

# Mein super Visualistik Projekt

Sascha Lemke  
Matr.Nr.: 11074933

Dennis Dubbert  
Matr.Nr.: 11089478

## ABSTRACT

Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. Ut wisi enim ad minim veniam, quis nostrud exerci tation ullamcorper suscipit lobortis nisl ut aliquip ex ea commodo consequat. Duis autem vel eum iriure dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat nulla facilisis at vero eros et accumsan et iusto odio dignissim qui blandit praesent luptatum zzril delenit augue duis dolore te feugait nulla facilisi.

## 1 EINLEITUNG

This is my introduction. You can write your documentation in English, but if you prefer German you can also write it in German.

## 2 RELATED WORK

Are there other visualization papers you might want to cite? You can cite like this [1], the references are located in the file references.bib.

## 3 AUFBAU DER VISUALISIERUNG

Die Visualisierung ist in zwei Abschnitte eingeteilt, welche vertikal voneinander Abgetrennt sind und in einer starken Synergie stehen. Die Linke Seite des Bildschirms wird von der eigentlichen Visualisierung eingenommen und die rechte Seite stellt einen Text-View dar. In der Visualisierung können verschiedenen Auswahl getroffen und hierdurch innerhalb des Text-Views markiert werden. Anhand der textuellen Darstellung und Nähe der markierten Begriffe kann der Nutzer nun Rückschlüsse bezüglich der Wissenschaftlichkeit seines Dokuments bilden und gegebenenfalls Verbesserungen an diesem vornehmen. Die Abschnitte sind durch einen Balken klar abgetrennt, welcher jedoch, je nach Anforderung des Nutzers, verschoben werden kann, um die Größenverhältnisse anzupassen. Wird gerade eine Auswahl getroffen, so kann es nützlich sein die Visualisierungsseite zu vergrößern. Wurde bereits eine Auswahl getroffen so kann die Textseite vergrößert werden.

Um eine geeignete und feingranulare Auswahl dieser Begriffe gewährleisten zu können, ist auch der Visualisierungsbereich in weitere, aufeinander aufbauende, Einzelvisualisierungen gegliedert. Somit beinhaltet sie:

- eine Navigation,
- eine Kapitelübersicht / -auswahl,
- und eine Detailansicht der Selektion

Jede dieser Visualisierung hat ein festes Seitenverhältnis, sodass eine Verzerrung durch Größenänderungen vermieden wird.

Innerhalb der Navigation wird dem Nutzer ein Drop-Down-Menü geboten, über welches sich die Filterungsmethode definieren lässt (Redundanzen, Satzzeichen, Satzlänge oder Füllwörter). Wurde hier eine Auswahl getroffen, so wird die gesamte Visualisierung auf diese abgestimmt sodass dem Nutzer eine gezieltere Suche ermöglicht wird.

Der zweite Abschnitt veranschaulicht die Hierarchie des ausgewählten Dokuments und wird für die Auswahl zu betrachtender Kapitel genutzt. Hier werden zwei alternative Ansichten geboten, eine in Form eines Sunburst und eine als Baumstruktur, welche teilweise unterschiedliche Ausprägungen hervorheben und sich somit gegenseitig ergänzen. Über Tabs kann der Nutzer zu jedem Zeitpunkt die für ihn nützlichere Ansicht selektieren. Wichtig ist hierbei, dass die, in den Ansichten getätigten, Auswahlen synchronisiert werden, wodurch ein fließender Wechsel unterstützt wird.

Wurde in der Kapitelübersicht eine Auswahl getroffen, so wird diese in einer Detailansicht aufgegriffen. Hier sind die Begriffe der Ausgewählten Kapitel anhand der Filterungsmethode eingegrenzt und in Form eines Blasen-Diagramms visualisiert. Wurden beispielsweise Redundanzen selektiert, so steht jede Blase für ein Wort, welches innerhalb der Kapitelauswahl redundant auftritt. Zu jedem dieser Begriffe wird auch die entsprechende Ausprägung der Filterungsmethode bereitgestellt, sodass eine persönliche Problemeinschätzung des Nutzers erfolgen kann. Ist nun eine Überarbeitung bestimmter Begriffe gewünscht, so können diese selektiert und hierdurch jedes Vorkommen, begrenzt durch die ausgewählten Kapitel, in dem Text-View markiert werden.

Das folgende Kapitel dient somit zunächst der Erläuterung dieser Einzelvisualisierungen. Hierbei werden besonders die Entscheidungsfindungen und Kompromisse beleuchtet, welche zu dem Aufbau und den Interaktionsmöglichkeiten der jeweiligen Visualisierung führten.

