

# Texte und Thesaurus

Notizen zu BAIK

Wolfgang Giere

30. September 2020

## Inhaltsverzeichnis

<b>1 Ziel dieser Notizen</b>	<b>1</b>
<b>2 Freiheit des ärztlichen Ausdrucks: unbegrenzt</b>	<b>2</b>
<b>3 GMDS-Arbeitsgruppe Klartextanalyse</b>	<b>2</b>
<b>4 Sortierwort bei MEDIUC: Symptom-Lexikon</b>	<b>2</b>
<b>5 IATROS mit der Abfragesprache IATINT</b>	<b>3</b>
<b>6 Übersetzung von DXplain</b>	<b>3</b>
<b>7 Operationenschlüssel</b>	<b>3</b>
<b>8 Übersetzung der Prozedurenliste von 3M</b>	<b>4</b>
<b>9 Diagnosenverschlüsselung und ICD-10-Diagnosethesaurus</b>	<b>4</b>
<b>10 Transatlantisches Forschungsvorhaben MuchMore</b>	<b>5</b>

## 1 Ziel dieser Notizen

Meine bisherigen BAIK-Seiten im Web ergeben zwar ein chronologisches Bild der Entwicklung, aber mir fehlen Querschnittsbetrachtungen zu mir wichtigen Themen, beispielsweise zur Textverarbeitung und Klassifikation. Hiermit möchte ich diese nachliefern – soweit möglich. Denn meine Ansichten und Arbeiten zu medizinischen Texten liegen mir besonders am Herzen. Deswegen diese kompakte Gesamtbetrachtung.

## 2 Freiheit des ärztlichen Ausdrucks: unbegrenzt

Untertützung des Arztes, um valide Dokumentation als Grundlage für die Forschung zu gewinnen war das Ziel aller meiner Arbeiten. Do ut des, gib mir Daten und ich erleichtere Dir die Arbeit. Das war das Prinzip der Elektronischen Krankenakte von vorneherein. Dabei galt immer: Kodes für Häufiges, Klartext für Seltenes – ohne jede Beschränkung. Ich habe stets Hochachtung vor der ärztlichen Kreativität gehabt, und deswegen war in BAIK Klartextverarbeitung vom ersten Tage an (seit 1967) konstitutiver Bestandteil<sup>1</sup>.

## 3 GMDS-Arbeitsgruppe Klartextanalyse

Über die Natur medizinischer Texte habe ich viel nachgedacht und zusammen mit den Kollegen Röttger, Wingert und anderen schon 1969 die Arbeitsgruppe Klartextanalyse der GMDS gegründet. Den damals angelegten multidimensionalen sog. AGK-Thesaurus habe ich dann bis zu meiner Emeritierung gepflegt. Er war die Grundlage für alle meine Beiträge zur medizinischen Nomenklatur vom *Sortierwort* bei MEDIUC über IATROS, die DXplain-Übersetzung, Operationenschlüssel bis hin zum ICD-10-Dignosethesaurus und schließlich MuchMore.

## 4 Sortierwort bei MEDIUC: Symptom-Lexikon

Das sogenannte Sortierwort bei MEDIUC hatte zum Ziel, ein systematisch geordnetes Verzeichnis der Symptome drucken zu können. Dazu wurden die Symptome klassifiziert. Es gab eine Hauptkategorie und multiple Unter-Aspekte, so dass insgesamt eine recht genaue multiaxiale Klassifikation der Symptome ermöglicht wurde. Diese Aktivität von 1968/69 blieb damals mit meinem Weggang<sup>2</sup> liegen. Ich habe das stets bedeuert, denn ein sytematisches Symptom-Lexikon fehlt meines Wissens weltweit bis heute. Den Entwurf eines ontologischen Rahmens zur Klassifikation von Symptomen habe ich seinerzeit in Stuttgart geliefert (siehe MEDIUC-Buch).

---

<sup>1</sup> Diese unabdingbare Forderung nach beliebig langen Texten und Ergänzungsmöglichkeiten mit Kommentaren bedeutete variable Längen und verbot die Benutzung von Lochkarten, die damals noch üblich waren. Deswegen entschieden sich Herr Baumann, der Systemprogrammierer, und ich, sein Lehrling, seinerzeit für Lochstreifen als Eingabemedium; und deswegen entwickelte ich die komplexe Datenstruktur und Feldsystematik von BAIK, die wiederholt beschrieben wurde (siehe BAIK-Buch), und die sich in vielfältiger Praxis sehr bewährt hat.

