

# Archivisches Suchen und moderne Informatik

*Von Andreas Kellerhals-Maeder*

## 1 Einleitung

Fast seit alters beklagen die Archivare die Flut von Schriftgut, welche in die Archive geschwemmt wird, um dort vielfältige Probleme zu verursachen<sup>1</sup>. Diese Aktenflut ist in den letzten Jahren nicht kleiner geworden - im Gegenteil. Die Informatik verspricht zwar bereits seit längerer Zeit das «papierlose Büro», hat aber bis heute sowohl die Informations- wie die Papierflut vermehrt. Die Hoffnung besteht aber, dass sie auch ohne «papierloses Büro» zur Lösung des Problems «Informationsüberflutung» beitragen wird, das sie, wenn nicht verursacht, so doch verschärft hat. Die Selbstverständlichkeit des EDV-Einsatzes im Büro - von der einfachen Textverarbeitung bis zu komplexen multifunktionalen Büroinformationssystemen als informationstechnischer Unterstützung aller Verwaltungsarbeiten - hat immerhin dazu geführt, dass man heute nicht mehr eine grundsätzlich mögliche Archivwürdigkeit maschinenlesbarer Daten gegenüber ihrer im voraus angenommenen Wertlosigkeit nachweisen muss<sup>2</sup>.

Die Archivierung maschinenlesbarer Daten ist denn auch seit einigen Jahren ein Dauerthema archivarischer Diskussionen im In- und Ausland, die sich allerdings häufig v. a. um die Archivierung gleichförmiger Mas sendaten (statistische Daten etc.) drehten. In Zukunft wird es aber in erster Linie um die Archivierung elektronischer Sachakten gehen, denn es

- 
- 1 PAPRITZ 1964 zitiert den Bamberger Archivvorstand Oesterreicher, welcher zu Beginn des 19. Jh.s geschrieben hat: "Nach den in meiner Anleitung zur Archivwissenschaft aufgestellten richtigen Grundsätzen gehören Akten ganz und gar nicht in ein Archiv."
  - 2 BUCHMANN 1988:74: "Auch bei sehr strenger Anwendung von Bewertungskriterien können maschinenlesbare Daten nicht pauschal als uninteressant etikettiert und dadurch den Archiven ferngehalten werden."

ist wegen der zunehmenden Verbreitung von Textverarbeitungs- und Büroautomationssystemen seit einiger Zeit absehbar, dass das Verwaltungsschriftgut seinen Charakter ändern und 'richtige' elektronische Akten nicht nur zu archivieren sein könnten, sondern sinnvollerweise in elektronischer Form archiviert werden müssen.

Die Informatik wird nicht nur den Rohstoff archivischer Arbeit verändern; wie schon bisher wird sie auch weiterhin die Arbeitsmittel der ArchivarInnen fundamental verändern. Dies alles verursacht neue Probleme, wirft neue Fragen auf; es eröffnet aber auch neue Perspektiven: technische wie methodische<sup>3</sup>.

In diesem Artikel sollen verschiedene Aspekte aus dem Problemkreis "Archiv und Informatik" beleuchtet werden: Nach einigen Bemerkungen zur Archivinformatik im allgemeinen und den Bemühungen des Bundesarchivs betreffend die Entwicklung einer Strategie zur Archivierung elektronischer Akten wird das Archivinformatiksystem EDIBAR des Bundesarchives im besonderen kritisch gewürdigt. Anschliessend wird ein neueres Forschungsprojekt zur Entwicklung eines Prototyps eines intelligenten archivischen Retrievalsystems vorgestellt; dieser Teil, anknüpfend an methodische Überlegungen zur spezifisch archivischen Suchstrategie bzw. den entsprechenden Ordnungsprinzipien, wird allerdings teilweise im Hypothetischen verbleiben.

## 2 Informatik im Archiv

Informatik im Archiv<sup>4</sup> ist kein isolierter oder isolierbarer Anwendungsbereich der EDV. Die Archiv-Informatik sollte Teil eines EDV-Gesamtkonzeptes einer Verwaltung oder eines Betriebes sein, so wie das Archiv als Aktenaufbewahrungsstelle eng verbunden ist mit allen anderen aktenproduzierenden und informationsverarbeitenden Verwaltungsstellen. Diese Verbundenheit des Archivs mit den Aktenproduzenten

---

3 Viele Anregungen und Ideen verdankt dieser Aufsatz DOLLAR 1992a und BEARMAN 1991, bzw. BEARMAN 1992. Diese Werke ebenso wie die Diskussionen mit D. Bearman im Bundesarchiv in Bern wirkten sehr inspirierend; sie sollen deshalb hier zu Beginn dieses Aufsatzes erwähnt werden, ohne dass bei jeder Gelegenheit noch einzeln auf sie verwiesen wird.

4 Vgl. zur Einführung allgemein GREEN 1991.

stellt eine Eigenheit dar, welche das Archiv von anderen Dokumentationsstellen unterscheidet: die grundsätzliche Möglichkeit ja Notwendigkeit, die aufzubewahrenden und zu vermittelnden Informationen bereits während deren Entstehung auszuwählen und ihre Struktur mitzuformen.

Beim Informatikeinsatz im Archiv sind die drei folgenden funktionalen Zielsetzungen unterscheidbar: die Haupteinsatzbereiche *Eigenverwaltung und Büroautomation*<sup>5</sup>, *Archivinformatik i.e.S.* als Automatisierung spezifischer Archivfunktionen, und die *Archivierung elektronischen 'Schriftgutes'*<sup>6</sup>.

Aus den Hauptaufgaben eines Archivs - Sicherung, Bewertung, Erschließung, Aufbewahrung und Vermittlung - leiten sich die konkreten Einsatzgebiete der Archivinformatik ab: Zur *Sicherung* gehören die Bemühungen, Archivierungsverfahren für elektronische Akten zu entwickeln und die Anstrengungen, von den Amtsstellen maschinenlesbare Registraturpläne und Ablieferungsverzeichnisse zu erhalten, welche die Funktion von *Erschließungshilfen* übernehmen können. Auf der Basis solcher maschinenlesbarer Ablieferungsverzeichnisse kann aufgebaut werden, um mit EDV-Unterstützung traditionelle archivische Findmittel zu erstellen. Die eigentliche *Bewertung* als intellektuelle Entscheidung über die Archivwürdigkeit von Schriftgut lässt sich dagegen mit der heutigen Informatik nur schwer automatisieren. Einfacher ist der Informatikeinsatz in den Bereichen *Aufbewahrung* und *Vermittlung*, wenn es um Fragen der Magazinverwaltung und Ausleihkontrolle geht. Die informatische Unterstützung der Vermittlung als fragenspezifische Erschließung der Archivbestände dagegen ist wieder schwieriger zu realisieren.

Zusammengefasst können allgemein zwei Anwendungsziele der Archivinformatik unterschieden werden:

- 
- 5 Auf den Aspekt der Büroautomation im Archiv soll im folgenden nicht näher eingegangen werden, da diese Problematik sich im Archiv nicht wesentlich anders präsentiert als bei anderen Verwaltungsstellen.
  - 6 Heute ist viel von Multimedia-Applikationen die Rede; der Begriff Schriftgut müsste notfalls entsprechend erweitert werden, z.B. zu Schrift-, Ton- und Bildgut. Zur Vermeidung sprachlicher Schwerfälligkeit wird im weiteren trotz allem immer wieder von 'Schriftgut' statt von Multimediatgut die Rede sein.

◆ Zur teilweisen oder vollständigen *Automatisierung* sind archivarische Funktionen wie die Magazinverwaltung, die Aktenausleihe, die Ausscheidung kassierter Akten etc. geeignet, kurz: Aufgaben, welche klar beschreibbar und nicht von wechselnden Faktoren abhängig sind, die sich leicht formalisieren lassen und regelmässig durchgeführt werden.

◆ Bei komplizierteren Aufgaben kann die Informatik eine sinnvolle *Unterstützung* anbieten. In diesen Bereich fallen v.a. Funktionen wie Retrieval und Bewertung, Aufgaben also, welche i.d.R. eine gewisse 'Intelligenz' erfordern. In unmittelbarer Zukunft wird es wohl noch nicht möglich sein, diese Aufgaben ganz zu automatisieren.

Unabhängig von diesen Anwendungszielen lassen sich auch Leistungsziele formulieren, die mittels informatischer Applikationen erreicht werden sollen, z.B.:

◆ *Sicherstellung herkömmlicher Aufgabenwahrnehmung*: Diese soll wenn möglich rationeller und effizienter als bisher erfolgen können.

◆ *Qualitative Verbesserung der Aufgabenwahrnehmung*: Ein modernes Informations-System muss benutzerfreundlich, wenn möglich sogar natürlichsprachlich sein und qualitativ bessere Suchergebnisse liefern. Sowohl das Werkzeug wie die Arbeitsleistung kann verbessert werden.