### 3.1 Kapitelübersicht: Sunburst-Diagramm

Dem Sunburst-Diagramm liegt der Gedanke zugrunde, dass der Nutzer den Aufbau des Dokuments bereits kennt, sich also im Vorhinein damit auseinander gesetzt hat und nun mithilfe der Visualisierung eine weitere Iteration vornehmen möchte. Im Fokus steht hierbei die bereits gut bekannte Kapitelhierarchie zentriert und auf einen Blick sichtbar darzustellen und somit eine schnelle Navigation und Auswahl zu ermöglichen. Da bei solch einem Diagramm nahezu keine Freiflächen zwischen den Elementen gegeben sind, kann den Elementen selber weitaus mehr Fläche zugeordnet werden ohne die Übersichtlichkeit der Visualisierung zu gefährden. Dies führt zu einer Maximierung der Datendichte. Die verfügbare Visualisierungsfläche wird also bestmöglich ausgenutzt, da lediglich das Zentrum und die Ecken des Diagramms Freiflächen bilden, welche jedoch für Informationstexte nutzbar sind. Bei einer Verkleinerung des Diagramms sind einzelne Abschnitte somit trotzdem klar erkenn- und trennbar, sodass die Visualisierung problemlos auf unterschiedlichen Bildschirmgrößen angezeigt und genutzt werden kann.

Ein weiterer Beweggrund für diese Art der Darstellung war, dass dem Nutzer ein Auswahlmedium bereitgestellt werden soll,

welches die Dokumenthierarchie sowie Kapitel-Abhängigkeiten veranschaulicht. Durch die Aufeinandererschichtung der unterschiedlichen Ringlagen dieser Diagrammart sind die genannten Abhängigkeiten schnell erkennbar. Es werden keine weiteren visuellen Mittel benötigt um die Zugehörigkeit eines Kapitels zu verdeutlichen, da die Position der Elemente ausreichend Aufschluss hierüber bietet. Somit wird auch die Data-Ink-Ratio erhöht und Chartjunk vermieden. Der mittlere Kreis stellt hierbei das Dokument dar, welches zunächst von den Hauptkapiteln (Abstract, Einleitung, Hauptteil, Fazit, etc.) umringt wird. Die Größe des jeweiligen Ringabschnitts, in Abhängigkeit des Radius, ist hierbei der prozentuelle Anteil des Kapitels im Verhältnis zum gesamten Dokument. Als Maß gilt hier die Wortanzahl eines Kapitels, welche sich aus dem eigenen Inhalt, sowie dem Inhalt aller Unterkapitel erschließen lässt und anschließend auf den entsprechenden Teilwinkel des Kreises umgerechnet wird. Dieses Schemata erstreckt sich auch auf die Unterkapitel eines Abschnittes, welche jedoch nun ihr jeweiliges Elternkapitel als Maximum ansehen (sowohl deren radialen Abschnitt in der Visualisierung als auch deren Wortanzahl) und nicht das gesamte Dokument.

Ein Vorteil der hierdurch entsteht ist, dass neben der sichtbaren Hierarchie auch ein Größenverhältnis zum jeweiligen Oberkapitel ersichtlich wird. Für wissenschaftliche Arbeiten bestehen häufig formale Anforderungen, sodass auch die Größe einzelner Kapitel vorgegeben sein kann. Somit könnte es also nützlich sein solche Verhältnisse aus der Visualisierung entnehmen zu können. Hierbei wurde jedoch bedacht, dass Größenverhältnisse in Kreisdiagrammen trügen können, da der radiale Abschnitt die Größe eines Ringabschnitts definiert und somit äußere Abschnitte aufgrund ihrer Fläche automatisch größer erscheinen als Innere (auch bekannt als Lie Factor oder Lügenfaktor). Da bei der Überarbeitung eines wissenschaftlichen Textes jedoch vorwiegend das direkte Verhältnis eines Kapitels zu Nachbarkapiteln des selben Oberkapitels betrachtet wird, wurde dieser Aspekt im Kontext des Projektes als vernachlässigbar eingeschätzt. Im Sunburst-Diagramm besitzen zudem alle Lagen die selbe Höhe, sodass der Flächeninhalt hier eine geringere Bedeutung als der dargestellte Radius besitzt und somit trotzdem eine akkurate visuelle Einschätzung getätigt werden kann. Damit der eigene Textanteil eines Oberkapitels ebenfalls mit den Textanteilen der Unterkapitel verglichen werden kann, wurde entschieden auch diesen Anteil als Unterkapitel aufzulisten und ihn somit auf deren Ebene zu verschieben. Diese werden jedoch als Einleitung betitelt und somit durch ihre Namensgebung speziell gekennzeichnet, sodass die Zugehörigkeit bestehen bleibt. Dies ist jedoch nur benötigt, wenn ein Kapitel weitere Kapitel umschließt. Ansonsten wird der Textinhalt weiterhin im eigenen Element dargestellt.

