Designing a domain-specific language for infographics

Tijmen Derikx (t.m.j.f.derikx@students.uu.nl)
Information and Computing Sciences

First Supervisor: Dr. Sergio España (s.espana@uu.nl)
Second supervisor: Dr. Sietse Overbeek (s.overbeek@uu.nl)

Abstract: Infographics are used for a long time already and their usage has grown more with the rise of social media. Infographics are also produced as the result of a social and environmental accounting. Within the process of creating infographics in the context of social and environmental accounting several pain points are identified. A domain-specific language is designed and engineered to improve the process by tackling the identified pain points. An interpreter of the domain-specific language is implemented in the OpenESEA tool. We validate our proposed domain-specific language by conducting a comparative experiment between original infographics and infographics generated by the OpenESEA tool. In addition, the domain-specific language will be validated by means of an expert assessment.

Keywords: Infographics, Social and Environmental Accounting, Domain-Specific Language, OpenESEA

Table of contents

List of figures

List of tables

1. Introduction

An infographic (short for information graphic) is a type of picture that blends data with design, helping individuals and organizations concisely communicate messages to their audience (see Figure 1) (Thatcher, 2012). An infographic is defined as a visualization of data or ideas that tries to convey complex information to an audience in a manner that can be quickly consumed and easily understood. The process of creating infographics is called data visualization, information design, or information architecture (Smiciklas, 2012). An infographic is suitable to dissect a complex subject and is able to sustain the attention of the reader while doing so. It might seem that infographics are a recent phenomenon growing in conjunction with the growth of the internet. However, we have been using icons, graphics, and pictures throughout history to tell stories, share information, and build knowledge. Many years ago, people drew on walls to communicate with one another which is essentially the earliest form of an infographic (Smiciklas, 2012).
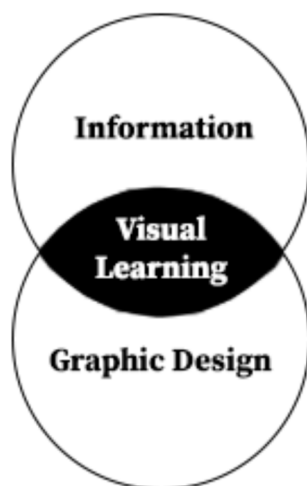
Figure 1: Infographic blends data with design to improve visual learning.

Infographics may be produced as a one-time artefact, for example a student which produces an infographic on the research he conducted as part of a university course. But infographics can also be produced as an artefact which may be reused multiple times, for example an infographic that is part of the annual reporting of a company that needs to be updated every year with the correct data. Infographics are also produced as a summary of the results of social and environmental accounting, an example of such an infographic may be found in Appendix A. This research considers social and environmental accounting as the process of assessing and reporting on the social and environmental effects caused by an organization's economic actions to particular interest groups within society and to society itself (España et al., 2016). During a social and environmental account an organization is accounted on social and environmental topics that are relevant and important to the organization and/or the stakeholders. The concept of social and environmental accounting will be further explained in Section 3.

Infographics are processed more easily by our brain than just plain text. This has to do with how the brain actually works. The brain processes data from images all at once but processes text in a linear manner (Smiciklas, 2012). Therefore, by using infographics to communicate, you make it physically easier for your readers to relate and connect to your information. Next to that, the information age has changed the way people process information and navigate the web. One behaviour pattern that developed is Continuous Partial Attention (Rose, 2010), in which web users are simultaneously connected to multiple digital channels in order to maximize their access to information. As a result, the web users are exposed to more content but at a more superficial level. Today, where time is a precious commodity, people are looking for nuggets of information. Infographics serve that need by presenting knowledge in an easy-to-digest format. Add the fact that social media fuels "shareability," and everything points to infographics becoming one of the most effective forms of content for communicating information in the digital age (Smiciklas, 2012).

1.1 Problem statement

Based on an interview with one person, which was responsible for creating infographics as a result of social and environmental accounting for multiple years, we believe the current situation to be the following when creating infographics in a data-centered process in the context of social and environmental auditing. All the data generated in a social and environmental account is stored in a database, this may be a simple excel sheet or a more advanced database system. The data collected will eventually lead to a social and environmental report. An infographic can be used to summarize the highlights of the social and environmental report in one clear overview. Now, whenever a company wants to produce such an infographic, they often hire a designer from a third party. This designer is located outside of the data-centered process what causes a disruption in the continuous process.
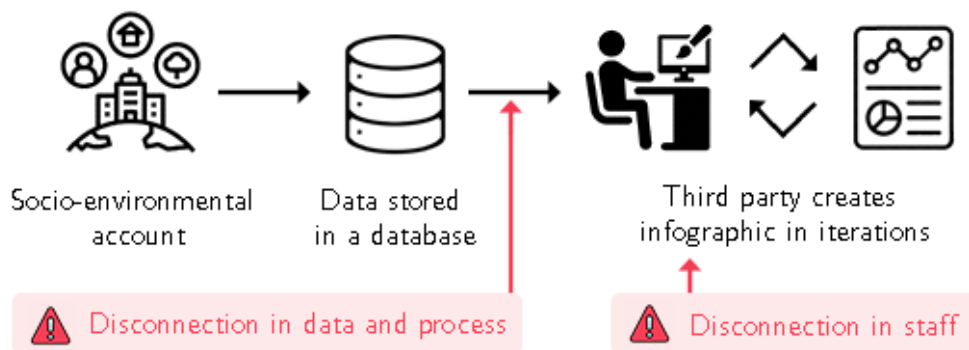
Figure 2: Current problematic situation of creating infographics in the context of social and environmental accounting.

Based on the interview, three major problems are identified within this current situation. First of all, there is a disconnection in staff while no-one inside of the data-centered process has the skills to create the infographic. Therefore, someone outside of the data-centered process is designated to design the infographic. A second disconnection exists in the process. While there is external staff needed to create the infographic, the continuous process is interrupted and may continue when the external staff finished creating the infographic. A third disconnection is identified in the data. The data used for creating the infographic is already available in the data-centered process but needs to be exported to be available for external staff and is therefore costing more time and more money. In addition, data might be constantly changing on a daily basis, so the information exported and presented on the infographic might be outdated. In addition, the exporting of data is done manually which higher the risk of errors in the data.

1.2   Research objective

The main goal of the project is to create a domain-specific language for infographics based on a model-driven engineering approach. A domain-specific language (DSL) is a high-level software implementation language that supports concepts and abstractions that are related to a particular (application) domain (Visser, 2007). Using the proposed domain-specific language, it is possible to describe an infographic in a YAML file, which we will call an infographic specification. YAML is a human-readable data-serialization technique and commonly used for configuration files and in applications where data is being stored or transmitted (Ben-Kiki et. al., 2009). Within Espana et. al.'s (2019) model-driven OpenESEA tool, an interpreter will be designed and will use the infographic specification as input to render the infographic. The proposed domain-specific language lowers the difficulty of creating infographics for people who are normally not designing infographics because domain-specific languages in general are easier to learn due to their narrow and well-scoped domain.

1.3   Research Questions

The problem we aim to solve with this study is decomposed into several research questions and sub questions which can be found below. This research project follows the design cycle as described by Wieringa (2010). The design cycle consists of three stages: problem investigation, treatment design and treatment validation. A fourth phase is included in the engineering cycle but this extra phase – treatment implementation – is omitted while it is considered out of the scope for this project.

Problem Investigation:
- **RQ 1:** What is state-of-the-art in engineering infographics in the context of a data-centered process?
  - o **RQ 1.1:** What is the current process of engineering infographics in the context of a data-centered process?

o **RQ 1.2:** What are the current pain points in the process of engineering infographics in the context of a data-centered process?

Treatment Design:
- **RQ 2:** Which requirements should a domain-specific language on infographics meet?
  Note: We are considering generic requirements for infographics and also requirements that come from the social and environmental accounting tool. We will solve conflicting requirements by prioritizing requirements, which are needed for the social and environmental accounting tool.
- **RQ 3:** What is a suitable domain-specific language for infographics and what are suitable tools to support this?
  - o **RQ 3.1:** Which tools do exist to support the creation of a domain-specific language for infographics?

Treatment validation:
- **RQ 4:** What are benefits and drawbacks of using this domain-specific language and tool?

1.4 Contributions

By answering the research questions described above we contribute to the information science and social sciences domain. At first, we will contribute to science by defining and describing the current process, including its pain points, of creating infographics in the context of social and environmental accounting. Analyses performed in this research will identify and define the different elements populating infographics in the context of social and environmental accounting. In addition, by developing a domain-specific language we will contribute to developing domain-specific languages by applying and validating techniques described by Čeh et. al. (2011). At last, because we are implementing an interpreter of the DSL in the OpenESEA tool, we will contribute to the realisation of the truly open social and environmental accounting tool.

1.5 Outline

This paper is structured as follows. In Section 2 we define how this research is executed by describing the research method. In Section 3 the conceptual model underlying this research is depicted and described. The subsequent section will provide details on the feature analysis of online infographic tools. Section 5 will discuss our proposed domain-specific language for infographics. In Section 6 the proposed domain-specific language was validated by means of an experiment. The results of this experiment can be found in the same section. Section 7 is the conclusion in which answers are provided for the research questions. To conclude we discuss the limitations and explore future research possibilities in Section 8.

2 Research Method

The Design Science framework (Wieringa, 2009) is used to structure this thesis and incorporates the three following phases. The Process Deliverable Diagram (PDD) in Figure 3 depicts the research method, each activity is listed with a unique identifier.

Problem investigation
This phase is intended for exploring the domain of the research project and the problem that we intend to treat. During this phase we will provide an answer to RQ1.
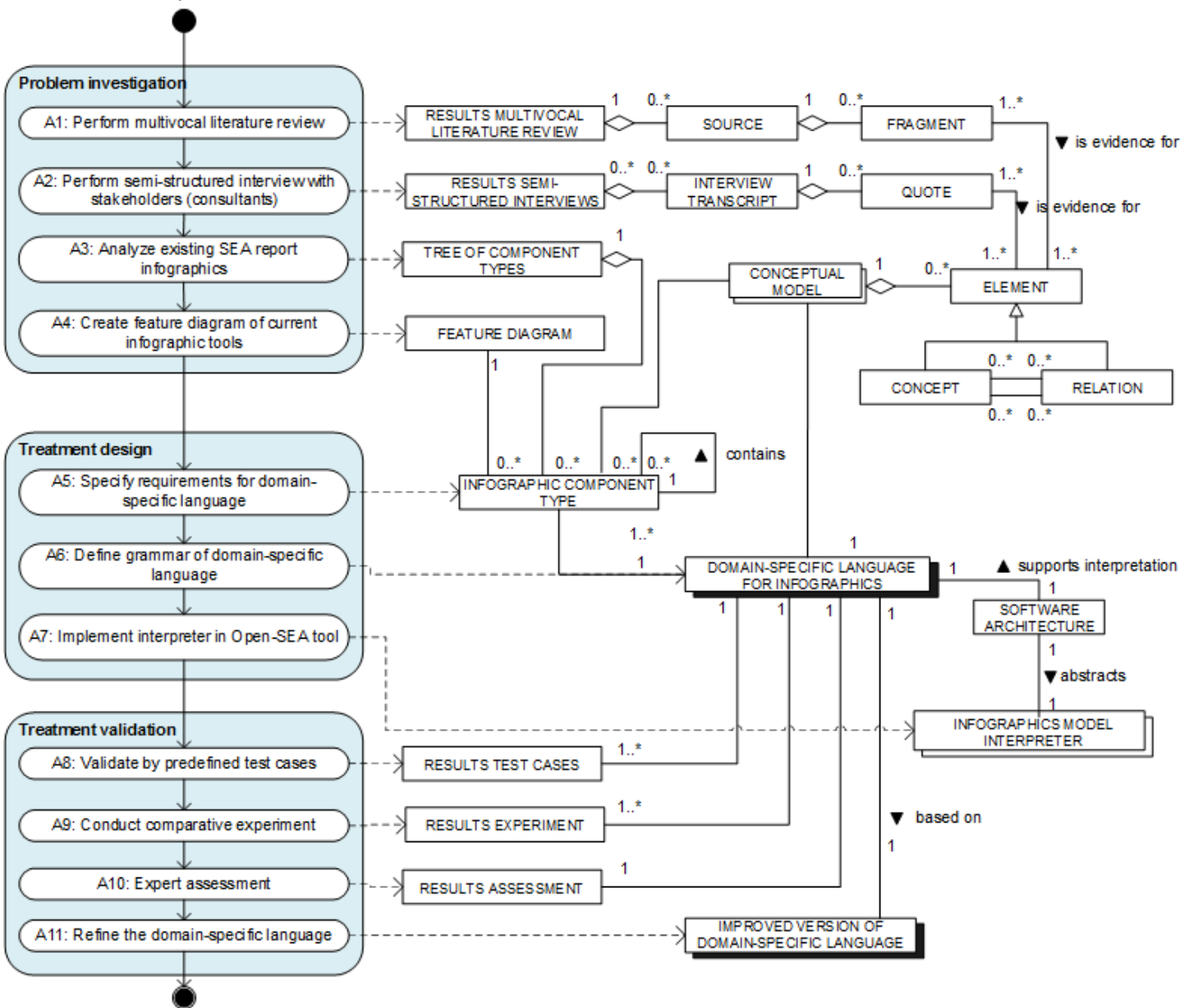
Figure 3: The Process-Deliverable Diagram of the research method.

To better understand the infographics domain, the purpose of infographics, the way infographics are designed, created and how tools can support the creation of infographics we will conduct a multivocal literature review (*A1*). A literature review was not deemed to be feasible while such a review does not include "grey" literature (non-published, nor peer reviewed sources of information) (Garousi et al., 2016). We will specifically be interested in infographics in the context of a data-centered process. In addition, the multivocal literature review will also encapsulate the domain of social and environmental accounting. The data found in the literature is coded in NVivo[1] and as a result, we will produce a conceptual model on the domain of social and environmental accounting. An initial conceptual model was constructed based on tacit knowledge of the researchers. The taxonomy of the so-called coding "nodes" was formed iteratively and represents entity classes from the conceptual model:

- AuditingTeam
- Indicator
    - DirectIndicator
    - IndirectIndicator
- Measure
- Method
- Network
- Organisation
- Report
- Social and environmental accounting
- Stakeholder
- StakeholderGroup
- Tool
- Topic

Figure 4: Coding scheme used in Nvivo

Fragments are extracted from the literature and coded, coding is the analytical process of categorising data, in the aforementioned coding scheme. A fragment is a word or phrase that summarises or captures the essence of a portion of data. An example of a fragment which is evidence for the element 'auditing team': Environmental audits can be conducted by internal auditors or by external independent auditors (Pahuja, 2013). Another example of a fragment which is evidence for the element 'direct indicator': An indicator that does not depend upon other indicators to evaluate or estimate a calculable concept (Martin, 2003).

In addition, we will execute a semi-structured interviews with a stakeholder (*A2*), specifically someone who has experience with developing infographics as a result of a social and environmental account in a data-centered process. The aim of this activity is to gain insights about the current process of creating infographics and

---

[1] https://www.qsrinternational.com/nvivo/nvivo-product

the pain points that occur within this process. The result of this multivocal review in combination with the semi-structured interviews will answer RQ1 and will be the basis for the requirements identified within RQ2.

We will also analyse existing SEA report infographics (*A3*), which can be found via Google or existing SEA reports. As a result, we will produce a tree of component types, displaying the different components an infographic consists of. The following search terms are applied in combination with the term 'infographic': social and environmental auditing, environmental auditing, social auditing, non-financial report and environmental sustainability. Only infographics reporting on real companies are included in this analysis.

In addition, we will analyse tooling that helps you create infographics online, based on this analysis we will create multiple feature diagrams (*A4*). The tools are found through the search engine Google using the keyword infographic combined with one of the following keywords: tool, creator or generator. For every search combination the first two pages are analysed. From the results of this search, every tool, which is explicitly stating that it can be used for creating infographics, has been selected. In addition, the tool must offer a free version which can be used for analysis.

Treatment Design
During this phase we designed a treatment for the problem that is explored within the first phase and providing an answer to RQ2 and RQ3.

This phase will consist of specifying the requirements a domain-specific language on infographics needs to meet (*A5*). The requirements will be derived from the literature study and analyses performed in the problem investigation phase. The requirements will be prioritized using the MoSCoW method. The domain-specific language will be engineered (*A6*) by following the method of Čeh et al. (2011). The domain-specific language will be engineered by using Xtext, which is a framework for developing programming languages and domain-specific languages. Xtext ensures that the grammar of the domain-specific language is clear and has the advantage that it functions as an interpreter and validator. At last, the interpreter will be implemented in the OpenESEA tool (*A7*). Since this is a web-based tool the infographic will also be rendered in a web-based environment.

Treatment validation
In this phase the proposed treatment for the identified problem needs to be validated, hereby providing an answer to RQ4.

To validate the proposed domain-specific language, we perform two steps to check whether the requirements identified in the problem investigation phase are met. At first, we will validate by defining our own test cases (*A8*). Ten testcases will be

constructed whether to check our proposed domain-specific language is expressive enough to render the infographics. Infographics found in activity A3 will be randomly selected and will be used as testcases. Secondly, we will perform an experiment (A9) in where test subjects are confronted with ten infographics. Half of the infographics are rendered by the OpenESEA tool, and the other half of the infographics will be the original infographic found in activity A3. In the experiment test subjects are questioned on the visual attractiveness of the infographics and whether the infographics are real or rendered by our tool. Another way of validating the proposed domain-specific language is by means of an expert assessment. The expert assessment will be executed with an expert in the field of social and environmental accounting and this expert is currently refactoring the OpenESEA tool described by Bik (2018). The following questions will be used to guide the expert assessment: What are the strengths and weaknesses of the infographic feature and its underlying domain-specific language? In addition, would you consider adopting this tool and at what point in time? The goal of this assessment is to define any benefits and drawbacks to the implemented interpreter and its domain-specific language. As a final step, we will update the domain-specific language (*A11*) based on the results of the former activities performed in this phase. This will result in an improved version of the domain-specific language.

3 Conceptual model on domain of social and environmental accounting

In this section, the conceptual model on the domain of social and environmental accounting is described and depicted in Figure 5. A textual description of the conceptual model is found in section 3.1. The conceptual model is the result of the multivocal literature review. For this literature review we identified 18 papers and as can be seen in Table 1 we identified a total of 95 fragments extracted from literature referring to elements in the conceptual model. In Appendix B it is displayed which 18 papers are used for the construction of the conceptual model including the number of fragments identified for each paper[2].

Table 1: Elements of the conceptual model including the number of fragments referring to each element.

Name	Fragments
AuditingTeam	8
Indicator	6
DirectIndicator	2
IndirectIndicator	2
Measure	6
Method	13
Network	4
Organisation	6
Report	5
Social and environmental accounting	12
Stakeholder	11
StakeholderGroup	10
Tool	2
Topic	8
Total:	95

3.1 Context of social and environmental accounting

Social and environmental accounting is the process of assessing and reporting on the social and environmental effects caused by an organization's economic actions to particular interest groups within society and to society itself (Grey, 2000) (España et al., 2016). Multiple terms exist meaning the same as social and environmental accounting these are non-financial reporting, social balance, and social auditing. The ultimate aim of social and environmental accounting is to ensure that the organization's environmental performance meets the goals set in its

---

[2] An automatically generated report by Nvivo containing all fragments of text related to each node may be found at: http://tiny.cc/NvivoCodingByNode

environmental policy and also to ensure compliance with standards and regulatory requirements. An organisation wants to be ronsponsible and sustainable about those interactions, due to norms, values and beliefs, and/or due to laws and regulations (Pahuja, 2013).
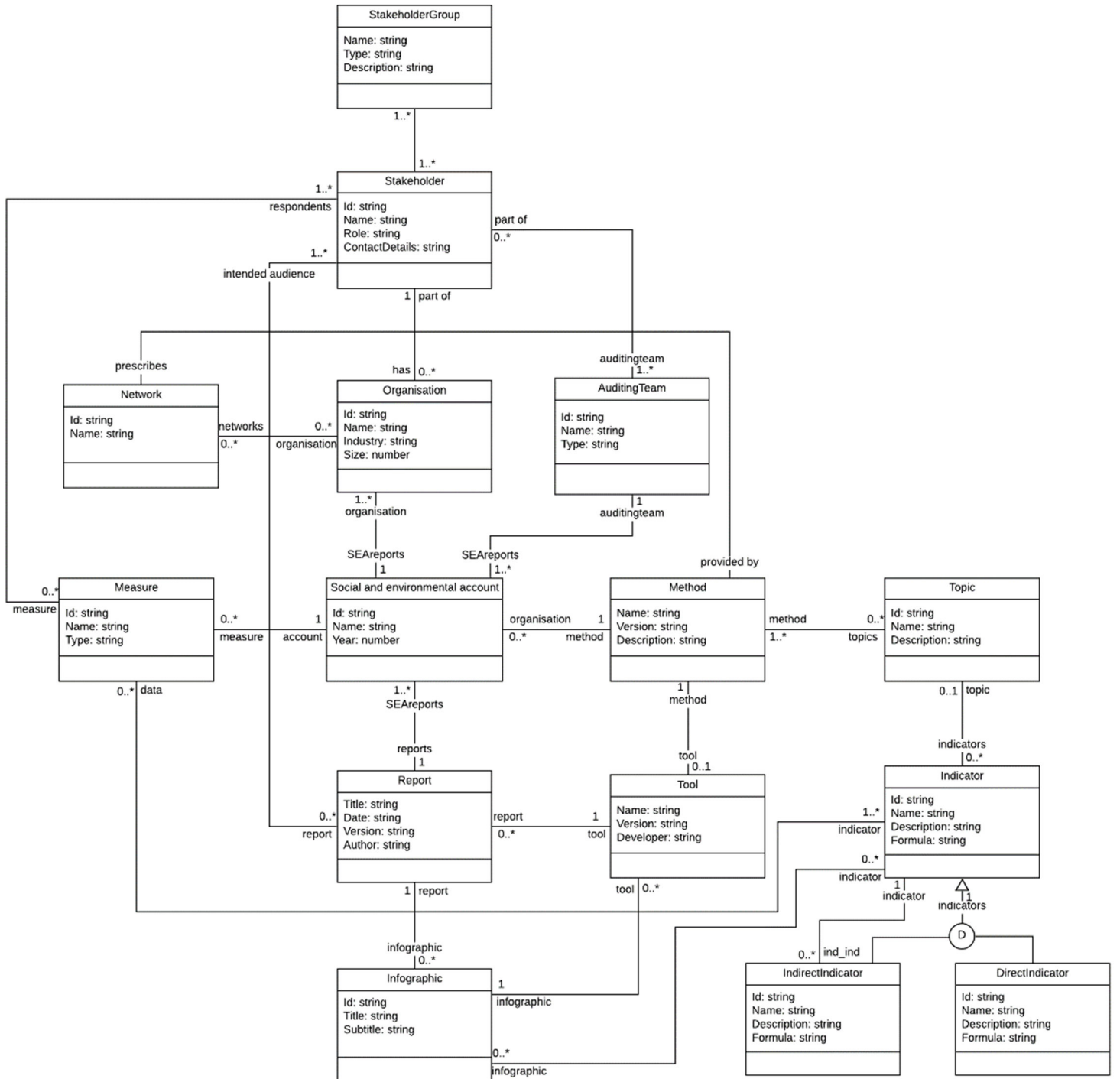
Figure 5: The conceptual model on the domain of social and environmental accounting

A *social and environmental account* is the result of social and environmental accounting, meaning that a social and environmental account only covers a certain time-period. A social and environmental account is executed by an auditing team. The *auditing team* may be an internal team but the audit may also be executed by an external team (Pahuja, 2013). A social and environmental account results in a collection of data which can be used to create reports or infographics. A *report* is defined as a document published by an organisation about the economic, environmental and social impacts caused by its everyday activities An *infographic* is defined as a visualization of data or ideas that tries to convey complex information to an audience in a manner that can be quickly consumed and easily understood (Thatcher, 2012).