◆ *Eröffnung neuer Möglichkeiten*: In weiteren Bereichen wird der Informatik-Einsatz neue Möglichkeiten überhaupt erst eröffnen, z.B. einen orts- und zeitunabhängigen Zugriff auf archivarische Hilfsmittel oder gar auf archiviertes Schriftgut über moderne Kommunikationsnetzwerke.

Unabhängig von Anwendungs- und Leistungszielen gilt die Archivinformatik als Werkzeug, das helfen soll, archivarische Aufgaben besser und effizienter zu erledigen. Einzelne Aufgaben lassen sich überhaupt nur noch oder gar erst dank der Archivinformatik bewältigen, andere können zumindest auf eine neue Art und Weise wahrgenommen werden.

Aus dem Werkzeugcharakter der Informatik leitet sich in der Regel ab, dass ihr Einsatz nicht einfach die Art und Weise diktieren darf, wie archivarische Arbeiten zu erledigen sind, sondern dass die Informatik sich

den bestehenden Methoden und Arbeitstechniken grundsätzlich anzupassen hat. Es scheint mir durchaus richtig, dass weder die ArchivarInnen noch die ArchivbenutzerInnen sich einfach nach der Technik richten müssen; man muss sicher versuchen, die neue Technik den erprobten Methoden anzupassen. Aber: Der Einsatz der EDV muss auch informatischen Grundsätzen Rechnung tragen, und wenn das Nachdenken über den Einsatz der Informatik zur Erkenntnis führt, dass neue Wege der Problemlösung oder Arbeitsbewältigung effizienter und realisierbar sind, so muss man solche Erkenntnisse fruchtbar umzusetzen versuchen. Es muss eine Übereinstimmung zwischen archivgerechtem und informatikgerechtem Arbeiten gesucht werden. Die Herausforderung durch die Informatik besteht gerade darin, dass man gegebenenfalls bereit sein muss, das Archiv in elektronischer Umgebung neu zu erfinden und einen Paradigmenwechsel zu vollziehen<sup>7</sup>.

### 3 Informatik im Bundesarchiv heute

#### 3.1 Archivierung elektronischen Schriftgutes

Nachdem bereits die *Weisung betreffend die Abgabe von Schriftgut an das Bundesarchiv* (1980)<sup>8</sup> die datenträgerunabhängige Archivierungspflicht für die Amtsstellen der Bundesverwaltung fixiert hat, haben die *Weisung betreffend die Archivierung und Benutzung der elektronisch gespeicherten Daten der Bundesverwaltung im Bundesarchiv* von 1984 und die entsprechenden archivinternen *Richtlinien für die Erfassung, Bewer-*

---

7 DOLLAR 1992b:2 schreibt unter der Überschrift "What Librarians and Archivists Have in Common": "Peggy Johnson [Automation and Organizational Change in Libraries, Boston 1991:156] argues there is a library paradigm shift underway in which new technologies are transforming how 'libraries are viewed, experienced, and organized.' As Pat Batin put it, libraries are being 'reinvented in the electronic environment.' [quoted in Peggy Johnson, p. 40]".

8 Weisung betreffend die Abgabe von Schriftgut an das Bundesarchiv, hg. vom Schweizerischen Bundesarchiv, 30.6.1970 (3. Aufl. 1980): Punkt 132 definiert Akten folgendermassen: "Der Begriff Akten umfasst: - das eigentliche Schriftwerk [...], - elektronische Datenträger [...]" und hält fest (Punkt 121), dass die Archivierung elektronischer Datenträger in einer speziellen Weisung noch genauer zu regeln sei..

ung und Archivierung von EDV-Datenbeständen der Bundesverwaltung im Bundesarchiv (1987)<sup>9</sup> die Struktur archivierungsfähiger Daten<sup>10</sup> ebenso wie diejenige der Datenträger<sup>11</sup> festgelegt und gleichzeitig die Bewertung der Datenbestände, das Übernahmeverfahren und den Unterhalt der Daten, die Erstellung und den Unterhalt der Findmittel im Archiv sowie die Benutzung der Daten geregelt.

Für die Archivierung sequentieller Daten ist die bestehende Lösung ein erster Ansatz. Die bestehenden Weisungen und Richtlinien bedürfen aber dringend einer (eigentlich: einer ständigen) Aktualisierung, um den sich wandelnden technischen Rahmenbedingungen angepasst zu bleiben. Eine solche Revision ist in näherer Zukunft auch zu erwarten. In diesem Zusammenhang sind auch erste Schritte gemacht, um das Bundesarchiv infrastrukturell aus seiner Abhängigkeit von externen Rechenzentren zu lösen. Mit einer archiveigenen Bandstation ist die hausinterne Verarbeitung (Qualitätskontrolle, Kontrolle der Übereinstimmung von Datenbeschreibungen und Daten, Kopierarbeiten etc.) flacher Dateien jetzt möglich.

Nun sind die Probleme zu lösen, die sich bei der Archivierung von elektronischen Sachakten und Datenbanken ergeben, nicht zu reden von den Problemen im Zusammenhang mit virtuellen, Hyper- oder Multimedia-Dokumenten (Compound Documents)<sup>12</sup>. Die Fragen betreffend die

---

9 ABEDIB-Weisung vom 2.5.1984, gemeinsam vom Bundesarchiv und dem Bundesamt für Organisation ausgearbeitet. Die dazugehörenden Richtlinien folgten am 27.5.1987.

Neben den technischen Fragen muss man sich aus archivischer Sicht auch Gedanken machen, über die Auswahl des Datenmaterials für die Überlieferungsbildung, vgl. dazu grundlegend BÜTIKOFER 1991b.

10 Punkt 2.1 der Weisung: "Datenbestände [...] die unabhängig von der Hauptanwendung und ohne besondere Voraussetzungen [...] ausgewertet werden können".

11 Punkt 2.2 der Weisung: "ausschliesslich Magnetbänder"

12 Virtuelle Dokumente bestehen nur auf dem Bildschirm oder als Papierkopien als physische Einheit; im Grunde sind sie zusammengesetzt aus Textfragmenten und Verweisen auf Datenbanken oder andere dokumentenfremde Informationsquellen.

Compound documents/Multimediadokumente verbinden verschiedene informations-tragende Medien zu einer Einheit, z.B. Text, (bewegtes) Bild und Ton.

Hypermediadokumente vereinigen Charakteristika von virtuellen und multimedialen Dokumenten. Sie haben neben einer linearen Struktur (Text) noch eine vordefinierte Zweitstruktur, welche ein Beziehungsnetz von einzelnen Textelementen zu anderen (textfremden) Informationen schafft (Text, Bild, Ton), die nach Bedarf ak-

Speichermedien können dabei zwar weitgehend als gelöst betrachtet werden. Es sind heute einerseits andere EDV-technische Probleme, wie sie sich aus der existierenden Systemvielfalt ergeben, andererseits Probleme der logischen Ordnung und einer möglichen Erschließung, welche einer zukunftsfrächtigen und richtungsweisenden Lösung harren.

Ausgehend von der reglementarisch fixierten Pflicht zur bundeszentralen Archivierung von Verwaltungsakten und der amtsstelleneigenen Beschaffungsautonomie sowohl im Hardware- als im Softwarebereich ergibt sich als wesentliches informatisches Problem die Portabilität elektronischer Akten aus den EDV-Systemen in den Verwaltungsstellen in ein Archiv-Informationssystem. Als Grundsatz könnte gelten: elektronische Akten müssen in einer Form archiviert werden, welche unabhängig ist von der ursprünglichen Hardware- und Softwareumgebung. Die Einhaltung allgemein anerkannter Standards und Normen scheint dabei die einfachste und kostengünstigste Lösung zu sein. Allerdings sind noch lange nicht in allen Bereichen entsprechende Standardisierungen erreicht, noch sind sie immer absehbar. Die Archive sind also gezwungen, sich um die Formulierung ihrer Bedürfnisse und deren Umsetzung in EDV-Normen aktiv zu bemühen<sup>13</sup>. Eine andere Lösung würde in der Aufhebung der zentralen Archivierung liegen: die Amtsstellen würden dann ihre eigenen Akten innerhalb ihrer EDV-Systeme selber aufbewahren und für ihre bleibende Lesbarkeit und Verstehbarkeit besorgt sein; eine zentrale Archivstelle würde die Wahrnehmung dieser Aufgaben nur noch beratend begleiten und die Durchführung überwachen<sup>14</sup>.

---

tiviert werden können, so dass neben einer herkömmlichen sequentiellen auch eine 'assoziative' Lektüre möglich ist.