Da in einem wissenschaftlichen Dokument jedes Kapitel eine Einleitung besitzen sollte und dieses nicht direkt aus der bisherigen Visualisierung hervorgeht, wurde zudem eine Möglichkeit gesucht, fehlenden Text aufzuzeigen. Sollte also ein Kapitel keinen eigenen Text beinhalten, so wird in der Mitte des zugehörigen Elements ein schwarzer Kreis gezeichnet. Dieser nimmt wenig Platz ein, wodurch die Data-Ink-Ratio nahezu unbeschadet bleibt. Durch die nun entstandene Unregelmäßigkeit innerhalb der Visualisierung ist er jedoch auf den ersten Blick klar erkennbar, sodass dem Nutzer das Problem direkt übermittelt wird.

Ein weiterer wichtiger Aspekt dieser Visualisierung findet sich in der Farbgebung der Abschnitte. Da die Farbe ein gutes visuelles Mittel zur Übermittlung der Qualität eines Elements ist, wird sie hier zu exakt jenem Zweck verwendet. Anhand der ausgewählten Filterungsmethode werden die jeweiligen Ausprägungen der Kapitel berechnet und diese anschließend

in dem zugehörigen Farbtönen dargestellt. Bei der Auswahl der Farbe gilt der höchste ermittelte Wert innerhalb eines Kapitels als maßgebend für diese Einfärbung. Dies ist damit begründet, dass besonders Ausreißer erfasst und sichtbar gemacht werden sollten, welche bei der Bildung eines Durchschnitts verborgen bleiben könnten. Wurde beispielsweise die Filterungsmethode bezüglich der Satzlänge ausgewählt, so könnte bei der Durchschnittsbildung ein sehr langer Satz übersehen werden, welcher jedoch die Qualität des Dokumentes beeinträchtigt. Diese Einfärbung wird für jedes Kapitel eigenständig vorgenommen. Dies ist damit begründet, dass es Worthäufigkeiten gibt, welche einen hohen Einfluss auf die Textqualität eines Unterkapitels haben, jedoch verglichen mit dem Oberkapitel unauffällig erscheinen. Ebenso könnte in einem Oberkapitel eine Worthäufigkeit als problematisch angesehen werden, diese jedoch gleichmäßig auf die Unterkapitel aufgeteilt sein und somit für diese nur eine geringe Auswirkung haben. Diese Aspekte sind nun direkt aus der Visualisierung entnehmbar, sodass die Trennung dem Nutzer eine gezieltere und schnellere Suche ermöglicht.

Bei der Farbwahl wurde von einer linearen Interpolation zwischen zwei Farbtönen abgesehen und sich stattdessen für eine klar unterteilte Skala mit den Farbtönen Grün, Gelb, Orange, Hellrot und Dunkelrot entschieden. Die Farbtöne wurden so gewählt, da die Endtöne im Komplementärkontrast stehen, also eine einfache Unterscheidung der Qualität stattfinden kann. Zudem soll dem Nutzer über diese Farbgebung eine direkte Assoziation mit der gewünschten Aussage ermöglicht werden, da ein Grünton allgemein positives Feedback symbolisiert und ein Rotton vorwiegend als Gefahrensignal aufgefasst wird (Ampelprinzip). Auch wenn die letzten beiden Farbtöne den Grundton Rot besitzen, sind sie dennoch durch ihre Sättigung klar trennbar, sodass eine Qualitätsminderung erkenntlich wird. Weiterhin wurde die Skala auf fünf Untertöne reduziert. Auch hier ist wieder ein Argument, dass die Töne klar trennbar sein sollten und dies bei einer größeren Anzahl von Unterfarben nicht mehr gegeben wäre. Weiterhin ist es Ziel dieser Visualisierung dem Nutzer eine klare Einteilung der Qualität zu bieten, ebenso wie eine kurze Eingewöhnungsphase. Eine höhere Anzahl der Unterteilungen könnte die Eingewöhnungszeit verlängern.

Nun kann es vorkommen, dass unterschiedliche Nutzer auch unterschiedliche Anforderungen an ihr Dokument stellen. Ist die Zeit knapp, so steht beispielsweise die schnelle Identifikation der schwerwiegendsten Probleme im Vordergrund. Hat der Nutzer mehr Zeit oder sind strengere Regelungen an das Dokument gekettet, so sollten auch geringere Problemstellen aufgezeigt und beseitigt werden. Aus diesem Grund wurde den Farben kein fester numerischer Wert zugeordnet. Für die Zuordnung dieser Werte wurde ein Slider integriert, mit welchem der Nutzer definieren kann, ab welcher Ausprägung ein Kapitel dunkelrot eingefärbt wird. Diese Obergrenze wird nun zudem in vier gleichmäßig verteilte Unterbereiche aufgeteilt und den einzelnen Farbwerten zugeordnet. Wurde beispielsweise der Wert zehn als Obergrenze definiert, so zeigen sich folgende Zuordnungen:

Grün: Werte von 0 bis 2,5

Gelb: Werte von 2,5 bis 5

Orange: Werte von 5 bis 7,5

Hellrot: Werte von 7,5 bis 10

Dunkelrot: Werte ab 10

Durch die Anpassung der Einfärbung und die Größe der Kapitel kann der Nutzer nun eine geeignete Strategie bzw. Reihenfolge der Textbearbeitung festlegen.

