

# CNCR - Cancer VIS

Lea Maria Joseffa Koinig\*  
Fakultät für Informatik  
Universität Wien

Dana Markovic†  
Fakultät für Informatik  
Universität Wien

Johannes Preisinger‡  
Fakultät für Informatik  
Universität Wien

## ZUSAMMENFASSUNG

Da es noch kein geeignetes Tool für die Visualisierung von österreichischen Krebsdaten gibt, präsentieren wir in diesem Paper unseren Lösungsansatz zur Visualisierung dieser Daten. Unser interaktives Tool soll den User dazu animieren, sich tiefer mit dem Thema zu beschäftigen. Anhand einer Fallstudie zeigen wir, wie die Nutzer mit dem Tool interagieren und dass sie dazu motiviert werden, die Daten zu erforschen und zu verstehen.

**Keywords:** Krebsvisualisierung, Informationsvisualisierung

## 1 MOTIVATION

Die Motivation, uns im Rahmen der VU Visualisierung mit der Visualisierung von Krebsdaten zu beschäftigen, kommt daher, dass wir im Rahmen unserer Recherche kein Tool finden konnten, das eine interaktive Visualisierung von Krebsdaten für Österreich bietet. Außerdem ist Krebs ein wichtiges Thema, das viele Leute betrifft und interessiert.

### 1.1 Tasks

Patienten und Angehörige können sich über die Häufigkeit der Krebsart informieren und wie viele Überlebende es gibt, indem sie die Krebsart auswählen und sich Sterbefälle bzw. Überlebende anzeigen lassen. Zudem können sie ein bestimmtes Jahr auswählen, oder sich über die Anzahl der Erkrankungen über die Jahre hinweg informieren. Mediziner können das Tool bei der Erklärung der Diagnose verwenden, um aufzuzeigen, wie viele Erkrankungen dieser Art es gibt und wie die Überlebenschancen stehen, indem sie die Art der Erkrankung auswählen und die Sterbefälle bzw. die Überlebenden für ein bestimmtes Jahr oder über die Jahre hinweg anzeigen lassen. Angehende Mediziner und Interessierte können sich interaktiv mit der Häufigkeit des Auftretens beschäftigen, indem sie zum Beispiel nach Jahren und / oder nach Krebsart filtern und die Graphiken betrachten. Forschende können das Tool benutzen, um herauszufinden, ob in manchen Jahren besonders viele Erkrankungen einer gewissen Art oder in einem bestimmten Bundesland aufgetreten sind, indem sie nach Jahren und Erkrankungen filtern oder sich die Erkrankungen für die Bundesländer anzeigen lassen. Sie können danach ohne das Tool weiterforschen, warum dieser Sachverhalt vorliegt. Menschen können sich informieren, um präventive Maßnahmen zu ergreifen, indem sie zum Beispiel nach der Anzahl der Erkrankungen in ihrem Bundesland filtern.

### 1.2 User

Zu den Usern zählen Patienten, Angehörige, Mediziner, Medizinstudenten, aber auch allgemein interessierte Personen, die sich informieren möchten.

\*e-mail: a1109391@unet.univie.ac.at

†e-mail: a1169281@unet.univie.ac.at

‡e-mail: a1106089@unet.univie.ac.at

## 1.3 Daten

Die Daten liegen auf der Webseite von Statistik Austria [1] im xls Format vor. Das Problem dabei ist, dass es zwar Daten für die unterschiedlichen Krebsarten von 1983 bis 2011 gibt, diese aber alle in separaten Files gespeichert sind. Für eine bessere Performance haben wir alle für unser Tool notwendigen Daten in ein File gespeichert und lesen aus diesem aus. Des Weiteren liegt für die Erkrankungen und Sterbefälle für die einzelnen Bundesländer leider nur der Durchschnitt der Jahre 2009 - 2011 vor.

## 2 RELATED WORK

Die medizinische Universität Erlangen hat ein Tool für die Visualisierung von Krebsdaten für Deutschland entwickelt [2], das uns als Inspiration bei der Entwicklung unseres eigenen Tools diente. Des Weiteren bietet Statistik Austria den sogenannten STATcube [3] an, mit dem man beispielsweise die Krebsdaten erforschen könnte, jedoch ist das Tool sehr umständlich und nicht sehr intuitiv zu bedienen.

