

Team Wallraff: IVENA – ein datenjournalistisches Projekt

Jörn Ratering



Jörn Ratering
infoNetwork GmbH
joern.ratering@
infonetwork.de

Innerhalb unserer Abteilung bestand schon seit längerem die Idee, unsere Kompetenzen als Medien-dokumentare in ein datenjournalistisches Projekt einfließen zu lassen. Nachdem ich ein Datenjournalismus-Seminar des VFM besucht hatte, traf ich mich im April 2015 mit Redakteuren des investigativen Enthüllungsformats „Team Wallraff“. In diesem arbeiten Undercover-Reporter als Mitarbeiter in Institutionen oder Wirtschaftsbetrieben und filmen ihre Erfahrungen mit versteckter Kamera. In der derzeitigen Produktion wollte man thematisch den teils katastrophalen Zustand in deutschen Krankenhäusern dokumentieren: unzureichende Maßnahmen zur Hygiene, zu wenig angestellte Pflegekräfte und Privatisierung von Krankenhäusern zur Gewinnmaximierung der Krankenhausbetreiber. Man präsentierte mir die aktuellen Recherchen zur schlechten Situation der Versorgung in den Notaufnahmen. Für diese gibt es für Städte in Hessen und in München über die IVENA-Internetseiten¹ die Möglichkeiten, aktuelle Behandlungs- und Versorgungsmöglichkeiten graphisch im 10-Minuten-Takt darzustellen (Abbildung 1). IVENA dient damit

„Zentralen Leitstellen graphisch für den Rettungsdienst, den Gesundheitsbehörden und anderen medizinischen Diensten (...) verletzte und/oder erkrankte Patienten rasch in das ihren Verletzungen entsprechend geeignete Krankenhaus zu führen, in denen sie behandelt werden können.“²

Aus datenjournalistischer Sicht können die IVENA-Daten als „Transparent Data“ bezeichnet werden, die zwar „frei zugänglich (...), jedoch nicht in Tabellenform und nicht in die gewünschte Relation gesetzt“ vorliegen.³

Problematisch an der Auswertung der anfallenden Daten war, dass immer nur die zurückliegenden fünf Stunden angezeigt werden. Auf der Seite ist kein

Archiv einsehbar. Daher erstellten die RedakteurInnen von Team Wallraff seit Februar 2015 im fünf-Studentakt Screenshots der IVENA-Seiten, um die teils frappierende Auslastung der Krankenhäuser zu dokumentieren. Diese wurden in ein Word-Dokument gespeichert und ausgedruckt. Ziel war es, die so anfallenden Daten statistisch auszuwerten, die Ergebnisse als Basis für Aussagen über den Zustand in den Krankenhäusern zu nehmen und für eventuelle justiziable Angelegenheiten nach der Ausstrahlung eine lückenlose und korrekte Dokumentation der Daten vorweisen zu können. Leider wurden die Screenshots mit verschiedenen Browsern und Endgeräten angefertigt. Sie waren also uneinheitlich formatiert und daher eine maschinenlesbare Auswertung nicht möglich. Ein weiteres Erfassen und Auswerten von Screenshots der Krankenhäuser samt ihrer verschiedenen Fachbereiche auf diese Art und Weise hätte enorm viel zeitliche Ressourcen verschlungen.

Wir schrieben also ein automatisches Screen-Scraping-Programm, das die Screenshots und die IVENA-Seite als HTML-Dokument von nun an automatisch zu einer festgelegten Uhrzeit speicherte⁴ (Abbildung 2). Dadurch wurden die RedakteurInnen entlastet und das „Überwachen“ der IVENA-Seiten vollautomatisch übernommen. Screen-Scraping (zu Deutsch „Bildschirm auskratzen“), ist eine Methode zum automatischen Auslesen von Texten von Computerbildschirmen.⁵ Mit einem ebenfalls geschriebenen Parser-Programm wurde eine Datenbank mit dem jeweiligen auswertbaren Status (1 = Grün, 2 = Rot, 3 = Grau) des Krankenhauses gefüllt. Um einem Ausfall des Computers, auf dem das Screen-Scraping Programm installiert war, vorzubeugen, ließen wir zusätzlich einen Backup-Rechner laufen.

Die Datenbank wurde so mit dem jeweiligen Status des Krankenhauses, der Uhrzeit, dem Datum, der Auslastung (wie viel Prozent der Krankenhäuser

¹ www.ivena-hessen.de und www.ivena-muenchen.de, Ivena steht für Interdisziplinärer Versorgungsnachweis.