Verändert sich jedoch der Kontext durch die Auswahl einer anderen Filterungsmethode, so werden auch andere Ausprägungen betrachtet. Tritt ein Wort redundant auf, sollte dessen Anzahl im Verhältnis zum jeweiligen Paragraphen betrachtet werden. Wird hingegen nach der Satzlänge oder der Anzahl von Satzzeichen gesucht, so ist eine Angabe der maximalen Wort- oder Satzzeichenanzahl eines kritischen Satzes sinnvoller. Da die Werte des Sliders somit nicht auf jede Filterungsmethode gleichzeitig abgestimmt werden kann, wurden jeweils Obergrenzen und Inkrementierungsschritte in Form von unterschiedlichen Skalen definiert. Durch diese Vorauswahl wird der Slider nun auf den jeweiligen Kontext angepasst, jedoch weiterhin eine starke Anpassung des Nutzers ermöglicht.

Damit dem Nutzer stets bekannt ist, welcher Farbe welcher Wertebereich zugeordnet ist, sind diese Zusammenhänge auf der linken Seite in Form einer Legende aufgelistet. Wird der Slider bewegt und losgelassen, so werden die neuen Unterabschnitte direkt berechnet und innerhalb der Legende angepasst. Neben der Farbgebung werden in dieser Legende zudem ungewöhnliche Aspekte dieser Visualisierung aufgegriffen. So findet sich hier beispielsweise eine kurze Erklärung der Punkte, welche fehlenden Text darstellen.

Weitere wichtiger Aspekt jeder Visualisierung sind dessen Interaktionsmöglichkeiten. Eine dieser Interaktionen bildet in dieser Visualisierung das Hervorheben eines Kapitels sobald der Mauszeiger das zugehörige Element betritt. Standardmäßig besitzen die Kapitelelemente eine niedrige Sättigung, welche nun erhöht wird und somit ein klare Trennung von selektierten und nicht selektierten Kapiteln bewirken soll. Hierdurch wird dem Nutzer eine direkte Rückmeldung und dadurch ein Auffordrungscharakter geboten, welcher den Nutzer zu weiteren Interaktionen und Erkundungen animieren soll. Um diesen Effekt zu verstärken wurde an den einzelnen Objekten ein schmaler schwarzer Rahmen angebracht. Auch wenn dieser wiederum Chartjunk darstellt, welcher in dieser Visualisierung weitestgehend umgangen wurde, so sorgt er durch die dunklen Abgrenzungen für eine stärkere Hervorhebung der Auswahl und wurde somit als sinnvoll erachtet. Durch den Mouse-Over werden zusätzlich die Texte im Zentrum des Sunburst-Diagramms angepasst, welche nun den Titel des Ausgewählten Kapitels, sowie dessen höchste Ausprägung im Bezug auf die Filterungsmethode und den prozentualen Anteil zum Oberkapitel anzeigen. Ist der Titel zu lang für den inneren Bereich des Diagramms, so wird er abgekürzt, dieses jedoch durch die beifügung dreier Punkte gekennzeichnet. Da dem Nutzer jedoch die Struktur bekannt ist, sollte dieser Ausschnitt für eine Zuordnung ausreichend sein. Somit erhält der Nutzer stets ausreichend Auskunft über dieses Kapitel, welche die angesprochenen optischen Verhältnisse komplettieren. Besitzt das ausgewählte Kapitel Unterkapitel, so werden auch diese Hervorgehoben und die Gesamtgröße dieser als Grundlage für die Informationstexte genommen. Auf diese Weise kann eine möglichst flexible Erkundung des Dokuments gewährleistet werden. Verlässt der Mauszeiger das Kapitel, so wird auch die Auswahl rückgängig gemacht.

Hat sich der Nutzer jedoch entschieden ein Kapitel genauer zu betrachten, so kann er es, zusammen allen Unterkapiteln, über einen Linksklick an seine Auswahl binden und den restlichen Visualisierungen zur Verfügung stellen. Verlässt der Mauszeiger nun das Kapitel, so wird dieses nicht mehr Abgewählt. Ist ein Element bereits ausgewählt, so wird dieses durch einen weiteren Klick aus der Auswahl entfernt. Auf diese Weise können nahe gelegene Abschnitte parallel in den anderen Visualisierungen analysiert werden.