An *organisation* is an organised group of people with a particular purpose. Due to this purpose, this organisation performs activities, has behaviours, exploits resources and makes choices for profits or non-profits. Multiple synonyms of organisation exist in literature such as business, corporation, enterprise or institution. An organisation consists of stakeholders, a *stakeholder* is a person with an interest or concern in something, especially an organisation (Black et al., 2012). Stakeholders are vital for the survival of an organisation. Organisations categorise stakeholders, based on their conflicting needs, into *Stakeholder groups*. A stakeholder may be part of one or more stakeholder groups. We identify the following stakeholder groups 'consumer', 'network', 'media', 'staff', 'supplier', 'regulator', 'third party consultancy'. The decomposition of the concept stakeholder group is found in Appendix C. A few examples of organisations performing social and environmental accounting are: Tony Chocolonely, Patagonia and Competa.

An organisation can be part of a network in the context of social and environmental accounting. A *network* is an organization that develop initiatives, principles and/or standards related to corporate social and environmental performance (Gao et al., 2006). Organisations willing to join a network may need to meet certain performance, legal, transparency or other requirements. Examples of such networks are: Economy for the Common Good, B Labs of the B Corporation and Fair Trade Software Foundation. A social and environmental account is performed according to a specific method. Each network is using their own method. The *method* prescribes how the social and environmental account should be executed, which questions are in the surveys and which items to report on. All methods are supported by their own *tool*, for some this comes down to a relatively simple Excel sheet, while other tools exist of larger (web) applications. Examples of such methods and accompanying tools are: Common Good Balance Sheet, B Impact assessment or Fair Trade Software Foundation certification method.

The social and environmental account is using measures to gather data. A *measure* is defined as a way of collecting data, for example surveys. A survey is a means of

gathering information about the characteristics, actions, or opinions of a large group of people, referred to as a population (Pinsonneault et al., 1993).

Materiality assessment is a tool in stakeholder engagement designed to gather insight on the relative importance of specific environmental, social and governance issues (King, 2013). Methods are using materiality assessment to identify, prioritize and select the most relevant topics for their system or organisation. To address all sustainability issues is difficult and therefore these issues are fragmented based on recognized important *topics*. Several examples of a topic are: diversity in the workforce, economic impact or gas emissions. *Indicators* are defined by algorithms or calculation methods which combines one or more metrics with associated decision criteria, such as thresholds, targets or patterns, supporting the information need (Olsina et al., 2004). Indicators are split up in *direct indicators* and *indirect indicators*. *Direct indicators* can be directly used without any mathematical operations, two examples would be the number of men in the workforce and the number of women in the workforce. *Indirect indicators* may be calculated by using a formula. An example of an indirect indicator using the direct indicators described before is: the ratio of women in the workforce. This indicator may be calculated by the number of women in the workforce divided by the number of men and women in the workforce.

3.2 Context of infographics

An *infographic* (short for information graphic) is a type of picture that blends data with design, helping individuals and organizations concisely communicate messages to their audience (Thatcher, 2012). The foundation of all infographics consists of three major parts, they are visual, content and knowledge (Thatcher, 2012). Visual meaning the design of the infographic, the colors and graphics used in the infographic. Designing a visual appealing infographic is important while data visualizations that have been rated as more aesthetically appealing have been found to have positive effects on engagement (Cawthon et al., 2007) and memorability (Borkin et al., 2013). Content elements are populating the infographic: texts, statistics, figures and charts. Knowledge meaning the story or message intended to convey with the infographic.

Multiple design styles exist when creating an infographic (Ritchie, 2018):
1. *Static infographics*, these are the most common type of infographics. Static infographics combine text and images by presenting concepts in the form of images, graphs, maps and other visual elements.
2. *Animated infographics*, as the name implies these infographics are animated. These infographics are often in a GIF format and are therefore often called infoGIFs. Because of their animated nature they are particularly engaging if you want to grab attention, making them ideal for social media (where you have to stand out in someone's content stream).

3. *Interactive infographics* are infographics that require some sort of action or input from the viewer. This kind of infographics achieve the "interactivity" part by providing details on mouseover, clicking and scrolling, swiping on a touch screen, embedding video or other interactions with the viewer. This interactivity allows users to explore the dataset for themselves and helps to engage the viewer even more than with static or animated infographics.

Due to the context of current social and environmental methods and tools, and to narrow down the scope of this research, we will only focus on static infographics in this research.

Infographics are also produced as a summary of the results of social and environmental accounting, an example of such an infographic produced by REAS Euskadi may be found in Figure 6. REAS Euskadi is a Spanish network founded in 1997, it brings together 67 companies and social entities whose mission is to develop a solidarity economy.
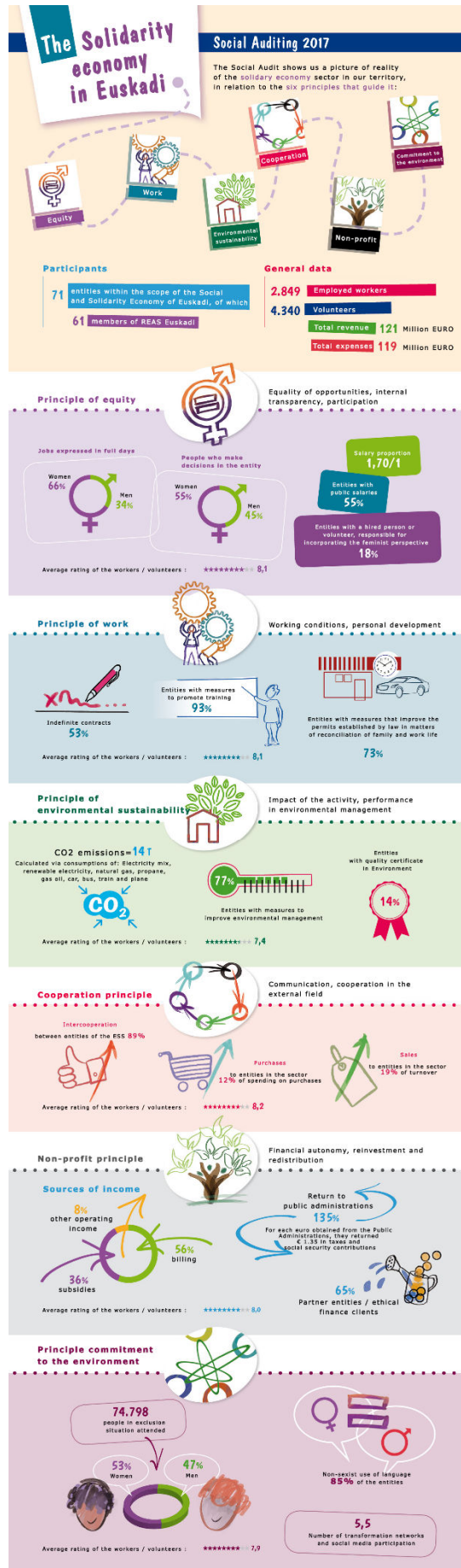
Figure 6: Example of an infographic produced by REAS Euskadi. The infographic is translated to English while the original was Spanish. The infographic depicts general information on entities in the network and depicts the six principles which are central in the network of REAS Euskadi.

3.3 Tree of component types

A total of 58 infographics were found, including infographics from Lego, Coca-Cola and Nestle. The goal of this analysis was to create a tree of component types (Figure 7), which displays the different components a social and environmental accounting infographic consists of. Each different component identified is described in Table 2. In Appendix D you may find a table displaying the mapping between component types to the analysed infographics.
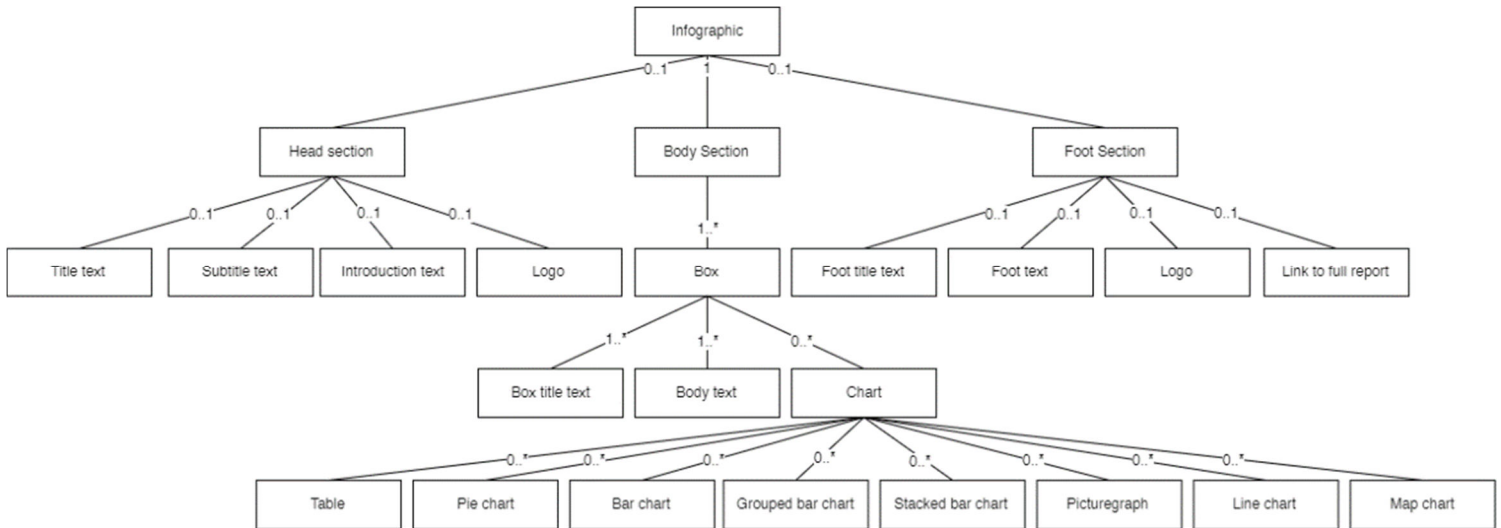
Figure 7: Tree of component types

Table 2: Components found in the analysis and their descriptions.

Component types	Description
Head section	The head section is the top part of the infographic often displaying the infographic title, subtitle, introduction text and logo.
Title text	Title text refers to the title of the infographic and the title is often found in the head section.
Subtitle text	Subtitle text refers to the subtitle of the infographic. Often displayed below the title text.
Introduction text	Introduction text refers to text found in the head section, which may spreads over multiple lines.
Logo	Logo refers to an image representing the logo of the company, which often was present in the head section.
Body section	The body section is the middle part of the infographic displaying the text, data and chart elements.
Box	A box is used to demarcate a certain width and height on the infographic. Within a box, it is possible to display text, data or chart elements.
Box title text	Box title text refers to the title of the box and may not spread over multiple lines.

Component types	Description
Body text	Body text refers to the text displayed in a box and is allowed to spread over multiple lines.
Chart	A chart is a graphical representation of data and multiple forms of charts exist.
Table	An arrangement of facts and figures, usually in rows and columns, that makes information easy to understand.
Pie chart	A pie chart is a circular statistical graph, which is divided into slices to illustrate numeric proportion.
Bar chart	A bar chart is a way of visualising categorical data with rectangular bars with heights/lengths proportional to the value they represent.
Grouped bar chart	A grouped bar chart differs from a normal bar chart since multiple bars are used for each categorical group.
Stacked bar chart	A stacked bar chart differs from a normal bar chart since stacked bar charts are displaying different groups on top of each other.
Picturegraph	A picturegraph represents data by using symbols and pictures. Variations in picturegraphs are unlimited while every symbol or picture may be used for representing data.
Line chart	A line chart is displaying information as a series of data points connected by straight line segments.
Map chart	A map chart allows the visualisation of spatial relationships in data by indicating data on a geographical map.
Foot section	The foot section refers to the bottom part of an infographic often displaying a foot title text, a foot text, logo and a link to the full report.
Foot title text	Foot title text refers to the title of the foot section and is not allowed to spread over multiple lines.
Foot text	Foot text refers to the text displayed in the foot section and is allowed to spread over multiple lines.
Logo	Logo refers to an image representing the logo of the company, which was less often present in the foot section.
Link to full report	Is there an URL present on the infographic linking to the full sustainability report.

The first ten infographics (excluding multiple infographics on one company) found were used to construct an initial tree of component types. Each component of the infographic is analysed and sorted in groups. After analysing the first ten infographics, 16 different elements were found. After the initial tree of component types was composed, the other infographics were analysed to check whether the tree of component types encapsulates every object found in the infographics. After analysing all the infographics, six new elements were found and eventually added to the tree of component types. Figure 8 depicts the data saturation throughout

the process, it shows when new elements are identified and added to the tree of component types. After analysing the thirtieth infographic the data was fully saturated and no new elements were observed.
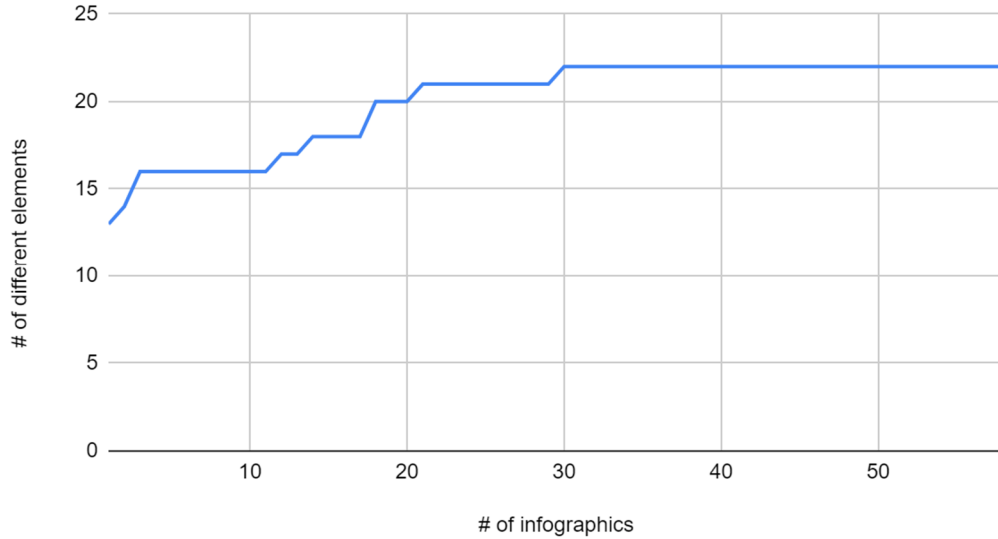
Figure 8: Data saturation in tree of components analysis
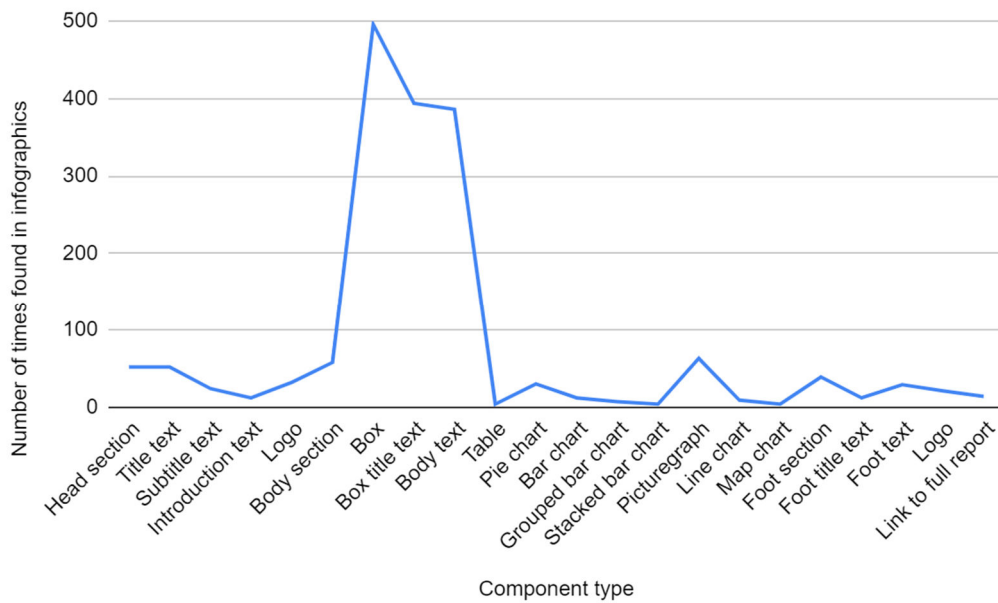
Figure 9: Number of times each component is found in existing infographics.

In figure 9 the number of times each component is found is depicted. It shows that the box component (496 times) is the component that is used the most closely followed by box title text component (394 timse) and body text component (386 times). Based on the analysis it was proven that a head section nor a foot section was present in all infographics. Naturally, a body section was present in all the infographics. In Figure 10 the number of each different chart identified is depicted.

It shows that picturegraphs (with 63 occurences) are the most used chart type on existing social and environmental accounting infographics. Pie charts were the chart type that were second most used with a total of 30 occurences and a total of 23 different bar charts were found in the infographics. The chart type that were used the least in social and environmental accounting infographics is the map chart. An example of an infographic partitioned into components may be found in Figure 11.
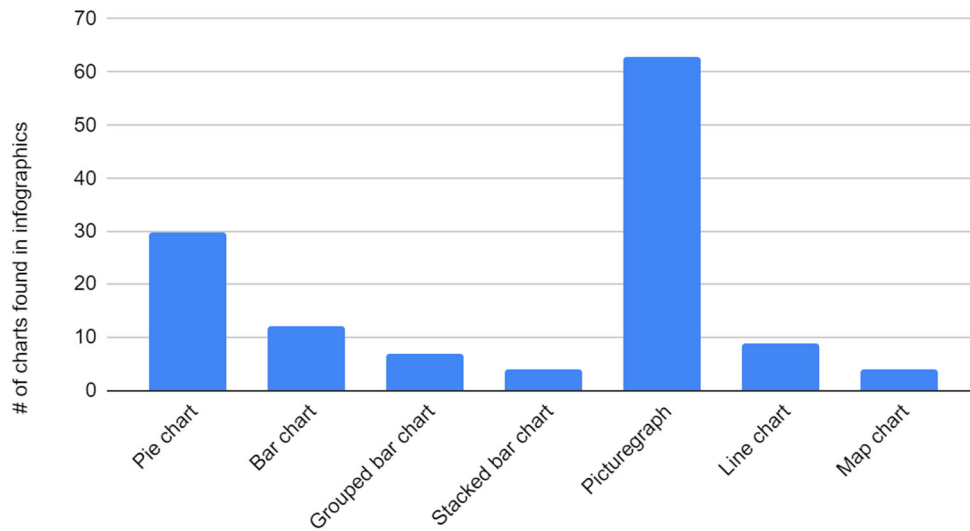
Figure 10: Distribution of the different charts in all infographics
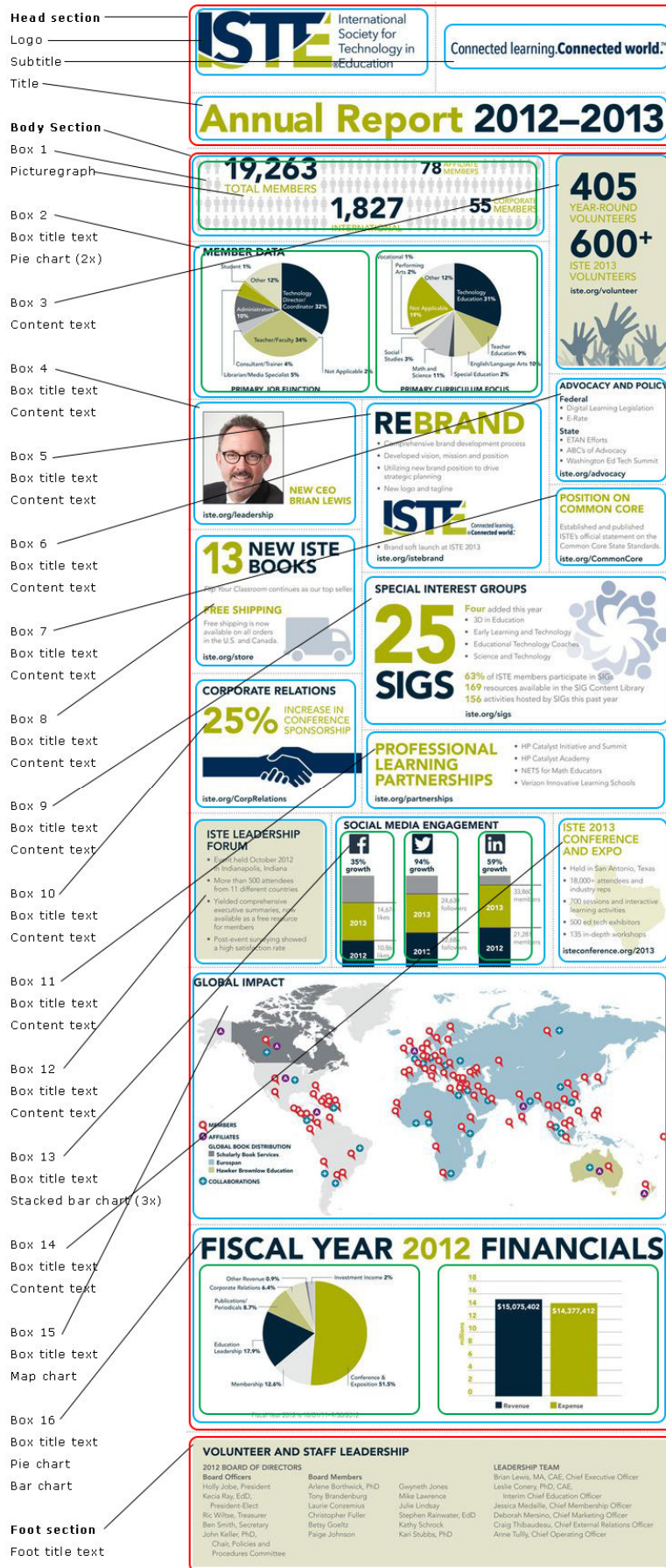
Figure 11: Example infographic partitioned into components

4 Feature analysis of existing tools for infographic creation

There are tools available online for the purpose of creating infographics. By analysing existing infographic tools, possible functions for our domain-specific language may be identified.

4.1 Introduction to the tools

A total of 13 tools were identified after analysing the search results. Three tools were excluded due to the fact that they only offered a paid version. The ten remaining tools are included in this analysis. For each tool that has been selected, the name, the URL and the description is listed in Table 3.

Table 3: Online infographic tools with their description

Name	URL	Description
BeFunky	https://www.befunky.com	BeFunky is an online and mobile photo editor launched in 2007. Only the online tool is analysed in this research. It enables the user to edit photos, add photo effects, create photo collages and also create infographics.
Canva	https://www.canva.com	Canva is a simplified graphic-design tool website, founded in 2012. It uses a drag-and-drop format and provides the user with an extensive amount of photographs, vectors, graphics and fonts.[3]
Easel.ly	https://www.easel.ly	Easel.ly is a web-based infographic platform used for representing and conveying ideas, concepts, reports, processes and projects through visual forms that are easy to follow and inspiring.[4]
Infogram	https://infogram.com	Infogram is a web-based data visualization and infographics platform. It allows people to make digital charts, infographics and maps. Infogram offers an intuitive WYSIWYG editor with the possibility to publish, embed or share the created infographics.[5]
Lucidpress	https://www.lucidpress.com	Lucidpress is a web-based desktop publishing software application developed by Lucid Software, which developed Lucidchart. It can be

---

[3] https://en.wikipedia.org/wiki/Canva

[4] https://reviews.financesonline.com/p/easel-ly/

[5] https://en.m.wikipedia.org/wiki/Infogram

Name	URL	Description
		used to create flyers, brochures, infographics, posters, magazines and presentations.
Mind the Graph	https://mindthegraph.com	Mind the Graph is an infographic maker for scientists, particularly focused on Life Sciences and Health. It's a useful tool to create figures for papers, infographics, slides and presentations for classes.[6]
Piktochart	https://piktochart.com	Piktochart is a web-based infographic application which allows users without intensive experience as graphic designers to easily create infographics and visuals using themed templates[7].
Snappa	https://snappa.com	Snappa is a cloud-based graphics editor for social media, personal, and marketing purposes. This makes creating visual content easy without the use of complex tools like those in Photoshop and other similar image editing applications.[8]
Venngage	https://venngage.com	Venngage is a free online infographic tool and their goal is to make design more accessible for everyone. Venngage lets users easily drag and drop icons and text widgets directly onto their canvas.[9]
Visme	https://www.visme.co	Visme is an intuitive online platform with a mission to allow anyone to easily visualize their ideas into engaging presentations, infographics, animations and more. It's entirely based on HTML5 and runs on any browser with simple easy to use drag and drop functions.[10]

For each tool a feature model is created, an example may be found in figure 12. The rest of the feature models can be found in Appendix E. Nine out of ten tools identified offered a free and a paid version. In most cases, the paid version includes more sample images, icons, templates, font pre-sets, different charts, and more export options. For the reason that purchasing all paid versions is expensive and that the difference in features between the free and paid version are not large, we chose to analyse the free version of each tool.

