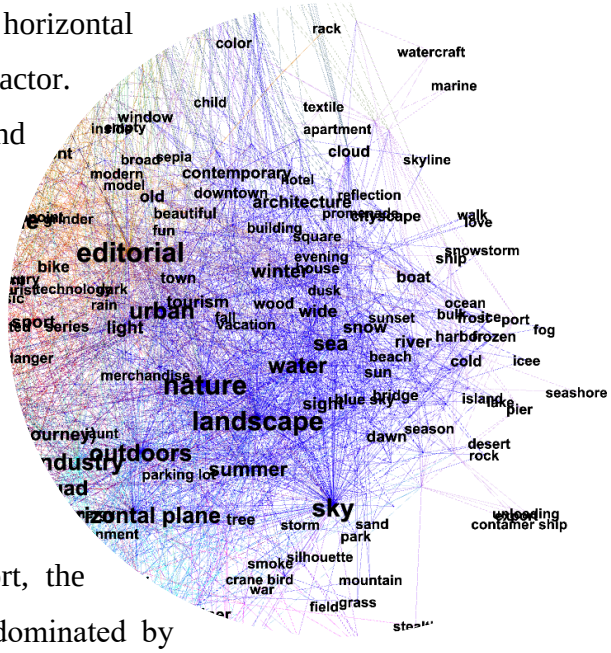



Figure 18: Transport and employment, Sweden; red cluster

According to gender research in the field of transport, the transport sector’s methods and planning process are dominated by masculine attributes that consider technical and economic rationality and a view of nature and human life as separate systems. Furthermore, the car is seen as a masculine attribute (Balkmar & Mellström 2018, 2020; Dahl, Henriksson & Levin 2012). Feminine attributes are focusing on health and care, lifestyles, and cultural values, and those are included to a smaller extent. Furthermore, equal power over the problem formulation in transport planning is important as shown in recent studies (Kronsell et al. 2016, 2020). Different gender-related approaches can affect whether the focus is on increased transport efficiency rather than travel opportunities for all, accessibility, and trip chaining. There are also criticisms of cost-benefit analyses, which are considered rather designed to respond to masculine attributes over feminine such as long-term sustainability.

The **green cluster** is more difficult to analyse since is rather isolated from the other clusters. Many of the concepts relate to media, graphic design, information, knowledge and editing and concepts are gathered in bouquet-like patterns joined by the bold (strong) concepts business and collage. It could be interesting to do more qualitative analysis of images that occur in this cluster. This might be an issue for future research.

The Swedish case shows for the concept of woman, connections to e.g. ”indoors”, ”restaurant”, ”service”, ”wedding”, ”fashion”, ”entertainment”, ”vintage”, ”famous person”, and ”exhibition”; meanwhile the connections for the concept of man are e.g. ”danger”, “modern”, “technology”, “wheel”, ”industry”, ”heavy”, ”landscape”, ”lorry”, and ”machine”. This reflects a rather traditional division between women and men in the labour market, which is especially valid for the transport sector in Sweden (SCB 2020).

## Transport and employment – trends in visualizations across Europe

Table 7: Comparison of clusters representing “Transport and employment”

| **Denmark** | **Romania** | **Sweden** |
|---|---|---|
| Office (30 %) | Land transport - professions (23 %) | Layout and documents (25 %) |
| Nature and the construction sector (23 %) | Representation and documents (22 %) | City, fast and slow-moving traffic (23 %) |
| People (17 %) | Everyday transport (18 %) | Persons and varied transport (14 %) |
| Public, private, and industry land transport (13 %) | Energy production (10 %) | Emergency, collective transport, trucks and vans (12 %) |
| Air transport (6 %) | Office and logistics (9 %) | Agriculture, industry, and construction (10 %) |
| Private and indoors (6 %) | Air transport (6 %) | Containers and shipping (5 %) |
| International (3 %) | Tubes and installation (6 %) | Trains and busses (5 %) |
| Flower (2 %) | Courthouse (3 %) | Air transport (5 %) |
| Not apparent (<1 %) | - | Not apparent (<0,1 %) |

In table 7, we see the clusters representing Transport and employment in Denmark, Romania, and Sweden. Romania stands out in the way that much land transport represented are different professions, such as ambulances, and the transport of these rather than public / private transport. A

qualitative inspection shows that Sweden stands out in representing more diverse forms of both fast- and slow-moving traffic in the images. More similarities were found between the networks. All had a minor separate cluster (5-6 %) representing air transport. This is the sector in transport in which most women in Denmark are employed. All networks included business related/ Office/ layout and documents / representation and documents-cluster that were relatively big, forming between 22-30 % of the tags.