13 NATIONAL ARCHIVES TECHNICAL PAPER 8/1990, DOLLAR 1992.

14 DOLLAR 1992a:71: Unter Berücksichtigung von Kostenfaktoren wird die Rolle von Archiven neu zu bestimmen sein. Die Archive "könnten sich auf die Entwicklung von Programmen, Werkzeugen, Richtlinien und Regulierungen konzentrieren, die den Zugang zu verschiedensten Datenbasen und Informationssystemen erleichtern [...] den Zugang regulieren. Aus diesen Gründen müssten zentralisierte Archiveinrichtungen als *finale Endarchive* definiert werden. Nur dann, wenn eine Organisation nicht länger für die Aufbewahrung und Übertragung auf neue Technologien eintreten kann, wären die zentralisierten Archiveinrichtungen der Ort für die physische Aufbewahrung von elektronischen Akten." Das würde m.E. die Rolle zentraler Archive im Lebenszyklus von Schriftgut nur zeitlich verschieben, ihre Aufgaben im Grunde aber nicht aufheben. Eine systemunabhängige Archivierung müsste einfach zu einem späteren Zeitpunkt realisiert werden.

Zu den archivarischen Anforderungen an elektronische Sachakten und die sie begleitenden Metadaten<sup>15</sup> gehört sicher die Erhaltung einer sachsystematischen, vorgangsorientierten Ordnung während des Prozesses der Informationsproduktion. Allerdings musste noch Ende der 1980er Jahre festgestellt werden, dass marktgängige Büroautomationssysteme die Anforderungen an "amtszentrale Registraturen oder an die bundeszentrale Langzeitarchivierung, insbesondere an die Dossierbildung nach einem sachsystematischen Registraturplan" nicht zu erfüllen vermochten. Da die geschäftsbezogene, vorausschauende Ordnung des Schriftgutes wesentlich ist, gilt bis heute die Regel, dass archivwürdiges Schriftgut in Papierform in einer herkömmlichen Registratur aufbewahrt und später in dieser Form ins Archiv abgeliefert werden muss<sup>16</sup>. Mit der Weiterentwicklung der Büroautomationssysteme, um die sich die Archivre ebenfalls aktiv kümmern sollten, wird sich die Situation ändern. Die heute bestehende Notlösung der Archivierung von Papierkopien muss überwunden werden.

Aus diesem Grund beschäftigt sich eine vom Bundesarchiv geleitete interdepartementale Arbeitsgruppe ARELDA (*Archivierung elektronischer Daten*) mit den Fragen rund um die Archivierung von Büroautomationsdaten. Ziel dieser Arbeitsgruppe ist es, ein Verfahren zu entwickeln für die Archivierung elektronischer Sachakten. Logischerweise muss dieses Verfahren bereits bei der Aktenproduktion ansetzen, damit die zu archivierenden Daten einmal auch in der gewünschten Form und mit den notwendigen quellengenerierten Metadaten zusammen vorliegen werden. Diese Metadaten sollen als Informationen über die archivierten elektronischen Sachakten den institutionellen und geschäftsbezogenen Entstehungszusammenhang dokumentieren und als erstes Findmittel dienen. Soweit möglich sollen sie automatisch während des Prozesses der Aktenentstehung gesammelt und nicht, wie die heutigen Findmittel, nachträglich zusammengestellt werden. Anders ist die Les- und Verstehbarkeit elektronischer Akten nicht sicherzustellen. Denn: "Überlieferungsbildung kann angesichts der Büroautomation nicht mehr unbewus-

---

15 Mit Metadaten sind Daten gemeint, welche die Struktur, den Inhalt der eigentlichen Daten beschreiben, also: Daten über Daten.

16 Vgl. das REALISIERUNGSKONZEPT 1989:23, zusammengefasst dargestellt in GRAF 1990.

Vgl. auch die grundsätzliche Auseinandersetzung mit diesem Thema in MENNEHARITZ 1988.

ster Nebeneffekt sein, sondern muss geplant und gezielt in den Arbeitsprozess im Büro eingegliedert werden<sup>17</sup>.

Die genannte Arbeitsgruppe ARELDA eruiert in diesem Zusammenhang zuerst einmal empirisch, welche Büroautomationssysteme überhaupt in Gebrauch sind, was für Daten und welche Mengen von Akten damit produziert werden, wie diese geordnet sind (Registrierungsplan, Indexierung durch die Sachbearbeiter), welche Zugriffshäufigkeiten festzustellen sind, wie die Speicherungsmöglichkeiten sind etc. Parallel zu dieser Informationssammlung wird versucht, die juristischen und archivisch-historischen Anforderungen an eine Langzeitarchivierung und spätere Nutzung elektronischer Akten zu formulieren und die technischen Möglichkeiten zu eruiieren, welche die Erfüllung dieser Anforderungen gewährleisten, damit die relevanten Schritte eines Geschäftsganges jederzeit rekonstruiert werden können. Absicht ist es, die Überlieferungsbildung in einem offenen System sicherzustellen, welches durch *Einhaltung allgemein anerkannter EDV-Normen und -Standards eine allgemeine Akzeptanz finden und einen (Meta-)Datenaustausch in alle Richtungen erlauben wird*<sup>18</sup>. Mit einem solchen Vorgehen wird die Sicherstellung künftigen Archivgutes (elektronische Sachakten und Metadaten), analog zur Bewertung an der Quelle, aus dem archivischen Bereich vorverlegt in den Bereich administrativer Geschäftserledigung und Aktenproduktion<sup>19</sup>.

Die Digitalisierung bereits archivierter Akten, wie sie in verschiedenen aufwendigen ausländischen Grossprojekten für klar umrissene Bestände mit Erfolg gewagt worden ist<sup>20</sup>, steht aber als beständeübergreifendes Unternehmen zur Zeit nicht zur Diskussion, selbst wenn sie aus konservatorischer Sicht interessant sein und für viele Bestände eine Alternative oder Ergänzung zu teuren Restaurierungs- und Konservierungsmassnahmen oder herkömmlichen Mikroverfilmungen darstellen könnte. Eine

---

17 MENNE-HARITZ 1990:94.

18 Die Anstrengungen im Rahmen der Bundesverwaltung sind natürlich nicht isoliert zu sehen. Vgl. etwa die ähnlichen Überlegungen, die in den USA angestellt werden: NATIONAL ARCHIVES TECHNICAL INFORMATION PAPER 8/1990.

19 Vgl. DOLLAR 1992

20 Vgl. u.a. Spanisches Nationalarchiv: Digitalisierung von Beständen aus dem Archivo General de Indias in Sevilla, National Archives der USA: Digitalisierung von Beständen aus der Zeit des amerikanischen Bürgerkriegs.

Digitalisierung einzelner Bestände(teile) im Rahmen von Forschungsprojekten dagegen ist nie auszuschliessen.

### **3.2 Teilweise Automatisierung der Verwaltung herkömmlichen Archivgutes: Das EDIBAR-System**

Der Einsatz der EDV hat auch im Bundesarchiv nicht mit der Archivierung maschinenlesbarer Daten begonnen, sondern mit der Verwaltung herkömmlichen Schriftgutes: Nach der Entwicklung eines beständeübergreifenden Personennamennachweissystems (PERSEIS: Personennamen: Ein Informationssystem) in den frühen 1970er Jahren<sup>21</sup> war es der Aktenzuwachs ohne gleichzeitige Personalzunahme, welcher den weiteren Einsatz der neuen Technologien in den 1980er Jahren wünschbar werden liess, damit die vorhandenen Mittel effizienter eingesetzt, die Arbeit rationalisiert werden konnte. Die eigentliche Planung für die Entwicklung eines archivischen Arbeitsinstrumentes begann vor rund 10 Jahren (1983 Voranalyse, 1985 Konzeptbericht). 1987 konnten erste Teile des angestrebten benutzerfreundlichen, "umfassenden, aufeinander abgestimmten Instrumentariums zur Verwaltung und Kontrolle der Bestände und ihrer Benutzung" in Betrieb genommen und der Öffentlichkeit vorgestellt werden<sup>22</sup>. Seither ist das EDIBAR-System weiterentwickelt und durch neue Komponenten ergänzt worden. Gleichzeitig hat sich aber auch der state-of-the-art im Informatik-Bereich wesentlich verändert, und die Vorstellungen von Benutzerfreundlichkeit oder Natürlichsprachlichkeit haben sich verschoben. Aus dem Wandel der Vorstellungen und Erwartungen ergibt sich, dass das vor fünf Jahren moderne und noch heute dem Archivpersonal in verschiedenen Bereichen äusserst hilfreiche EDIBAR-System weiterentwickelt werden muss.