## 3 ANSATZ

Zu Beginn bekommt der Nutzer eine Österreichkarte angezeigt, per Mouseover werden dem Benutzer dann die Zahlen für die Mortalitäten oder Inzidenzen (je nach Auswahl) für die einzelnen Bundesländer angezeigt. Diese Darstellung soll dem Benutzer vermitteln, dass es sich um ein Tool handelt, das Krebsdaten für Österreich visualisiert und dazu animieren, sich länger mit dem Tool zu beschäftigen. Für die Farbauswahl benutzten wir den Color Brewer. [4]

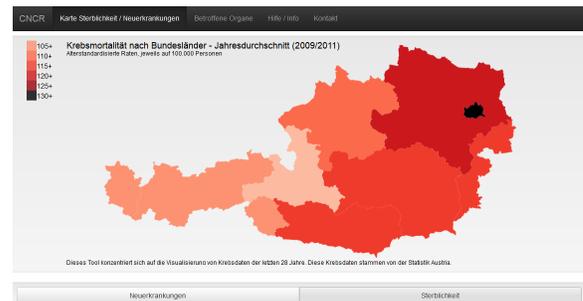


Abbildung 1: Startansicht

Außerdem steht dieser Ansatz in völligem Kontrast zu unserem ersten Entwurf (siehe Abbildung 2), bei dem der User zuerst die Daten, die er angezeigt haben wollte, auswählen musste, was sich aber im Nachhinein als keine gute Idee entpuppte.

Die zweite Ansicht zeigt ein Balkendiagramm mit allen verfügbaren Krebsarten, wobei der Benutzer dann über einen Slider ein Jahr auswählen kann, für das er die Daten sehen möchte. Wir haben uns für ein Balkendiagramm entschieden, da es sich um kategoriale Daten [5] handelt und Balkendiagramme allgemein bekannt sind und somit vom User leicht verstanden werden können. Die Inzidenzen sind im Balkendiagramm und im Liniendiagramm blau gehalten, während die Mortalitäten rot gehalten sind. Im Gegensatz zu unserem ersten Entwurf haben wir die Anzahl der Buttons

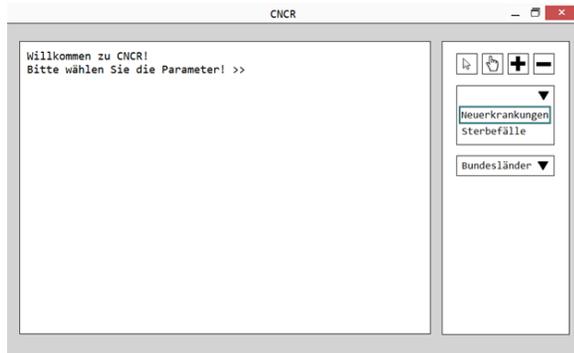


Abbildung 2: Erster Prototyp

auf zwei reduziert, so wird das Tool noch einfacher und kann intuitiv bedient werden. Auf der rechten Seite kann man den Verlauf der Inzidenz- und Mortalitätsrate für ein bestimmtes Organ beobachten, indem man links im Balkendiagramm auf den Balken des entsprechenden Organs klickt. Das Liniendiagramm haben wir gewählt, da es sich gut eignet, um den zeitlichen Verlauf darzustellen [5], außerdem sind auch Liniendiagramme verbreitet genug, um gut verstanden werden zu können. Zur besseren Übersicht haben wir im Balkendiagramm die Jahreszahl inkludiert und im Liniendiagramm das ausgewählte Organ, sodass der User immer weiß, welche Daten er gerade betrachtet. Die Achsen beginnen jeweils bei Null und reichen bis zum jeweiligen Maximum, dadurch, dass die Achsen immer bei Null beginnen, soll der Vergleich der Diagramme für den User erleichtert werden.



Abbildung 3: Zweite Ansicht

Im Gegensatz zur ursprünglichen Idee, wo das Balkendiagramm und das Liniendiagramm auf getrennten Seiten waren (siehe Abbildung 4), befinden sich die beiden Ansichten nun auf einer Seite und "interagieren" miteinander. Wir haben uns für diese Ansicht entschieden, da die ursprüngliche Variante mit den getrennten Seiten für das Balken- und das Liniendiagramm wie zwei verschiedene Tools wirkte und dem User der Zusammenhang verloren ging.

## 4 IMPLEMENTIERUNG

In diesem Abschnitt gehen wir auf die Herausforderungen bei der Implementierung sowie auf die verwendeten Toolkits, Sprachen und Plattformen ein.

### 4.1 Toolkits, Sprachen, Plattformen

Für die Erstellung der Webseite haben wir HTML 5 in Verbindung mit CSS genutzt. Zur Implementierung unseres Tools verwendeten wir Javascript sowie D3.