The Swedish case shows that for the concept of 'woman', connections were found to e.g. "indoors", "restaurant", "service", "wedding", "fashion", "entertainment", "vintage", "famous person", and "exhibition." Meanwhile, the connections for the concept of 'man' are e.g. "danger", "modern", "technology", "wheel", "industry", "heavy", "landscape", "lorry", and "machine". This reflects a rather traditional division between women and men in the labour market, which is valid for the transport sector in Sweden (see explanations in the Swedish case).

The Romanian case represented different professions and the transport of these, e.g. ambulances, trash collector, mechanics, police or farming. More tags in this cluster was with a qualitative inspection shown to represent images of both women and men at the wheel which requires deeper analysis e.g. of the historical development and the equality tradition from the socialist period is at stake.

In the Danish case, we see that the highest occurred tags, which have not been excluded to show a clearer pattern, are "business", "street", and "nature," which forms a triangle. The biggest cluster (30 %) included the tag "Business", which was connected to office-related employment or representations, as this included tags of illustration, vector and design. Tags connected to "nature" were often related to construction related tags and tags connected to "street" were related to cars and tramways. Neither of these three clusters showed tags of persons in "general concepts". Persons were nested in a separate cluster, and not especially connected to either of the clusters. This could indicate a more impersonal representation.

Concluding

The TInnGO digital analysis focuses on s the visual representations, which average European citizens – in and across countries – encounter when searching for web-based information on smart transport as well as transport employment. This report has exploratively analysed contextualized representations of 'smart car and passenger' and 'transport and employment' in Denmark, Sweden, Romania, Portugal and UK. In the visualized clusters, and in the qualitative inspections, we found both similarities and differences between the representations in the respective countries.

On the representation of *Smart car and passenger* more networks included some less clear clusters and a low modularity score. This finding suggests that the representations of 'smart car and passenger' are in dialogue and not separate or isolated communities. The majority of the images representing 'smart car and passenger' in Portugal, Denmark, Romania, and UK are in general not represented with persons. A qualitative inspection of the Romanian images behind the tags showed that though both male and female car users were present in some images, it was still technology, "who was at the wheel". Though it was not the general picture of the cluster, a qualitative inspection from the Danish material showed that there were examples of both men and women as actors and users of technology, without one being subordinated to the other. UK explored explicitly feminine and masculine tags and no immediate gendered patterns were found as both genders were represented across all clusters. A closer inspection is though recommended as the preliminary qualitative inspection found that "person" was identified well by the recognition tool, but problems in identifying male and female still consisted. A qualitative inspection of the Portuguese material showed that more images, especially in the cluster "Landscape / tourism", were not related to smart cars. The lack of images could signify that smart cars is still not a subject that raises interest to the majority of the population in Portugal.

The representation of *Transport and employment* showed clearer clusters representing some of the same themes, e.g. a bigger cluster of graphics, or tags relating to documents, representation and office, a bigger cluster related to city and nature, and a smaller cluster related to air transport. The material from the representation of transport and employment in UK showed that gender, race, and age were infrequently shown in the images. When examining images behind feminine and masculine tags, it appeared that women were mainly shown as passengers on public transport and riding bikes and not as transport workers. The data from Portugal showed Clarifai demographic

tags, that were quite similar for men and women: 4 % were tagged female and 3,2 % were tagged male. The biggest clusters, and thereby the clusters that are more represented in Transport and employment in a Portuguese context, are the groups of supplies transport by road mode and passenger transport, mostly associated with tourism. Romania stands out in the way that much land transport represented are different professions, such as ambulances, and the transport of these rather than public/private transport. Moving to the Swedish case, a qualitative inspection shows that Sweden stands out in representing more diverse forms of both fast- and slow-moving traffic in the images. We found more similarities were between the networks as both Romania, Sweden, and Denmark had networks included an Office/ layout and documents / representation and documents-cluster that were relatively big forming between 22-30 % of the tags.

## The TInnGO digital analysis: Ways forward.

The Digital method enables us to work with a large amount of data in an explorative manner. Media archives carry many images, which represent transport related themes and discourses throughout Europe and at a global scale. Images are powerful . The Clarifai recognition tool, and the visualization programme Gephi makes it possible to work with these data both with qualitative (close) readings, and quantitative (distant) patterns. This report has introduced TInnGO-partners to the method, and the intuitive and user-friendly programmes that can be accessed for free. The digital method offers a new approach to transport studies that can be conducted without great coding skills and which opens up for a large amount of data to be taken under the scope/loop.