#### **3.2.1 Der Ist-Zustand**

Das EDIBAR-System ist ein für die Benutzung durch das Archivpersonal konzipiertes Archiv-Informatiksystem für die Verwaltung von Papierakten, das aus verschiedenen, einander ergänzenden Komponenten

---

21 CADUFF 1978.

22 CADUFF 1989:213-272, hier: S. 218.

besteht. Tabelle 1 zeigt die wichtigsten archivischen Aufgaben, die mit diesen verknüpften Ordnungsmittel (Findmittel) und die entsprechenden EDV-Applikationen (fett = EDIBAR-Teil, normal = sonstige Applikationen [des Bundesarchivs oder anderer Anbieter]).

Kernstück eines Archivverwaltungssystems sollte eine relationale Datenbank sein, deren verknüpfte Tabellen die Informationen enthalten, welche für die logische und physische Ordnung der Archivbestände notwendig sind.

Grundlage für das EDIBAR-System ist das ORACLE-Datenbank-Managementsystem. Die Teile *BAK* (Bestände- und Akzessionsverwaltung<sup>23</sup>), *BAKMAG* (Magazinverwaltung) und *BAKBV* (Benützerverkehr) enthalten die wichtigsten Daten aus dem Archiv und die dazu passenden Recherche-Instrumente (-RE)<sup>24</sup>. Bei *KOMP* handelt es sich um eine Übersicht über die Kompetenzen und Aufgaben des Bundes, wie sie aus den organisations- oder anderen bundesrechtlichen Erlassen oder auch aus Geschäftsberichten etc. eruiert werden können. In zeitraubender Kleinarbeit wird die Verbindung zwischen den rechtsordnungsbezogenen KOMPentenzentabellen und den beständebezogenen BAKtabellen hergestellt; diese Arbeiten sollten in näherer Zukunft abgeschlossen werden können<sup>25</sup>.

- 
- 23 Akzession: Ablieferung von Akten ans Archiv. Zur Begrifflichkeit vgl. MENNEHARITZ 1992.
- 24 Eine aktualisierte Beschreibung des EDIBAR-Systems findet sich in CADUFF et. al. 1990.
- 25 SCHÄRER 1976.

**TAB. 1: (VOR-) ARCHIVISCHE TÄTIGKEITEN UND ENTSPRECHENDE ORDNUNGSMITTEL**

| <b>Tätigkeit</b>                            | <b>Ort der Ausübung</b>        | <b>Ordnungsmittel (logische)</b>  | <b>Ordnungsmittel (physische)</b>                          |
|---|--------------------------------|---|--|
| <i>Geschäftserledigung, Sachbearbeitung</i> | <i>Amtsstelle/ Registratur</i> | <i>Registaturplan [ZAC], Geschäftskontrolle [GEKO-BV]</i>   |  |
| <i>Ablieferung/ Akzession</i>               | <i>Zwischenarchiv</i>          | <i>Abgabeverzeichnisse [ZAC]</i>  | <i>Magazinverwaltung (Standort grundsätzlich [BAKMAG])</i> |
| <i>Bewertung/ Kassation</i>                 | <i>Archiv</i>                  | <i>Kassationsliste (Abgabeverzeichnis)</i>  |  |
| <i>Erschliessung</i>                        |                                | <i>Abgabeverzeichnisse Repertorien [REP]<br/>Inventare, Indices [SYBU, KOMP, BAK(-RE), BAK-BV-RE, DONA-BAR, NFP 23, EX PERTENSY-STEM]</i> |  |
| <i>Vermittlung</i>                          |                                | <i>Benutzerverwaltung (aktueller Standort) [BAK-BV]</i>   |  |

Neben diesen eigentlichen EDIBAR-Teilen verfügt das Bundesarchiv über das bereits erwähnte Personennamen-Nachweis-System PERSEIS und über verschiedene, selbständige Kleinapplikationen, welche beispielsweise die klassische archivische Erschliessung (REP als Arbeitshilfe für die Erstellung von Repertorien) oder, im vorarchivischen Bereich, die Erfassung von Registraturplänen oder Abgabeverzeichnissen (ZAC) unterstützen. Ebenfalls für den Einsatz im vorarchivischen Bereich ist das vom Bundesarchiv entwickelte komplette Geschäftskontroll-

system (GEKO-BV) gedacht. Neben dieser Eigenentwicklung beteiligt sich das Bundesarchiv auch an entsprechenden Projekten für die Ausarbeitung ähnlicher Geschäftskontrollsysteme auf unterschiedlichsten Systemplattformen.

Betrachten wir allein das EDIBAR-System, so verfügt das Bundesarchiv schon mit diesem über ein Informatik-System, welches es erlaubt, archivierte herkömmliches Schriftgut physisch unter Kontrolle zu halten (von der Akzessionierung bis zur Ausleihe) und logisch zugänglich zu machen.

Mit EDIBAR sind verschiedene Arbeitsschritte automatisiert worden (z.B. Fristenkontrolle und Mahnwesen bei Aktenausleihen, Magazinraumbewirtschaftung). Die Kontrolle des Benützerverkehrs ermöglicht zudem eine einfache Ermittlung und Auswertung der Benutzungsfrequenzen für die einzelnen Archivbestände. Das EDIBAR-System ermöglichte es weiter - gleichsam als Nebenprodukt -, eine *Systematische Beständeübersicht*<sup>26</sup> zusammenzustellen, welche jederzeit nach Bedarf aktualisiert, d.h. um die Neuzugänge ergänzt werden kann. Ihre Integration ins Retrievalsystem von EDIBAR bleibt allerdings noch zu realisieren.

### 3.2.2 Unmittelbare Ausbaupläne: DONABAR - IZBAR

Bevor wir uns der Retrievalmethode des EDIBAR-Systems und möglichen methodischen Erweiterungen zuwenden, sollen die EDIBAR-Erweiterungsabsichten des Bundesarchivs vorgestellt werden, die sich teils im Stadium konkreter Detailplanung, teils im Stadium konzeptioneller Entwürfe befinden.

Das DONABAR-Projekt ist in der Planung am weitesten fortgeschritten. Es steht vor seiner Realisierung. In Rahmen dieses Projektes sollen in den nächsten Jahren alle Detailfindmittel (z.B. Abgabe-/Zugangsverzeichnisse) im Bundesarchiv digitalisiert und als maschinenlesbare Findmittel im Rahmen des EDIBAR-Systems zugänglich gemacht werden.

---

26 BÜTIKOFER 1991a.

In BAK sind bis heute die wichtigsten inhaltlichen und formalen Angaben zu den Beständen (Signatur), Generationen (Geltungsdauer eines Registraturplans) und Akzessionen (Zeitraum enthaltener Akten und Gruppentitel vertretener Registraturplanpositionen je Aktenablieferung) enthalten. Die letzten Schritte einer Recherche, d.h. insbesondere die Ermittlung der gesuchten Akteneinheiten innerhalb einer Akzession, müssen immer noch in den papierenen Detailfindmitteln gemacht werden. Nach deren Digitalisierung müsste eine Recherche im EDIBAR-System nicht mehr auf der Ebene der Gruppentitel der einzelnen Abgabeverzeichnisse abgebrochen, sondern sie könnte am Bildschirm bis zur Ermittlung einzelner Dossiertitel vollendet werden.

Dank der zunehmenden Verbreitung der im Auftrag des Bundesarchivs entwickelten Einzelplatz-Applikation ZAC (Zugangs- und Abgabeverzeichnis-Computer), welche im Bedarfsfall zusammen mit der nötigen Hardware einzelnen Amtsstellen zur Vorbereitung von Aktenablieferungen und zur Erstellung der begleitenden Abgabeverzeichnisse leihweise abgegeben wird, erhielt das Bundesarchiv in den letzten Jahren eine ständig wachsende Zahl maschinenlesbarer Abgabeverzeichnisse. Diese könnten im Rahmen des DONABAR-Projektes direkt ins EDIBAR-System integriert werden.

Diesen modernen Abgabeverzeichnissen, welche in der Regel auch qualitativ heutigen archivischen Ansprüchen zu genügen vermögen, steht aber eine riesige Menge papierener Abgabeverzeichnisse unterschiedlicher Qualität gegenüber. Diese müssen durch Abschreiben oder mit einem Scanner und anschliessender OCR-Bearbeitung in ein brauchbares Datenformat überführt werden, bevor sie in die EDIBAR-Datenbank integriert werden können. Diese Arbeit wird etliche Arbeitsjahre dauern, und es hängt wesentlich von den finanziellen Mitteln und den zur Verfügung stehenden Arbeitskapazitäten ab, in welchem zeitlichen Rahmen dieses Projekt vollendet werden kann.