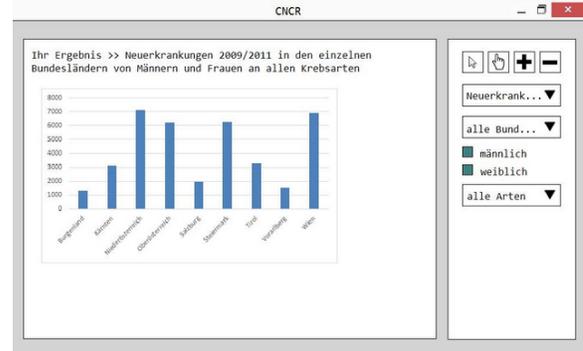
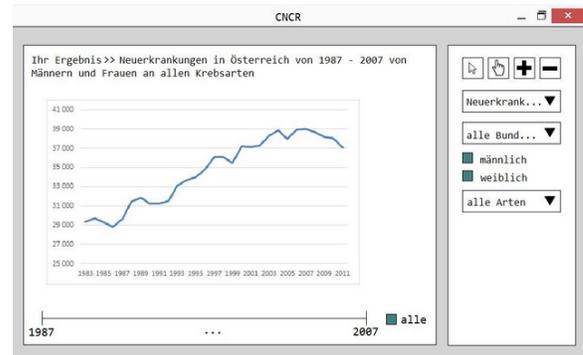


Abbildung 4: Die ursprünglich getrennt geplanten Fenster

## 4.2 Herausforderungen bei der Implementierung

Da wir die Daten zu einem File zusammengefasst haben, stellte das Filtern der Daten für die einzelnen Ansichten unseres Tools eine Herausforderung dar. Mithilfe der D3 Dokumentation [6] konnten wir dieses Problem jedoch schnell lösen. Des Weiteren stellte uns die Anzahl der Jahre bei der Ansicht mit dem Balkendiagramm und dem Liniendiagramm vor eine Herausforderung, da die Achsenbeschriftung mit den Jahren bei unserem ersten Versuch nicht sehr leserlich war. Dieses Problem lösten wir, indem wir nur die Beschriftung für sechs Jahre anzeigen.

## 5 ERGEBNISSE

In diesem Abschnitt zeigen wir anhand eines Use Case Szenarios, wie das Tool verwendet werden kann, außerdem gehen wir auf das Feedback, welches uns im Zuge der Evaluation gegeben wurde, ein.

### 5.1 Use Case Szenario

Elisabeth Schneider, die vor einigen Jahren an Brustkrebs erkrankt ist, möchte sich über das Auftreten und die Sterbefälle von Brustkrebs in Österreich informieren. Sie gelangt beim Aufruf der Seite zuerst zur Österreichkarte, auf der die Inzidenzen und Sterbefälle für die Bundesländer abgebildet sind (siehe Abbildung 5).

Da sie sich jedoch für Brustkrebs interessiert, klickt sie auf den Menüpunkt "Betroffene Organe". Sie erhält die Defaultansicht mit der Bauchspeicheldrüse (siehe Abbildung 6).

Elisabeth S. wählt nun den Balken "Brustkrebs"(Inzidenzen) aus.

Rechts wird ihr nun der zeitliche Verlauf der Inzidenzen (blau) und der Mortalitäten (rot) für Brustkrebs angezeigt (siehe Abbildung 7).

### 5.2 Performanz des Systems

Da wir alle Daten in einem File zusammengefasst haben und nur noch dieses gelesen und gefiltert werden muss, ist das Tool recht

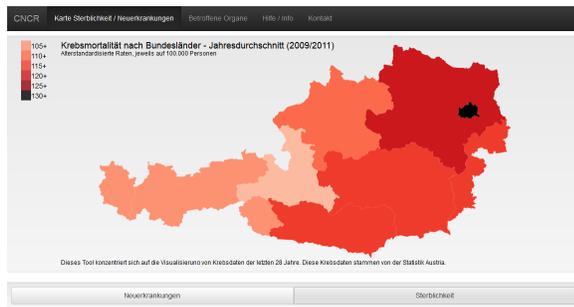


Abbildung 5: Startansicht

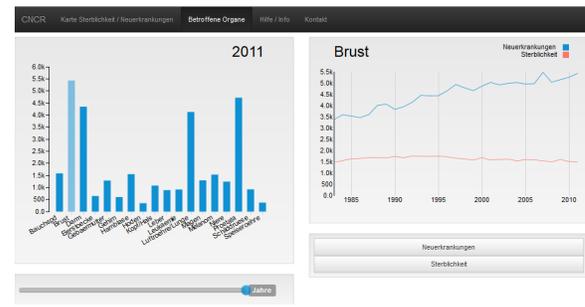


Abbildung 7: Auswahl von Brustkrebs



Abbildung 6: Defaultansicht

performant. Leider hängt die Performanz des Tools jedoch sehr stark von der Qualität der Internetverbindung des Users ab, wie wir im Zuge unserer Evaluierung feststellen mussten.