Eine parallele Auswahl mehrerer Hauptkapitel wird hierbei

jedoch nicht unterstützt. Dies liegt zunächst an der Form der Datenstruktur. Solch eine Auswahl hätte den Nachteil, dass eine Menge Vorberechnungen auf Seiten des Front-Ends durchgeführt werden müssten, was zu ungewollten Ladezeiten führen könnte. Weiterhin ist eine Einzelselektion sinnvoll, da Unterkapitel meist das selbe Thema aufgreifen und beleuchten. Hierbei besteht eine hohe Wahrscheinlichkeit für ungewollte Wiederholungen, welche es nun in dem ganzen Abschnitt zu beseitigen gilt. Wird jedoch ein Unterkapitel eines anderen Oberkapitels ausgewählt, so ist dieser Zusammenhang oft nicht gegeben. Um Verfälschungen oder Verdeckungen von Problemstellen entgegenzuwirken, wird in diesem Falle die vorherige Auswahl verworfen und der angeklickte Abschnitt als neue Auswahl gespeichert. Findet der Klick außerhalb oder im Zentrum des Diagramms statt, so wird die gesamte Auswahl zurückgesetzt. Hierdurch werden dem Nutzer alle Möglichkeiten geboten, welche er für eine erste grobe Einschätzung und Auswahl der Kapitel benötigt.

Möchte der Nutzer nun ein weiteres Unterkapitel des gleichen Hauptkapitels (dargestellt durch die innerste Ringschicht) über einen Klick an die Auswahl binden, so wird dieses der bereits getätigten Sammlung beigelegt.

Eine letzte Interaktion betrifft die eigentliche Navigation durch das Textdokument. Wird ein Kapitel innerhalb des Diagramms angeklickt, so wird auch der Text-Viewer auf die Auswahl abgestimmt. Da sich das Sunburst-Diagramm auch bei geringer Größe nutzen lässt, kann es zudem als eine Art Inhaltsverzeichnis fungieren, über welches nun das Dokument erkundbar ist. In jedem Falle kann die Kapitelübersicht und somit auch dieses Sunburst-Diagramm als Brückenstück zwischen dem Textdokument und der Visualisierung gesehen werden, welches den Startpunkt jeglicher Interaktion darstellt.

### 3.2 Kapitelübersicht: Baum-Diagramm

Ähnlich dem Sunburst-Diagramm steht auch bei dieser Visualisierung die Darstellung der Kapitelhierarchie im Vordergrund. Dem Nutzer soll ein Überblick über das Dokument geboten werden, sodass eine strukturierte Bearbeitung ermöglicht wird. Der größte und maßgebliche Unterschied zur vorherigen Visualisierung findet sich jedoch in der fokussierten Zielgruppe. Das Sunburst-Diagramm ist auf jene Personen ausgelegt, welche ihr eigenes Dokument überarbeiten wollen und sich somit exzellent in diesem Auskennen. Das ist der Grund, warum solch eine zentrierte Form und minimalistische Beschriftung als Ausreichend erachtet wurde. Wird sie jedoch einer Drittperson angeboten, so benötigt diese eine längere Einarbeitungsphase um sich zunächst mit den einzelnen Kapiteln vertraut zu machen. Da jedoch auch diese Personengruppe zu berücksichtigen ist, wurde an einer weiteren Visualisierung gearbeitet, welche nun Zweitprüfern oder anderen Korrekturlesern dienlich ist. Hierbei wurde versucht möglichst gleiche Visualisierungsmittel und Interaktionsmethoden zu verwenden, sodass ein fließender Übergang zwischen den zwei Ansichten stattfinden kann. Hat sich ein Nutzer nun beispielsweise durch die Baumstruktur ausreichend mit den Kapiteln auseinander gesetzt, so kann dieser ohne Probleme zu der Sunburst-Ansicht wechseln. Aufgrund der weitgehenden Übereinstimmungen befasst sich der folgende Abschnitt nun vorwiegend mit den Maßgebenden Änderungen, welche auf diese neue Zielgruppe abgestimmt sind.

Die Baum-Struktur wurde hier gewählt, da sie die Vernetzung der Kapitelstruktur übersichtlich aufgliedert. Ähnlich dem Sunburst spielt auch hier die Position der Knoten eine wichtige Rolle. Diese ist optisch am einfachsten wahrzunehmen und somit der Orientierung des Nutzers sehr dienlich. Dieser Aspekt wird durch die Anzeige der Äste unterstützt. Auch wenn solche Äste einen hohen Anteil des Chartjunks der Visualisierung ausmachen, bieten sie in dem Kontext mehr positive als negative Aspekte. Der

unwissende Nutzer findet sich schnell zurecht, da er lediglich den einzelnen Verästelungen folgen muss um das gewünschte Kapitel zu erreichen. Hierdurch bieten sie bereits auf den ersten Blick ausreichend Orientierungspunkte, sodass Zusammenhänge zwischen den Kapiteln schnell erkannt werden können. Um diesen Aspekt weiterhin zu unterstützen, wurden dezente, jedoch klar voneinander trennbare Farben ausgewählt. Diese sind in Form einer ordinalen Skala eingebunden, sodass sie eindeutig einem Hauptkapitel und somit dessen Verästelungen zugeordnet werden können.