This type of digital analysis might have wider potentials for the provision of critical accounts of existing visual digital representations and patterns, which are often taken for granted or seen as "innocent " noise compared with the analysis of texts and words. The significance has been to provide knowledge of patterns of connections and visual clusters of representations, which might inform our perceptions, behaviours and practices. ( Christensen 2014) Enhanced critical awareness and new and more balanced visualizations, could be gainful for both for the development of democratic policy making as well as for higher market potentials. [15] At the same time it must be considered whether the price of stepping up "fairness" in digital visualization including the provision of more sensitive API image tagging platforms and image scraping tools, might come as a side product of enhanced surveillance and control.

[15] See also Hoffmann 2019

## Additional literature – inspiration for working with digital analysis

For an introduction to Gephi use these tutorials:

- https://medium.com/@EthnographicMachines/visual-network-analysis-with-gephi-d6241127a336
- martingrandjean.ch/gephi-introduction

A concrete case on how to read and understand visual networks, written by the digital supervisor team from Aalborg University:
Elgaard Jensen, T., Kleberg Hansen, A. K., Ulijaszek, S., Munk, A. K., Madsen, A. K., Hillersdal, L., & Jespersen, A. P. (2019). Identifying notions of environment in obesity research using a mixed-methods approach. *Obesity Reviews*, *20*(4), 621-630.

TInnGO digital – see guidelines in enclosed powerpoint presentation and process guide. All files for the workshop can be found here:
https://docs.google.com/presentation/d/1Q4LPwwrxUt6b3aldqJzB9YTK1cSn7o1ySrdKkXuJvpc/edit?usp=sharing

You can read more about Force2Atlas layout here:
Jacomy, M., Venturini, T., Heymann, S., & Bastian, M. (2014). ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PloS one*, *9*(6).

Richard Rogers, fra International Policy Review. What we are doing can be described by ”digital methods” on page 75 in the lower right corner:
https://pure.uva.nl/ws/files/2321162/156893_Rogers_R..pdf

A comparative study of image tagging algorithms (here among Clarifai) and human tagging:
https://dl.acm.org/doi/pdf/10.1145/3320435.3320442

A benchmarking of the various commercial algorithmic image taggers (including Clarifai):
https://www.aaai.org/ojs/index.php/ICWSM/article/view/3232

A chapter that describes how we use Google as a research tool. By Richard Rogers:
https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118900772.etrds0076

# Bibliography

Agência Portuguesa do Ambiente (APA). (2019). Relatório do estado do ambiente em Portugal. Retrived from: https://rea.apambiente.pt/

Amacker, M., Isabelle Schlaepfer,I.,Bigler, C, & Graf. A. ( 2017): The development of gender-responsive indicators: towards a participatory approach in *Women, Gender and Research* 1. 2017 .https://tidsskrift.dk/KKF/issue/view/6751

Balkmar, D. and Mellström, U. (2020). Towards an anthropology of transport affect: the place of emotions, gender, and power in smart mobilities. In Priya Uteng, T., Rømer Christensen, H. and Levin, L 2020. *Gendering Smart Mobilities*, pp. 57–74. London and New York: Routledge.

Balkmar, D. and Mellström, U. (2018). Masculinity and autonomous vehicles. *Transfer* 8(1): 44–63.

Breengaard, M., Christensen H.R., Levin, L.: ( 2019). TInnGO EDUCATIONAL MANUAL From Gender-Blind to Gender Smart Mobility. TInnGO D. 4.2. 2019.

Joy Buolamwini, J.& Gebrum T: (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Proceedings of Machine Learning Research 81:1–15, 2018 Conference on Fairness, Accountability, and Transparency http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf

Christensen, H.R ( 2014). Citizenship as Communicative Events: Visual politics of The European Year of Equal opportunities for All. (FEMCIT EU FP 6/ University of Copenhagen. Research paper.)

Christensen, H.R & Breengaard, M.H (2019). Gender Smart Transport. Translating theory into practice. TInnGO Road Map. ( Innovation Gender Observatory mobility. EU Horizon 2020 project) forthcoming 2020/Routledge.

Christensen, H.R; Henningsen I & Steffensen,T (eds) (2017). Quantitative methods, big data a gender. Special issue of *Women, Gender and Research* 1. 2017 pp 1-90. https://tidsskrift.dk/KKF/issue/view/6751

Conceição, O., Costa, V., (2017). Entrepreneurship in Tourism Industry: Evidence of Business dynamics in Portugal. Turismo & Desenvolvimento. Nº 27/28, 35-50. Available from: https://pdfs.semanticscholar.org/1ed6/417e2bc5394f8e3110028a7b92f451cfb67c.pdf (Accessed 27th March 2020)

Dahl, E., Henriksson, M. and Levin, L. (2012). Konstruktioner av maskulinitet i samtal om kollektivtrafik. [Constructions of masculinities in conversations about public transport]. *NORMA* 2012, 7(2): 160–181.