---

[6] https://www.producthunt.com/alternatives/mind-the-graph-2

[7] https://en.wikipedia.org/wiki/Piktochart

[8] https://reviews.financesonline.com/p/snappa/

[9] https://www.youtube.com/user/VenngageInc/about

[10] https://www.linkedin.com/company/visme/about/

Most of the features identified are not unique for a single tool, but rather appear across many tools. Therefore, we identify which features appeared in more than one tool, we named them with the same label, and we added an identifier in the form of Fn for ease of reference. We can now refer to features as generic features (meaning that they generalise concrete features of tools resulting in a similar functionality). Such generic features are listed and described in the next section in Table 4. An example of a feature model of one of the tools may be found below, the other feature models are displayed in Appendix E.
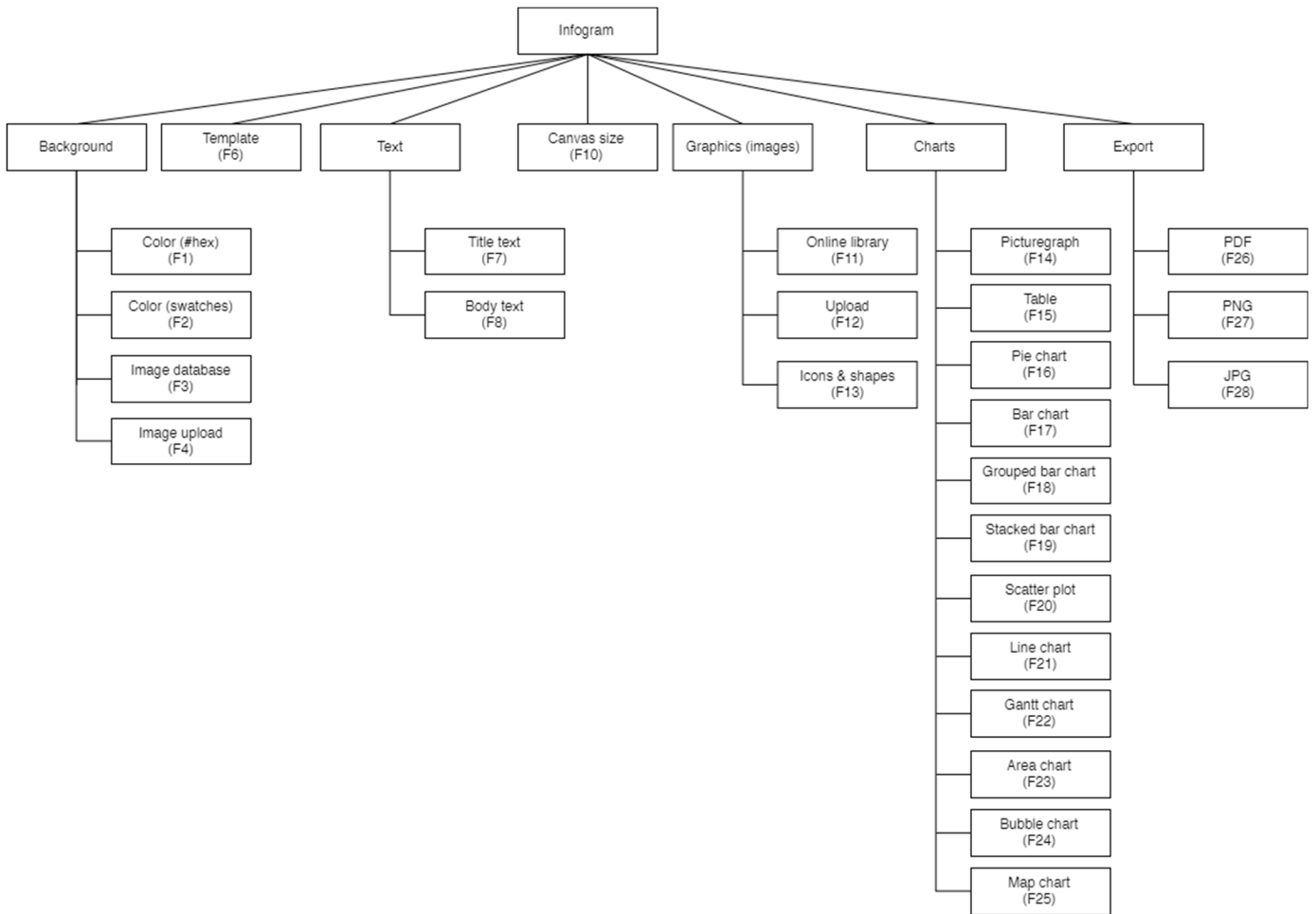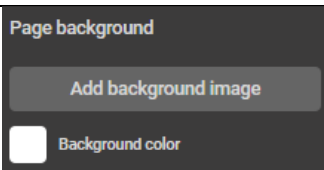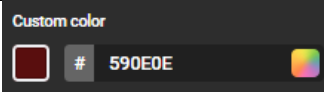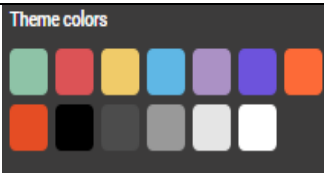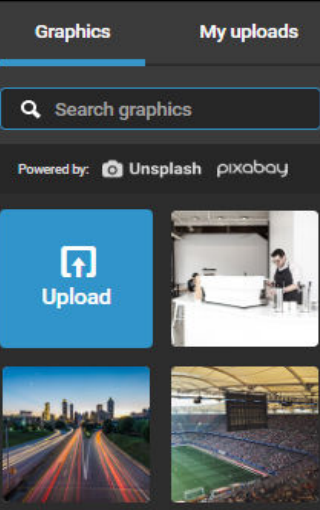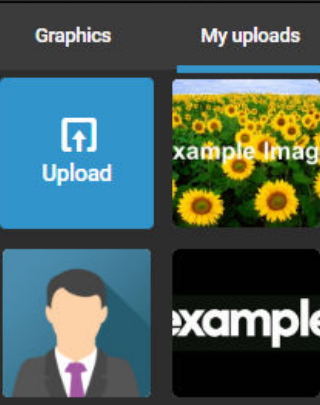
Figure 12: Feature model of Infogram
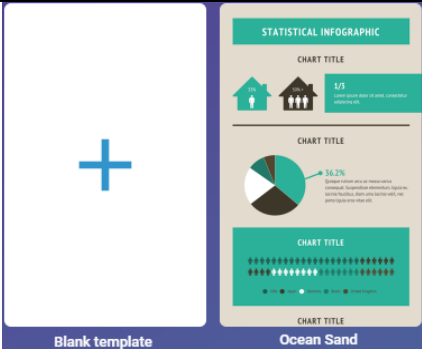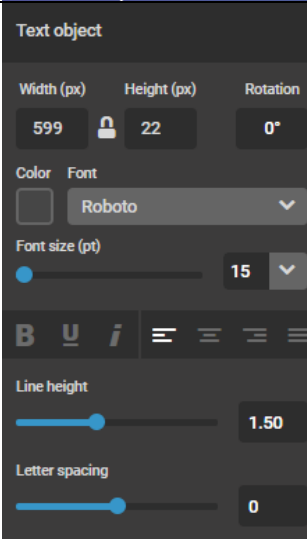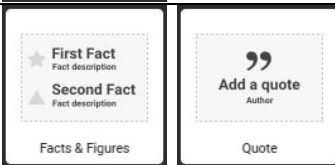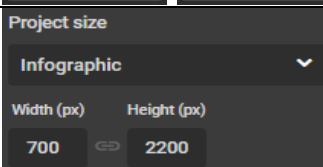
4.2 Identification of different generic features
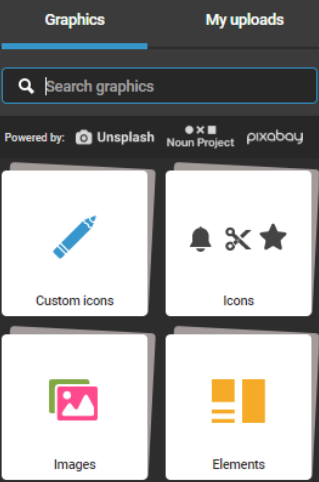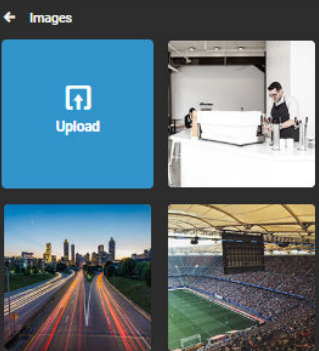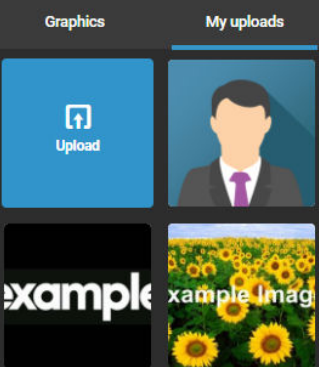
Each different generic feature and feature group identified is described in Table 4 below. When multiple features are revolving around a single concept then the features are grouped, for example background, charts or export. The screenshots are captured in the tool of Infogram except for F5 (pattern database) which is captured in the tool of Venngage. For traceability purposes, a traceability matrix between the different tools and features is included and displayed in Table 5.
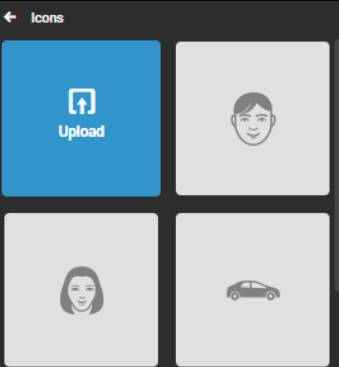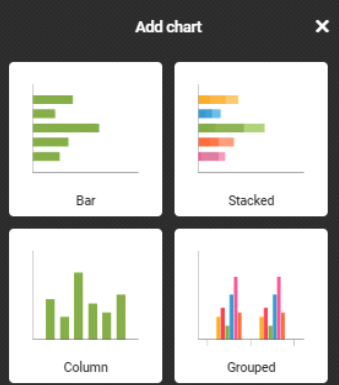
Table 4: Different features identified in existing tools with their description and screenshot.
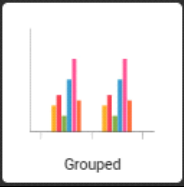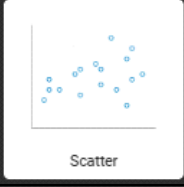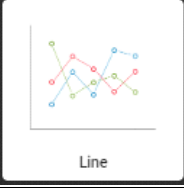
Features	Description	Screenshot	Number of tools incorporating the feature
Feature group: Background	The part of the infographic representing what lies behind objects in the foreground. Five different features (F2 to F6) were identified to create the background of an infographic.		10
Color (#hex) (F1)	The ability to define the background with a color hex code that describes the composition of a certain color in a specific color space, usually RGB. For example, the color hex code for white is #FFFFFF; and at the opposite end is black #000000.		8
Color (swatches) (F2)	The ability to define the background with sample colors, as displayed in the screenshot.		10

Features	Description	Screenshot	Number of tools incorporating the feature
Image database (F3)	The ability to select and use a sample image as the background of the infographic.		6
Image upload (F4)	The ability to use an uploaded image as the background of the infographic.		5
Pattern database (F5)	The ability to select a pattern and use this as the background of the infographic.		3

Features	Description	Screenshot	Number of tools incorporating the feature
Template (F6)	The ability to select a blank template or a pre-defined template to start with when designing an infographic.		10
Feature group: Text	The ability to place a text box and change its settings, such as font, font-size, color, text-box size and alignment among other settings. Three different features (F7 to F9) were identified to insert and use text on the infographic.		10
Title text (F7)	The ability to place a text box with pre-set settings so it will display a title text.		8
Body text (F8)	The ability to place a text box with pre-set settings so it will display a body text.		9
Font presets (F9)	The ability to place additional different text boxes with pre-set settings other than F9 and F10.		6
Canvas size (F10)	The ability to adjust the canvas size according to the user's preference. This might be done with pre-set sizes like A3 or A4,		10

31

Features	Description	Screenshot	Number of tools incorporating the feature
	or by filling in a number of pixels.		
Feature group: Graphics (images)	The ability to use images on the infographic. The user is able to resize and rotate the images on the infographic. Three different features (F11 to F13) were identified to select and use images on the infographic.		10
Online library (F11)	The ability to select an image from a database and use this on the infographic.		10
Upload (F12)	The ability to upload an image from your computer and use this on the infographic.		10

Features	Description	Screenshot	Number of tools incorporating the feature
Icons & shapes (F13)	The ability to select an icon or shape from a database and use this on the infographic.		10
Feature group: Charts	The ability to insert a chart to the infographic. Nine different features (F14 to F25) were identified to insert and use charts on the infographic.		8
Picturegraph (F14)	The ability to display picturegraphs on the infographic. This feature may also be called 'pictorial' or 'picture chart'.		4
Table (F15)	The ability to display tables on the infographic.		5
Pie chart (F16)	The ability to display pie charts on the infographic.		7
Bar chart (F17)	The ability to display bar charts on the infographic.		7

Features	Description	Screenshot		Number of tools incorporating the feature
Grouped bar chart (F18)	The ability to display grouped bar charts on the infographic.	Grouped		7
Stacked bar chart (F19)	The ability to display stacked bar charts on the infographic.	Stacked		7
Scatterplot (F20)	The ability to display scatterplots on the infographic.	Scatter		4
Line chart (F21)	The ability to display line charts on the infographic.	Line		7
Gantt chart (F22)	The ability to display Gantt charts on the infographic.	Gantt		1
Area chart (F23)	The ability to display area charts on the infographic.	Area		3
Bubble chart (F24)	The ability to display bubble charts on the infographic.	Bubble		3
Map chart (F25)	The ability to display map charts on the infographic.	World map		4
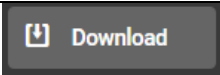
Features	Description	Screenshot	Number of tools incorporating the feature
Feature group: Export	The ability to export the infographic from the tool so you will obtain a downloadable copy. This feature may also be called 'publish' or 'download'. Three different features (F26 to F28) were identified to export the infographic.		10
PDF (F26)	The ability to export the infographic from the tool to a PDF format.		8
PNG (F27)	The ability to export the infographic from the tool to a PNG format.	See screenshot F26.	9
JPG (F28)	The ability to export the infographic from the tool to a JPG format.	See screenshot F26.	7

Table 5: Traceability matrix between the online infographic tools and the identified features

	Tools									
	BeFunky	Canva	Easel.ly	Infogram	Lucidpress	Mind the Graph	Piktochart	Snappa	Venngage	Visme
F1	x	x	x	x	x		x	x	x	
F2	x	x	x	x	x	x	x	x	x	x
F3		x	x	x			x	x		x
F4			x	x	x			x		x
F5						x		x	x	
F6	x	x	x	x	x	x	x	x	x	x
F7			x	x	x	x	x	x	x	x
F8			x	x	x	x	x	x	x	x
F9		x		x		x	x		x	x
F10	x	x	x	x	x	x	x	x	x	x
F11	x	x	x	x	x	x	x	x	x	x
F12	x	x	x	x	x	x	x	x	x	x
F13	x	x	x	x	x	x	x	x	x	x
F14				x			x		x	x
F15				x	x		x		x	x
F16		x	x	x		x	x		x	x
F17		x	x	x		x	x		x	x
F18		x	x	x		x	x		x	x
F19		x	x	x		x	x		x	x
F20				x			x		x	x
F21		x	x	x		x	x		x	x
F22				x						
F23				x			x		x	
F24				x			x		x	
F25				x			x		x	x
F26		x	x	x	x	x	x		x	x
F27	x	x		x	x	x	x	x	x	x
F28	x	x	x	x	x					x

4.3 Analysing the features
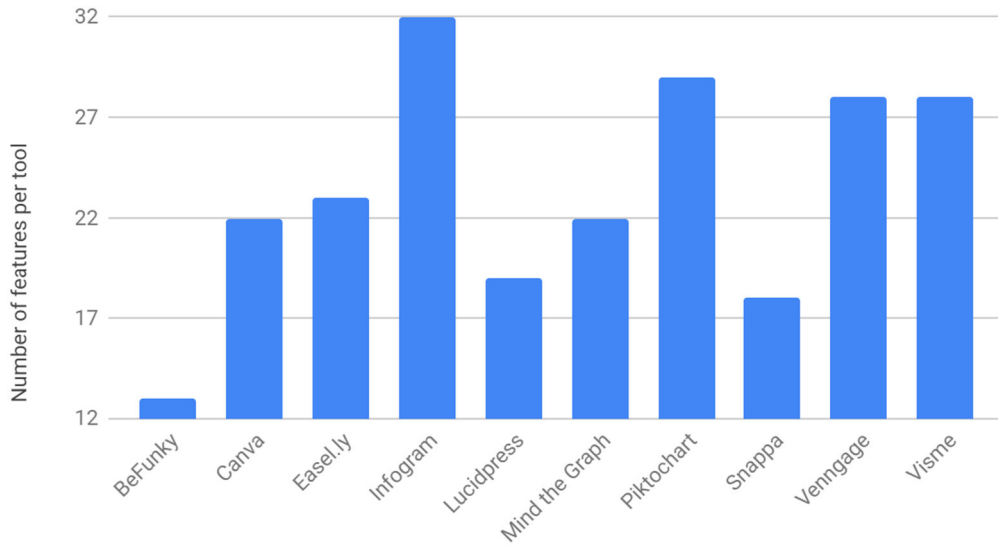
Figure 13: Total features per tool

In the bar chart above the number of features per tool is depicted. It was not possible to insert charts with the tools BeFunky and Snappa, hence they offer the least number of features.
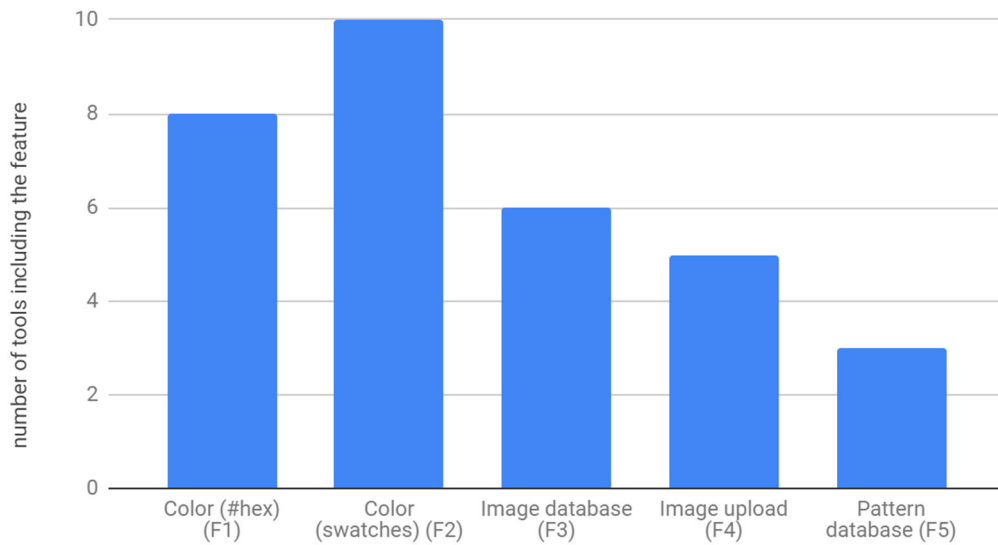
Figure 14: Distribution of feature group background

In the bar chart above the number of tools including F1 to F5 is depicted. These features are part of the background feature group. In all ten tools it was possible to define the background with color swatches. Setting the background color with a hex code was the second most implemented feature. In all ten tools it was possible to make use of templates (F6). In all ten tools it was possible to set the canvas size

(F10). In addition, every tool provided the possibility of using the feature group Graphics (images) which includes the features F11 to F13.
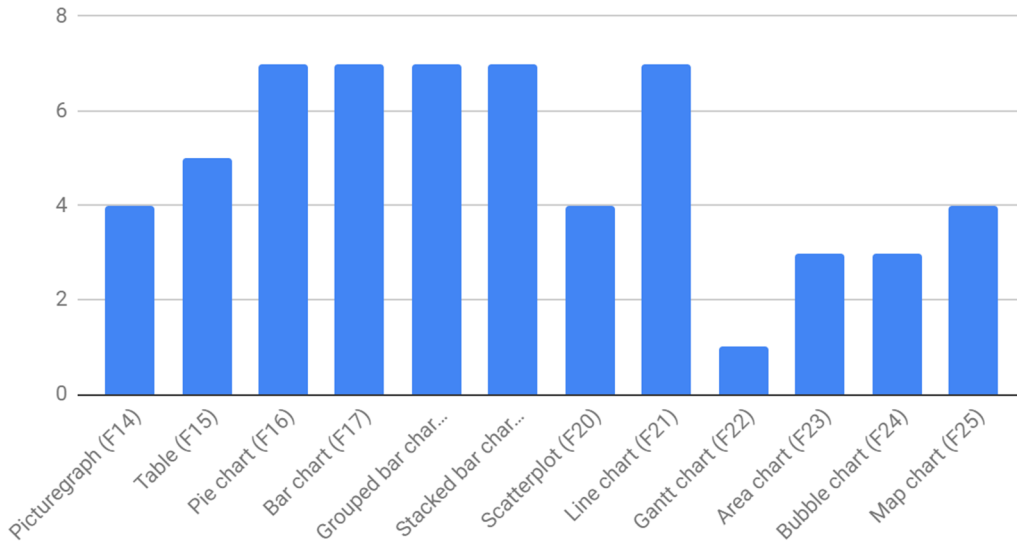
Figure 15: Distribution of feature group charts

In the bar chart above the number of tools including F14 to F25 is depicted. These features are part of the chart feature group, a total of 8 tools were able to use charts on the infographic. Pie charts, different bar charts and line charts are the most common chart types provided by the different tools.
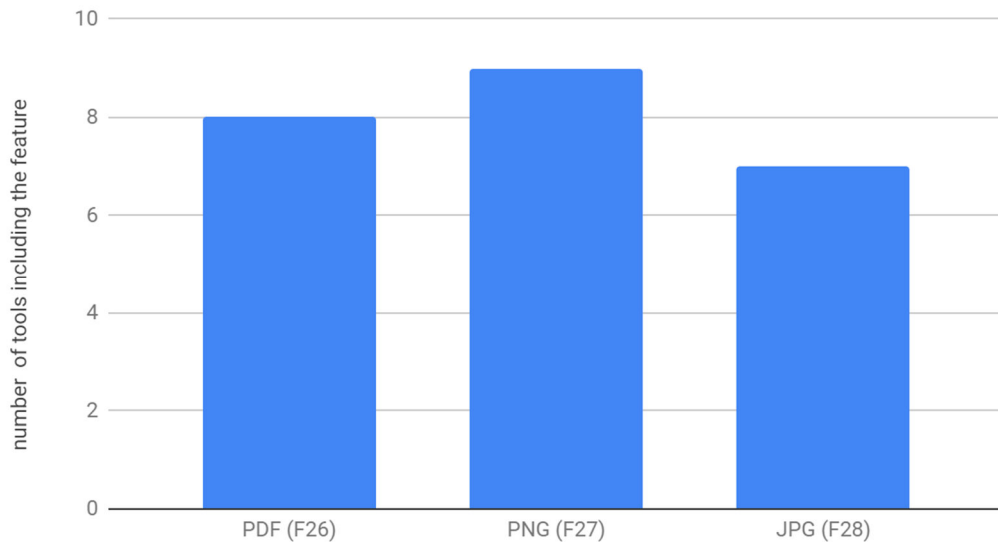
Figure 16: Distribution of feature group export

In the bar chart above the number of tools including F26 to F28 is depicted. These features are part of the export feature group. PNG is the most common export function and JPG is the least common export function.