Weiterhin wurde ein horizontaler Aufbau der Baumstruktur gewählt, welcher die gewohnte Leserichtung aufgreift und dem Nutzer somit unterbewusst eine Hilfestellung bietet. Ein weiterer maßgeblicher Unterschied findet sich in der Kapitelbeschriftung. Im Sunburst wurde diese lediglich im Zentrum angebracht, da einzelne Ringe unter Umständen zu klein für eine ausgiebige Beschriftung sein und die Krümmung den Lesefluss beeinflussen könnten. Da diese Aspekte jedoch nicht auf die Knoten eines Baumes zutreffen, sind hier die Beschriftungen stets innerhalb der Knoten angebracht. Im Gegensatz zu dem stark zentrierten Aufbau des Sunburst-Diagramms ist eine Baumstruktur von Natur aus breitflächiger. Da die x-Achse bei horizontalen Bäumen deren Tiefe bestimmt und diese in einem wissenschaftlichen Dokument eine maximale Ausprägung von vier oder (in Ausnahmefällen) fünf besitzt, kann den einzelnen Knoten eine ausreichende Breite zur Verfügung gestellt werden, wodurch auch die Titelbeschriftung profitiert. Sollte jedoch der Fall eintreten, dass ein Titel über den Rand des Knotens hinaus gehen würde, so wird dieser mithilfe eines Clip-Path eingeschränkt. Für solche Fälle besteht die Möglichkeit, sich den vollständigen Titel über einen Hover-Effekt anzeigen zu lassen. Hierzu muss die Maus lediglich für kurze Zeit auf dem Element ruhen.

Im Gegensatz zum Sunburst wird die Kapitelgröße in dieser Visualisierung nicht beachtet. Einer Einfindung in die Datenstruktur wird hier der höchste Stellenwert zugesprochen. Die repräsentativste Veranschaulichung der Kapitelgrößen wäre anhand der proportionalen Größe oder Position des jeweiligen Knotens. Dies würde jedoch die Kapitelhierarchie oder Anzeige der Titel beeinträchtigen, welche als Hauptorientierungspunkte gelten. Fehlt einem Kapitel jedoch die Einleitung oder allgemein Text, so wird dies auch hier mit einem Kreis markiert. Dieser befindet sich jedoch nicht im Zentrum des Elements, sondern an dessen Ende, sodass auch hierdurch kein Titel beeinträchtigt wird.

Wie zuvor erwähnt ist in diesem Projekt lediglich eine Tiefe von maximal vier bis fünf Ebenen zu erwarten (Dokument / Root, Kapitel, Unterkapitel, Unterunterkapitel und ggf. Unterunterunterkapitel). Die Höhe jedoch wird von der Anzahl an Nachbarkapiteln einer Ebene bestimmt, welche je nach Dokument sehr unterschiedlich ausfallen können. In der Vollansicht stellt dies kein Problem dar, wird diese jedoch verkleinert, so können einzelne Elemente einer breit gefächerten Dokumentstruktur gegebenenfalls nur noch schwer auseinander gehalten werden. Vor allem hier zeigt sich wiederum der Vorteil einer Aufteilung in zwei Visualisierungen. Aufgrund der hohen Data-Density des Sunbursts würde nun diese Ansicht ausgewählt werden um weitere Navigationen zu vollziehen. Der Baum hingegen besitzt eine geringe Datendichte und ist somit für solch eine Navigation weniger geeignet. Der Fokus des Baumes liegt jedoch auch mehr in der strukturierten, als in einer zentrierten Darstellung. Dies ist auch einer der Gründe, warum das Sunburst-Diagramm als Hauptansicht und die Baumstruktur lediglich als Erweiterung gewählt wurde.

Abgesehen von diesen Aspekten finden sich nur geringe Änderungen. Die Farbgebung der Knoten wird anhand der selben

Aspekte ermittelt und auch der Slider zur explorativen Spezifikation der Ausprägungen wird beibehalten. Auch die Legende findet sich an der selben Stelle. Lediglich die Auswirkungen der Interaktion mit dem Diagramm wurden geringfügig angepasst. Wird der Mauszeiger über ein Element bewegt, so folgt weiterhin die Anzeige des gefilterten Wertes dieses Kapitels, da er für eine Verfeinerung der Auswahl benötigt wird. Dieser ist hier jedoch an den Hover-Text des jeweiligen Elementes gebunden, welcher zudem den Titel des Kapitels Anzeigt. Das Größenverhältnis zum jeweiligen Oberkapitel wird hierbei jedoch nicht angebracht, da der Größe hier im Allgemeinen zu Gunsten der Leserlichkeit weniger Beachtung geschenkt wird. Eine letzte Änderung findet sich in der Markierung von Kapiteln. Im Gegensatz zum Sunburst-Diagramm wird hier nicht nur das ausgewählte Element zusammen mit dessen Unterkapiteln hervorgehoben, sondern auch der Pfad vom Root-Element zu diesem Knoten sowie alle zugehörigen Verästelungen der Auswahl. Dies ist wiederum damit begründet, dass diese Visualisierung vorwiegend von dokumentfremden Personen genutzt wird und ihnen die Auswahl auf diese Weise nachvollziehbarer dargestellt wird.