<sup>2</sup> Von der Med. Biol. Forschungsstelle des Robert-Bosch-Krankenhauses in Stuttgart (Leiter Dr. med. R. Pirtkien) wechselte ich zum Aufbau-Team der Deutschen Klinik für Diagnostik (DKD), Wiesbaden, das in Frankfurt arbeitete.

## 5 IATROS mit der Abfragesprache IATINT

In der DKD entwickelten wir für das Kode-Klartextgemisch der BAIK-Befunde die Klassifikationsmöglichkeit *Informations-Aufbereitendes Text-Retrieval-Orientiertes System (IATROS)* und die interaktive Abfragesprache *IATROS-INTERpreter*. Wir haben IATROS erstmals auf der GMDS-Jahrestagung 1973 in Bielefeld vorgestellt. Die Abfragesprache hat sich schon an der DKD bewährt, z.B. bei der Lösung der Frage, ob das Sandgefühl im Auge mit der Einnahme der Pille zu tun hätte. Die Leistungsfähigkeit der Systematik zeigte sich später in Frankfurt besonders eindrucksvoll an der Klärung der Frage, ob es schon *vor* der Benennung von *AIDS* bzw. *HIV* Fälle gab, die seziiert wurden<sup>3</sup>. Das Ergebnis war positiv: es wurden tatsächlich Fälle gefunden, die aber natürlich nicht als "HIV" oder "AIDS" klassifiziert waren, weil die Diagnose zum Zeitpunkt der Sektion noch nicht bekannt war. Wir waren stolz auf die Leistungsfähigkeit unseres Systems IATROS. Welches andere System hätte diese Recherche erlaubt? Noch dazu online, interaktiv.

Übrigens habe ich später in einem Sabbaticum am Mass. General Hospital (MGH) im Laboratory of Computer Science (LCS) bei G. Octo Barnett seine *COSTAR*-Abfragesprache an die BAIK-Datenstrukturen adaptiert und so die *BAIK-Abfrage-Sprache mit Integrierter Statistik (BASIS)* geschaffen.

## 6 Übersetzung von DXplain

DXplain ist das Diagnose-unterstützende System von G.O. Barnett aus dem Mass. Gen. Hospital bzw. der Harvard Universität. Ich hatte es bei meinen Besuchen kennengelernt. Außerdem kannte ich G. W. (Bill) Moore im Johns Hopkins Hospital in Baltimore und sein Übersetzungsprogramm *Transsoft*. Er hatte es für die Übersetzung des Pathologiebuches von Letterer<sup>4</sup> aus dem Deutschen ins Englische geschrieben. Ich wollte es nun für die Übersetzung vom Englischen ins Deutsche benutzen. Dafür mussten die Lexika und Regelwerke neu angelegt werden. Dazu arbeiteten Bill und ich erfolgreich zusammen. Die damals günstigste Kommunikationsmöglichkeit für große Listen war Fax<sup>5</sup>. Das Ergebnis, meine DXplain-Übersetzung, konnte ich auf der MEDICA präsentieren.

## 7 Operationenschlüssel

Auf derselben Ausstellung bekam ich übrigens den Auftrag vom DIMDI, einen Operationenschlüssel zu entwickeln. Der Leiter, Dr. Fritz, hatte mich schon vorher telefonisch gefragt, ob ich bereit sei, die *International Classification of Procedures in Medicine (ICPM)*

<sup>3</sup> Alle Sektionsprotokolle am Uniklinikum wurden mit BAIK geschrieben. So wurde auch der AGK-Thesaurus ständig erweitert und gepflegt.