Dant, T: (2004). The Driver-Car, in *Theory, Culture and Society*. Special issue on Automobilities: 71-79

D'Ignazio, C & Klein, L F. ( 2020). *Data Feminism.* The MIT Press 2020 https://data-feminism.mitpress.mit.edu/

Dobbs, L. (2007). Stuck in the slow lane: reconceptualizing the links between gender, transport and employment. *Gender, Work and Organization*, 14(2): 85–108.

Ferreira, V., (2010). A Igualdade de Mulheres e Homens no Trabalho e no Emprego em Portugal: Políticas e Circunstâncias. Comissão para a Igualdade no Trabalho e no Emprego, Lisboa 2010. Available from http://cite.gov.pt/asstscite/downloads/publics/Igualdade_CITE_NET.pdf (Accessed 30th March 2020)

Hildebrand, J.M. & Scheller, M. (2018). Media Ecologies of Autonomous Automobility: Gendered and Racial Dimensions of Future Concept Cars. *Transfers*. 8(1).

Hoffmann. A.L. ( 2019). Where fairness fails: data, algorithms, and the limits of antidiscrimination discourse. In *Information, Communication, & Society*. Volume 22, 2019 900-915. https://doi.org/10.1080/1369118X.2019.1573912

Hughes. C and Cohen . R.L. (eds.) (2012). *Feminism counts: quantitative methods and researching gender*. Abingdon: Routledge.

Jacomy, M., Venturini,T.,Heymann, S., Bastian, M. (2014). ForceAtlas2, a Continuous Graph Layout Algorithm for Handy Network Visualization Designed for the Gephi Software. *Plos ONE*, 2014 9(6). DOI: 10.1371/journal.pone.0098679

Kronsell, A., Dymén. C., Smidfelt Rosqvist, L. and Winslott Hiselius, L. (2020). Masculinities and femininities in sustainable transport policy: a focus on Swedish municipalities. *NORMA: International Journal for Masculinities Studies*, 1/2020: 1–17. DOI: 10.1080/18902138.2020.1714315

Kronsell, A., Smidfelt Rosqvist, L. and Winslott Hiselius, L. (2016): Achieving climate objectives in transport policy by including women and challenging gender norms: The Swedish case, *International Journal of Sustainable Transportation*, 10(8): 703–711, DOI: 10.1080/15568318.2015.1129653.

Kyriakou, K., Barlas, P., Kleanthous, S., & Otterbacher, J. (2019). Fairness in Proprietary Image Tagging Algorithms: A Cross-Platform Audit on People Images. *Proceedings of the International AAAI Conference on Web and Social Media*, *13*(01), 313-322. https://www.aaai.org/ojs/index.php/ICWSM/article/view/3232

Lovink, Geert (2012). Networks Without a Cause, A Critique of Social Media, Cambridge, Polity Press.

McCall. L: (2005). The complexity of intersectionality. In *Signs*, Vol. 30, no. 3: 1771-1800.

Mellström, U. (2017). Masculinity, Power and Technology. London: Routledge.

SCB. (2020). På tal om kvinnor och män. [Speaking of women and men]. https://www.scb.se/contentassets/98bad96ec1bd44aea171b2fb4f7b3f64/le0201_2019b20_br_x10br2001.pdf

Munk, A,K. (2019). Four Styles of Quali-Quantitative Analysis. *Nordicom Review; Gothenburg* Vol. 40, Iss. s1, (2019): 159-176.: 159) DOI:10.2478/nor-2019-0020

Oakley , A (1998). Gender, methodology and people's ways of knowing: some problems with feminism and the paradigm debate in social science. In *Sociology*, Vol. 32, no. 4: 707-731.

Oldenziel, R. (1999). *Making Technology Masculine: Men, Women and Modern Machines in America*. Amsterdam. Amsterdam University Press

Priem, Groth,J, Camero.P ( 2011). Altmetrics. A manifesto. September 2011.

Rogers. R. ( 2014). Political research in the digital age. Special issue. *International public policy review*. June 2014. Vol. 8. No 1.

Saraiva, A. (2018): Speech of the Pres.dent of the Portuguese Business Confederation (CIP), 107th International Labour Organization Conference, Geneva, [online] Available from: https://www.youtube.com/watch?v=-AW_PB-EJJ0 (Accessed 20th January 2020)

Scharff, V. (1991). *Taking the Wheel. Women and the Coming of the Motor Age*. Albuquerque: University of New Mexico Press SCB.