5 Domain-specific language for infographics

In this section the requirements for the domain-specific language for infographics are described and prioritized using the MoSCoW method. Furthermore, the grammar of the domain-specific language defined in Xtext is described. Afterwards, a dictionary of the domain-specific language is displayed, depicting the different coding possibilities. Afterwards, the implementation in OpenESEA is described. Finally, two examples of a YAML file used as input by the interpreter in the OpenESEA tool are described.

5.1 Requirements for the domain-specific language

To construct the proposed domain-specific language for infographics, we first need to identify the requirements the domain-specific language has to adhere. The requirements are based on the needs of the OpenESEA tool and on the literature study described earlier, including the conceptual model, tree of component types and feature analysis. The requirements are expressed as user stories, short descriptions of functionality –told from the perspective of a user- that are valuable to any stakeholder of the project (Cohn, 2004). The user stories are prioritized by using the MoSCoW method (Waters, 2009) and are listed below. In Appendix F, a user story matrix diagram may be found, depicting where each requirement originated for traceability purposes. For example, the user story regarding the size of the infographic (US7) originates from the conceptual model (Section 3.1 & 3.2), the tree of component types analysis (Section 3.3) and feature seven, eight and nine of the feature analysis of existing tools for infographic creation (Section 4).

Requirements originating from the OpenESEA tool were mandatory while we are designing an interpreter for this tool. Therefore, requirements originating from this source are defined as must haves. Based on the different analyses performed in previous sections we identified that certain features or elements are more important than others. For example, we must have piecharts or barcharts incorporated in our domain-specific language since the tree of component types analysis showed that these are the most common charts used on a social and environmental accounting infographic. Therefore the user stories regarding piecharts and barcharts are included in the must haves and user stories regarding map charts or scatter plots are in the won't haves. The requirements are also prioritised based on the engineering possibilities within the timeframe of this research meaning that requirements listed as could haves and won't haves might change to must haves in the future.

Must have

US1: As a user, I want to create an infographic summarizing the social and environmental accounting report, so I can easily share the results of the social and environmental accounting report as an image with my stakeholders.

US2:   As a user, I want to upload my own infographic specification, so that I can tweak the infographic based on my needs.

US3:   As a user, I want to select which infographic specification I will use, so that I can adhere to the design of my organisation or network.

US4:   As a user, I want to remove the infographic specification, so that the infographic specification is deleted from the database.

US5:   As a network user, I want to upload an infographic specification, so that users of organisations in my network can use this infographic specification.

US6:   As an organisation user, I want to upload an infographic specification, so that users in my organisation can use this infographic specification.

US7:   As a user, I want to define the background of the infographic by selecting a color hex code, so I can manipulate the color of the background.

US8:   As a user, I want to add text to the infographic, so that the infographic is populated with text.

US9:   As a user, I want to change the size of the canvas of the infographic, so that I can tweak the infographic based on my needs.

US10:   As a user, I want to create pie charts on my infographic, so I can display data in a visual manner.

US11:   As a user, I want to create bar charts on my infographic, so I can display data in a visual manner.

US12:   As a user, I want to export the infographic as a PNG image, so I can use this to share with my stakeholders.

US13:   As a user, I want to use the values of indicators (direct and indirect) to populate my charts, so that my charts contain valuable data.

US14:   As a user, I want to use the values of indicators (direct and indirect) to populate parts of my text, so that my text contains valuable data.

US15:   As a user, I want to be able to use a head, body and foot section, so that my infographic is divided into sections.

US16:   As a user, I want to be able to create boxes, so that my infographic is divided into configurable sections.

US17:   As a user, I want to add a type basic element, so that my infographic is styled the same way as reports are style in OpenESEA.

Should have

US18:   As a user, I want to create picturegraphs on my infographic, so I can display data in a visual manner.

US19:   As a user, I want to define the background of the infographic by selecting color swatches, so I can manipulate the color of the background.

Could have

US20:   As a user, I want to define the background of the infographic by selecting an image from a database, so I can manipulate the looks of the background.

US21: As a user, I want to define the background of the infographic by uploading an image, so I can manipulate the looks of the background.

US22: As a user, I want to define the background of the infographic by selecting a pattern, so I can manipulate the looks of the background.

US23: As a user, I want to be able to select an image from a database and use this image on the infographic, so that I can manipulate the looks of the infographic.

US24: As a user, I want to be able to upload an image and use this image on the infographic, so that I can manipulate the looks of the infographic.

US25: As a user, I want to be able to select icons and shapes and use these on the infographic, so that I can manipulate the looks of the infographic.

US26: As a user, I want to create tables on my infographic, so I can display data in a visual manner.

US27: As a user, I want to create grouped bar charts on my infographic, so I can display data in a visual manner.

US28: As a user, I want to create stacked bar charts on my infographic, so I can display data in a visual manner.

US29: As a user, I want to export the infographic as a PDF file, so I can use this to share with my stakeholders.

US30: As a user, I want to export the infographic as a JPG image, so I can use this to share with my stakeholders.

Won't have

US31: As a user, I want to create scatterplots on my infographic, so I can display data in a visual manner.

US32: As a user, I want to create line charts on my infographic, so I can display data in a visual manner.

US33: As a user, I want to create Gantt charts on my infographic, so I can display data in a visual manner.

US34: As a user, I want to create area charts on my infographic, so I can display data in a visual manner.

US35: As a user, I want to create bubble charts on my infographic, so I can display data in a visual manner.

US36: As a user, I want to create map charts on my infographic, so I can display data in a visual manner.

5.2 Grammar of the domain-specific language

The grammar of our domain-specific language is specified in Xtext (Behrens et. al., 2008), a framework for developing programming languages and domain-specific languages. The entire grammar of the domain-specific language may be found in Appendix G. In Appendix H Table 11 and Table 12 are displaying which components of the tree of component analysis and which features of the analysis of online tools are incorporated in the domain-specific language and its interpreter.

```
1   grammar org.xtext.example.mydsl.IDSL with org.eclipse.xtext.common.Terminals
2   generate iDSL "http://www.xtext.org/example/mydsl/IDSL"
3
4   Infographic : InfographicElementType;
5
6   InfographicElementType:
7         ('type' ':' type='basic')?
8         & (('bgcolor' ':' bgcolor=Color) | ('bgpattern' ':'
   bgpattern=Pattern) | ('bgimage' ':' bgimage=ImageSrc))
9         & 'bgsize' ':' bgsize=SIZE_POS
10        & head=Head?
11        & boxes+=Box*
12        & texts+=Text*
13        & images+=Image*
14        & piecharts+=Piechart*
15        & barcharts+=Barchart*
16        & picturegraphs+=Picturegraph*
17        & foot=Foot?
18  ;
```

Figure 17: Basic infographic elements definition in Xtext

In line 6 to 18 the basic infographic element types are defined as depicted in Figure 17. It is possible to add 'type: basic' (line 7), add a background color, background pattern or background image(line 8) and adjust the infographic size (line 9). Next to that, it is possible to include a head section (line 10), boxes (line 11), text (line 12), images (line 13), piecharts (line 14), barcharts (line 15), picturegraphs (line 16) and a foot section (line 17). Four different cardinalities may be used in Xtext, no operator meaning exactly one (example at line 9), ? as operator meaning zero or one (example at line 7), * as operator meaning zero or more (example at line 11) and + as operator meaning one or more (not used in the grammar).

```
21  Head:
22        name="head" ':' ('off' |
23        ( BEGIN (
24              (('bgcolor' ':' bgcolor=Color) | ('bgpattern' ':'
   bgpattern=Pattern) | ('bgimage' ':' bgimage=ImageSrc))?
25              & ('size' ':' size=SIZE_POS)?
26              & ('position' ':' position=SIZE_POS)?
27              & ('logo' ':' logo=Logo)?
28              & ('title' ':' title=Title)?
29              & ('subtitle' ':' subtitle=Title)?))
30        END );
```

Figure 18: Head section defined in Xtext

In line 21 to 30 the head section is defined (see Figure 18), the head section is similar to the foot section at line 148 to 156 (see Figure 19). It is possible to turn the head section off by using 'head: off' (line 22). At line 23 and at line 30 you may find respectively a BEGIN and an END tag. This implies that everything between those two tags should be indented with one tabular space. These BEGIN and END tags are used multiple times in the grammar. Within the head section it is possible to define a background color, background pattern or a background image (line 24). In addition, it is possible to adjust the size (line 25) and the position (line 26) of

the head section. Next to that, you can include a logo (line 27), a title (line 28) and a subtitle (line 29) in the head section. The foot section differs from the head section because it does not offer the possibility of adding a title, subtitle or logo, however it does offer the possibility of using text (line 154) in the foot section.

```
148 Foot:
149        name="foot" ':' ('off' |
150        ( BEGIN
151            (('bgcolor' ':' bgcolor=Color) | ('bgpattern'
     ':' bgpattern=Pattern) | ('bgimage' ':' bgimage=ImageSrc))?
152            & ('position' ':' position=SIZE_POS)?
153            & ('size' ':' size=SIZE_POS)?
154            & ('text' ':' text=STRING)? )
155        END)
156 ;
```

Figure 19: Foot section defined in Xtext

The title and subtitle in the head section are following the same grammatical rules and may be found at line 33 to 42, as depicted in Figure 20. The element may be turned off (line 34). If the element is not turned off, it offers the possibility of adding alignment (line 36), a color (line 37), a font (line 38), a position (line 39), maximum width (line 40) and a value (line 41) in the title or subtitle element. A logo in the head section is defined at line 45 to 49. For the logo a size (line 47) and a position (line 48) is required.

```
33 Title:
34        ({Title} 'off' |
35        ( BEGIN
36            {Title} (('align' ':' align=AlignPossibilities)?
37            & ('color' ':' color=Color)?
38            & ('font' ':' font=FONT)?
39            & ('position' ':' position=SIZE_POS)?
40            & ('maxwidth' ':' maxwidth=INT)?
41            & ('value' ':' value=STRING)?))
42        END );
```

Figure 20: Title element and subtitle element defined in Xtext

Boxes are defined at line 52 to 58 (See Figure 20), they are used for formatting the infographic into distinguishable parts with each box having their own background color, background image or background pattern (line 55). In addition, the size and position may also be defined for each box.

```
52 Box:
53        name=BOXID ':'
54        BEGIN
55            ((('bgcolor' ':' bgcolor=Color) | ('bgpattern' ':'
   bgpattern=Pattern) | ('bgimage' ':' bgimage=ImageSrc))?
56            & ('size' ':' size=SIZE_POS)?
57            & ('position' ':' position=SIZE_POS))
58        END ;
```

Figure 21: Box element defined in Xtext

It is possible to add text to the infographic and this is defined at line 61 to 81 as depicted in Figure 22. Two options of text are available, content text (line 62) or titletext (line 72). The difference between those two is that content text is able to spread content over multiple lines where titletext can only use one line. While this is the only difference, it almost follows the same grammatical rules. Within titletext it is possible to define the following elements: alignment (line 74), a color (line 75), a font (line 76), a position (line 77), maximum width (line 78) and a value (line 79). The content text has one more element called lineheight (line 67). At line 84 to 90 the element image is defined. There are three elements mandatory when defining an image, namely: size, position and src (URL to the image). The definition of the image element may be found in Appendix G.

```
61  Text:
62        (name=TEXTID ':'
63        ( BEGIN
64              (('align' ':' align=AlignPossibilities)?
65              & ('color' ':' color=Color)?
66              & ('font' ':' font=FONT)?
67              & ('lineheight' ':' lineheight=INT)?
68              & ('position' ':' position=SIZE_POS)
69              & ('maxwidth' ':' maxwidth=INT)?
70              & ('value' ':' value=STRING)?)
71        END )) |
72        (name=TITLETEXTID ':'
73        ( BEGIN
74              (('align' ':' align=AlignPossibilities)?
75              & ('color' ':' color=Color)?
76              & ('font' ':' font=FONT)?
77              & ('position' ':' position=SIZE_POS)
78              & ('maxwidth' ':' maxwidth=INT)?
79              & ('value' ':' value=STRING)?)
80        END ))
81  ;
```

Figure 22: Text element and titletext element defined in Xtext

Furthermore, piecharts (line 93), barcharts (line 111) and picturegraphs (line 130) may be used in this domain-specific language. Only the definition of the piechart element is depicted in Figure 23, the definition of the barchart element and the picturegraph element may be found in Appendix G. All three charts are almost following the same grammatical rules. All three have the following elements in common: background color (line 96), colors used in the chart (line 97), data (line 98), padding (line 100), position (line 101), show legend (line 102), show title (line 103), size (line 105), title (line 106) and type (line 107). Within piecharts and barcharts it is also possible to adjust the legendstyle (line 99). Solely with piecharts it is possible to use the element showpercentage (line 104). Only with barcharts it is possible to use the elements gridscale (line 117) and showgrid (line 121). Solely with picturegraphs it is possible to use the element label (line 136).

```
93  Piechart:
94      name=PIECHARTID ':'
95      BEGIN
96          (('bgcolor' ':' color=Color)?
97          & ('colors' ':' colors=COLOR_CHARTS)?
98          & ('data' ':' BEGIN ChartData (piedata+=ChartData)* END)
99          & ('legendstyle' ':' legendstyle=LegendStyle)?
100         & ('padding' ':' padding=INT)?
101         & ('position' ':' position=SIZE_POS)
102         & ('showlegend' ':' showlegend=ShowOptionsOff)?
103         & ('showtitle' ':' showtitle=ShowOptionsOff)?
104         & ('showpercentage' ':' showpercentage=ShowOptionsOff)?
105         & ('size' ':' size=INT)?
106         & ('title' ':' title=STRING)?
107         & ('type' ':' type=PieType)?)
108     END ;
```

Figure 23: Piechart element defined in Xtext

Line 159 to 180 are the basic rules for the validation of the input. Within these lines it is defined among other things which input is allowed for SIZE_POS, COLOR and FONT. From line 183 to 189 the different ID's used for the numbering of the different elements are defined. Finally, at line 192 and 193 two synthetic tokens are defined which are used for expressing white spacing in the domain-specific language. Elements described in this paragraph are partially depicted in Figure 24. The entire grammar specification may be found in Appendix G.

```
159 terminal SIZE_POS : INT'x'INT;
161 Color : {Color} color=Color_Basic | {Color} COLOR_HEX;
162 enum Color_Basic : c1='green' | c2='red' | c3='blue' | c4='yellow' | c5='black' |
    c6='white' | c7='orange' | c8='purple' | c9='pink';
163 terminal COLOR_HEX :
    ('0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F')(
    '0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F');
164 terminal COLOR_CHARTS : (COLOR_HEX',')* COLOR_HEX;
169 FONT : (weight=Font_Weight)? FONT_SIZE family=Font_Family;
170 enum Font_Weight : b1='bold ' | b2='100 ' | b3='200 ' | b4='300 ' | b5='400 ' |
    b6='500 ' | b7='600 ' | b8='700 ' | b9='800 ' | b10='900 ';
171 terminal FONT_SIZE : ('0'..'9')+  'px ';
172 enum Font_Family : font1='Arial' | font2='Verdana' | font3='Times New Roman' |
    font4='sans-serif' | font5='Gotham';
173 enum AlignPossibilities: left | right | center;
174
183 terminal BOXID : 'box'INT;
184 terminal TEXTID : 'text'INT;
185 terminal TITLETEXTID : 'titletext'INT;
186 terminal IMAGEID : 'image'INT;
187 terminal PIECHARTID : 'piechart'INT;
188 terminal BARCHARTID : 'barchart'INT;
189 terminal PICTUREGRAPHID : 'picturegraph'INT;
190
192 terminal BEGIN: 'synthetic:BEGIN';
193 terminal END: 'synthetic:END';
```

Figure 24: Basic validation rules, ID numbering and white space tokens defined in Xtext

45

5.3 Domain-specific language dictionary

Within Table 6, a dictionary for our domain-specific language is depicted. Every possible input for our domain-specific language is described in this dictionary including their possible values. YAML files are used as input for the interpreter in the OpenESEA tool. For the reason that YAML files are whitespace sensitive, our domain-specific language is also whitespace sensitive. Tab indents are used to structure our domain-specific language and this comes back in our dictionary as an empty cell. An example of this phenomenon may be found below the input 'head:'. In addition, in some cases it is possible to have multiple instances of the same element. This can be established by using a numeric suffix behind the input definition, for example barchart(n) where (n) is an absolute number starting at 1.

Table 6: Domain-specific language for infographics dictionary

Input		Description	Possible values
type:		type: basic will render a default infographic with the looks and feel of an ordinary OpenESEA tool report.	basic
bgsize:		The size of the infographic defined in pixels.	widthxheight
bgcolor:		The background of the infographic is displayed with a single color.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
bgpattern:		The background of the infographic is displayed with a single image in a repeating pattern.	patternx (x is a number) refering to an existing pattern in the database or the direct URL to an image stored online in the format of .png, .jpg or .gif.
bgimage:		The background of the infographic is displayed with a single image spreading to the size of the infographic.	direct URL to an image
head:		The head section containing the title and subtitle which may be turned off.	off
	bgcolor:	The background of the head section is displayed with a single color.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
	bgpattern:	The background of the head section is displayed with a single image in a repeating pattern.	patternx (x is a number) refering to an existing pattern in the database or the direct URL to an image stored online in the format of .png, .jpg or .gif.

Input			Description	Possible values
bgimage:			The background of the head section is displayed with a single image spreading to the size of the infographic.	The direct URL to an image stored online in the format of .png, .jpg or .gif.
size:			The size of the head section defined in pixels.	widthxheight
position:			The position of the head section relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
title:			The possibility to use a title on the infographic.	off
	align:		The way the title is horizontally aligned.	left, right, center
	color:		The color in which the title is displayed.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
	font:		The font that will be used to display the title.	INTpx font-name
	position:		The position of the title relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
	maxwidth:		The maximum width the title may be defined in pixels.	Any number
	value:		The text that is displayed in title. +organisation+ will be replaced by the organisation's name. +report+ will be replaced by the name of the report belonging to the infographic.	Any text
subtitle:			The possibility to use a subtitle on the infographic.	off
	align:		The way the subtitle is horizontally aligned.	left, right or center
	color:		The color in which the subtitle is displayed.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
	font:		The font that will be used to display the subtitle.	INTpx font-name
	position:		The position of the subtitle relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
	maxwidth:		The maximum width the subtitle may be defined in pixels.	Any number

Input			Description	Possible values
		value:	The text that is displayed in subtitle. +organisation+ will be replaced by the organisation's name. +report+ will be replaced by the name of the report belonging to the infographic.	Any text
box(n):			The possibility to display a box on the infographic. (n) meaning a number starting at 1. Position is mandatory.	-
	bgcolor:		The background of box(n) is displayed with a single color.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
	bgpattern:		The background of box(n) is displayed with a single image in a repeating pattern.	patternx (x is a number) refering to an existing pattern in the database or the direct URL to an image stored online in the format of .png, .jpg or .gif.
	bgimage:		The background of the box(n) is displayed with a single image spreading to the size of the infographic.	The direct URL to an image stored online in the format of .png, .jpg or .gif.
	size:		The size of box(n) defined in pixels.	widthxheight
	position:		The position of box(n) relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
titletext(n):			The possiblity to display titles of paragraphs on the infographic. (n) meaning a number starting at 1. Position is mandatory.	-
	align:		The way the titletext(n) is horizontally aligned.	left, right, center
	color:		The color in which the titletext(n) is displayed.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
	font:		The font that will be used to display the titletext(n).	INTpx font-name
	position:		The position of titletext(n) relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
	maxwidth:		The maximum width the titletext(n) may be defined in pixels.	Any number
	value:		The text that is displayed in titletext(n).	Any text

Input	Description	Possible values
text(n)	The possibility to display text on the infographic. (n) meaning a number starting at 1. Position is mandatory.	-
align:	The way the text(n) is horizontally aligned.	left, right, center
color:	The color in which the text(n) is displayed.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
font:	The font that will be used to display the titletext(n).	INTpx font-name
position:	The position of text(n) relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
maxwidth:	The maximum width the text(n) may be defined in pixels.	Any number
value:	The text that is displayed in text(n).	Any text
lineheight:	The line-height when text is displayed over multiple lines.	Any number or auto
image(n):	The possibility to display an image on the infographic. (n) meaning a number starting at 1. Position, size and src are mandatory.	-
position:	The position of image(n) relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
size:	The size of image(n) defined in pixels.	widthxheight
src:	The direct URL to an image stored online for image(n).	The direct URL to an image stored online in the format of .png, .jpg or .gif.
piechart(n):	The possiblity to display piecharts of paragraphs on the infographic. (n) meaning a number starting at 1. Position and data are mandatory.	-
bgcolor:	The background of piechart(n) is displayed with a single color.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
colors:	The colors used for displaying pie slices in piechart(n).	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc) for each pie slice
data:	The data used populating piechart(n).	(String / Indicator : String / INT / Indicator)n

Input		Description	Possible values
	legendstyle:	The way the legend of piechart(n) is displayed.	1 for two columns or 2 for one column
	padding:	Padding is the space between the contents of piechart(n) and its border.	Any number
	position:	The position of piechart(n) relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
	showlegend:	Whether the legend of piechart(n) is displayed.	off
	showtitle:	Whether the title of piechart(n) is displayed.	off
	showpercentage:	Whether the percentages of piechart(n) are displayed.	off
	size:	The radius of piechart(n) in pixels.	radius
	title:	The text that is displayed in the title of piechart(n)	Any text
	type:	The type of piechart(n) that is displayed.	pie or donut
barchart(n):		The possiblity to display barcharts of paragraphs on the infographic. (n) meaning a number starting at 1. Position and data are mandatory.	-
	bgcolor:	The background of barchart(n) is displayed with a single color.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
	colors:	The colors used for displaying bars in barchart(n).	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc) for each bar
	data:	The data used populating barchart(n).	String / Indicator : String / INT / Indicator
	gridscale:	Adjusting the gridscale will influence the distance between grid lines.	Any number
	legendstyle:	The way the legend of barchart(n) is displayed.	1 for two columns or 2 for one column
	padding:	Padding is the space between the contents of barchart(n) and its border.	Any number
	position:	The position of barchart(n) relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
	showgrid:	Whether the grid of barchart(n) is displayed.	off

Input		Description	Possible values
	showlegend:	Whether the legend of barchart(n) is displayed.	off
	showtitle:	Whether the title of barchart(n) is displayed.	off
	size:	The size of barchart(n) in pixels.	widthxheight
	title:	The text that is displayed in the title of barchart(n)	Any text
	type:	The type of barchart(n) that is displayed.	horizontal or vertical
picturegraph(n):		The possiblity to display picturegraphs of paragraphs on the infographic. (n) meaning a number starting at 1. Position and data are mandatory.	-
	bgcolor:	The background of barchart(n) is displayed with a single color.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)
	colors:	The colors used for displaying bars in barchart(n).	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc) for each bar
	data:	The data used populating barchart(n).	String / Indicator : String / INT / Indicator
	padding:	Padding is the space between the contents of barchart(n) and its border.	Any number
	position:	The position of barchart(n) relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
	showlegend:	Whether the legend of barchart(n) is displayed.	off
	showtitle:	Whether the title of barchart(n) is displayed.	off
	size:	The size of barchart(n) in pixels.	widthxheight
	title:	The text that is displayed in the title of barchart(n)	Any text
	type:	The type of picturegraph(n) that is displayed.	bin or electric
	label:	Defines the label displayed in the legend right to each value.	Any text
foot:		The foot section which may be turned off.	off
	bgcolor:	The background of the foot section is displayed with a single color.	6-digit hex code (000000, 0000ff, etc) or color name (green, blue, etc)

Input	Description	Possible values
bgpattern:	The background of the foot section is displayed with a single image in a repeating pattern.	patternx (x is a number) refering to an existing pattern in the database or the direct URL to an image stored online in the format of .png, .jpg or .gif.
bgimage:	The background of the foot section is displayed with a single image spreading to the size of the infographic.	The direct URL to an image stored online in the format of .png, .jpg or .gif.
size:	The size of the foot section defined in pixels.	widthxheight
position:	The position of the foot section relative to the top left corner of the infographic defined in pixels.	X-coordinatexY-coordinate
text:	The text that is displayed in foot section. +organisation+ will be replaced by the organisation's name. +report+ will be replaced by the name of the report belonging to the infographic.	Any text

5.4 Engineering of the feature in OpenESEA

An interpreter of the domain-specific language is implemented in the OpenESEA tool. The software architecture is adapted in a way so that it is possible to use, store and retrieve infographic specifications. Creating an infographic follows the same approach as creating a company, a network or a report. At the left side of the tool the possibility exist to create infographics and clicking on it will lead to the screen depicted in Figure 25. When creating an infographic, you will have to specify the organisation, select a report from exisiting reports and define a name for the infographic.