Nach Abschluss des DONABAR-Projektes, der logischen Weiterführung von EDIBAR, wäre das archiv-informatische Arbeitsinstrument für die Verwaltung archivierter Papierakten gemäss Anfangskonzept grundsätzlich vollendet. Ein weiterer Schritt wird noch sein, die archivierten elektronischen Akten selbst in ein erweitertes Archiv-Informatik-System einzubinden.

Die datenmässige Ausweitung des EDIBAR-Systems durch Integration maschinenlesbarer Detailfindmittel und - später - elektronischer Sachakten kann nicht die einzige Zukunftsperspektive sein. Im Rahmen einer Ideenskizze IZBAR ist bereits der Frage nachgegangen worden, wie das Bundesarchiv zu einem fachbereichsspezifischen Informationsversorgungszentrum entwickelt werden könnte<sup>27</sup>. Ein solches fachbereichsspezifisches Informationsversorgungszentrum müsste selbstverständlich den Zugang zu den Akten der Bundesverwaltung eröffnen. Gleichzeitig sollte auch der Zugang zu anderen Archiven - im Rahmen eines Archivverbundes ähnlich dem bereits bestehenden Bibliotheksverbund - bzw. zu Bibliotheken, (kommerziell betriebenen) öffentlich zugänglichen Datenbanken (Nachrichtenagenturen, H[istorisches]L[exikon der ]S[chweiz], Kulturgüterdatenbank, Informationssammlungen aus Forschungsprojekten etc.) oder anderen Informationsquellen ermöglicht werden. Umgekehrt müsste von Informationsversorgungszentren anderer Fachbereiche aus nach Bedarf auf die Akten und Daten der Bundesverwaltung zugegriffen werden können. Damit könnten BenutzerInnen zwischen Literatur- und Quellenstudium hin und herwechseln, könnten Bundes- und Kantonsakten, die häufig eine materielle Einheit bilden, im gleichen Arbeitsgang konsultiert bzw. den EndbenutzerInnen zu einer computerunterstützten Analyse und Auswertung on-line zur Verfügung gestellt werden<sup>28</sup>.

Im eigentlichen Bereich der Vermittlung elektronischen Schriftgutes von eidgenössischen Amtsstellen müssten in einem zukünftigen Informationsversorgungszentrum neue Retrievalmöglichkeiten geschaffen werden, welche auf der Grundlage einer umfassenden Sammlung von Metadaten zur Dokumentation der Entstehungsprozesse von Akten vielfältige Zugänge zu diesen erlauben würden. Die Erschliessung als Bearbeitung definitiv bewerteter, abgeschlossener Archivbestände müsste sich in den Entstehungsprozess des Schriftgutes vorverschieben, wie es in einer elektronischen Umgebung eigentlich naheliegend ist. Die eigentlichen Sachakten würden dann nicht mehr nur als archivischer Rohstoff dienen,

---

27 Zur Idee des Informationsversorgungszentrums vgl. den Bericht der COMMISSION FEDERALE POUR L'INFORMATION SCIENTIFIQUE 1988.

28 Vgl. das Recherchebeispiel in Tabelle 2, das eine Kompetenz berührt, die teilweise von den Kantonen wahrgenommen wird.

Allg. zu den Bedürfnissen der ForscherInnen: MICHELSON 1992.

sondern sie könnten gleichzeitig selber auch Findmittelfunktionen übernehmen<sup>29</sup>.

#### 4 Das EDIBAR-Retrievalkonzept

Die Erschliessung von Archivbeständen und die Benutzung der Findmittel während eines Suchprozesses stehen naheliegenderweise in einem engen wechselseitigen Verhältnis. Nur was während der Erschliessungsarbeit an Informationen in ein Findmittel einfließt, kann später während einer Recherche die Informationssuchenden weiterführen. Während der Erschliessung werden Informationen über die Herkunft, den Charakter und Inhalt von Beständen und deren Teilen gesammelt. In den archivischen Verzeichnissen entsteht ein konzentriertes Abbild der Akten selbst<sup>30</sup>, ohne allerdings deren eigene ursprüngliche Informationsfülle bewahren zu können. Diese Konzentrierung der Information setzt Entscheidungen voraus - ähnlich der Bewertung - über die Trennung des Wichtigen vom Unwichtigen.

Die Benützungsbefürfnisse ihrerseits wecken Wünsche nach spezifischer Erschliessung des Archivmaterials. Der ständige Wandel der Interessen von ForscherInnen verweist aber auf eine grundsätzliche Problematik: Es ist schwierig vorauszusagen, nach welchen Kriterien Archivmaterial erschlossen werden sollte, damit auch künftigen Generationen von Archivbenutzern der Zugang zu ihnen entsprechend ihren eigenen

---

29 Vgl. BEARMAN 1992.

Bereits der EDIBAR-Konzeptbericht von 1985 formuliert weitsichtig eine solche Perspektive: "Das BAR ist möglicherweise gezwungen, in Zukunft auf traditionelle und bewährte Arbeitsverfahren und -ergebnisse zu verzichten."

Es wird auch von den Ergebnissen der Arbeitsgruppe ARELDA abhängen, wie viele Informationen aus dem Schriftgutproduktionsprozess gesichert und als archivische Erschliessungshilfen genutzt werden können. Als Beispiel für Dokumente, welche eine solche Doppelfunktion haben, können die Geschäftsberichte genannt werden: Als Verwaltungsdokumente gehören sie, obwohl publiziert, sicher auch ins Archiv. Sie können aber auch als Findmittel benutzt werden, geben sie doch Aufschluss über die tatsächlich wahrgenommenen Kompetenzen und Aufgaben (vgl. SCHÄRER 1976).

30 Nach INTERNATIONAL COUNCIL ON ARCHIVES 1990:1.

Interessenlagen sinnvoll erleichtert werden kann. Heute sind die archivischen Findmittel angebotsorientiert und verlangen von den BenutzerInnen spezifische eigene Suchanstrengungen mit ungewissem Ausgang: erst am Schluss eines Suchprozesses zeigt sich, ob Akten gefunden werden können, welche die Beantwortung einer Frage ermöglichen<sup>31</sup>.

Dazu kommt: Die archivischen Suchstrategien unterscheiden sich nach archivarischem Selbstverständnis methodisch grundsätzlich von bibliothekarischen oder dokumentarischen Suchstrategien, weil Archivgut organisch gewachsen und nicht nach sachlichen Kriterien bewusst gesammelt worden ist. Methodisch muss deshalb ein archivisches Retrievalkonzept auf dem allseitig anerkannten *Provenienzprinzip* und dem *Respect des fonds* aufbauen - so auch im Rahmen des EDIBAR-Systems.

Recherchieren setzt deshalb bei den Fragenden immer ein spezifisches Wissen um die Struktur des archivierten Materials voraus, welches primär nach abliefernden Informationsproduzenten (Verwaltungsstellen) und sekundär, innerhalb der Gruppen von Informationen gleicher Herkunft, nach einer hierarchisch-geschäftsbezogenen Sachgliederung (Registaturplan) geordnet ist, welche die den informationsproduzierenden Stellen übertragenen Aufgaben (Kompetenzen) widerspiegelt.

Eine themenbezogene Recherche im EDIBAR-System läuft folgendermassen ab (vgl. Tabelle 2): In einem ersten Schritt muss eine Frage in Kompetenz-Begriffe übersetzt werden, d.h. im Rahmen des EDIBAR-Teiles KOMP, einer Art vorgangsbezogener Pertinenzordnung, muss die der Frage entsprechende Kompetenz eruiert werden. Diesem Vorgehen liegt die Überlegung zugrunde, dass nur der Aufgabenbezug Grundlage für eine sinnvolle Ordnung archivierten oder archivwürdigen Schriftgutes sein könne.

In einem zweiten Schritt muss festgestellt werden, welcher Kompetenzträger (mit einer Kompetenz/Aufgabe beauftragte Amtsstelle) während eines bestimmten Zeitraumes mit der eruierten Kompetenz betraut war. Das Schriftgut eines Kompetenzträgers formt zusammen einen Bestand. Systemtechnisch entspricht dieser zweite Schritt der Recherche dem Übergang von KOMP zu BAK, logisch demjenigen von einer Handlungsordnung zu einer Organisationsordnung.