Woman and men in Sweden. Facts and figures 2020. Available from: https://www.scb.se/en/finding-statistics/statistics-by-subject-area/living-conditions/gender-statistics/gender-statistics/pong/publications/women-and-men-in-sweden2.-facts-and-figures-2020/

Svensk kollektivtrafik. (2019). Kollektivtrafikbarometern. [Public transport barometer]. https://www.svenskkollektivtrafik.se/globalassets/svenskkollektivtrafik/dokument/aktuellt-och-debatt/publikationer/kollektivtrafikbarometern-arsrapport-2019.pdf

Transport Analysis. (2015a). Uppföljning av de transportpolitiska målen [Follow-up of transport policy objectives]. Report 2015:7. https://www.trafa.se/globalassets/rapporter/2010-2015/2015/rapport_2015_7_uppfoeljning_av_de_transportpolitiska_maalen.pdf

Transport Analysis. (2015b). Kollektivtrafikens utveckling – en analys av den nationella statistiken [The development of public transport - an analysis of national statistics]. https://www.trafa.se/globalassets/rapporter/2010-2015/2015/rapport-2015_15-kollektivtrafikens-utveckling---en-analys-av-den-nationella-statistiken1.pdf

Thylstrup,N., & Veel, K: (2017). Data visualization from a feminist perspective: Interview with Catherine D'Ignazio in *Women, Gender and Research* 1. 2017. https://tidsskrift.dk/KKF/issue/view/6751

# Annex I

## The TInnGO project [16]

The TInnGO project meets the need to address smart mobility in innovative ways in order to improve the access of mobility for all. In so doing, the TInnGO project addresses transport related areas, such as production, consumption and governance as integrated in the notion of *Gender Smart Mobility.* TInnGO aims to:

- Sharpen the critical analysis of gender-blind approaches in the planning, production and policy-making of Smart Transport as well as in the usage and consumption of smart transport services.
- Develop proactive Gender Smart approaches to mobility governances.
- Feed into new and visionary examples and approaches of Gender Smart Mobility.

The TInnGO project will apply the broad, inclusive and dynamic concept of *Gender Smart Mobility,* that will be elaborated and advanced during the project.[17] The key concepts of smart transport should meet the following requirements:

INCLUSIVE TRANSPORTATION: Smart transport systems should address various groups of citizens (men, women, gender-neutral, old, young, child, ethnic) in non-stereotyping ways from start to end. This includes both technical and non-technical issues from the inception of ideas to the end-products. It means that various groups must be included in processes of design, accessibility, safety, communication and marketing, living labs, and end-products.
AFFORDABLE TRANSPORTATION: Public and public private investments should be addressing robust and stable public transit provisions. E.g. investments supporting the innovation of smart

[16] https://www.TInnGO.eu/

[17] This (normative) concept of gender smart mobility as broad, inclusive and dynamic is inspired by a merge of various perspectives found in Marsden & Reardon 2017; Lyons 2018; Gendered innovation project: https://genderedinnovations.stanford.edu/

small cars for all rather than luxury cars for the few. Keep in mind, the gender pay gap and that women in general are less resourceful than men.

EFFECTIVE TRANSPORTATION: Seamless transport should be provided for all. Also, smart mobility provisions should include smart biking and walking. Market stakeholders are required to produce smart and efficient public transport devices rather than smart luxury cars for individual use.

ATTRACTIVE TRANSPORTATION: Transport planning should provide safe, accessible and liveable spaces in all parts of the cities. Smart solutions for shared transport and various non-motorized transport modalities for more, broader and diverse groups of people.

SUSTAINABLE TRANSPORTATION: Non-motorized transport should be included in smart transport ideas and practices. Actions should be taken in order to motivate and to socialize different groups of citizens to prefer non-motorized modes of transport for the last mile transit. E-bikes and shared sustainable forms of transport should be cheap and accessible for all, including parents with kids, people who travel with luggage or goods, elderly people, people with disabilities.

# Annex II

## A step-by-step guide

In the following you will find a step-by-step guide to the conduction of the digital media analysis, hereunder the shared and comparative framework for collecting data, a guide to google trends, creating an API-key, attaching tags to the images, and generating a file that can be opened and analysed in Gephi.

### Digital analysis design: Shared and comparative framework

The aim of the TInnGO analysis is to provide for a critical analysis of the "google world of images". Below you find the updated aims of the analysis, as well as a description of the changed conduct of the analysis to a meta-analysis.

Aims of the digital analysis

*On the analytical and methodological level the aims were:*

a. To explore an analysis of new types of (visual) digital material aiming at exploring emerging smart mobilities and gendered and stereotyped discourses of (STEM) entrepreneurship and transport employees.
b. To launch methods where quantitative (distant) and qualitative (close) readings of the images and visuals could be explored as a new method for transport and mobility analysis.

*On the technical level the aims were:*

c. To refine the comparative dimensions, to use google Trends in order to find relevant national specific as well as comparative search queries.
d. To combine the search queries with the tools: The google image scarper, the clarifai recognition /tagging data service and the visualizing Gephi tool.