When an infographic is created you will be redirected to the screen depicted in Figure 26. Within this screen the user is asked if he wants to use the infographic specification defined by the organisation, or the infographic specification defined by the network or to use a new

Figure 25: Creating an infographic in OpenSEA

52

Figure 26: Selecting an infographic specification in OpenSEA

infographic specification which the user can upload right away. The software architecture is adapted so that on each level (organisational level, network level or infographic level) an infographic specification can be defined in order to be used when generating an infographic. Adapting the software architecture in this way ensures that organisations or networks are able to produce multiple infographics with the same data elements populating the infographic. Users are provided with a list summarizing all infographics created for a company as depicted in Figure 27. When opening an infographic the infographic will be generated on the fly as you visit the webpage. To construct the infographic we're making use of the canvas element present in HTML5[11]. Several disadvantages exist (described in Section 8.2 on future research) when making use of the canvas element. However, it proved to be a suitable technique to create a working proof of concept.

Infographics

Infographic	Report	Created by	Last updated
Infographic 2019 - 3rd	2019	Tijmen Derikx	11-02-2020
Testcase GSI	2019	Tijmen Derikx	05-12-2019
Testcase FirstHorizon	2019	Tijmen Derikx	30-11-2019

Figure 27: A list of all infographics created for a company.

---

[11] https://www.w3schools.com/html/html5_canvas.asp

5.5 Infographic specification examples for OpenESEA tool

Two infographic specifications that can be used in the OpenESEA tool are explained below. The first specification is an example making use of the type: basic element. This element is used whenever you want a default infographic with the looks and feel of an ordinary OpenESEA tool report. The second specification is more advanced explaining the general use of the domain-specific language and the interpreter in the OpenESEA tool.

The only rule that is required to generate a basic infographic is the following (note: all elements defined before or after the type: basic declaration are discarded):

```
1  type: basic
```

Figure 28: Type: basic declaration in an infographic specification.

Based on this declaration the infographic below will be generated. Direct and indirect indicators are used to populate this infographic:
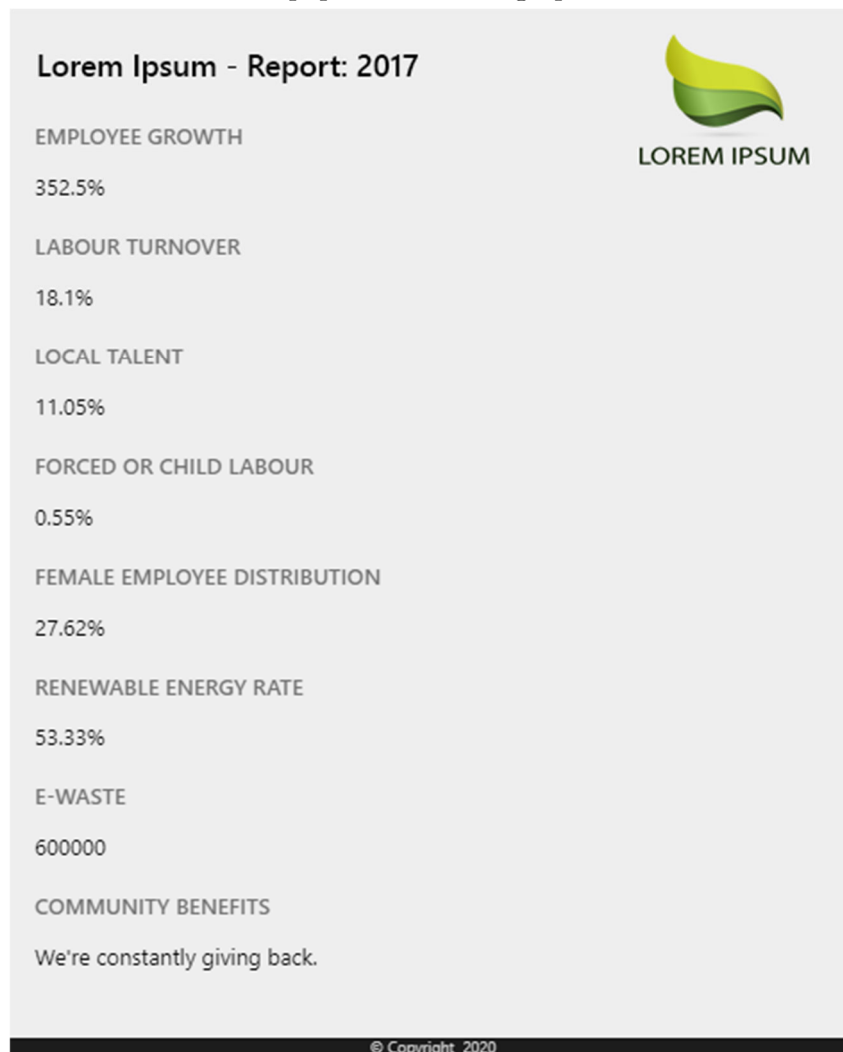
Figure 29: Basic infographic generated by the OpenESEA tool.

In Figure 30, the generated infographic of a more advanced infographic specification is depicted. The infographic specification may be found in Appendix I and interesting parts will be highlighted below.
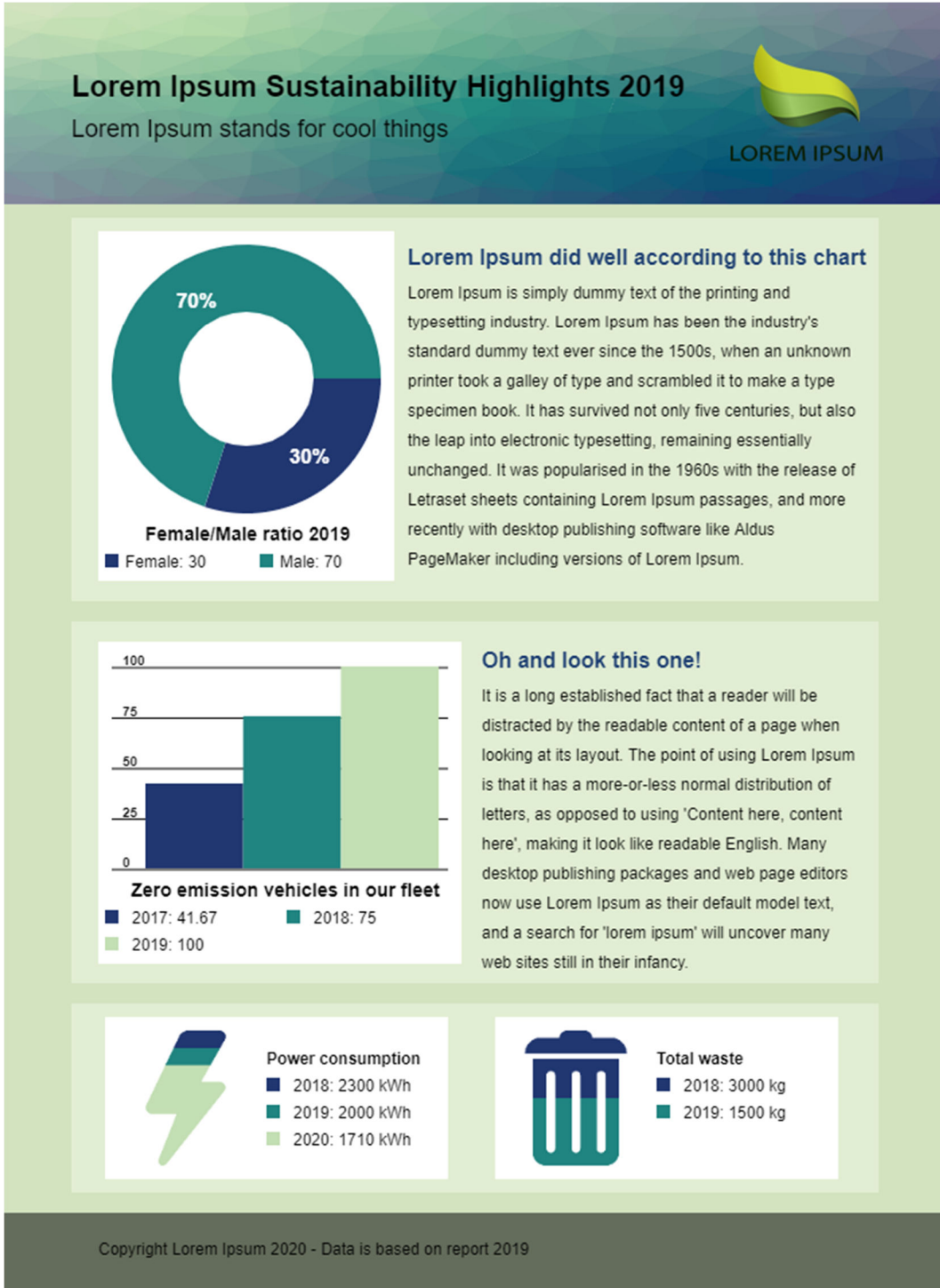
Figure 30: The generated infographic from the more advanced infographic specification depicted in Appendix I.

At first, the background color (line 1) and background size (line 2) is defined. These two elements are fundamental for each infographic, whenever they are not defined a predefined value from OpenESEA will be used. The infographic contains three boxes (line 3 to 14) which are defined afterwards. For each box a background color, position and size is defined.

```
1   bgcolor: d3e3c0
2   bgsize: 700x960
3   box1:
4     bgcolor: e2eed3
5     position: 50x160
6     size: 600x285
```

Figure 31: Background color, background size and one box definition

The head section depicted in Figure 32 is defined at line 15 to 25 and is using a background image as can be seen in line 15. Besides that, it is displaying a logo (line 17 to 19), a title (line 23 to 25) and a subtitle (line 20 to 22).

```
15  head:
16    bgimage: 'https://www.umib.net/wp-content/uploads/2019/06/ranger-4df6c1b6.png'
17    logo:
18      position: 535x48
19      size: 100x60
20    subtitle:
21      position: 50x100
22      value: 'Lorem Ipsum stands for cool things'
23    title:
24      position: 50x70
25      value: 'Lorem Ipsum Sustainability Highlights +report+'
```

Figure 32: Head section defined including a logo, subtitle and title.

Afterwards a bar chart depicted in Figure 33 is defined at line 30 to 39. A different color scheme has been used and is defined at line 31. Three data elements are defined from line 32 to 35, including one data element that is using a value from an indicator. Next to that, it is also defining gridscale (line 36), position (line 37), size (line 38) and a title (line 39) for the barchart.

```
30  barchart1:
31    colors: 213771, 208480, c3dfb4
32    data:
33      2017: S38
34      2018: '75'
35      2019: '100'
36    gridscale: '25'
37    position: 70x475
38    size: 250x170
39    title: 'Zero emission vehicles in our fleet'
```

Figure 33: Barchart derfinition including a different color scheme, data and a title.

6 Validation

In this section, the activities A8 (validate by predefined test cases), A9 (conduct experiment) and A10 (expert assessment) are described.

6.1 Validate by predefined test cases

To validate our domain-specific language for infographics we selected ten infographics from the tree of component types analysis (Section 3.3). For each infographic, an infographic specification is constructed. With this infographic specification the interpreter in the OpenESEA tool is able to generate a similar infographic to the original one. These ten predefined test cases[12] are used in the experiment described in the next section. One test case, including the original infographic and its corresponding generated infographic is displayed on the next page in Figure 34. Minimal changes may be observed in box sizes, fonts and positioning of elements.

Several shortfalls were detected while constructing the ten test cases. One example of such a shortfall was that it was not possible to create donut charts which is present in one of the test cases. The difference between pie charts and donut charts is only one extra circle and therefore we introduced a new element to the piechart definition in our domain-specific language namely type. With this element it is possible to define whether it is a regular pie chart or a donut chart. Another shortfall was that it was not possible to adjust the style the legend of barcharts, piecharts and picturegraphs that were defined. Since this was needed to construct one of the test cases we incorporated a new element in the domain-specific language namely legendstyle.

In our tree of component types analysis we have observed many infographics using picturegraphs. As described earlier, a picturegraph represents data by using symbols and pictures. Variations are therefore unlimited while every symbol may be used for representing data. In the test cases we noticed we applied a quick and dirty way of specifying infographics regarding picturegraphs using text and images instead of the creation of a new type of picturegraph.

6.2 Experiment definition

This section introduces the experimental subject, experimental instruments and the experiment protocol. After that, we discuss the validity of the experiment by describing what threats are observed and potentially avoided. Finally, we present how the data will be analysed.

---

[12] The ten test cases including the original infographic, the rendered infographic and the infographic specification may be found on the following url:
http://tiny.cc/TestcasesInfographics

Left infographic (original):

2017/18
SUSTAINABILITY REPORT

32% Reduction of Scope 1 and 2 greenhouse gas emissions since 2010

Perfect score from the Human Rights Campaign Foundation's Corporate Equality Index for two consecutive years — **100%**

+6.7% Efficiency improvement in transportation CO_2 emissions due to innovative, bio-based packaging materials—like bamboo and sugar cane—driving down packaging weight

Named one of the 100 Best Companies for Working Mothers by *Working Mother* magazine — Best Company for Working Mothers

CO_2 59 MT Total CO_2 emissions saved by Lenovo's Low-Temperature Solder process at the end of FY 2017/18

Global Week of Service directly impacted 33,000 consumers across 19 countries, benefiting minority populations and increasing access to STEM education

Lenovo.

Right infographic (generated):

2017/18
SUSTAINABILITY REPORT

32% Reduction of Scope 1 and 2 greenhouse gas emmisions since 2010

Perfect score from the Human Rights Campaign Foundation's Corporate Equality Index for two consecutive years — **100%**

+6.7% Efficiency improvement in transportation CO2 emissions due to innovative, bio-based packaging materials - like bamboo and sugar cane - driving down packaging weight

Named one of the 100 Best Companies for Working Mothers by Working Mother magazine — Best Company for Working Mothers

CO2 59MT Total CO2 emissions saved by Lenovo's Low-Temperature Solder process at the end of FY 2017/18

Global Week of Service directly impacted 33,000 consumers across 19 countries, benefiting minority populations and increasing access to STEM education

Lenovo.

Figure 34: One of the testcases displaying the original infographic on the left and its corresponding generated infographic on the right.

6.2.1 Experimental subjects

As seen in Figure 35 and Figure 36, the experimental subjects (N=40) consisted of 25 male subjects (62.5%) and 15 female subjects (37.5%). The majority of the experimental subjects were aged between 25 and 34 years old, namely 28 subjects. Ten subjects were aged between 18 and 24 years old and two subjects were aged between 35 and 44. Almost half of the experimental subjects successfully completed a master's degree, 12 experimental subjects successfully completed a bachelor's degree, eight experimental subjects had high school degree or equivalent as highest degree of education completed and one experimental
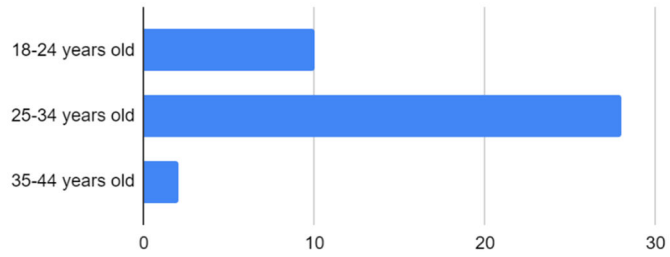
Figure 35: Bar chart of the ages of the experimental subjects.
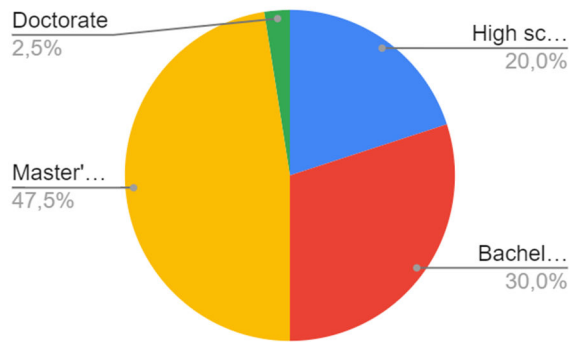
Figure 36: Pie chart of participants highest degree of education completed.

subject completed their doctorate. All participants are randomly assigned to version one or two ensuring that each version has the same number of participants, namely 20.

6.2.2 Experimental instruments

Three instruments are used throughout the experiment. The first instrument is a document with ten infographics, five infographics are the original infographics found in the tree of component types analysis and five infographics are generated by the OpenESEA tool. The infographics used for this experiment may be found in the online version of the survey. The second instrument is a document explaining the different terms used in rating the infographics, described in the next section. The third instrument is a survey consisting of two parts. A demographic part, which provides insights on the background of the experimental subjects. The second part of the survey consist of 11 questions for each infographic on the visual attractiveness and whether the infographic is manually designed or whether it is generated by a software tool. The full version of the survey may be viewed online[13]. Based on question four in the survey the experimental subject will fill in version 1 (question 6 to 35) or version 2 (question 36 to 65).

---

[13] The full survey can be found on the following url: http://tiny.cc/SurveyInfographics

Visual attractiveness

We will measure the visual attractiveness of the infographics by using a questionnaire that was subjected by Lavie and Tractinsky (2004). Lavie and Tractinsky found the following items generated the most reliable and valid measure of "classical aesthetics". "Classical aesthetics" stands for orderly and clear design that is very close to the views of many usability professionals. Subjects rate the following items on a five-point, agree-disagree Likert scale: aesthetic, symmetrical, pleasant, organized (clear) and clean. Lavie and Tractinsky also described the concept of "expressive aesthetics", representing creativity and originality of designers. The following items are proposed to survey this: creative, using special effects, original, sophisticated and fascinating. Subject also rate these items on a five-point, agree-disagree Likert scale. The sum of the ten aforementioned items will be used as the visual attractiveness score. We will be using the following definitions for each item within this experiment:

- Aesthetic design refers to the visual attractiveness of a product.
- Symmetrical design refers to beauty due to balanced proportions.
- Pleasant design refers to an enjoyable feeling due to the design.
- Organized design refers to a design arranged or structured in a systematic way.
- Clean design refers to a design that is precise and consistent.
- Creative design refers to the inventivity of the design.
- Sophisticated design refers to elegant, well-thought out and carefully planned design.
- Original design refers to the uniqueness of the design.
- Using special effects refers to the use of illusions or visual tricks to enhance the design.
- Fascinating design refers to a design that is extremely interesting.

6.2.3 Experiment protocol

The overall design type of this experiment is 'completely randomized design' meaning that all experimental subjects are randomly assigned to the treatments (Wohlin et. al., 2012). The experiment starts with an introduction, where the subject receives a short introduction to the experiment and the topic of research. In addition, they will also receive a version number. The subject starts by filling in the demographic survey, including questions on age, sex, level of education and their version number. To ensure every experimental subject has the same definitions in mind when rating the infographic, we provided the experimental subjects with explanations for each term in the survey. Afterwards, the experimental subjects will answer 11 questions for each infographic concluding the experiment. The experiment protocol is depicted in Figure 37. The order of original and generated infographics is randomised and is the same for all subjects across a version. In addition, the two versions are a complete opposite of one another meaning that

when version one starts off with an original infographic that version two starts off with a generated infographic.
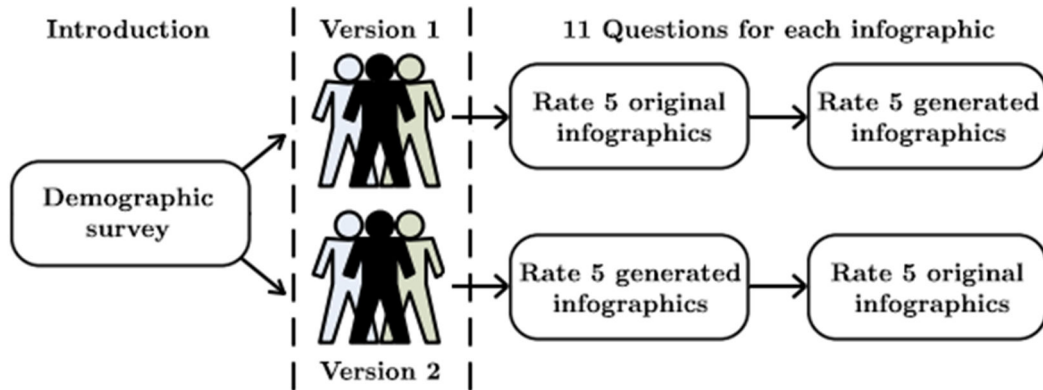
Figure 37: Visualisation of the experiment protocol

6.2.4 Validity

The experiment described comes with several threats which must be elaborated. Table 7 shows a summary of those threats, categorised into four different types, being conclusion validity, construct validity, internal validity and external validity. The threats and categorisations are based on the work of Wohlin et. al. (2012) and Panach et. al. (2015). For each threat it is mentioned if it is mitigated, suffered or whether it is not applicable to this experiment. In addition, we provide an explanation that clarifies how we dealt with each threat. The most critical threat we have suffered is maturation within 'Internal validity'. It may have been that participants felt slightly bored or less concentrated toward the end of the experiment since the experiment was set up with a large amount of repeating elements.

Table 7: Threats to the validity of the experiment; adapted from Wohlin et. al. (2012) and Panach et. al. (2015)

Type of threat	Status	Threat	Due to	How we dealt with it
Conclusion validity	Mitigated	Fishing for results	The observers search for a specific result during the research.	All available data generated is analysed and used, and the outcome is not directly relevant for the observers.
	Mitigated	Random disturbances in experimental setting	External elements from the experimental setting can interfere with the results.	Based on observations, no remarkable disturbances occurred during the experiment.
	Mitigated	Random heterogeneity of subjects	The background of subjects can be too heterogeneous due to random selection.	The experiment was sent out to friends and family in the same age group, and we also gathered participants on the university of Utrecht in the same age group. Therefore, we mitigated this threat.
	Mitigated	Reliability of measures	The experiment's validity depends on the reliability of the measures.	By clearly formatting the questions and instructions, the experiment is explained extensively. When subjects conduct the same experiment, their ratings are not likely to be different.
	Suffered	Low statistical power	The sample size is too small.	Significant results can be obtained with this sample size, but for it to be more powerful, more respondents are required.
Construct validity	Suffered	Mono-operation bias	A single type of measures or observations can lead to measurement bias.	While we are asking the subjects to evaluate ten infographics in the same way, we might have suffered this threat.

Type of threat	Status	Threat	Due to	How we dealt with it
	Suffered	Mono-method bias	Measurement bias can occur due to experiments containing a single type of measure.	We only measure scores for visual appearance for each infographic and therefore we suffered this threat.
	Mitigated	Hypothesis guessing	Subjects guess and act on the supposed hypothesis of the experiment.	This threat is mostly mitigated by not including the goal of the research and its research questions. However, the final question on the design method of the infographic may have provided a clue on the research goal and its research questions.
	Mitigated	Interaction of treatment and testing	The testing itself may make the subjects more sensitive or receptive to the treatment.	Mostly mitigated because the experiment was conducted without a time constraint.
	Mitigated	Interaction of different treatments	It is possible that treatments of various studies may interact with each other.	This experiment only focuses on single treatments, so the results are not affected by multiple treatments.
	Mitigated	Evaluation apprehension	Subjects are anxious about being assessed.	This threat is mostly mitigated since the subjects could participate in solitude and completely anonymous. We as observers had no incentive to evaluate the subjects.
Internal validity	Mitigated	Mortality	Subjects drop out before completion of the experiment	None of the participants left before ending the experiment and there were no missing answers in the responses.