---

31 DOLLAR 1992a:84ff.

Im dritten Schritt ist das Schriftgut des ermittelten Kompetenzträgers genauer zu befragen. Dieses Schriftgut ist seit dem Stadium seiner Produktion gemäss einem Registraturplan geordnet; Auszüge aus diesem Registraturplan begleiten als Abgabeverzeichnisse die einzelnen Aktenablieferungen dieser Amtsstelle, und die kumulierten Abgabeverzeichnisse stellen eine Art Rohdetailfindmittel für einen Bestand dar. Ordnungsprinzip für die Registraturpläne sollte wiederum die Aufgabenverteilung sein. Konkret: jeder Registraturplan sollte in seiner Gliederung die wichtigsten Aufgaben einer Verwaltungsstelle spiegeln. Entsprechend dem Kompetenzprinzip müsste man eigentlich zwischen einzelnen Positionen eines Registraturplanes und den einzelnen Kompetenzen einer Amtsstelle eine klare Verbindung herstellen können. Gestützt auf eine solche Beziehung, selbst wenn diese in der Praxis nicht so eindeutig herzustellen ist, müsste im weiteren Verlauf des Rechercheprozesses nur noch auf diejenigen Akzessionen eines Bestandes verwiesen werden, welche auch entsprechendes Schriftgut enthalten. Im Rahmen der BAKRE-Applikation wird aber auf eine solche, in der Praxis nur in den seltensten Fällen sauber herzustellende Verknüpfung von Kompetenzen mit Registraturplanpositionen ganz verzichtet, d.h. dass während des Rechercheprozesses immer auf das ganze Schriftgut einer Verwaltungsstelle verwiesen wird. Es bleibt den BenutzerInnen vorbehalten, eine Feinausscheidung möglicherweise relevanten Schriftgutes vorzunehmen. Hilfreich sind dabei v.a. die Randdaten der einzelnen Akzessionen, welche auf den Zeitraum hinweisen, aus dem das Schriftgut stammt. Anhand dieser Randdaten und anhand der Haupttitel der einzelnen Abgabeverzeichnisse kann eine erste Feinausscheidung in Frage kommenden Archivmaterials im Rahmen des EDIBAR-Systems vorgenommen werden. Zudem verweist EDIBAR in diesem Stadium auch auf Bestände, welche erst theoretisch bestehen, zu denen aber noch nie Akten abgeliefert worden sind<sup>32</sup>; ein Suchprozess kann so in mehrfacher Hinsicht in Sackgassen führen (vgl. die schematische Zusammenstellung in Tabelle 3). Die Resultate werden nach der im Rahmen von DONABAR zu realisierenden Integration der bestehenden Detailfindmittel ins EDIBAR-System wesentlich positiver ausfallen.

Nach der Eruiierung potentiell wichtiger Bestände und eventuell einzelner Akzessionen muss die weitere Recherche vor der im Rahmen des DONABAR-Projektes geplanten Digitalisierung der Abgabeverzeichnisse in den papierernen Detailfindmitteln zu Ende geführt werden.

---

32 1966 ist im Bundesarchiv eine Zusammenstellung aller Amtsstellen erarbeitet worden (Basis: Staatskalender etc.) und jeder Amtsstelle ist eine Bestandessignatur zugeordnet worden. Diese dauernd ergänzte Bestandesübersicht ist in EDIBAR integriert worden.

EDIBAR liefert dann wieder die Angaben zu den Standorten der einzelnen Akzessionen im Magazin.

**Tab. 2: RECHERCHE IM EDIBAR-SYSTEM I**

---

---

|                  |   |
|------------------|---|
| Frage            | "Ich interessiere mich für die Geschichte des Kokainkonsums in der Schweiz."  |
| Stichwort        | Kokain  |
| Kompetenz        | Betäubungsmittelverkehr 1925-52 [1]<br>Betäubungsmittelgesetz 1952- [2]   |
| Kompetenz-träger | 1 Gesundheitsamt 1925-52 [11]<br>Zollverwaltung 1925-52 [12]<br>Kommission für die Ausführung des Betäubungsmittelgesetzes 1925-52 [13]<br>2 Gesundheitsamt 1952-79 [21]<br>Bundesamt für Gesundheitswesen 1979- [22]<br>Bundesanwaltschaft 1952- [23]<br>Schweiz. Zentralstelle für die Bekämpfung illegalen Betäubungsmittelverkehrs 1952- [24]<br>Zollverwaltung 1952-79 [25]<br>Eidg. Zollverwaltung 1979- [26]<br>Kantone 1952- [27] |

---

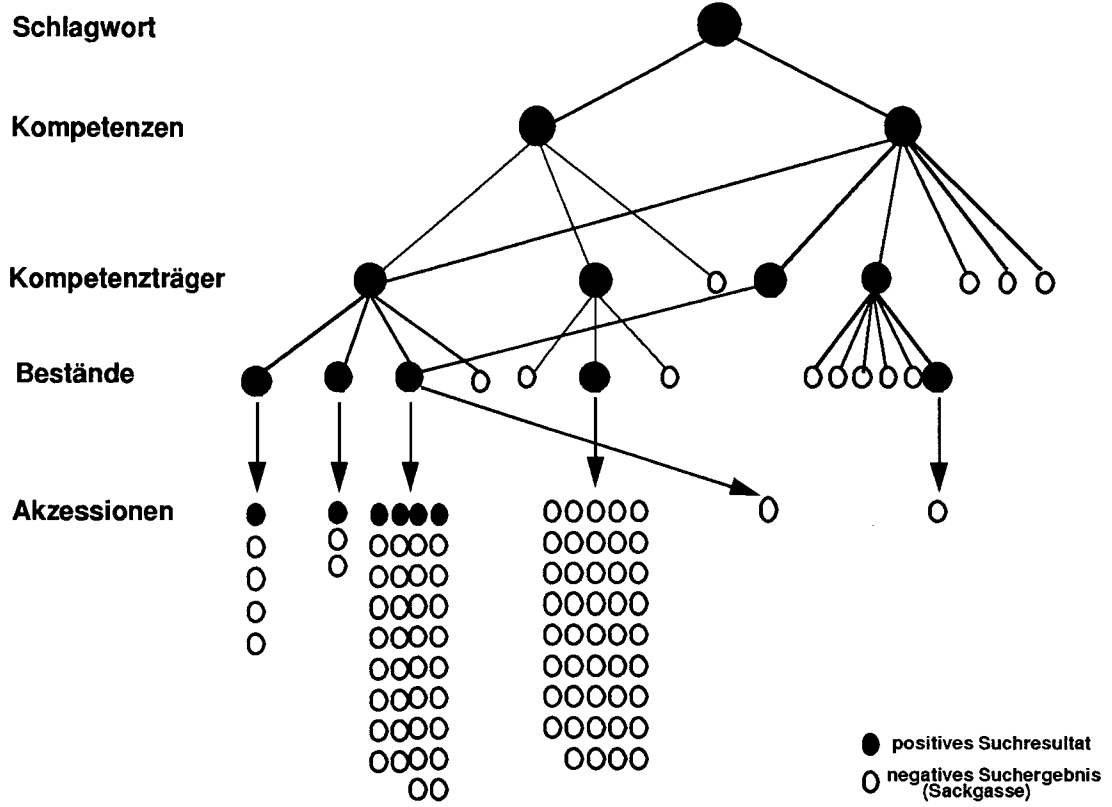
---

---

|                        |                         |   |   |
|------------------------|-------------------------|---|---|
| Bestände<br>(Signatur) | 11                      | E 3300 (A) Gesundheitsamt 1896-1935 / 1921-79 [111]                 |   |
|                        |                         | E 3300 (B) Gesundheitsamt 1936-39 / 1921-79 [112]                   |   |
|                        |                         | E 3300 (C) BA Gesundheitswesen 1940- / 1979- [113]                  |   |
|                        |                         | E 3301 (-) [7] Personen-<br>dosimetrie                              |   |
|                        | 12                      | Zollverwaltung -> Ober-<br>zolldirektion                            |   |
|                        |                         | E 6350 (A) [7] Amtsbücher   |   |
|                        |                         | E 6350 (B) [121]  |   |
|                        |                         | E 6355 (-) [7] Schweiz.<br>Zolltarif                                |   |
|                        | 13                      | --  |   |
|                        | 21                      | wie [11]  |   |
|                        | 22                      | Teil von [11] +<br>E EDI, 1004 (-) Abt. Gifte,<br>Daten aus RAGIDOS |   |
|                        | 23                      | E 4320 (A) [7] Polizeidienst  |   |
|                        |                         | E 4320 (B) [7] Polizeidienst  |   |
|                        |                         | E 4320 (C) [7] Polizeidienst  |   |
|                        |                         | E 4321 (A) [7] Rechtsdienst   |   |
|                        |                         | E 4325 (-) [7] Zentrale Dienste                                     |   |
|                        |                         | E 4327 (-) Bundesanwaltschaft<br>1889 - [231]                       |   |
|                        | 24                      | --  |   |
|                        | 25                      | wie [12]  |   |
|                        | 26                      | --  |   |
|                        | 27                      | --  |   |
|                        | Akzessionen<br>(Anzahl) | 111   | 5 - 1 davon zum Thema Betäu-<br>bungsmittel |
|                        |                         | 112   | 3 - 1                                       |
|                        | 113                     | 38 - 4  |   |
|                        | 121                     | 44 - 0  |   |
|                        | 231                     | 1 - 0   |   |