## Google Trends: How to find relevant search queries for the network analysis

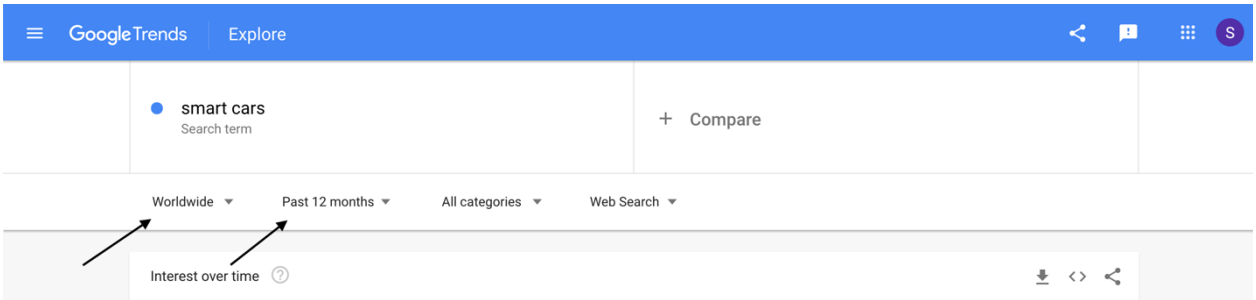
We need to find the exact search words, that each partner needs to type in their search queries. Google Trends can be used for comparative keyword research and provide keyword-related data including search volume index and geographical information about search engine users. This section will guide you to find search queries on the following two tracks for the meta-analysis:

- *"smart car" AND passenger*
- *transport AND employment*

The following example will show how to find the specific search query for the first track (*"smart car" AND passenger*). Google Trends provides access to a largely unfiltered sample of actual search requests made to Google. In this way, Google Trends can help us shed light on what it is Google users are curious about or interested in when they search for words such as transportation and mobility.

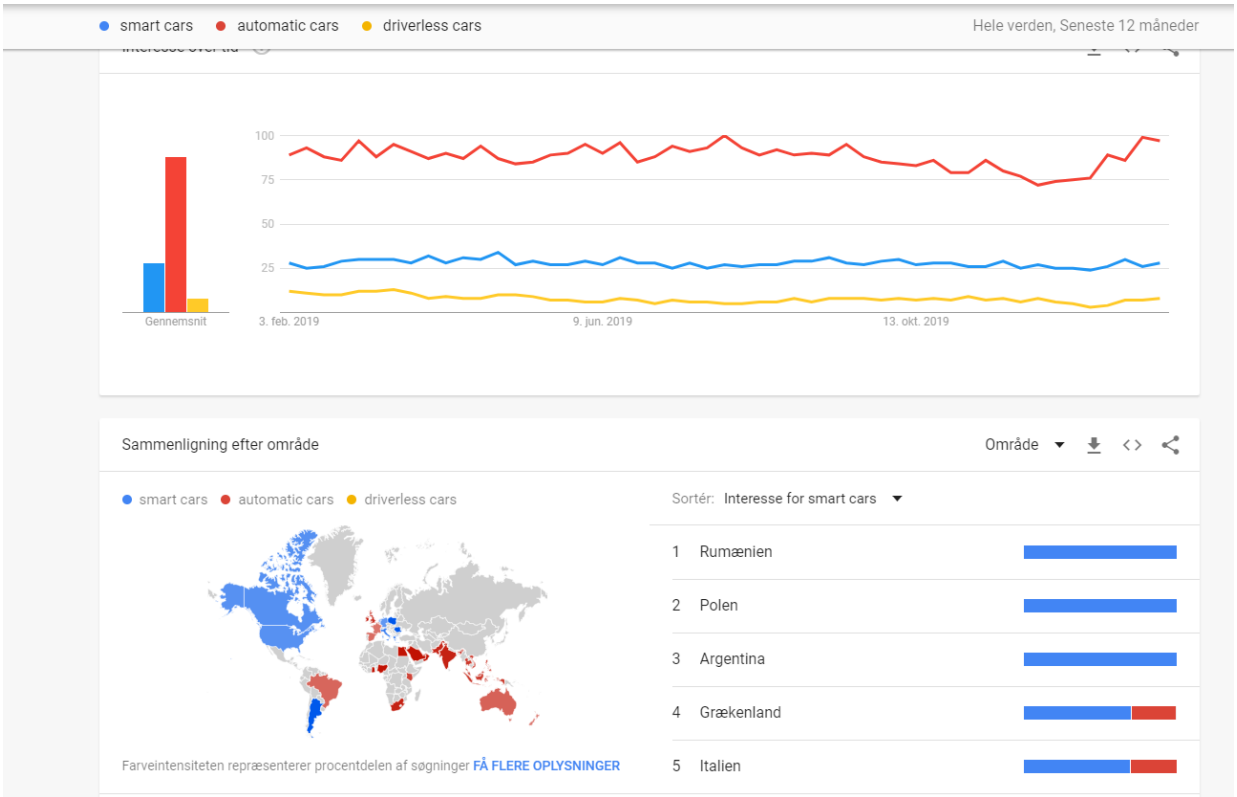
Example: A Danish search query on Smart Cars