Type of threat	Status	Threat	Due to	How we dealt with it
	Suffered	Subject motivation	Subjects that are less motivated can potentially achieve worse results than highly motivated subjects.	There was no incentive for the subjects to perform well, therefore we suffered this threat.
	Not applicable	History	The same subject receives different treatments at different times.	There is one treatment for each subject, so this threat is not applicable to this experiment.
	Mitigated	Selection	The selection of subjects can affect the results of this research.	This threat is mostly mitigated by sending the experiment to friends and family in the same age group, and by gathering participants on the university of Utrecht with similar backgrounds.
	Mitigated	Compensatory rivalry	Subjects obtaining less desirable treatments can be motivated to adjust the outcomes of the experiment.	Both treatments are equally desirable and therefore this threat is avoided. In addition, the subjects don't know which treatment they perform.
	Mitigated	Resentful demoralisation	Subjects receiving treatments that are less desirable may perform worse.	Both treatments are equally desirable and therefore this threat is avoided. In addition, the subjects don't know which treatment they perform.
	Suffered	Maturation	The subjects react differently as time passes.	This threat has been suffered. It may have been the participants felt slightly bored or less concentrated towards the end of the experiment.

Type of threat	Status	Threat	Due to	How we dealt with it
	Not applicable	Learning of objects	Subjects can obtain knowledge coming from the first treatment and use it for other treatments.	There is one treatment for each subject, so this threat is not applicable to this experiment.
External validity	Mitigated	Interaction of treatment and history	The experiment is conveyed on a particular day or time which potentially affect the results.	This threat is avoided while there are no days or times that could affect the results.
	Suffered	Interaction of treatment and selection	The subject population is not representative to generalise.	While the experiment was conducted on the university and was sent out to friends and family with the same background, we might have suffered this threat.
	Mitigated	Interaction of setting and treatment	The objects used in the experimental setting makes it not representative for the real world.	We are using real examples representing the real world in our experiment.

6.2.5 Data analysis

The first ten questions are used to calculate the visual attractiveness score for each infographic. First, for each subject and infographic the sum of the answers to the first ten questions will be calculated. Afterwards, for each subject the sum of the answers for the original infographics will be calculated and the sum of the answers for the generated infographics will be calculated. After the data is converted, the data will be imported in SPSS where our statistical tests are performed. We will apply a Kolmogorov-Smirnov test to check whether the data gathered is distributed normally. If the data is distributed normally, we will apply a paired samples t test to check whether there are any statistical differences between the scoring of original infographics or the scoring of the corresponding generated infographics. We will be using the following hypotheses to test whether there are any statistical differences between the original infographics and the generated infographics:

H_0: The visual attractiveness score for original infographics is similar to the visual attractiveness score for the generated infographics.

H_1: The visual attractiveness score for original infographics differs from the visual attractiveness score for the generated infographics.

When the results of the paired sample t test are significant ($p<0.05$), we will check the effect size by calculating Cohen's d (Cohen, 2013). Cohen's d is defined as the difference between two means divided by a standard deviation of the data. According to Cohen, the meaning of the effect size is as follows: more than 0.8 is a large effect; from 0.79 to 0.5 is a moderate effect; from 0.49 to 0.2 is a small effect.

The eleventh question for each infographic will be handled as a classification problem. Subjects can provide an answer to this question which is right or wrong. Based on this premise we can calculate the classification accuracy. By checking the results to the eleventh question and by checking the calculated classification accuracy we might have evidence that there is no difference in perceiving original infographics or their corresponding generated infographics.

6.3 Experiment results

A Kolmogorov-Smirnov test indicates that the visual attractiveness scores for original infographics and their corresponding generated infographics follow a normal distribution, $D(40)=0.093$, $p=0.200$.

A paired samples t test was conducted to compare the visual attractiveness score of original infographics and generated infographics. There was no significant difference in the visual attractiveness score of original infographics (M=138.7, SD=29.6) and the visual attractiveness score of generated infographics (M=140.4, SD=32.7); $t(39)=-0.406$, $p=0.687$. Therefore, we retain the null hypothesis that the visual attractiveness scores of original infographics and their corresponding generated infographics are similar. Using the tool G*power tool (Faul et. al., 2007)

we found that for a paired sample t test, we need a sample size of 32 units for an effect size of 0.6 to get a power of 95%. Meaning that when another researcher wants to replicate this experiment, the experiment needs to be conducted with a sample size of at least 32 units in order to get significant results.

Within Table 8 the results for the eleventh question on the perceived design method is displayed. This table shows that the design method of original infographics and generated infographics are perceived similarly. There is almost an equal ratio on how the experimental subjects perceived the original and the generated infographics. The classification accuracy between design methods is the complete opposite of each other. This is logical, since we observed an equal ratio between manually designed infographics and generated by a tool for both original infographics and generated infographics.

	Manually designed	Generated by using a tool	Classification accuracy
Original infographics	78	122	39%
Generated infographics	80	120	60%

Table 8: Results to the eleventh question of the survey regarding the perceived design method.

6.4 Expert assessment

Another way of validating our proposed domain-specific language and its interpreter will be performed by executing an expert assessment. The expert is someone who is active in the field of social and environmental accounting and is concerned with refactoring the OpenESEA tool. The goal of this expert assessment is to gain insight about benefits and drawbacks of the proposed domain-specific language and its interpreter. The following questions are used to guide the expert assessment: What are the strengths and weaknesses of the infographic feature and its underlying domain-specific language? In addition, would you consider adopting this tool and at what point in time?

One of the advantages of this feature according to the expert is that infographics are a good way of visualising data for every stakeholder in an organisation. You can easily generate one infographic regarding a financial view and generate another infographic regarding a social view. In addition, the expert mentioned that infographics are a good fit for displaying certifications. At this moment, certifications is a feature that is incorporated in OpenESEA. Certifications are used to display which steps an organisation has to make to improve the current situation.

One of the disadvantages of this feature according to the expert is that the feature is not validating the input. Without a proper validation on the data, it is possible that incorrect infographics are generated.

The expert mentioned that the infographic feature will be part of the refactored OpenESEA tool since it's a valuable addition to the tool. At what point in time they are going to adopt this feature was hard to tell. Depending on the roadmap and resources available this feature will be incorporated in the refactored tool in one to one and a half years.

7 Conclusion

Based on the interview we identified the current process of creating infographics as a result of social and environmental accounting. The process is the following: the data of a social and environmental account is stored in a database. The data is exported to a third party which produces an infographic. The interview also provided us insights on the pain points in the current process. The skills to create infographics are not present in the company and therefore an external party is hired to create the infographics. The data used for populating the infographic needs to be exported to the external party hereby interrupting the continuous process.

Besides the popular image-editing software (Adobe Photoshop or Gimp) to create infographics with, there are tools available online to create infographics. A total of 13 tools were identified and due to exclusion criteria only ten were analysed in this research. The analysis of online tools led to 28 distinguishable features. In addition, the analysis of social and environmental accounting infographics found online yielded a tree of component types, depicting the different elements populating an infographic. Those two analyses were the basis for the construction of 36 user stories used as the requirements for our domain-specific language. The domain-specific language was constructed in Xtext, which proved to be a suitable framework for the development of a domain-specific language.

Unfortunately, the statistical tests performed in the experiment we conducted and described in Section 6 rendered no significant results. Therefore, we cannot conclude that there are any statistical differences between perceiving original infographics and their corresponding generated infographics. In addition, we also do not have evidence that there are no statistical difference. Based on the answers given to the eleventh question regarding the design method we concluded that the design method of original infographics and their corresponding generated infographics are perceived similarly. However, we do not have any statistical evidence since no statistical test was deemed fit to analyse the data. Based on the expert assessment we conclude that the infographic feature is a valuable feature for the OpenESEA tool. A disadvantage of the feature is that there is no validation on the data used in infographics and therefore may lead to incorrect generated infographics.

8 Discussion

Within this section the limitations and future research for this research are described.

8.1 Limitations

The functionalities of Xtext are highly configurable and within this research we did not use all the functionalities of Xtext. We did offer some validation rules to validate input in our domain-specific language[14]. However, Xtext with its generator classes also offers the generation of additional file types. Right now, these generator classes are not used in our domain-specific language for the reason that several bugs emerged when using these classes. Right now, Xtext is able to validate the infographic specification and when it is valid, you need to manually copy the specification and create a YAML file yourself in Notepad or some other text editor. This can be improved by generating a YAML file by Xtext, which can be instantly used by the OpenESEA tool.

Another limitation on the DSL is the broadness of the language. A total of 28 features were identified in the feature analysis of existing tools and only 11 features were incorporated in the domain-specific language and its interpreter. From the 12 different form of charts identified in the different analyses only piecharts, barcharts and two forms of picturegraphs were incorporated in the domain-specific language and its interpreter. From the three different export features identified in the different analyses only one (F27: export as PNG) is incorporated in the domain-specific language and its interpreter.

At this moment, when using images in the infographic specification the user needs to specify an URL linking to the image. It is not possible to upload images directly to the OpenESEA tool and therefore images used in the infographic specification are hosted on external websites, e.g. Imgur or Flickr. Ideally, users are able to upload images directly to the OpenESEA tool and specify a name so that users are able to use these images and their file names (instead of an URL) on the infographic specification. This limitation corresponds with features F4 and F12 identified in section 4.

A final limitation to the interpreter of the domain-specific language is that infographics at this moment are rendered by making use of the HTML5 canvas element. Several disadvantages are observed during the development of the interpreter of the domain-specific language. At first, the HTML5 canvas element is heavily dependent on javascript. While there are plans to refactor the OpenESEA tool from Javascript to Python, this design choice may not be future proof. In

---

[14] The full Xtext project including the validation classes may be found at Github (https://github.com/TijmenDe/idsl).

addition when rendering large infographics with the HTML5 canvas element the loading times are high.

A final limitation to this research is that all analyses and research activities performed are regarding infographics in the context of social and environmental accounting. We assume infographics used in this context are incorporating the same elements as infographics that are designed for other purposes. Therefore, we assume that the domain-specific language developed in this research is generalizable for other infographics as well. Since we have not researched this, this is only an assumption and future research is needed to validate this assumption.

8.2 Future research

In this research we created a domain-specific language to declare infographics with. The user is obliged to install Xtext and Eclipse to use and run the domain-specific language. When using Xtext it is possible to define an infographic by coding each element in the infographic by hand. A more elegant solution would be that the user could design an infographic making use of a drag-and-drop editor. This would lower the difficulty of using our domain-specific language even more. Eventually, we envision a tool that is able to generate infographic specifications based on what the user created in a drag-and-drop editor and to use these specifications in OpenESEA without interference of Xtext, Eclipse or the user. Further research is needed to explore this research possibility. Additional research may regard design suggestions supplied by the OpenESEA based on benchmarking. For example, the OpenESEA tool notices that the result on a specific indicator is better than the average result for that specific indicator, therefore the OpenESEA tool suggests that this indicator is displayed on the infographic.

Based on the analyses performed in this research the OpenESEA tool first described by Bik (2018) is extended upon. In concurrence with this research project another student from the University of Utrecht is investigating and implementing a feature to include surveys in the OpenESEA tool. Therefore, the OpenESEA tool is becoming a more complete tool to use when performing social environmental accounting. However, the foundation of the OpenESEA tool may be improved by executing a research regarding the different topics and different indicators in the different methods. Based on observations it seems that there is a lot of overlap in topics and indicators between the different methods however, this is not yet properly researched and described.

9 References

Adams, C. A., Hill, W. Y., & Roberts, C. B. (1998). Corporate social reporting practices in Western Europe: legitimating corporate behaviour?. The British accounting review, 30(1), 1-21.

Ahmed, A. C. A. (2012). Social and Environmental Auditing; Some Basic Concept. Some Basic Concept (October 25, 2012).

Becker, H. A. (2001). Social impact assessment. European Journal of Operational Research, 128(2), 311-321.

Behrens, H., Clay, M., Efftinge, S., Eysholdt, M., Friese, P., Köhnlein, J., Wannheden, K. & Zarnekow, S. (2008). Xtext user guide. Retrieved from: http://www. eclipse. org/Xtext/documentation/1_0_1/xtext.pdf.

Ben-Kiki, O., Evans, C., & Ingerson, B. (2009). Yaml ain't markup language (yaml™) version 1.1. Working Draft 2008-05, 11.

Bik, N. (2018). Developing a Model-driven Socio-environmental Auditing Tool (Master's thesis, University of Utrecht, Utrecht, The Netherlands).

Black, J., Hashimzade, N., & Myles, G. (Eds.). (2012). A dictionary of economics. OUP Oxford.

Castka, P., & Balzarova, M. A. (2008). ISO 26000 and supply chains—On the diffusion of the social responsibility standard. International journal of production economics, 111(2), 274-286.

Cawthon, N., & Moere, A. V. (2007, July). The effect of aesthetic on the usability of data visualization. In 2007 11th International Conference Information Visualization (IV'07) (pp. 637-648). IEEE.

Čeh, I., Crepinšek, M., Kosar, T., & Mernik, M. (2011). Ontology driven development of domain-specific languages. Computer Science and Information Systems, 8(2), 317-342.

Cohen, J. (2013). Statistical power analysis for the behavioral sciences. Routledge.

Cohn, M. (2004). Advantages of user stories for requirements. InformIT Network, available at: http://www. informit. com/articles.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior research methods, 39(2), 175-191.

Gao, S. S., & Zhang, J. J. (2006). Stakeholder engagement, social auditing and corporate sustainability. Business process management journal, 12(6), 722-740.

Garousi, V., Felderer, M., & Mäntylä, M. V. (2016, June). The need for multivocal literature reviews in software engineering: complementing systematic literature reviews with grey literature. In Proceedings of the 20th International Conference on Evaluation and Assessment in Software Engineering (p. 26). ACM.

Gray, R. (2000). Current developments and trends in social and environmental auditing, reporting and attestation: a review and comment. International journal of auditing, 4(3), 247-268.

Gray, R. (2006). Social, environmental and sustainability reporting and organisational value creation?. Accounting, Auditing & Accountability Journal.

Guthrie, J., & Abeysekera, I. (2006). Content analysis of social, environmental reporting: what is new?. Journal of Human Resource Costing & Accounting.

Hazlett, S. A., McAdam, R., Sohal, A., Castka, P., & Balzarova, M. A. (2007). A critical look on quality through CSR lenses. International Journal of Quality & Reliability Management.

Idowu, S. O., Zu, L., & Gupta, A. D. (2013). Encyclopedia of corporate social responsibility (Vol. 21). N. Capaldi (Ed.). Berlin, Germany: Springer.

King, B. (2013). Materiality assessments: The missing link for sustainability strategy. Retrieved from: GreenBiz.com.

Lavie, T., & Tractinsky, N. (2004). Assessing dimensions of perceived visual aesthetics of web sites. International journal of human-computer studies, 60(3), 269-298.

Lehman, G. (1999). Disclosing new worlds: a role for social and environmental accounting and auditing. Accounting, Organizations and society, 24(3), 217-241.

Line, M., Hawley, H., & Krut, R. (2002). The development of global environmental and social reporting. Corporate Environmental Strategy, 9(1), 69-78.

Martin, M., & Olsina, L. (2003, November). Towards an ontology for software metrics and indicators as the foundation for a cataloging web system. In

Proceedings of the IEEE/LEOS 3rd International Conference on Numerical Simulation of Semiconductor Optoelectronic Devices (IEEE Cat. No. 03EX726) (pp. 103-113). IEEE.

Morhardt, J. E. (2010). Corporate social responsibility and sustainability reporting on the internet. Business strategy and the environment, 19(7), 436-452.

Olsina, L., & Martin, M. D. L. A. (2004). Ontology for software metrics and indicators. J. Web Eng., 2(4), 262-281.

Pahuja, S. (2013). Environmental audit. Encyclopedia of Corporate Social Responsibility, 969-979.

Panach, J. I., España, S., Dieste, O., Pastor, O., & Juristo, N. (2015). In search of evidence for model-driven development claims: An experiment on quality, effort, productivity and satisfaction. Information and Software Technology, 62, 164-186.

Perrini, F. (2005). Building a European portrait of corporate social responsibility reporting. European management journal, 23(6), 611-627.

Pinsonneault, A., & Kraemer, K. (1993). Survey research methodology in management information systems: an assessment. Journal of management information systems, 10(2), 75-105.

Planken, B. C. (2013). Global Reporting Initiative.

Ritchie, J. (2018, November 14). Infographics. Retrieved from https://www.columnfivemedia.com

Rose, E. (2010). Continuous partial attention: Reconsidering the role of online learning in the age of interruption. Educational Technology, 50(4), 41-46.

Smiciklas, M. (2012). The power of infographics: Using pictures to communicate and connect with your audiences. Que Publishing.

Thatcher, B. (2012). An overview of infographics, Webinar. Illinois Central College Teaching and Learning Center.

Visser, E. (2007, July). WebDSL: A case study in domain-specific language engineering. In International Summer School on Generative and Transformational Techniques in Software Engineering (pp. 291-373). Springer, Berlin, Heidelberg.

Waters, K. (2009). Prioritization using moscow. Agile Planning, 12, 31.

Wieringa, R. (2009). Design science as nested problem solving. In Proceedings of the 4th international conference on design science research in informationsystems and technology (p. 8).

Wieringa, R., & Moralı, A. (2012, May). Technical action research as a validation method in information systems design science. In International Conference on Design Science Research in Information Systems (pp. 220-238). Springer, Berlin, Heidelberg.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. Springer Science & Business Media.

10 Appendix

Appendix A: Example of infographic

This example of infographic in the context of social and environmental auditing was published by REAS Euskadi and it was translated to English while the original was Spanish.
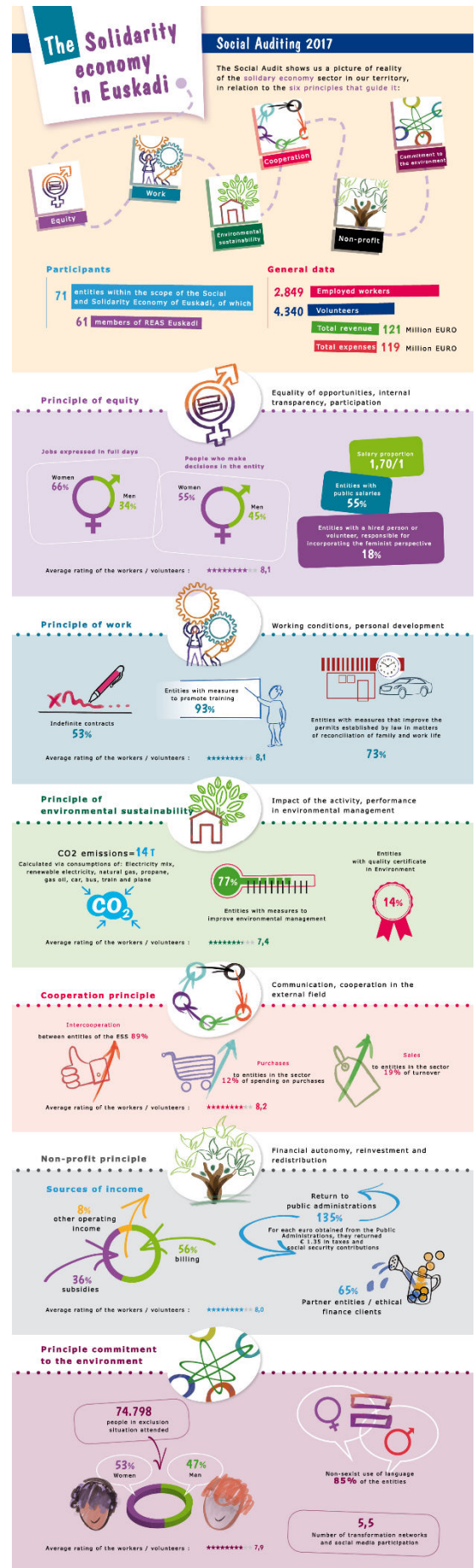
Appendix B: Table displaying the origin of each element in the conceptual model

Table 9: Matrix table displaying the origin of each element in the conceptual model.

	AuditingTeam	Indicator	DirectIndicator	IndirectIndicator	Measure	Method	Network	Organisation	Report	Social and environmental accounting	Stakeholder	StakeholderGroup	Tool	Topic
Adams, C. A., Hill, W. Y., & Roberts, C. B. (1998)						1							1	1
Ahmed, A. C. A. (2012)						1				1	2	1		
Becker, H. A. (2001)												1		
Castka, P., & Balzarova, M. A. (2008)						1				1		2		1
Gao, S. S., & Zhang, J. J. (2006)						2	2			2	3	3		
Gray, R. (2000)	1								1			1		
Gray, R. (2006)										1				
Guthrie, J., & Abeysekera, I. (2006)						3								
Hazlett, S. A., McAdam, R., Sohal, A., Castka, P., & Balzarova, M. A. (2007)	1					1				1		1		
Idowu, S. O., Zu, L., & Gupta, A. D. (2013)							1							
Lehman, G. (1999)	1									2				

	AuditingTeam	Indicator	DirectIndicator	IndirectIndicator	Measure	Method	Network	Organisation	Report	Social and environmental accounting	Stakeholder	StakeholderGroup	Tool	Topic
Line, M., Hawley, H., & Krut, R. (2002)	1								2					
Martin, M., & Olsina, L. (2003)		1	1	1		1								
Morhardt, J. E. (2010)								1						2
Olsina, L., & Martin, M. D. L. A. (2004)		1	1	1		1								
Pahuja, S. (2013)	4				6					3			1	
Perrini, F. (2005)		4				1		5	1	1	6	1		4
Planken, B. C. (2013)						1	1	1						

# Appendix C: Concept of StakeholderGroup decomposed
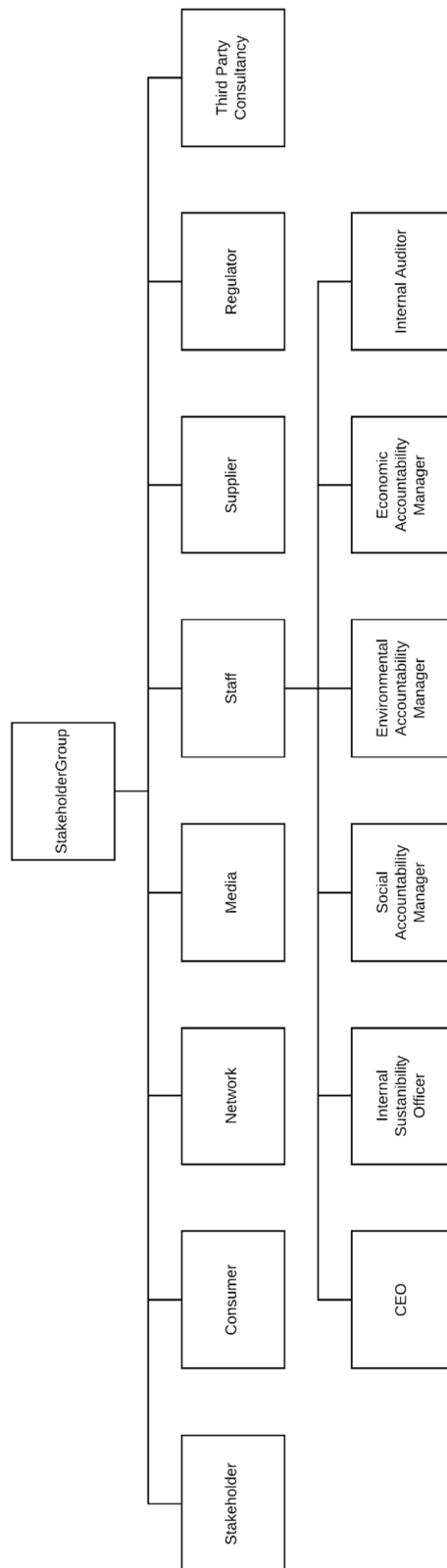
Appendix D: Table displaying the mapping between components identified and analysed infographics

| Infographic	Head section	Title text	Subtitle text	Introduction text	Logo	Body section	Box	Box title text	Body text	Table	Pie chart	Bar chart	Grouped bar chart	Stacked bar chart	Picturegraph	Line chart	Map chart	Foot section	Foot title text	Foot text	Logo	Link to full report
1	1	1	0	1	1	1	6	4	6	0	0	0	0	0	7	0	0	1	0	1	1	1
2	1	1	1	0	1	1	10	10	10	0	0	0	0	0	0	0	0	1	0	1	0	1
3	1	1	1	0	1	1	8	8	8	0	1	1	0	0	0	0	0	1	0	1	0	0
4	1	1	1	0	1	1	9	9	9	0	0	0	0	0	0	0	0	1	0	1	1	1
5	1	1	0	0	1	1	5	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0
6	1	1	1	0	1	1	9	9	9	0	0	0	0	0	0	0	0	1	0	1	0	1
7	1	1	1	1	1	1	9	9	9	0	0	1	0	0	1	0	0	1	0	1	0	1
8	1	1	0	0	1	1	11	11	11	0	2	1	0	0	0	0	0	1	0	1	0	0
9	1	1	0	0	0	1	15	0	15	0	0	0	0	0	0	0	0	1	0	1	0	0
10	1	1	1	1	0	1	17	8	17	0	4	0	0	0	9	0	0	1	0	1	1	0
11	1	1	0	0	0	1	8	7	8	0	0	0	0	0	0	0	0	1	0	1	0	1
12	0	0	0	0	0	1	14	0	14	0	1	0	0	0	0	0	0	1	1	1	0	0
13	0	0	0	0	0	1	13	0	13	0	1	0	0	0	0	0	0	1	1	1	0	0
14	1	1	1	0	1	1	7	4	6	1	0	0	0	0	0	0	0	1	1	1	1	0
15	1	1	1	0	1	1	7	4	6	1	0	0	0	0	0	0	0	1	1	1	1	0
16	1	1	1	0	1	1	7	4	6	1	0	0	0	0	0	0	0	1	1	1	1	0
17	1	1	1	1	0	1	6	3	6	0	0	0	0	0	5	0	0	1	1	1	1	0
18	1	1	1	0	1	1	16	14	11	0	3	1	0	3	0	0	1	1	1	1	0	0
19	1	1	0	0	1	1	16	16	16	0	1	3	0	0	6	0	1	0	0	0	0	0
20	1	1	0	0	1	1	12	12	12	0	7	0	0	0	0	0	0	1	1	0	1	0
21	1	1	1	0	1	1	7	7	6	0	2	0	7	1	16	0	1	1	0	1	0	1
22	1	1	1	0	0	1	7	7	7	0	0	0	0	0	0	0	0	1	0	0	1	0
23	1	1	0	0	0	1	10	10	10	0	0	0	0	0	0	0	0	1	0	0	1	0
24	1	1	1	1	0	1	9	3	9	0	0	0	0	0	6	0	0	0	0	0	0	0
25	0	0	0	0	0	1	9	9	0	0	0	0	0	0	0	0	0	1	0	0	1	0
26	0	0	0	0	0	1	11	11	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	1	5	5	5	0	0	0	0	0	0	0	0	1	0	1	0	0
28	1	1	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
29	1	1	0	0	0	1	8	8	8	0	2	2	0	0	0	0	0	0	0	0	0	0
30	1	1	1	0	0	1	9	8	9	0	0	1	0	0	0	7	0	0	0	0	0	0