---

Tab. 3: RECHERCHE IM EDIBAR-SYSTEM II: SUCHBAUM



Der oben beschriebene Suchweg zeigt eine komplette Recherche in unserem heutigen elektronischen Suchsystem. Für einen erfahrenen Archivaren ist es nicht notwendig, alle diese Schritte wirklich im EDIBAR-System nachzuvollziehen. In den meisten Fällen weiss er bereits aufgrund einer gestellten Frage bzw. seiner Rückfragen, welche Amtsstellen mit ihrem Schriftgut für die Beantwortung zu berücksichtigen sind; er wird also erst auf der Ebene der Bestände ins elektronische Retrievalsystem einsteigen und mit dem Überblick über die vorhandenen Akzessionen weiterarbeiten. In KOMP und in demjenigen Teil von BAK, welcher die Übersicht über die Amtsstellen enthält, ist viel archivarisches Expertenwissen formalisiert, das heute auch noch in den Köpfen der langjährigen Archivmitarbeiter gespeichert ist. Liegen für den Archivar-Experten die Vorteile von EDIBAR also nicht in erster Linie in dessen Retrievalkomponenten, so sind diese für die ArchivbenutzerInnen ohne Vermittlung durch das Archivpersonal noch zu kompliziert. Hier besteht ein Ausbaubedarf, und im Hinblick auf eine solche notwendige Erweiterung des BenutzerInnenkreises sind auch die nachfolgenden methodischen Überlegungen zu verstehen.

## 5 Mögliche Erweiterungen der Retrievalstrategie

Grundsätzlich denken wir an zwei Arten der Erweiterung anzubietender Retrievalmöglichkeiten:

1. Das bei ArchivbesucherInnen häufig fehlende methodische und sachliche Wissen um archivistische Suchmöglichkeiten verursacht einen grossen Betreuungsaufwand - eine repetitive Vermittlung von elementarem Basiswissen. Dieser liesse sich vermindern durch ein intelligentes wissensbasiertes System, welches die Teilergebnisse einer Recherche kommentieren und weitere Schritte vorschlagen bzw. deren Konsequenzen erläutern könnte. Dies würde es BenutzerInnen erlauben, ihren Kenntnissen entsprechend selbständig durch die vorhandenen Findmittel zu navigieren, ohne dauernd von der Unterstützung des Archivpersonals abzuhängen. Sie würden bei längeren Arbeiten dabei eine Vertrautheit mit den Archivbeständen gewinnen, welche sie sich wegen des heute

fehlenden direkten Zugangs auf ungleich schwierigere Art erwerben müssen.

2. Beschränkungen des Informationsgehaltes, wie sie sich aus der herkömmlichen Erschliessung durch Informationsverdichtung in Aktensurrogaten ergeben, könnten überwunden werden, wenn man die Akten selbst auch als "Findmittel" befragen könnte. Mit einem kombinierten Suchinstrument, welches die Vorteile des bestehenden, hierarchisch organisierten EDIBAR-Systems mit den neuen Möglichkeiten eines intelligenten Volltext-Retrievalsystems verbinden würde, d.h. mit einem System, welches sich sowohl auf strukturierte Informationen *über* - wie auf unstrukturierte Informationen *in* den Akten stützt, liessen sich erhebliche Verbesserungen und eine zeitgemässe Gestaltung eines nach heutigen Vorstellungen benutzerfreundlichen, natürlichsprachlichen Informationssystems realisieren.

### 5.1 Sprachanalytischer Ansatz (NFP 23)

Seit gut zwei Jahren beteiligt sich das Schweizerische Bundesarchiv, zusammen mit dem Institut dalle Molle pour les études sémantiques et cognitives (ISSCO) in Genf und dem Istituto dalle Molle di studi sull'intelligenza artificiale (IDSIA) in Lugano mit dem Projekt *Specification and Prototyping of a System for the Intelligent Management of Information* am Nationalen Forschungsprogramm 23, *Artificial Intelligence and Robotics*.

Ausgangspunkt unserer Überlegungen für den Entwurf des Forschungsplanes ist die Vorstellung gewesen, dass in Zukunft Massen maschinenlesbarer Informationen zur Verfügung stehen werden, welche sinnvollerweise auch in dieser Form archiviert werden sollten. Die technischen und wirtschaftlichen Fragen, die mit der Aufbewahrung grosser Datenmengen verbunden sind, können heute grundsätzlich als gelöst gelten. Völlig offen ist es dagegen noch, welche Form diese maschinenlesbaren Dokumente haben und wie sie geordnet sein werden. Die Eigenschaft der Maschinenlesbarkeit elektronischer Akten eröffnet aber auf alle Fälle neue Möglichkeiten der Analyse und des Zugriffs - sei es auf ungeordnet abgespeicherte und nur indexierte oder auf klar strukturierte und mit herkömmlichen Aktenzeichen versehene Dokumente.

Im Rahmen des genannten NFP-23-Projektes sind zwei methodische Ansätze ausgewählt worden zur Sicherstellung oder gar Erweiterung und Verbesserung archivischer Retrievalmöglichkeiten in einer elektronischen Umgebung: die automatische Verarbeitung natürlichsprachlicher Dokumente und der Einsatz wissensbasierter Systeme. Die Expertensystemteile würden dabei Funktionen des heutigen EDIBAR-Systems in verbesserter Form übernehmen; die sprachanalytischen Teile müssten die auf herkömmliche Art nicht erreichbare Präzision der Suche durch Berücksichtigung von Dokumenteninhalten erhöhen.

Tabelle 4 zeigt die geplante Struktur des ganz in Prolog programmierten Prototyps, welcher ursprünglich eine enge Verflechtung der beiden Elemente - Expertensystem und Sprachverarbeitung - vorsah, die sich auch funktional bestens ergänzen können.

**Tab. 4: Struktur des Prototyps**

|                  |              |   |              |                                   |              |                                    |
|------------------|--------------|---|--------------|-----------------------------------|--------------|------------------------------------|
| <b>DOKUMENT</b>  | <b>-&gt;</b> | <b>DOKUMENT-<br/>ZERGLIE-<br/>DERUNG</b>      | <b>-&gt;</b> | <b>ERKENNEN</b>                   | <b>-&gt;</b> | <b>DOKUMENTEN-<br/>STRUKTUR</b>    |
|                  | <b>-&gt;</b> | <b>TEXT-<br/>ANALYSE</b>                      | <b>-&gt;</b> | <b>INDEXIE-<br/>RUNG</b>          | <b>-&gt;</b> | <b>DOKUMENTE</b>                   |
|                  | <b>-&gt;</b> | <b>EXPERTEN<br/>SYSTEM</b>                    | <b>-&gt;</b> | <b>ANALYSE<br/>DOSSIERS</b>       | <b>-&gt;</b> |                                    |
|                  |              |   |              |                                   |              | <b>FRAGE<br/>BEANT-<br/>WORTEN</b> |
|                  | <b>-&gt;</b> | <b>EXPERTEN<br/>SYSTEM</b>                    | <b>-&gt;</b> | <b>PRÄZISIERUNG<br/>DER FRAGE</b> | <b>-&gt;</b> |                                    |
|                  | <b>-&gt;</b> | <b>FRAGEZER-<br/>GLIEDERUNG/<br/>-ANALYSE</b> | <b>-&gt;</b> | <b>INDEXIERUNG<br/>DER FRAGE</b>  | <b>-&gt;</b> |                                    |
| <b>RECHERCHE</b> | <b>-&gt;</b> | <b>BENUTZER-</b>                              | <b>-&gt;</b> | <b>FRAGEFORMU-<br/>OBERFLÄCHE</b> | <b>-&gt;</b> | <b>LIERUNG</b>                     |
| <b>-&gt;</b>     |              |   |              |                                   |              |                                    |

Im Rahmen des NFP-23-Projektes sind, den zur Verfügung stehenden Personalkapazitäten entsprechend, zuerst der Dokumentenparser und die

Textanalyse ausgearbeitet worden<sup>33</sup>. Die Analyse der Fragen wird sich analog der Textanalyse vollziehen; sie stellt deshalb keine weiteren Probleme. Die Ausarbeitung einer ergonomischen Benutzeroberfläche stellt v.a. ein technisches Problem dar; sie war deshalb nie von grosser Dringlichkeit. Die Expertensystemteile sollten erst in einer zweiten Phase des Projekts in Angriff genommen werden. Da die vorgesehene Weiterfinanzierung nicht bewilligt worden ist, können sie im Rahmen dieses Forschungsprojektes nun nicht mehr realisiert werden und bleiben als Pläne auf dem Papier bestehen.