To start the search please go to website: https://trends.google.com/trends/ . To describe "smart car" in our search, we have looked into the three terms that nuanced the concept, that were suggested at workshop in October 2019: "smart cars", "automatic cars" and "driverless cars". Once you search for a word, for example "smart cars", the search will appear like this:



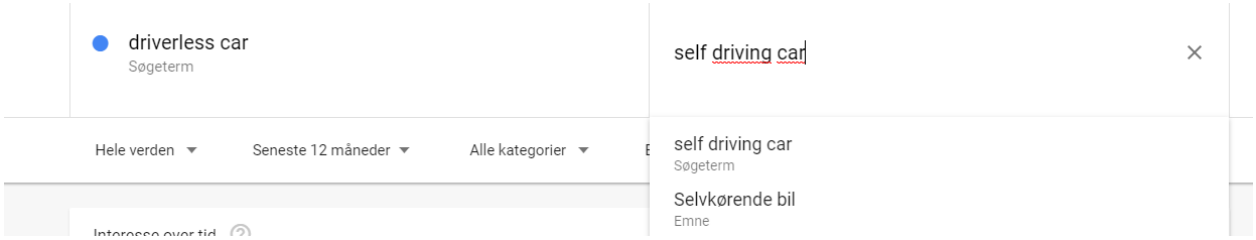

Be sure that the settings for the search is printed to be "Worldwide" and "Past 12 months", as pinpointed with arrows on the above screenshot. There can be big differences in the search queries

on google that people use to find the same concept. Below you see which search queries are often used and where:

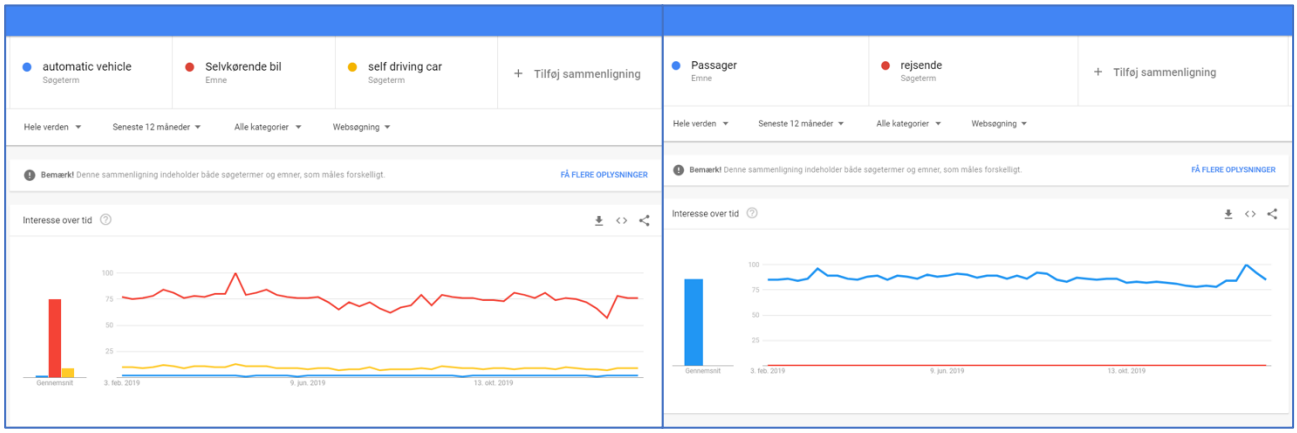



*Language specific search queries*

In order to do a comparative analysis, one can be aware of the importance of following the premises of the media, which requires a language specific search queries in google trends. In short, you have to search for the terms in your own language in Google Trends, but try to find the equivalent word, not only the translated word. On the next page, you see an example with a Danish search. As we type the different terms, google will suggest related terms in Danish. If you have problems with this step in your own language on your search on Google Trends, please check the language preferences in your Google account.




Every time we searched on different terms, the Danish equivalent ‘selvkørende bil’, came up as a relevant alternative suggestion. As you see in the graph below, this term was also the most frequently search query (red colour).




As we would like to have the Danish equivalent to both ”smart car” AND ”passenger”, we have gone through the same steps to examine, which term for “passenger” is the best descriptive and most used term in Denmark. Here we compared the Danish terms “passager” and “rejsende” and found that “passager” (Danish for ‘passenger’) occured much more frequently in the searches.