Infographic	Head section	Title text	Subtitle text	Introduction text	Logo	Body section	Box	Box title text	Body text	Table	Pie chart	Bar chart	Grouped bar chart	Stacked bar chart	Picturegraph	Line chart	Map chart	Foot section	Foot title text	Foot text	Logo	Link to full report
31	1	1	0	0	0	1	4	4	0	0	0	0	0	0	0	0	0	1	0	1	0	1
32	1	1	0	0	1	1	7	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0
33	1	1	0	0	0	1	4	4	4	0	0	0	0	0	0	1	0	1	0	0	1	0
34	1	1	0	1	1	1	7	7	7	0	0	0	0	0	0	0	0	0	0	0	0	0
35	1	1	0	1	1	1	7	7	7	0	0	0	0	0	0	0	0	1	0	1	0	0
36	1	1	1	0	1	1	24	24	21	0	0	0	0	0	2	0	0	1	1	0	1	0
37	1	1	0	1	1	1	7	7	7	0	0	0	0	0	0	0	0	0	0	0	0	0
38	1	1	0	0	1	1	2	2	2	0	2	0	0	0	0	0	0	0	0	0	0	0
39	1	1	1	0	1	1	7	7	7	0	1	0	0	0	8	0	0	1	0	1	1	1
40	1	1	1	0	1	1	8	8	8	0	1	0	0	0	0	1	0	1	1	1	1	1
41	1	1	1	0	0	1	5	5	5	0	0	0	0	0	0	0	0	1	0	1	1	0
42	0	0	0	0	0	1	14	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
43	1	1	1	0	1	1	8	8	0	0	0	0	0	0	0	0	0	1	0	1	0	0
44	1	1	0	1	1	1	8	8	8	0	1	0	0	0	0	0	0	0	0	0	0	0
45	1	1	0	1	1	1	3	3	3	0	0	0	0	0	0	0	1	0	0	0	0	0
46	1	1	1	0	0	1	12	11	11	0	1	2	0	0	0	0	0	1	0	0	1	0
47	1	1	0	0	0	1	16	16	16	0	0	0	0	0	0	0	0	1	1	1	1	1
48	1	1	0	0	0	1	6	6	4	0	0	0	0	0	0	0	0	0	0	0	0	0
49	1	1	0	0	1	1	3	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
50	1	1	0	0	1	1	4	4	4	0	0	0	0	0	0	0	0	1	0	1	0	1
51	1	1	0	1	1	1	5	5	5	0	0	0	0	0	0	0	0	0	0	0	0	0
52	1	1	1	1	0	1	4	4	4	0	0	0	0	0	0	0	0	1	0	0	1	0
53	1	1	0	0	0	1	6	0	6	0	0	0	0	0	0	0	0	1	0	0	1	0
54	1	1	0	0	0	1	3	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0
55	1	1	0	0	1	1	8	8	8	0	0	0	0	0	0	0	0	1	1	1	1	1
56	1	1	1	0	0	1	13	13	0	0	0	0	0	0	2	0	0	0	0	0	0	0
57	1	1	0	0	1	1	12	12	0	0	0	0	0	0	0	0	0	1	0	1	0	0
58	1	1	0	0	0	1	8	3	0	0	0	0	0	0	0	0	0	1	0	0	0	1

Appendix E: Feature models of the different online infographic tools

Feature model BeFunky
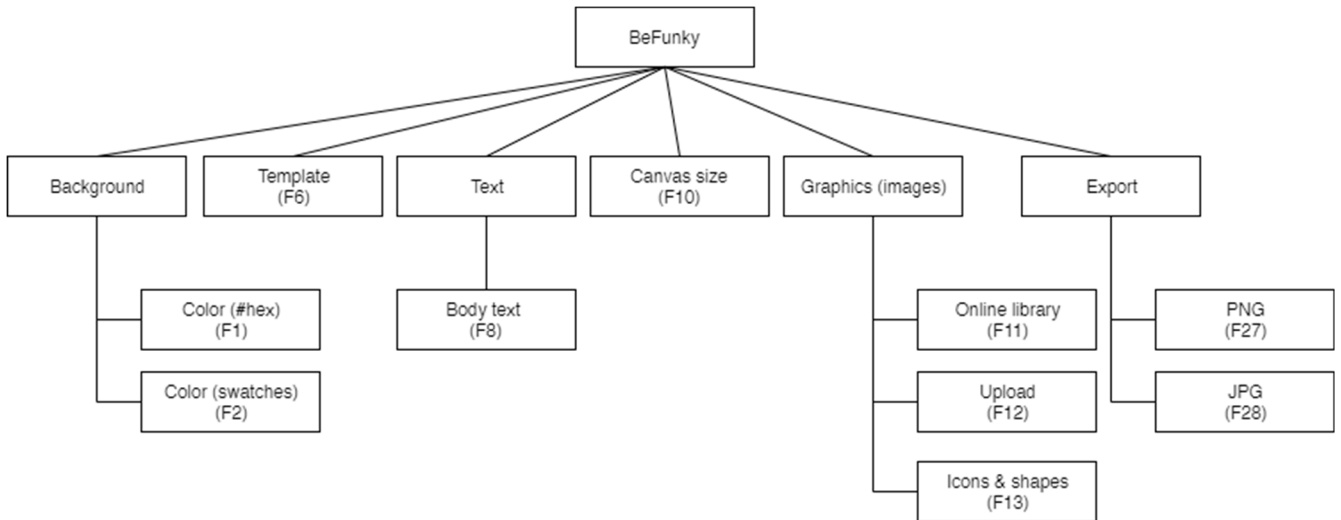
Feature model Canva
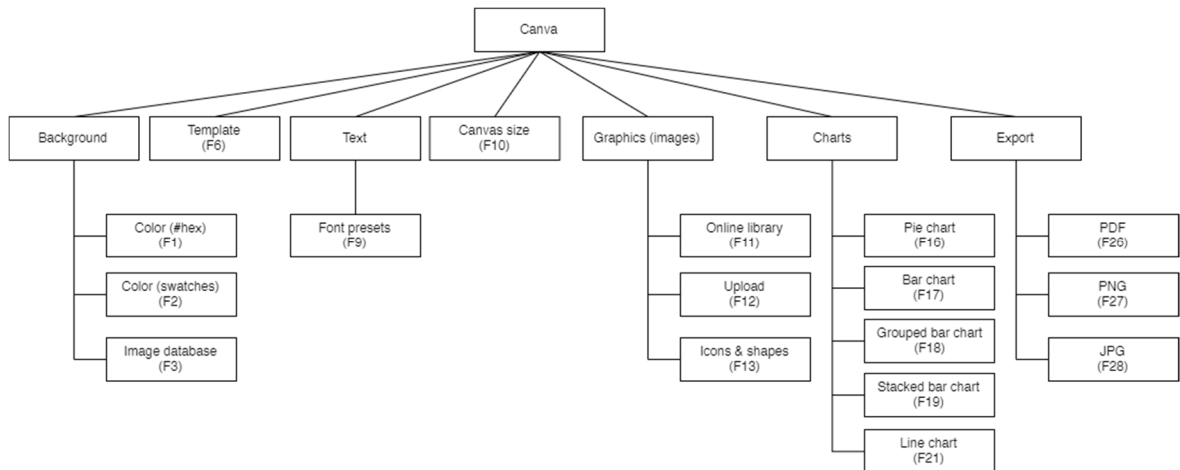
Feature model Easel.ly
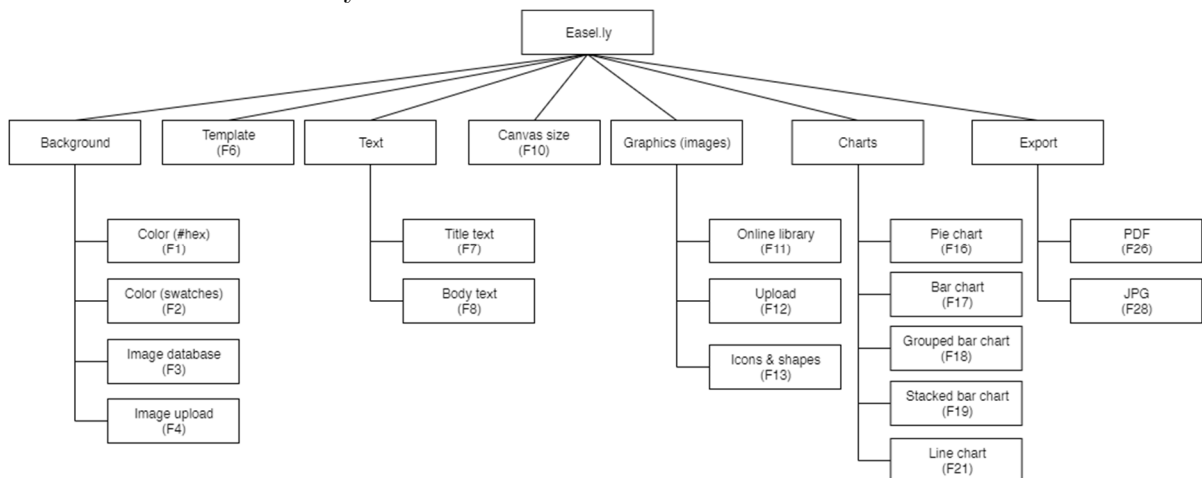
Feature model Infogram
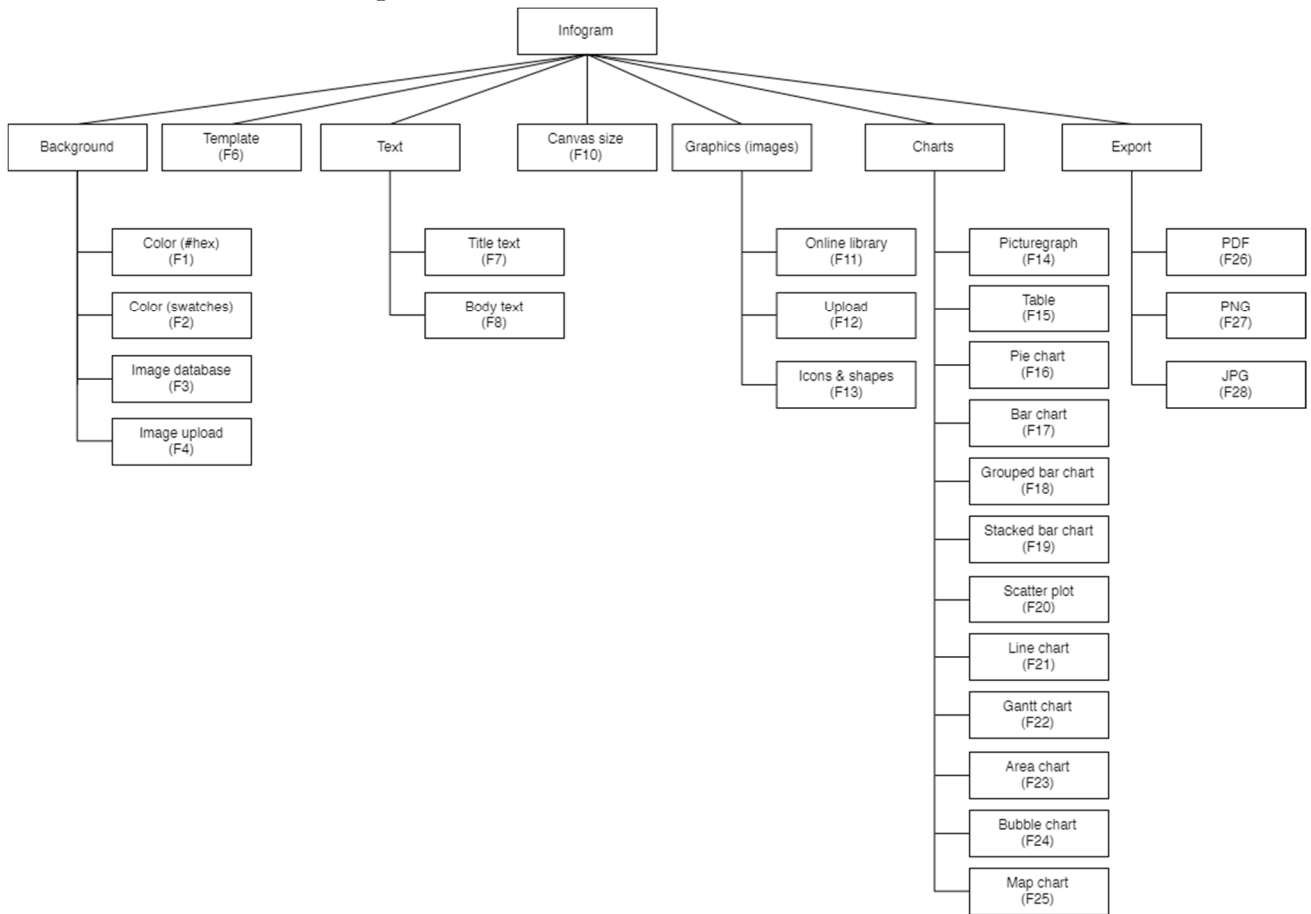
Feature diagram Lucidpress
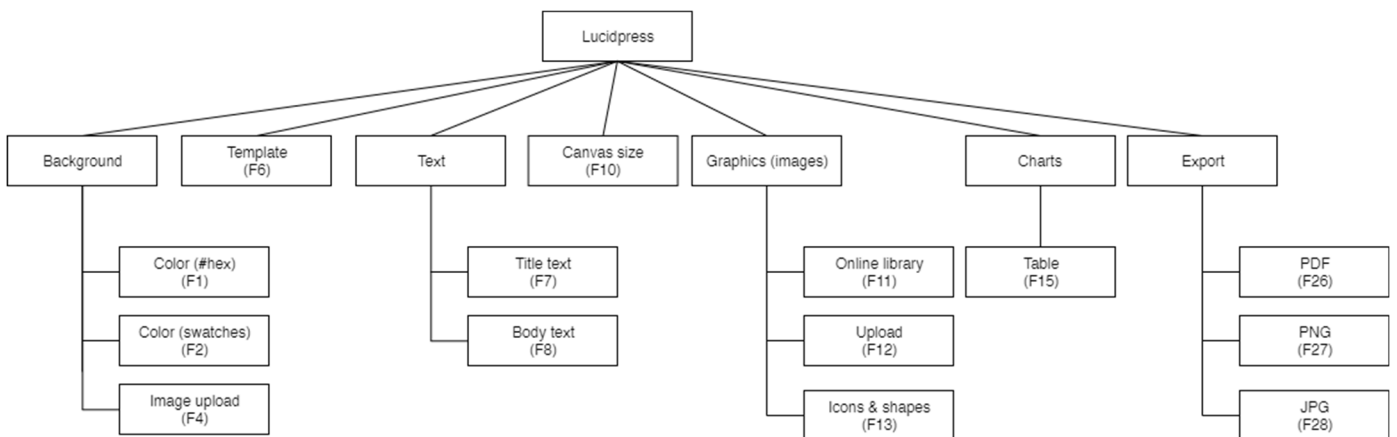
Feature diagram Mind the Graph
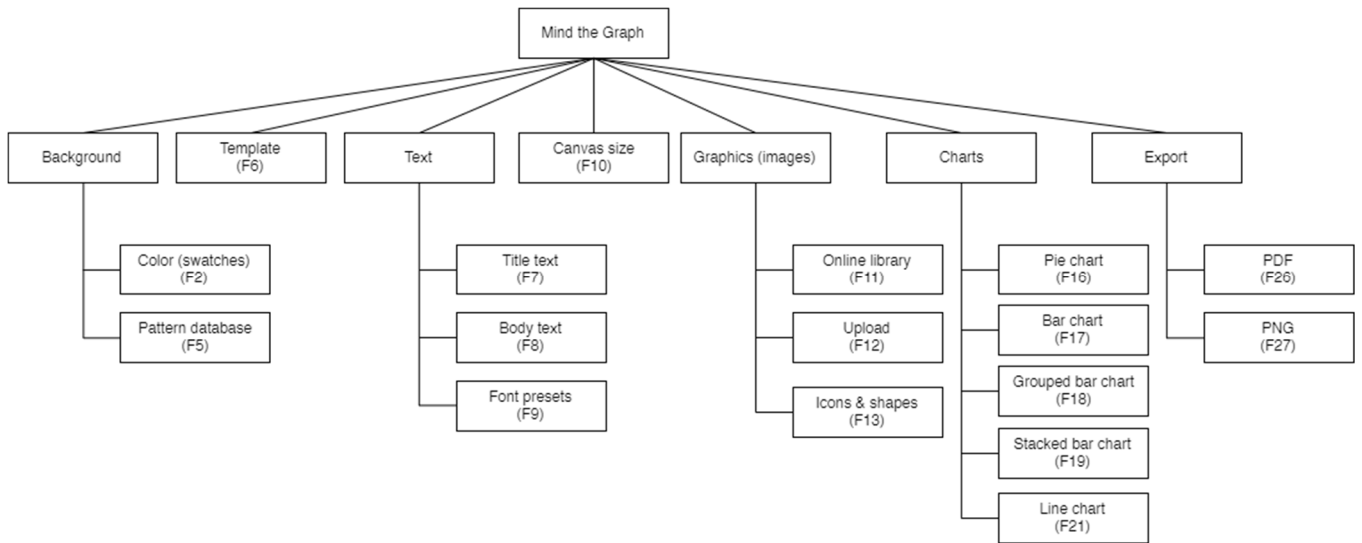
Feature diagram Piktochart

Feature diagram Snappa

Snappa
- Background
  - Color (#hex) (F1)
  - Color (swatches) (F2)
  - Image database (F3)
  - Image upload (F4)
  - Pattern database (F5)
- Template (F6)
- Text
  - Title text (F7)
  - Body text (F8)
- Canvas size (F10)
- Graphics (images)
  - Online library (F11)
  - Upload (F12)
  - Icons & shapes (F13)
- Export
  - PNG (F27)
  - JPG (F28)
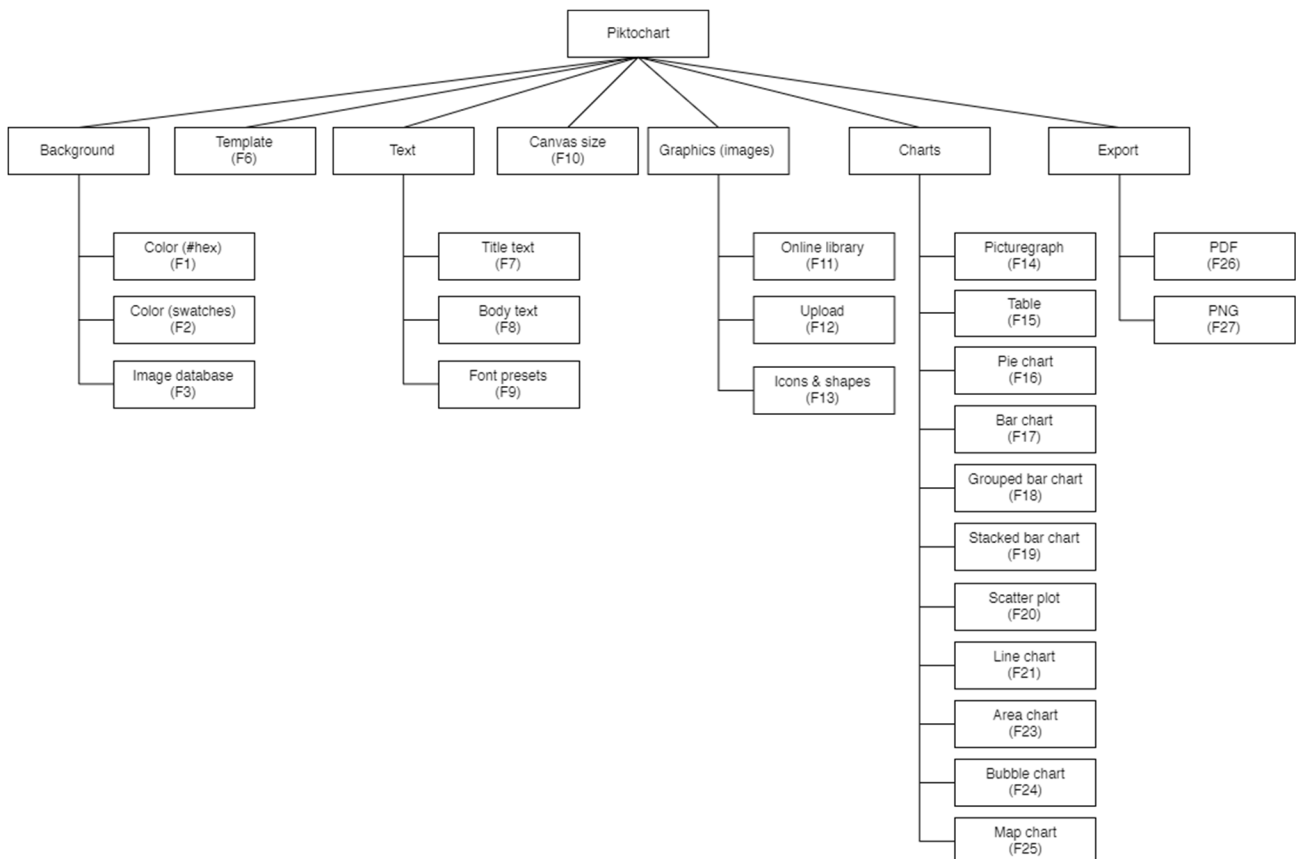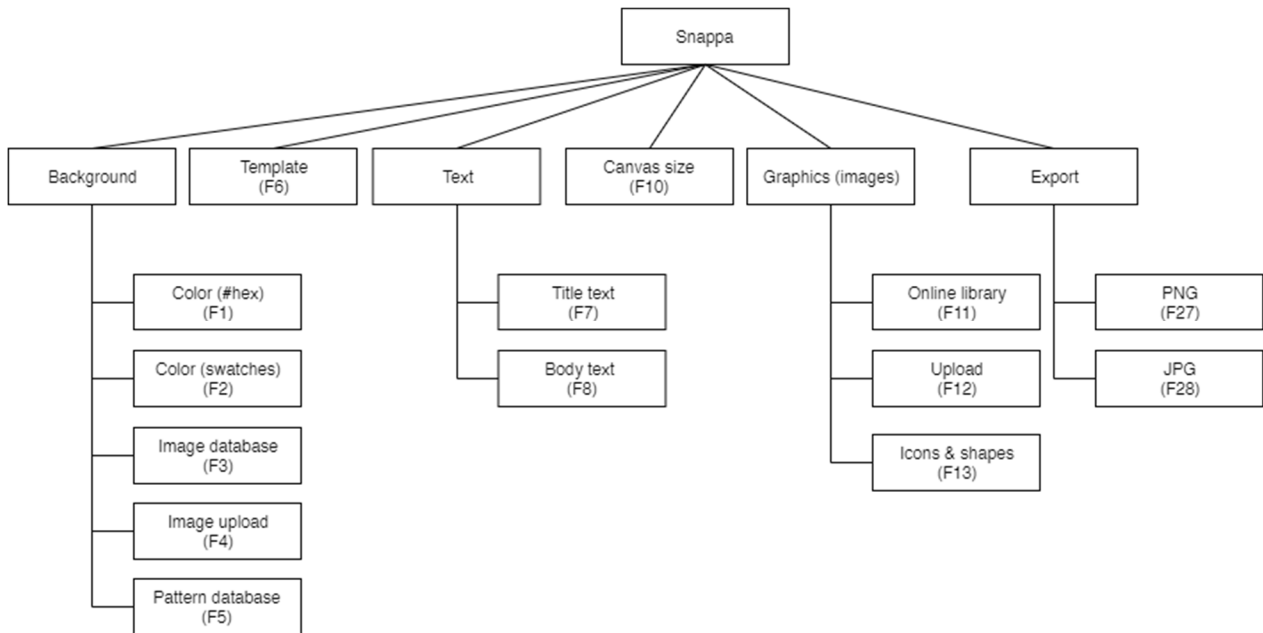
Feature diagram Venngage

