

## Web-PaDaWaN: Eine Web-basierte Benutzeroberfläche für ein klinisches Data Warehouse

Georg Dietrich<sup>1</sup>, Friedrich Fell<sup>1</sup>, Georg Fette<sup>1,2</sup>, Jonathan Krebs<sup>1</sup>, Max Ertl<sup>1,2</sup>, Mathias Kaspar<sup>2</sup>, Stefan Störk<sup>2</sup>, Frank Puppe<sup>1</sup>,

<sup>1</sup> Universität Würzburg, Institut für Informatik, Würzburg, Deutschland.

<sup>2</sup> Universität Würzburg, Deutsches Zentrum für Herzinsuffizienz, Würzburg, Deutschland.

Schlagworte: Data Warehouse; Data Mining; Studienunterstützung

Ein klinisches Data Warehouse macht die Daten eines Krankenhausinformationssystem für klinische Forschung verfügbar und muss daher über eine gute Benutzungsoberfläche verfügen. Beispiele für letztere sind Medical Research Insights von SAP [1], die Benutzungsoberfläche von I2B2 [2] oder PaDaWaN (Patienten-Data-Warehouse-Navigator) [3]. Anforderungen für die Benutzeroberfläche umfassen eine komfortable und ausdrucksstarke Abfragesprache sowohl für einfache als auch komplexe Anfragen, schnelle Antwortzeiten auch bei großem Datenvolumen und eine übersichtliche, weiterverwendbare Darstellung der Ergebnisse. In [3] wurde die grundsätzliche Struktur der Würzburger Data Warehouse Oberfläche PaDaWaN vorgestellt. Die Anforderung, schnelle Antwortzeiten bei großem Datenvolumen einschl. Suche in Textdaten zu gewährleisten, wurde durch eine Indexierung und Verwendung der open source Suchmaschine Lucene erreicht. Im Kontext der Weiterentwicklung zu einer Web-Anwendung des PaDaWaN wurde die Anfrageoberfläche grundlegend überarbeitet. Anfragen können sich auf Häufigkeiten oder auf Eigenschaften einzelner Patienten beziehen. In diesem Beitrag gehen wir nur auf Häufigkeiten ein. Ein Beispiel zeigt Abb. 1

The screenshot displays the Web-PaDaWaN interface. On the left is a 'Parameter' tree with categories like 'Stammdaten (4551281)', 'Station (4528592)', and 'Diagnose (3103286)'. The main area is the 'Anfrage' (Query) builder, which includes a search bar, a dropdown for 'Brieftext', and radio buttons for 'als Zeile', 'als Filter', and 'als Spalte'. The query is: 'I50 : Herzinsuffizienz'@Diagnose UND 'Medizinische Klinik'@Station. Below the query, there are filters for 'Alter <= 60', 'Alter 60...80', and 'Alter > 80'. A 'Zeitraum' (Time Range) is set from 2010 to 2012, with 'als Spalten' (as columns) selected and 'pro Jahr' (per year) as the aggregation. At the bottom, there are buttons for 'Suchen', 'Speichern', 'Laden', 'Excel', and 'CSV'. The 'Ergebnistabelle' (Results Table) shows the following data:

	Alle (...)	Patienten	2010	2011	2012
Alle ((I50 : Herzinsuffizienz UND Medizinische Klinik))	13 621	3 196	1 083	1 205	1 546
Alter <= 60	2 297	570	202	226	275
Alter ]60, 80]	7 678	1 732	589	659	840
Alter > 80	4 177	938	292	320	431

**Abb. 1:** Beispiel für Anfrage und Ergebnistabelle im Web-PaDaWaN (Erklärung im Text).

Eine typische Anfrage besteht aus drei Schritten: Definiere Kontext (Filter), Zeilen und Spalten des gewünschten Ergebnisses. Die resultierende Tabelle zeigt die Häufigkeiten von Patienten für alle Zeilen/Spalten-Kombinationen an, die alle Kontextbedingungen erfüllen. Um die Anfrage kompakt darstellen zu können, wird die Parametrisierung jeder Eingabezeile durch Radio-Buttons als Zeile, Spalte oder Filter gekennzeichnet (oberer gelber Bereich in Abb. 1). Eine Eingabezeile kann dabei ein mit UND, ODER oder NEGATION verknüpfter geschachtelter logischer Ausdruck sein, z. B. in der ersten Zeile in Abb. 1 "Diagnose = I50 (Herzinsuffizienz) UND Station = Medizinische Klinik", die als Filter gekennzeichnet ist. Die weiteren Zeilen (Alter ...) beschreiben Zeilen der Ergebnistabelle. Die Spalten mit Aufteilung nach Jahren sind der Bequemlichkeit halber direkt eingebbar. Sie hätten auch in zusätzlichen Eingabezeilen über den Parameter "Aufnahme-Datum" im linken Fenster in Abb. 1, das einen aufklappbaren Katalog aller Parameter mit Anzeige der Häufigkeit des Vorkommens im Data Warehouse enthält, spezifiziert und als "Spalte" gekennzeichnet werden können. Im unteren Bereich von Abb. 1 wird die berechnete Ergebnistabelle angezeigt (die Zahlen sind fiktiv). Die Eigenschaften der Patienten können sich auf den Katalog beziehen oder als Freitextsuche aus Textdokumenten wie Arztbriefen oder Befundberichten stammen. Die Eingabe wird durch Autovervollständigung mit Vorschlägen der häufigsten Parameter bzw. Wörter unterstützt. Im Vergleich zur Eingabeoberfläche Medical Research Insights und von I2B2 zeichnet sich der Web-PaDaWaN durch eine wesentlich kompaktere Eingabeoberfläche aus, in der auch komplexe Anfragen mit vielen Einzelementen noch übersichtlich eingegeben werden kann. Allerdings stehen systematische Usability Studien mit verschiedenen Nutzergruppen und Anfragetypen noch aus.

Diese Arbeit wurde unterstützt durch die Förderung des Bundesministeriums für Bildung und Forschung (BMBF01 EO1004)

Referenzen:

[1] Medical Research Insights: <https://icn.sap.com/projects/sap-medical-research-insights.html>

[2] I2B2: <https://www.i2b2.org/>

[3] Fette, G., Dietrich, G., Ertl, M., Toepfer, M., Kaspar, M., Störk, S., Puppe, F.: Die grafische Benutzeroberfläche PaDaWaN für das klinische Data Warehouse für Routinedaten an der Universitätsklinik Würzburg, Abstract auf der GMDS 2015, <http://www.egms.de/static/en/meetings/gmds2015/15gmds023.shtml>