<sup>4</sup> Erich Letterer, Allgemeine Pathologie: Grundlagen und Probleme. Thieme, Leipzig 1959

<sup>5</sup> Einige Monate haben wir intensiv daran gearbeitet. Die dicken Faxstapel mit Tests und Verbesserungen der Tabellen hatte ich jahrelang in meinem privaten Arbeitszimmer im Regal.

der WHO zu übersetzen. Im Hinblick auf das funktionierende Übersetzungsprogramm sagte ich damals zu. Nun aber sollte ich unter Leitung vom DIMDI mit Beteiligung der Fachgesellschaften in kurzer Zeit (wenn ich mich richtig erinnere in vier Monaten bis Ende März) einen deutschen Neuentwurf liefern. Ich sagte zu und wir<sup>6</sup> lieferten Ende März pünktlich die *Internationale Klassifikation der Prozeduren in der Medizin (IKPM)* ab. Weil die ICPM seit Jahrzehnten veraltet war, hatte ich vorgegeben: Wir bilden neue Nummernkreise mit mehr "Luft" in den Bereichen, in denen die Fortschritte der Chirurgie besonders ausgeprägt waren; es muss nur die algorithmische Aufwärts-Kompatibilität zur ICPM gesichert sein. Anmerkung: Diesen absolut eigenständigen Entwurf halte ich auch heute noch für besser, als den aktuellen OPS, der auf Wunsch des Gesundheitsministeriums in Kooperation mit ID, Berlin, entstand<sup>7</sup>.

## 8 Übersetzung der Prozedurenliste von 3M

Auch für die Firma 3M haben ich<sup>8</sup> übersetzt: Bei der Einführung der *Diagnosis Related Groups (DRG)* als Abrechnungsprinzip der Krankenhäuser habe ich deren Prozedurenbeschreibung ins Deutsche übersetzt.

## 9 Diagnosenverschlüsselung und ICD-10-Diagnosethesaurus

Der ICD-10-Diagnosethesaurus ist mein erfolgreichstes Produkt und das ist eigentlich ein Treppenwitz der Geschichte: Als ich von der DKD aus Wiesbaden weg nach Frankfurt berufen wurde, gab es im Klinikum der J.W.Goethe-Universität in einigen Kliniken eine Diagnosedokumentation unter Benutzung von ICD-9. Ich persönlich bevorzugte Freitext-Dokumentation und Klassifikation ad hoc abhängig vom Forschungsziel, wie es IATROS erlaubte und habe die Bemühungen um eine allgemeingültige Diagnosenklassifikation nicht unterstützt. Aber bei einem Kongress der Telekom in Berlin traf ich den Geschäftsführer des *Zentralinstitut für die kassenärztliche Versorgung der Bundesrepublik Deutschland (ZI)*, Herr Brenner, wieder. Der fragte mich, was man denn für die Niedergelassenen Ärzte tun könne, die jetzt die Diagnosen verschlüsseln sollten. Da berichtete ich ihm von der Existenz des AGK-Thesaurus, in dem ja alle Diagnose-Wörter enthalten seien, die in der DKD und in der Pathologie der Uniklinik Frankfurt benutzt wurden. Dieses Gespräch führte zur Gründung einer Arbeitsgruppe Diagnosethesaurus unter meiner Leitung. Das Produkt wurde dann in zwei Bundesländern, einem west-

---

<sup>6</sup> Beim CBI in Wiesbaden hatte ich gerade einen Weiterbildungskurs in Medizinischer Informatik für arbeitslose Ärzte beendet und konnte kurzfristig zwei gute Schülerinnen für diese Aufgabe rekrutieren.

<sup>7</sup> Das Gesundheitsministerium lehnte unseren Entwurf ab und zwang ID und uns, das ZInfo, zur gemeinsamen Formulierung eines neuen Entwurfs ohne Änderung der Systematik des ICPM. Viele Monate tagten wir wöchentlich gemeinsam im DIMDI und feilten in langen Diskussionen an Klassifikation und Benennung von chirurgischen Prozeduren. Nur mein Vorschlag der x/y-Systematik fand Anklang und wurde von der IKPM übernommen.

<sup>8</sup> Übersetzungen habe ich meist persönlich gefertigt.

und einem ost-deutschen, erfolgreich getestet. Und dann hat das DIMDI dem ZI den Diagnosethesaurus abgekauft zur Benutzung auch in den Krankenhäusern. Heute wird der ICD-10-Diagnosethesaurus in allen deutschsprachigen Ländern eingesetzt (Deutschland, Österreich und Schweiz) und vom DIMDI jährlich neu herausgegeben<sup>9</sup>.