² www.ivena.de

³ vgl. Grieb, Andreas 2011: Die Datenzweibel. URL: <http://www.youdaz.com/2011/06/20/die-datenzweibel/>

⁴ Um bei den Betreibern der Seite nicht dadurch aufzufallen, dass störisch exakt zur gleichen Zeit stündlich ein Zugriff mit einer IP-Adresse aus Köln erfolgte, bauten wir eine Varianz von ein paar Minuten ein.

⁵ vgl. Hoely, Daniel 2011: Scraping: Ist das Kopieren von Daten aus dem Netz erlaubt? URL: <http://blog.kooptech.de/2011/09/scraping-ist-das-kopieren-von-daten-aus-dem-netz-erlaubt/>

Leitstelle Offenbach		07:00	08:00	09:00	10:00	11:00
Ketteler Krankenhaus						
Klinik Dr. Frühauf						
Sana Klinikum Offenbach	TV VAV					

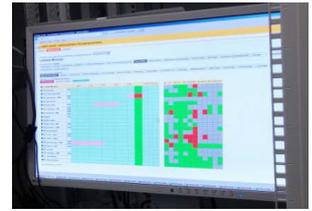


Abb. 1: Darstellung der Möglichkeit zur Anfahrt der Notaufnahme der Krankenhäuser der Stadt Offenbach auf der IVENA-Seite ab 6:20 Uhr zum Zeitpunkt x, Rot = ausgelastet, Grün = anfahrbar, Grau = keine Daten.

Abb. 2: Computer auf dem dauerhaft das Screen-Scraper-Programm lief.

ID	Datum	Region	Fachgebiet	Fachbereich	Dringlichkeit	24h	Krankenhaus	Status	Kennziffer	Auslastung	Wochentag	Filename	Filepath	Aktualisieru
2414376	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:00:00	Bethanien Krai	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414377	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:10:00	Bethanien Krai	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414378	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:20:00	Bethanien Krai	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414379	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:30:00	Bethanien Krai	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414380	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:40:00	Bethanien Krai	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414381	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:50:00	Bethanien Krai	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414382	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:00:00	Elisabethen Kr	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414383	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:10:00	Elisabethen Kr	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414384	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:20:00	Elisabethen Kr	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414385	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:30:00	Elisabethen Kr	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414386	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:40:00	Elisabethen Kr	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414387	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:50:00	Elisabethen Kr	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414388	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:00:00	Hospital zum F	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2414389	10.09.2015	Frankfurt	Innere Medizin	Allgemeine Ini	Stationäre Ver	00:10:00	Hospital zum F	1	12/11/1/0	92	Donnerstag	20150910_0015	C:\test\Data_k	00:15:59
2297656	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:40:00	Asklepios Paul	2	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297657	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:50:00	Asklepios Paul	2	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297658	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:00:00	Horst Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297659	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:10:00	Horst Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297660	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:20:00	Horst Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297661	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:30:00	Horst Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297662	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:40:00	Horst Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297663	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:50:00	Horst Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297664	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:00:00	Horst-Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297665	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:10:00	Horst-Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297666	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:20:00	Horst-Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297667	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:30:00	Horst-Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297668	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:40:00	Horst-Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297669	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:50:00	Horst-Schmidt	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297670	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:00:00	St. Josefs-Hosp	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297671	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:10:00	St. Josefs-Hosp	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297672	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:20:00	St. Josefs-Hosp	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297673	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:30:00	St. Josefs-Hosp	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297674	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:40:00	St. Josefs-Hosp	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297675	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	00:50:00	St. Josefs-Hosp	1	4/3/1/0	75	Donnerstag	20150910_0025	C:\test\Data_k	00:25:09
2297676	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	01:00:00	Asklepios Paul	2	4/3/1/0	75	Donnerstag	20150910_0124	C:\test\Data_k	01:24:33
2297677	10.09.2015	Wiesbaden	Innere Medizin	Innere Intensiv	Notfallversorg	01:10:00	Asklepios Paul	2	4/3/1/0	75	Donnerstag	20150910_0124	C:\test\Data_k	01:24:33

in der Stadt sind zum Zeitpunkt noch erreichbar), dem Wochentag und dem Speicherort des Screenshots gefüllt (Abbildung 3).