### 3.3 Detailansicht: Blasen-Diagramm

Wurde eine Auswahl in der Kapitelübersicht getroffen, werden dem Nutzer nun innerhalb der Detailansicht genauere Informationen geboten. Hier sind nun, in Abhängigkeit der Filterungsmethode, sämtliche Begriffe in Form eines Blasen-Diagramms aufgelistet. Wird nach den Redundanzen gesucht, so symbolisiert jede Blase ein Wort, welches redundant in dem ausgewählten Bereich auftritt. Wurden Satzlänge oder Satzzeichen selektiert, so bildet jede Blase einen Satz ab. Die Suche nach Füllworten hebt sich hier jedoch durch dessen Präsentation etwas von den Restlichen Filterungen ab. Hierbei steht jede Blase für einen Paragraphen innerhalb der Kapitelselektion. Diese Variante wurde gewählt, da Füllwörter eine eigene Kategorie bilden, welche durch die Masse aller beinhalteter Einzelwörter schädlich sein kann. Mehrere Füllwörter häufig verwendet worden, so ist dies ebenso schlimm, wie das Auftreten eines einzelnen Füllwortes in der selben Anzahl. Die Aufteilung nach Paragraphen wurde hier also gewählt, um die Suche weiterhin spezifizieren zu können.

Damit die einzelnen Blasen voneinander unterschieden werden können, wurden diese mit einem Schriftzug des jeweiligen Inhaltes versehen. Dieser wird in Abhängigkeit der Größe und Inhaltslänge jedoch verkürzt worden, damit er innerhalb der jeweiligen Blase dargestellt werden kann. Auf diese Weise wird dem Nutzer eine erste Zugehörigkeit ersichtlich gemacht. Die Inhalte sind innerhalb der einzelnen Kategorien wie folgt aufgebaut:

Redundanz: Das jeweilige Wort gefolgt von der Worthäufigkeit.

Satzlänge: Der vollständige Satz gefolgt von der Wortanzahl.

Satzzeichen: Der vollständige Satz gefolgt von der Anzahl an Satzzeichen.

Füllwörter: Der Pfad zu dem Paragraphen gefolgt von der Anzahl an Füllwörtern.  
(Kapitel > Unterkapitel > ... > Paragraphennummer + Anzahl)

Da die Blasengröße jedoch von der Anzahl an Worten abhängt, können diese unter Umständen sehr klein ausfallen, wodurch diese Beschriftung unerkennlich wäre. Aus diesem Grund wurde weiterhin eine Hover-Animation implementiert. Wird der Mauszeiger

über eine Blase bewegt, so erhöht sich dessen Radius und Textgröße auf einen festgelegten Anteil der Visualisierung. Zudem wird diese in den Vordergrund gehoben, sodass Sie nicht von den umliegenden Blasen verdeckt wird und der Text problemlos wahrgenommen werden kann und der, ansonsten weiße Text wird durch eine schwarze Einfärbung hervorgehoben. Bewegt sich der Mauszeiger nun von der Blase weg, so wird auch diese Vergrößerung rückgängig gemacht. Als Ankerpunkt für diesen Effekt wurde jedoch nicht die Blase selbst, sondern der textliche Inhalt ausgewählt. Wäre hier die Blase selbst das Ziel, so könnten keine umliegenden Knoten betrachtet werden, da diese von der momentanen Auswahl überdeckt sind.

Da bei langen Sätzen hier jedoch nur der Anfang präsentiert werden kann, wurde die Visualisierung um ein Textfeld ergänzt, welches sich direkt über dem Blasen-Diagramm aufstreckt. Innerhalb dieses Textfelds werden nun, parallel zu der Vergrößerung der Blasen, durch die Hover-Bewegung der jeweilige Inhalt weitaus größer und lesbarer präsentiert. Dieses Textfeld wird lediglich durch drei Punkte gekennzeichnet, welche den zunächst fehlenden Inhalt repräsentieren. Von Umrandungen, Hintergrundfarbe oder anderen Hervorhebungsmethoden wurde hier abgesehen, da diese vorwiegend Chartjunk darstellen, und somit der Data-Ink-Ratio schaden würden. Zudem ist das Textfeld auch ohne diese Hervorhebung klar erkennbar, spätestens bei der ersten Interaktion, innerhalb welcher es zudem erst an Bedeutung gewinnt. Dem Textfeld wurde zudem eine feste Größe zugeordnet. Wäre dies nicht der Fall, so würde sich die Position des Blasen-Diagramm stets in Abhängigkeit der Textlänge verschieben. Wäre das Textfeld unterhalb des Diagramms angeordnet, so könnte der Text aus dem Sichtbereich hinaus ragen. Für den Fall, dass nun ein langer Satz zu groß für dieses Textfeld ist, wird stets die Textlänge überprüft und dieser Satz gegebenenfalls abgeschnitten. Diese Verkürzung wird wiederum mit drei Punkten gekennzeichnet, gefolgt von dem Ausprägungswert in Klammern, sodass dem Nutzer dennoch alle Informationen gegeben werden.