### 5.3 Feedback aus der Evaluierung

Aus dem Feedback, das uns die Testuser im Zuge der Evaluierung gegeben haben, ging hervor, dass sie zwar dank der Überschriften sofort wussten, um welche Daten es sich handelt, sie sich aber eine Seite wünschen würden, auf der das Projekt kurz beschrieben wird. Aufgrund dieser Kritik, die alle Tester geäußert haben, bekommt man nun unter dem Menüpunkt "CNCR" allgemeine Informationen zum Projekt und die Information, woher die Daten stammen, da dies auch mehrere User interessiert hat. Zwei Drittel der User wussten bei der Verwendung des Tools nicht, dass sie die Balken anklicken können, um das Liniendiagramm für ein spezifisches Organ darstellen zu lassen. Aufgrund dieser Tatsache und dem einstimmigen Wunsch aller User für eine Hilfeseite, haben wir nun eine solche eingerichtet, die durch ein kurzes Video die Verwendung des Tools demonstriert. Außerdem erscheint nun ein Hand-symbol anstatt des Mauszeigers, wenn man den Zeiger über die Balken bewegt. Des Weiteren haben wir den Slider für die Jahreszahlen unter das Balkendiagramm verschoben (anstatt der ursprünglichen Darstellung unter dem Liniendiagramm), da so besser ersichtlich ist, dass damit die Daten für die unterschiedlichen Jahre im Balkendiagramm dargestellt werden. Zudem war sich ein Drittel der User nicht ganz im Klaren über die Begriffe "Mortalität" und "Inzidenz", weswegen wir anstelle dieser Begriffe nun "Sterblichkeit" und "Neuerkrankungen" verwenden. Die Darstellung der Daten wurde jedoch von allen Usern als gut empfunden und sie würden das Tool wieder verwenden.

## 6 DISKUSSION

### 6.1 Stärken und Schwächen des Ansatzes und der Implementierung

#### 6.1.1 Stärken

Die Stärken unseres Tools sind vor allem die gut verständliche Darstellung der Daten und die einfache Bedienung, sodass auch User, die im Umgang mit dem PC nicht so sicher sind, das Tool ohne Schwierigkeiten bedienen können. Sollten trotzdem Schwierigkeiten auftreten, kann jederzeit unsere Hilfeseite aufgerufen werden, die alle Schritte genau erklärt. Da Balkendiagramme, Liniendiagramme und Landkarten auch durch Zeitungen und andere Medien weitläufig bekannt sind, ist diese Darstellung ebenfalls eine Stärke unseres Tools, da man davon ausgehen kann, dass ein Großteil der User keine Probleme haben wird, die Darstellungen zu verstehen, dies hat auch unsere Evaluierung bestätigt.

#### 6.1.2 Schwächen

Die Schwächen unseres Tools liegen darin, dass wir Informationen, die für manche interessant sein könnten, nicht darstellen (z.B. die Aufteilung nach Frauen und Männern). Dies kann aber gleichzeitig auch als Stärke gesehen werden, da der User so nicht überfordert ist von zu vielen Auswahlmöglichkeiten. Manche User könnten außerdem durch das Klicken auf einen anderen Menüpunkt das Gefühl erhalten, dass es sich nicht um ein Tool handelt, sondern um mehrere. Wir haben versucht, dieses Gefühl zu minimieren, indem wir einheitliche Farben und Schriftarten verwenden, sodass der Zusammenhang beziehungsweise die Zusammengehörigkeit klar ersichtlich sind.

### 6.2 Was wir gelernt haben

Wir haben aus diesem Projekt gelernt, dass weniger manchmal mehr ist. Von den vielen ursprünglichen Ideen haben wir nur jene umgesetzt, die zur besseren Verwendung des Tools beitragen und daher auch sinnvoll sind. Außerdem haben wir gelernt, nicht einfach unsere eigenen Wünsche umzusetzen, sondern die der möglichen User.

## LITERATUR

- [1] [http://www.statistik.at/web\\_de/statistiken/gesundheit/krebserkrankungen/index.html](http://www.statistik.at/web_de/statistiken/gesundheit/krebserkrankungen/index.html)
- [2] <http://www.ekr.med.uni-erlangen.de/GEKID/Atlas/CurrentVersion/Inzidenz/atlas.html>
- [3] [http://www.statistik.at/web\\_de/services/datenbank\\_superstar/](http://www.statistik.at/web_de/services/datenbank_superstar/)
- [4] <http://colorbrewer2.org/>
- [5] Munzner, T.: Visualization Analysis and Design
- [6] <https://github.com/mbostock/d3/wiki>