Meine Aufgabe war außerdem die Aufbereitung der bislang vom Team Wallraff manuell erstellten Screenshots, da in diesem Zeitraum die investigativen Rechercharbeiten mit versteckter Kamera vor Ort stattfanden und diese Daten daher von besonderer Relevanz waren. Eindrücke, die die Redakteurin mit versteckter Kamera (z.B. von überfüllten und überforderten Notaufnahmen) in den Krankenhäusern sammelte, konnten so im Nachhinein mit den IVENA-Daten abgeglichen werden. Zu diesem Zweck programmierte ich ein Excel-Makro, das die Farbwerte eines Kästchens erkennt. Über ein so entstehendes Auswertungssheet wurden die analog vorliegenden ausgedruckten Screenshots, mit Hilfe der Excel-Schaltfläche „Format übertragen“ digital reproduziert (nachgemalt).⁶ Das Makro generierte dann automatisch die Werte über eine Formel in das maschinenlesbare Datenformat (Abbildung 4). Daraufhin wurden die Daten, identisch formatiert wie die vom Screen-Scraping-Programm automatisch erfassten Werte, in die gemeinsame Datenbank eingefügt. Mit diesen Methoden gelang es uns die Auslastung

für Krankenhäusern in sieben Städten⁷, in verschiedensten Fachgebieten, im manuell erschlossenen (von Februar bis Mai 2015) und im automatischen erschlossenen Zeitraum (von Juni 2015 bis zur Ausstrahlung am 11. Januar 2016) zu dokumentieren. Dabei fand ein enger Austausch zwischen uns Mediendokumentaren und den Redakteuren des „Team Wallraff“ statt, um die Ziele des Projekts immer wieder neu zu formulieren. So konnten wir die IVENA-Daten mit Google Grippe-Trends und Wetterdaten abgleichen, um weitere Schlüsse zu ziehen.

Für die Auswertung und Aufbereitung der Daten verwendete ich Pivot-Tabellen. Diese können flexibel angepasst werden und bei einer großen Datenmenge an ihren Kreuzungspunkten Summen und Durchschnittswerte generieren und Daten anschaulich visualisieren (Abbildung 5). Die Bedienung ist für den Redakteur leicht verständlich und für die jeweilige Fragestellung beliebig anpassbar.

Abschließend verfasste ich eine Auswertung über die Ergebnisse, welche die Redakteure als Basis für Aussagen über den Zustand in den Krankenhäusern nahmen. Investigative Rechercheergebnisse konnten anhand der Daten überprüft und neue Erkenntnisse gewonnen werden.

Abbildung 3: Ausschnitt aus der Datenbank. Markiert ist die Notaufnahme der Asklepios Paulinen Klinik in Wiesbaden um 00:40 Uhr am 10.09.2015. Sie hat den Status 2 (Rot). Zu diesem Zeitpunkt sind 75% aller Notaufnahmen in Wiesbaden anfahrbar.

⁶ An dieser Stelle gilt der Dank unter anderem fleißigen Studenten, die diese Aufgabe durchführten.

⁷ München, Wiesbaden, Frankfurt, Gießen, Kassel, Offenbach, Darmstadt

Abbildung 4: Auswertungssheet zur Ermittlung der Daten aus den manuell angelegten Screenshots für Kassel. Rechts im maschinen-lesbaren Format, das in die Datenbank übernommen wird (1 = Grün, 2 = Rot, 3 = Grau). Oben der „nachgemalte“ Screenshot. Das Feld Pinsel dient als Vorlage für die Funktion „Format übertragen“.