Venngage
- Background
  - Color (#hex) (F1)
  - Color (swatches) (F2)
  - Pattern database (F5)
- Template (F6)
- Text
  - Title text (F7)
  - Body text (F8)
  - Font presets (F9)
- Canvas size (F10)
- Graphics (images)
  - Upload (F12)
  - Icons & shapes (F13)
- Charts
  - Picturegraph (F14)
  - Table (F15)
  - Pie chart (F16)
  - Bar chart (F17)
  - Grouped bar chart (F18)
  - Stacked bar chart (F19)
  - Scatter plot (F20)
  - Line chart (F21)
  - Area chart (F23)
  - Bubble chart (F24)
  - Map chart (F25)
- Export
  - PDF (F26)
  - PNG (F27)
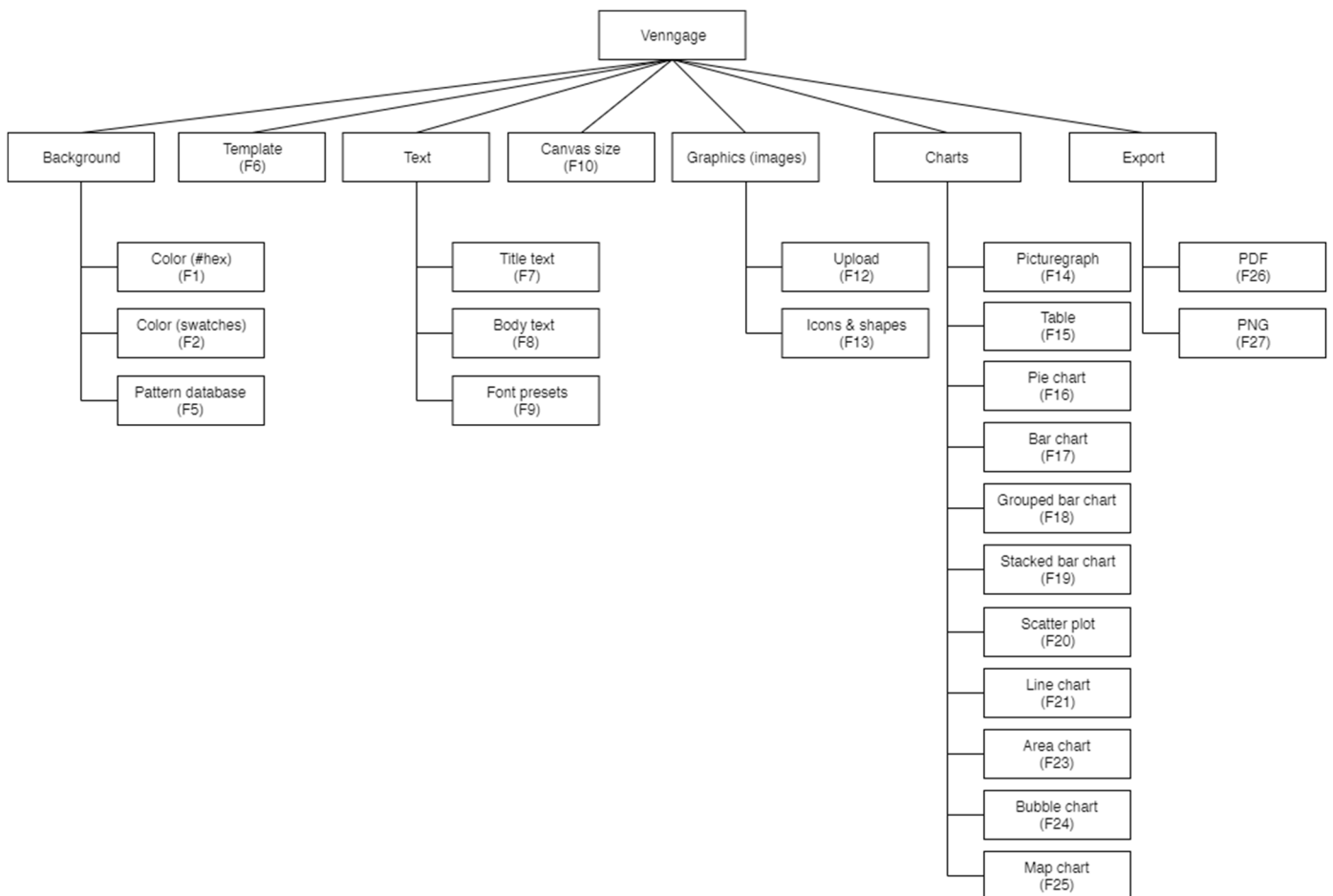
Feature diagram Visme
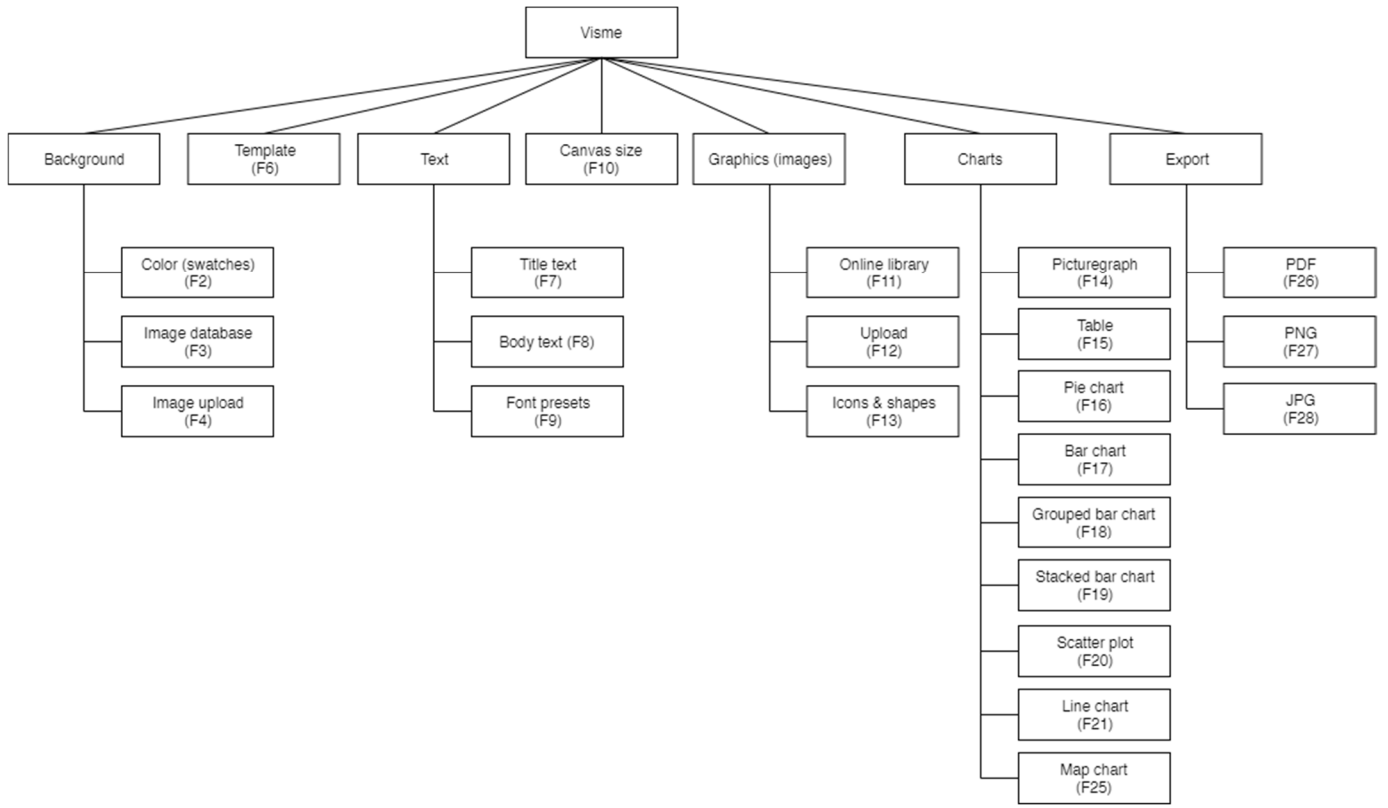
Table 10: User story matrix depicting where each user story originated. CM stands for conceptual model as described in Section 3.1 and TOCT stands for tree of component types as described in Section 3.3 and the feature analysis is described in Section 4.

	Open-SEA tool	CM	TOCT	Feature analysis																											
				F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28
US1	x																															
US2	x																															
US3	x																															
US4	x																															
US5	x																															
US6	x																															
US7				x																											
US8		x	x							x	x	x																				
US9													x																			
US10			x																x													
US11			x																	x												
US12																															x	
US13		x																														
US14		x																														
US15		x	x																													
US16		x	x																													

	Open-SEA tool	CM	TOCT	Feature analysis																											
				F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28
US17	x																														
US18			x														x														
US19					x																										
US20						x																									
US21							x																								
US22								x																							
US23												x																			
US24													x																		
US25														x																	
US26			x															x													
US27			x																			x									
US28			x																				x								
US29																													x		
US30		x																													x
US31																					x										
US32			x																				x								
US33																							x								
US34																								x							
US35																									x						
US36			x																							x					

```
1   grammar org.xtext.example.mydsl.IDSL with org.eclipse.xtext.common.Terminals
2   generate iDSL "http://www.xtext.org/example/mydsl/IDSL"
3
4   Infographic : InfographicElementType;
5
6   InfographicElementType:
7         ('type' ':' type='basic')?
8         & (('bgcolor' ':' bgcolor=Color) | ('bgpattern' ':' bgpattern=Pattern) | ('bgimage' ':'
    bgimage=ImageSrc))
9         & 'bgsize' ':' bgsize=SIZE_POS
10        & head=Head?
11        & boxes+=Box*
12        & texts+=Text*
13        & images+=Image*
14        & piecharts+=Piechart*
15        & barcharts+=Barchart*
16        & picturegraphs+=Picturegraph*
17        & foot=Foot?
18  ;
19
20  // Head: off or nothing required
21  Head:
22        name="head" ':' ('off' |
23        ( BEGIN (
24              (('bgcolor' ':' bgcolor=Color) | ('bgpattern' ':' bgpattern=Pattern) | ('bgimage' ':'
    bgimage=ImageSrc))?
25              & ('size' ':' size=SIZE_POS)?
26              & ('position' ':' position=SIZE_POS)?
27              & ('logo' ':' logo=Logo)?
28              & ('title' ':' title=Title)?
29              & ('subtitle' ':' subtitle=Title)?))
30        END );
31
32  // Title & Subtitle: off or nothing required
33  Title:
34        ({Title} 'off' |
35        ( BEGIN
36              {Title} (('align' ':' align=AlignPossibilities)?
37              & ('color' ':' color=Color)?
38              & ('font' ':' font=FONT)?
39              & ('position' ':' position=SIZE_POS)?
40              & ('maxwidth' ':' maxwidth=INT)?
41              & ('value' ':' value=STRING)?))
42        END );
43
44  // Logo: Size & position required
45  Logo:
46        BEGIN
47              (('size' ':' size=SIZE_POS)
48              & ('position' ':' position=SIZE_POS))
49        END;
50
```

```
51  // Box(n): Position required
52  Box:
53        name=BOXID ':'
54        BEGIN
55              ((('bgcolor' ':' bgcolor=Color) | ('bgpattern' ':' bgpattern=Pattern) | ('bgimage'
    ':' bgimage=ImageSrc))?
56              & ('size' ':' size=SIZE_POS)?
57              & ('position' ':' position=SIZE_POS))
58        END ;
59
60  // Text(n): Position required
61  Text:
62        (name=TEXTID ':'
63        ( BEGIN
64              (('align' ':' align=AlignPossibilities)?
65              & ('color' ':' color=Color)?
66              & ('font' ':' font=FONT)?
67              & ('lineheight' ':' lineheight=INT)?
68              & ('position' ':' position=SIZE_POS)
69              & ('maxwidth' ':' maxwidth=INT)?
70              & ('value' ':' value=STRING)?)
71        END )) |
72        (name=TITLETEXTID ':'
73        ( BEGIN
74              (('align' ':' align=AlignPossibilities)?
75              & ('color' ':' color=Color)?
76              & ('font' ':' font=FONT)?
77              & ('position' ':' position=SIZE_POS)
78              & ('maxwidth' ':' maxwidth=INT)?
79              & ('value' ':' value=STRING)?)
80        END ))
81  ;
82
83  // Image(n): Position, size and src required
84  Image:
85        name=IMAGEID ':'
86        BEGIN
87              (('position' ':' position=SIZE_POS)
88              & ('size' ':' size=SIZE_POS)
89              & ('src' ':' src=ImageSrc))
90        END ;
91
92  // Piechart(n): Data & position required
93  Piechart:
94        name=PIECHARTID ':'
95        BEGIN
96              (('bgcolor' ':' color=Color)?
97              & ('colors' ':' colors=COLOR_CHARTS)?
98              & ('data' ':' BEGIN ChartData (piedata+=ChartData)* END)
99              & ('legendstyle' ':' legendstyle=LegendStyle)?
100             & ('padding' ':' padding=INT)?
101             & ('position' ':' position=SIZE_POS)
102             & ('showlegend' ':' showlegend=ShowOptionsOff)?
103             & ('showtitle' ':' showtitle=ShowOptionsOff)?
104             & ('showpercentage' ':' showpercentage=ShowOptionsOff)?
105             & ('size' ':' size=INT)?
106             & ('title' ':' title=STRING)?
107             & ('type' ':' type=PieType)?)
108       END ;
109
```

```
110 // Barchart(n): Data & position required
111 Barchart:
112        name=BARCHARTID ':'
113        BEGIN
114                (('bgcolor' ':' color=Color)?
115                & ('colors' ':' colors=COLOR_CHARTS)?
116                & ('data' ':' BEGIN ChartData (bardata+=ChartData)* END)
117                & ('gridscale' ':' gridscale=INT)?
118                & ('legendstyle' ':' legendstyle=LegendStyle)?
119                & ('padding' ':' padding=INT)?
120                & ('position' ':' position=SIZE_POS)
121                & ('showgrid' ':' showgrid=ShowOptionsOff)?
122                & ('showlegend' ':' showlegend=ShowOptionsOff)?
123                & ('showtitle' ':' showtitle=ShowOptionsOff)?
124                & ('size' ':' size=SIZE_POS)?
125                & ('title' ':' title=STRING)?
126                & ('type' ':' type=BarType)?)
127        END ;
128
129 // Picturegraph(n): Data & position required
130 Picturegraph:
131        name=PICTUREGRAPHID ':'
132        BEGIN
133                (('bgcolor' ':' color=Color)?
134                & ('colors' ':' colors=COLOR_CHARTS)?
135                & ('data' ':' BEGIN ChartData (picturegraphdata+=ChartData)* END)
136                & ('label' ':' label=STRING)?
137                & ('padding' ':' padding=INT)?
138                & ('position' ':' position=SIZE_POS)
139                & ('showlegend' ':' showlegend=ShowOptionsOff)?
140                & ('showtitle' ':' showtitle=ShowOptionsOff)?
141                & ('size' ':' size=SIZE_POS)?
142                & ('title' ':' title=STRING)?
143                & ('type' ':' type=PicturegraphType)?)
144        END
145 ;
146
147 // Foot: off or nothing required
148 Foot:
149        name="foot" ':' ('off' |
150        ( BEGIN
151                (('bgcolor' ':' bgcolor=Color) | ('bgpattern' ':' bgpattern=Pattern) | ('bgimage' ':'
   bgimage=ImageSrc))?
152                & ('position' ':' position=SIZE_POS)?
153                & ('size' ':' size=SIZE_POS)?
154                & ('text' ':' text=STRING)? )
155        END)
156 ;
157
158 // SET BASIC RULES
159 terminal SIZE_POS : INT'x'INT;
160
161 Color : {Color} color=Color_Basic | {Color} COLOR_HEX;
162 enum Color_Basic : c1='green' | c2='red' | c3='blue' | c4='yellow' | c5='black' | c6='white' |
   c7='orange' | c8='purple' | c9='pink';
163 terminal COLOR_HEX :
   ('0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'.
   .'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F')('0'..'9'|'a'..'f'|'A'..'F');
164 terminal COLOR_CHARTS : (COLOR_HEX',')* COLOR_HEX;
165 Pattern : ImageSrc | PATTERN_DB;
166 ImageSrc : STRING;
167 terminal PATTERN_DB : 'pattern'INT;
168
```

```
169  FONT : (weight=Font_Weight)? FONT_SIZE family=Font_Family;
170  enum Font_Weight : b1='bold ' | b2='100 ' | b3='200 ' | b4='300 ' | b5='400 ' | b6='500 ' |
     b7='600 ' | b8='700 ' | b9='800 ' | b10='900 ';
171  terminal FONT_SIZE : ('0'..'9')+  'px ';
172  enum Font_Family : font1='Arial' | font2='Verdana' | font3='Times New Roman' | font4='sans-serif'
     | font5='Gotham';
173  enum AlignPossibilities: left | right | center;
174
175  enum LegendStyle : legend1='1' | legend2='2';
176  enum ShowOptionsOff : show1='off' | show2="n";
177  enum PieType : type1="pie" | type2="donut";
178  enum BarType : type1="horizontal" | type2="vertical";
179  ChartData: STRING ':' STRING;
180  enum PicturegraphType : type1="bin" | type2="electric";
181
182  // Set ID numbering for each element
183  terminal BOXID : 'box'INT;
184  terminal TEXTID : 'text'INT;
185  terminal TITLETEXTID : 'titletext'INT;
186  terminal IMAGEID : 'image'INT;
187  terminal PIECHARTID : 'piechart'INT;
188  terminal BARCHARTID : 'barchart'INT;
189  terminal PICTUREGRAPHID : 'picturegraph'INT;
190
191  // White spacing
192  terminal BEGIN: 'synthetic:BEGIN';
193  terminal END: 'synthetic:END';
```

Appendix H: Elements incorporated in DSL and its interpreter

Table 11: Table displaying which of the different components found in the tree of component types analysis are incorporated in the domain-specific language and its interpreter.

Component				Incorporated in DSL and its interpreter
Head section				x
	Title text			x
	Subtitle text			x
	Introduction text			
	Logo			x
Body section				x
	Box			x
		Box title text		x
		Body text		x
		Charts		
			Table	
			Pie chart	x
			Bar chart	x
			Grouped bar chart	
			Stacked bar chart	
			Picturegraph	x
			Line chart	
			Map chart	
Foot section				x
	Foot title text			
	Foot text			x
	Logo			x
	Link to full report			x

Table 12: Table displaying which of the different features identified in the feature analysis of online tools are incorporated in the domain-specific language and its interpreter.

	Incorporated in DSL and its interpreter		Incorporated in DSL and its interpreter
F1	x	**F15**	
F2		**F16**	x
F3		**F17**	x
F4	x	**F18**	
F5	x	**F19**	
F6		**F20**	
F7	x	**F21**	
F8	x	**F22**	
F9		**F23**	
F10	x	**F24**	
F11		**F25**	
F12	x	**F26**	
F13		**F27**	x
F14	x	**F28**	

Appendix I: More advanced infographic specification

```
 1   bgcolor: d3e3c0
 2   bgsize: 700x960
 3   box1:
 4     bgcolor: e2eed3
 5     position: 50x160
 6     size: 600x285
 7   box2:
 8     bgcolor: e2eed3
 9     position: 50x460
10     size: 600x270
11   box3:
12     bgcolor: e2eed3
13     position: 50x745
14     size: 600x140
15   head:
16     bgimage: 'https://www.umib.net/wp-content/uploads/2019/06/ranger-4df6c1b6.png'
17     logo:
18       position: 535x48
19       size: 100x60
20     subtitle:
21       position: 50x100
22       value: 'Lorem Ipsum stands for cool things'
23     title:
24       position: 50x70
25       value: 'Lorem Ipsum Sustainability Highlights +report+'
26   foot:
27     bgcolor: 646c5b
28     size: 700x60
29     text: 'Copyright Lorem Ipsum +year+ - Data is based on report +report+'
30   barchart1:
31     colors: 213771, 208480, c3dfb4
32     data:
33       2017: S38
34       2018: '75'
35       2019: '100'
36     gridscale: '25'
37     position: 70x475
38     size: 250x170
39     title: 'Zero emission vehicles in our fleet'
40   picturegraph1:
41     colors: 213771, 208480, c3dfb4
42     data:
43       '2018': '3000'
44       '2019': '1500'
45     padding: '10'
46     position: 375x765
47     size: 100x100
48     title: 'Total waste'
49     type: 'bin'
50   picturegraph2:
51     colors: 213771, 208480, c3dfb4
52     data:
53       '2018': '2300'
54       '2019': '2000'
```

```yaml
    '2020': '1710'
  padding: '10'
  position: 85x765
  size: 100x100
  title: 'Power consumption'
  type: 'electric'
piechart1:
  colors: 213771, 208480, c3dfb4
  data:
    Female: '2018#S38'
    Male: '100 - 2018#S38'
  padding: '10'
  position: 70x170
  size: '100'
  title: Female/Male ratio 2019
  type: donut
text1:
  color: 0
  font: 12px sans-serif
  lineheight: 0
  maxwidth: '340'
  position: 300x220
  value: "Lorem Ipsum is simply dummy text of the printing and typesetting industry. Lorem Ipsu
m has been the industry's standard dummy text ever since the 1500s, when an unknown printer too
k a galley of type and scrambled it to make a type specimen book. It has survived not only five
 centuries, but also the leap into electronic typesetting, remaining essentially unchanged. It
was popularised in the 1960s with the release of Letraset sheets containing Lorem Ipsum passage
s, and more recently with desktop publishing software like Aldus PageMaker including versions o
f Lorem Ipsum."
text2:
  position: 355x520
  value: "It is a long established fact that a reader will be distracted by the readable conten
t of a page when looking at its layout. The point of using Lorem Ipsum is that it has a more-
or-
less normal distribution of letters, as opposed to using 'Content here, content here', making i
t look like readable English. Many desktop publishing packages and web page editors now use Lor
em Ipsum as their default model text, and a search for 'lorem ipsum' will uncover many web site
s still in their infancy."
titletext1:
  position: 300x195
  color: 1d3f72
  value: 'Lorem Ipsum did well according to this chart'
titletext2:
  position: 355x495
  color: 1d3f72
  value: 'Oh and look this one!'
```