Außerdem arbeiteten wir für BAIK weiter an der automatischen Diangosekodierung mit Xmed. Für einen Auftrag der Bundesärztekammer wurden drei Verfahren an Millionen Datensätzen getestet. Herr Graubner, ein Experte, bewertete die Ergebnisse ohne zu wissen, um welche drei Verschlüsselungssysteme es sich handelte. Unser Xmed schnitt am besten ab. Ich vermute, die beiden anderen Systeme stammten von ID, Berlin, und Frau Kolodzig<sup>10</sup>.

## 10 Transatlantisches Forschungsvorhaben MuchMore

Wortgleiche Anträge<sup>11</sup> bei der *Europäischen Union (EU)* und der *US-National Science Foundation (NSF)* wurden beidseits des Atlantiks genehmigt und führten zu dem spannenden Vorhaben *Multilingual Concept Hierarchies for Medical Information Organization and Retrieval (MuchMore)*, an dem das ZInfo beteiligt war als medizinischer Partner. Letztlich ging es um Vergleich der Leistungsfähigkeit von Klassifikation mit ontologiebasierten und statistischen Methoden, Thesaurus versus big-data Evaluation.

---

<sup>9</sup> Meine ehemalige Schülerin und hochgeschätzte Mitarbeiterin Dr. B. Krause leitet inzwischen die Pflege beim DIMDI.

<sup>10</sup> Frau Dr. Christine Kolodzig war bei der Entwicklung des Operationenschlüssels die Experte vonseiten der Firma ID. Sie hatte sich aber zum Zeitpunkt des Vergleichs schon von ID getrennt und ihre eigene Firma (SBG Software und Beratung im Gesundheitswesen GmbH, Berlin) aufgemacht.

<sup>11</sup> Auf der Deckseite der Anträge befand sich das BAIK-Modell.

## Abkürzungsverzeichnis

- 3M: 3M Medica, 3M Deutschland GmbH, Neuss, 4
- AGK: Arbeitsgruppe Klartextanalyse der GMDS, 2
- AIDS: Acquired Immune Deficiency Syndrome, 3
- BAIK: Befunddokumentation und Arztbriefschreibung im Krankenhaus, 1
- BASIS: BAIK-Abfrage-Sprache mit Integrierter Statistik, 3
- CBI: Computer-Bildungs-Institut GmbH in Wiesbaden, 4
- DIMDI: Deutsches Institut für Medizinische Dokumentation, Köln, 4
- DKD: Deutsche Klinik für Diagnostik, Wiesbaden, 2
- DRG: Diagnosis Related Groups, 4
- EU: Europäische Union, 5
- GMDS: Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie GMDS e.V. (ex Gesellschaft für Medizinische Dokumentation und Statistik), 2
- HIV: Human immunodeficiency Virus, 3
- IATROS: Informations-Aufbereitendes Text-Retrieval-Orientiertes System, 3
- ICD: International Classification of Diseases der WHO, 2
- ICPM: International Classification of Procedures in Medicine, 3
- ID: ID Information und Dokumentation im Gesundheitswesen GmbH & Co. KGaA, 4
- IKPM: Internationale Klassifikation der Prozeduren in der Medizin, 4
- LCS: Laboratory of Computer Science, 3
- MEDICA: Medizin-Fachmesse, jährlich im November in Düsseldorf, 3
- MEDIUC: Medizinische Diagnose Unterstützender Computer, 2
- MGH: Massachusetts General Hospital, 3
- MuchMore: Multilingual Concept Hierarchies for Medical Information Organization and Retrieval, 5
- NSF: US National Science Foundation, 5
- OPS: Operationen-Schlüssel, 4
- SBG: Software und Beratung im Gesundheitswesen GmbH, Berlin, 5
- WHO: World Health Organization, 2
- Xmed: BAIK-Klassifikationssystem, 5
- ZI: Zentralinstitut für die kassenärztliche Versorgung der Bundesrepublik Deutschland, damals Köln, 4
- ZInfo: Zentrum der Medizinischen Informatik des Klinikums der J.W.Goethe-Universität, 4