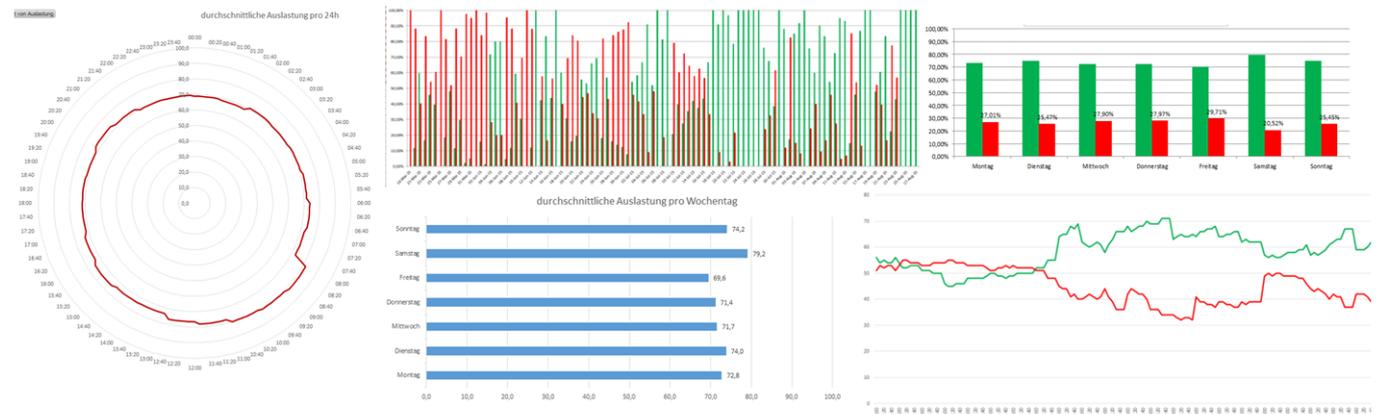
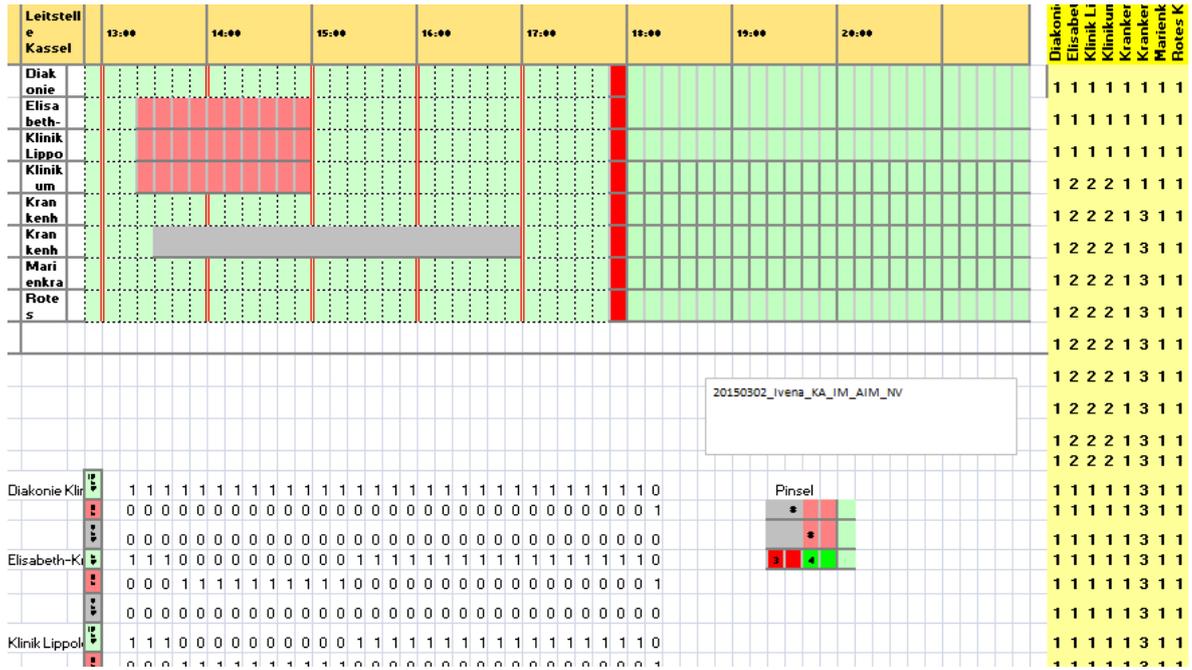


Abbildung 5: graphische Auswertungen (u.a. Anzahl an Status Rot und Grün im Krankenhaus x pro Tag, pro Wochentag, im Zeitraum 24h; Auslastung für das gesamte Stadtgebiet).

Zu guter Letzt waren wir Mediendokumentare sogar noch vor der Kamera und Bestandteil der ausgestrahlten Sendung (Abbildung 6).

Innerhalb des Datenjournalismus entwickelt sich meiner Ansicht nach ein zukünftiges Betätigungsfeld für die Berufsgruppe der Mediendokumentare. Das „Team Wallraff“-Projekt der Auswertung von IVENA-Daten ist in diesem Sinne konsequent umgesetzt worden: Von der Datenextraktion, über die strukturierte Auswertung bis hin zur anschaulichen Visualisierung. Daten repräsentieren dabei die Wirklichkeit wie Video-Rohmaterial und müssen im Produktionsprozess schnell verfügbar sein. Dazu müssen sie zuvor systematisch aufbereitet und sinnvoll ausgewertet werden; neue Schnittstellenkompetenzen erworben, neues Fachwissen in Sachen Statistik, Quellenkompetenz und Datenformat aufgebaut werden.

*Vortragsmanuskript (gehalten auf der Frühjahrstagung des vfm am 26. April 2016)



Abbildung 6: Ausschnitt aus der Sendung "Team Wallraff" vom 11.01.2016. Von links nach rechts: Jörn Ratering, Pia Osterhaus (Redakteurin), Robert Zängerle (Projektpartner).