## Concluding

The TInnGO partners have made search queries on two different tracks that are country-specific in the sense that the search words are not only translated search words, but equivalent search words for “smart car AND passenger” are chosen so that they reflect the language use in their specific context/country.

## Step-by-step guide to generating data

The following guidelines have been introduced and followed by the TInnGO partners. All partners have submitted their search findings to the lead of this task and report, UCPH.

**Step 1: Prepare the Clarifai recognition service:**

- Visit website: https://portal.clarifai.com/
- Sign in with your personal username and password
- You press "Create an application" in the upper right corner
- You click on the application/ app, that you just made
- You are shown a headline that says: "API keys". Below you press "Create new API key".
- You copy the key to clipboard by clicking on the following icon
- Visit website: https://jacomyma.github.io/google-image-dmi-clarifai/
- Paste the API-key in the box, and press "Generate Bookmarklet"
- Pull the bookmarklet manually and attach it to the Bookmarks Bar.

**Step 2: The Search – Generating Data:**

- Visit Google Image Scraper: https://tools.digitalmethods.net/beta/googleImages/
- Type "500" in the box that says, "*Number of Results".* We prefer to have more images to compare.
- Type your search query (see earlier section "Google Trends: How to find relevant search queries for the network analysis" to find the search query relevant for your country).
- Once you have searched, you have to wait until the search is fully downloaded. You know it has finished downloading, when you can scroll down the page, and see the all of the numbered queries listed below.

**Step 3: Tagging data with the Clarifai recognition service**

- When you are happy with your search, you can start tagging the data. Remember that the service is free up to 5000 images a month, so while step 2 was free of counting, the counting starts now. Click on the bookmarklet that you made in step 1. You have to wait while it generates a file.

- *It is important that you do not open the file on your computer.* Instead, you upload the file on Google Docs without opening the file: https://docs.google.com/spreadsheets. Once you have uploaded the file in Google Docs, you can open the csv-file through a spreadsheet on the internet.
- You are now ready to work with the tagged images.

**Step 4: Working with tagged Images**

- You now have a spreadsheet with images, and tagged concepts. You can discover data in many ways, and the principles are similar to working with data in excel. Mathieu Jacomy and Anders Kristian Munk suggested on the workshop November 19th 2019 that you could:
  - Display the images by creating a column, and use the formula ”=IMAGE(E2)” , if E2 is the coordinate of the image ”thumbnail url”.
  - Count the concepts of a row by typing ”=SUM(START:END)” of a chosen area, e.g. ”=SUM(J2:J101)”
  - Count specific terms, by typing ” =COUNTIF(START:END, "*string*") ”, where START and END are the coordinates of the cell and where "*string*" is the term you are counting e.g. =COUNTIF(K2:K200, "*office*")
  - Display filters by clicking on the funnel in the toolbar

There are many ways to discover data, but we will ask you to make two comments on each dataset, that you have extracted. Please write notions of the following two points:

1) Acceptable recognition

We would like to have an idea of the precision and correctness of the concepts that are tagged to describe the images. You are therefore asked to (briefly) compare the pictures with the column “General concepts”. Please take a few minutes to jot down if there are any major discrepancies between the concepts and the images, as well as what the concept describes, and how common the error appears. This will give us opportunity to comment on the acceptability of the network, once we move on to step 6.

2) Description of data

Please write a description of the tagged images based on your work with data in the spreadsheet. The short description shall include statistical frequencies of relevant categories e.g. gender, ethnicity. Please do also make comments on immediate themes and patterns, you have come to think of, when you worked with *Reliability of the recognition tool.* Prominent actors (which websites appeared (frequently), are also of great interest.

**Step 5: Generate a file that can be imported in Gephi**

- We will now create a file that Gephi can read. You visit website Table 2 Net: http://medialab.github.io/table2net/.
- You click ”Upload File”, and here you upload the csv-file, that you have *not opened* on your computer.
- Once you have uploaded the csv-file, you need to manage selections about the file you are creating. One way to go through with this is:
- Choose network: ”Normal”.
- Chose nodes: a) “General concepts” and 2) “comma-separated ‘,’”
- Choose Links a) “Row number”, b) “One expression per cell”
- Additional settings: No weight needed.
- Click “Build the network”
- You have now created a file, that can be uploaded in Gephi.

You have now generated a data set in a format that can be opened, explored and analysed in Gephi. This is an example of generating a file with general concepts. It is also possible to choose other steps at Table 2 Net in order to generate a file that contain other tags, such as demographic ones. If you are more interested in Gephi analysis, you are welcome to look closer at the material suggested by the cooperating researchers at Aalborg University in “Notes”.