Neben dem textuellen Inhalt sind die Blasen zudem durch eine Hintergrundfarbe gekennzeichnet, welche auch hier Aufschluss über die Qualität bzw. Problematik des Inhaltes gibt. Diese Farben richten sich, ebenso wie die der Kapitelauswahl, nach dem Spezifizierten Wert des Nutzers, sodass seine Entscheidungen und Ansprüche auch hier aufgegriffen werden. Durch die einheitliche Hervorhebung von Eigenschaften wird auch für ungeschulte Nutzer ein möglichst schneller Gebrauch der Anwendung ermöglicht.

Weiterhin werden die Knoten vorsortiert, sodass sich die Begriffe mit dem höchsten Problempotential in der Mitte des Diagramms ansammeln und dieses Risiko bis hin zum äußersten Ring abnimmt. Anhand der Position eines Elements und dessen Farbe kann der Nutzer also gezielt nach Problemstellen suchen. Durch diese Kombination bilden sich zudem unterschiedliche Lagen, deren Breite Aufschluss über die allgemeine Qualität des ausgewählten Dokumentbereichs liefern, sodass der Nutzer bereits auf den ersten Blick Einschätzungen tätigen kann.

Da bereits diese beiden Aspekte eine Aussage über die Qualität der Elemente ermöglichen, wurde zudem auf eine Anpassung der Blasengrößen anhand der Wert-Ausprägung verzichtet. Nutzer kennen die Bedeutung der Farbe bereits aus vorherigen Abschnitten der Visualisierung und assoziieren rote Elemente mit Problemstellen. Bringt man nun zudem die Dimension der Größe ein, so wachsen diese Gefahrenzonen weiter an und werden subjektiv als noch gefährlicher wahrgenommen. Unproblematische Elemente, welche nun um einiges kleiner sind, würden in diesem Zusammenhang untergehen, sodass der Textbereich insgesamt

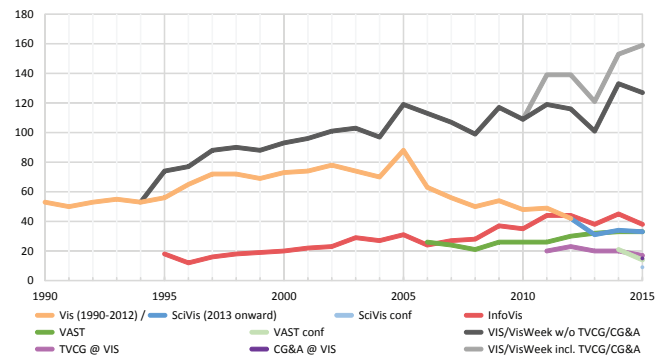


Figure 1: A visualization of the data from ???. The image is from [1] and is in the public domain.

schlechter eingeschätzt wird, als er in Wirklichkeit ist. Neben der Vermeidung solch eines Lügenfaktors wurde zudem bedacht, dass die Veränderung der Größe eines Kreises von Menschen schwer zu interpretieren ist. Steigt beispielsweise die Masse eines Kreises, so hat dieses eine geringe Auswirkung auf den Radius. Wird jedoch der Radius angepasst, so steigt die Masse unproportional stark an oder nimmt ebenso stark ab. Da auch dies einen Lügenfaktor darstellen kann, wurde sich dazu entschieden, einen identischen Radius für jeden Kreis zu nutzen, und dessen Eigenschaften über die Position und Farbe darzustellen.

Eine negative Auswirkung dieser Wahl besteht jedoch darin, dass nun Elemente mit unterschiedlichen Werten als gleichmäßig kritisch betrachtet werden könnten. Wurde beispielsweise eine Obergrenze von 10 definiert, so werden alle Elemente die darüber liegen dunkelrot eingefärbt und ein Wert von hundert hat das selbe Erscheinungsbild wie ein Wert von zehn. Wird jedoch ein Wort in einem Kapitel hundert Mal genannt und ein anderes nur zehn Mal, so stellt das Erste eine weitaus größere Problematik dar und sollte direkt ausgebessert werden. Auch wenn die Position bereits Aufschluss darüber bietet, wurde eine weitere

### 3.4 Text-Viewer

## 4 DISCUSSION

Discuss the benefits and drawbacks of your project.

## 5 PERSONAL CONTRIBUTION

Describe for each team member what her/his contribution to the project was and give a personal reflection.

### 5.1 Team member 1

blablabla

### 5.2 Team member 2

blablabla

### 5.3 Team member 3

blablabla

## 6 CONCLUSION

Write your conclusion and outlook here. How can your project be improved further?

## REFERENCES

- [1] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. Stolper, M. Sedlmair, J. Chen, T. Möller, and J. Stasko. vispubdata.org: A Metadata Collection about IEEE Visualization (VIS) Publications. *IEEE Transactions on Visualization and Computer Graphics*, 23, 2017. To appear. doi: 10.1109/TVCG.2016.2615308