Der Hauptteil unserer Arbeit hat sich also auf die Verarbeitung von maschinenlesbaren Dokumenten konzentriert. Damit eine automatische Fragebeantwortung in einigermaßen akzeptabler Zeit zu erreichen ist, müsste die Analyse des Schriftgutes im Rahmen der Akzessionierungsarbeiten durchgeführt werden. Die Resultate dieser Analyse - die Dokumentenprofile, eine Art komplexer Indexeinträge - würden anschliessend in einer Datenbank aufbewahrt.

Um eine Frage zu beantworten, genügt es, diese auf die gleiche Weise zu analysieren wie die archivierten Dokumente. Das Resultat der Frageanalyse - das Frageprofil - kann dann mit den Dokumentenprofilen verglichen werden. Finden sich vergleichbare Profile, so ist die Frage positiv beantwortet. Es bleibt dabei offen, wie weit sprachliche Deckungsgleichheit zwischen Fragen und Quellen überhaupt erwartet werden kann (man denke nur an die Änderungen des Sprachgebrauchs: Atomkraftwerk > Kernkraftwerk, technische Zusammenarbeit > Entwicklungshilfe etc.). Führt dieser Ansatz aber - wie wir glauben - zu positiven Ergebnissen, so liesse sich ein enorm flexibles Retrievalinstrument entwickeln, weil der Inhalt von Dokumenten plötzlich als "Findmittel" zur Verfügung stehen würde.

---

33 Am Projekt haben drei Forscher mitgearbeitet: Der Computerlinguist Michael Hess für die ISSCO, der Systemanalytiker Jean-Luc Cochard für die IDSIA und der Verfasser als Historiker und Vertreter des Bundesarchivs.

### 5.1.1 Dokumentenparser<sup>34</sup>

In diesem Kapitel soll der Verlauf der Analyse der rohen Dokumente bis zur Erstellung eines passenden Dokumentenprofils beschrieben werden. Diese Dokumentenanalyse besteht aus mehreren Arbeitsschritten. Sie beginnt mit der Analyse der Dokumentenstruktur und des Dokumententyps.

**Tab. 5: ORIGINALDOKUMENT**

---

---

DIRECTION FEDERALE DES FORETS

3001 Berne  
Laupenstrasse 20  
Case postale 5662  
Tél.: 031 / 61 80 91  
n.réf.: 267 AB/ef

Aux  
Inspections cantonales  
des forêts  
-----

Berne, mai 1989

Inventaire complet des vues aériennes des dégâts causés par  
les tempêtes  
-----

Messieurs

.....

---

---

Die Dokumente stehen uns in einem 'hart'-formatierten ASCII-Format zur Verfügung (keine Tabulatoren, keine Paragraphenbezeichnungen, nur Leerschläge und alle Zeilen mit <CR> abgeschlossen)<sup>35</sup>. Das uns

---

34 Parser, dt. Zerteiler; Programm, das eine Zeichenfolge einliest und diese syntaktisch analysiert bis ein Ableitungsbaum als Output ausgegeben werden kann. Die Verwendung des Begriffs hier verweist auf die strukturelle Ähnlichkeit dieses Arbeitsschrittes der Dokumentenerkennung mit demjenigen der Syntaxanalyse.

35 Aus methodischen Gründen ist der Ansatz dieses Forschungsprojektes auf französischsprachige Dokumente beschränkt, da die französische Sprache nicht so frei zusammengestellte Nominalkomposita kennt wie das Deutsche und da auch die Syntaxstruktur klarer definiert ist.

Da aktuelle Dokumente der Verwaltung dem Amtsgeheimnis unterliegen, war es schwieriger als ursprünglich geplant, eine ausreichende Zahl von maschinenlesbaren Dokumenten in französischer Sprache zu finden. Die Sektion Förderungsmassnahmen der Abteilung Walderhaltung und Jagd der Eidg. Forstdirektion (Bundesamt für Umwelt, Wald und Landschaft) hat schliesslich Dokumente aus ih-

zur Verfügung stehende elektronische Schriftgut zeigt, dass es bereits für einfachste Dokumente eine enorme Formvielfalt gibt. Dieser Umstand hat uns gezwungen, die Erkennung des Dokumententyps zu beschränken auf die Erkennung von Briefen und von Protokollen. Werden in den Büroinformationssystemen vermehrt internationale Normen und Standards eingehalten, so wird sich die Dokumentenarchitektur besser definieren und bei einer automatischen Analyse auch leichter wiedererkennen lassen<sup>36</sup>.

---

rem Büroautomationssystem zur Verfügung gestellt. Wir danken Herrn R. Gerber, Sektionschef, für dieses Entgegenkommen, dank dem wir uns eine Arbeitsbasis von rund 500 ASCII-Dokumenten beschaffen konnten, 120 davon in französischer Sprache. Die deutschsprachigen Dokumente konnten wir v.a. für die Strukturanalyse benutzen. Daneben verfügen wir noch über eine parallele Datenbank, welche die strukturierten Informationen enthält, wie sie zusammen mit den Dokumenten im Büroautomationssystem verwendet werden (Titel, Verfasser, Datum, Stichworte). Auch aus diesen Informationen konnten wir interessante Einsichten über die Nutzungsgewohnheiten gewinnen.

- 36 Wir denken etwa an folgende Standards: SGML (Standard Generalized Markup Language, ISO 2709) zur Standardisierung des Dokumentenaufbaus, nur für Textdokumente geeignet. ODA/ODIF (Office Document Architecture/Interchange Format, ISO 8613), ebenfalls eine Norm zur strukturierten Dokumentengestaltung. Die ODA/ODIF-Regeln könnten an sich in SGML ausgedrückt werden. Vgl. etwa APPELT 1990.

**Tab. 6:****ZERGLIEDERTES DOKUMENT**

Bloc: orig(1,1), 29 \* 1,  
DIRECTION FEDERALE DES FORETS

*orig(Spalte, Zeile  
[Position der linken  
oberen Blockecke]),  
n\*m [Blocklänge \*  
Anzahl Zeilen].*

Bloc: orig(37,1), 23 \* 4,  
Aux  
Inspections cantonales  
des forêts

-----

Bloc: orig(1,3), 20 \* 5,  
3001 Berne  
Laupenstrasse 20  
Case postale 5662  
Tél.: 031 / 61 80 91  
n.réf.: 267 AB/ef

Bloc: orig(40,9), 15 \* 1,  
Berne, mai 1989

Bloc: orig(1,12), 49 \* 3,  
Inventaire complet des vues aériennes des dégâts  
causés par les tempêtes

-----

Eine sequentielle Lesart würde unverständliche Informationen liefern: ... 'cantonales' '3001' 'Berne' 'des' 'forêts' ... Während der Dokumentanalyse muss also die menschliche Fähigkeit simuliert werden, blockartig angeordnete, sinnmässig zusammengehörige Informationen als Einheiten zu erfassen. Dieser Vorgang ähnelt dem Vorgehen von OCR-Systemen, die Zonen mit hoher Pixeldichte von anderen trennen; der Unterschied liegt darin, dass für unseren Prototyp die Buchstaben die 'Pixels' sind. Der Dokumentenparser liefert als Resultat die Angaben, die in Tabelle 6 enthalten sind.

Tab. 7:

TEXTGRAMMATIK<sup>37</sup>

| Objekt        | -> | Definition<br>(Beschreibung jeden Objektes durch die Aufzählung der obligatorischen, fakultativen, einmal oder mehrmals auftauchenden Elemente) |
|---------------|----|---|
| Textsorte     | -> | {...}   Brief   {...}   |
| Brief         | -> | Normaler_Brief   Rundschreiben  |
| Rundschreiben | -> | (Briefkopf),<br>(Telefonnummer),<br>(Adresse),<br>Datierung +,<br>Referenz<br>Betreff,<br>...<br>(Post_Scriptum),<br>...                        |
| Adresse       | -> | (Verteiler)   Amt   Firmen-<br>name   Personennamen, (Strasse), Ort.  |
| Post_Scriptum | -> | (['PS'])  |
| ...           |    |   |

Da eine automatische linguistische Analyse ganzer Dokumente die Möglichkeiten heutiger Computerlinguistik übersteigt, muss die sprachliche Analyse auf Dokumententeile mit hoher Informationsdichte beschränkt werden können, z.B. auf die Betreff-Angaben. Um dieses Ziel zu erreichen, muss jedem Block eine Etikette 'angeheftet' werden, welche dessen funktionale Bedeutung beschreibt. Dieser Vorgang basiert auf einer «Textgrammatik», welche das Resultat der Analyse von rund hundert Briefen darstellt. Die Grundform dieser «Textgrammatik» ist in Tabelle 7 dargestellt.

Durch die Angabe von Relationen (die Absenderangabe findet sich links von der Adressangabe, das Datum findet sich unter der Adressangabe

37 (Begriff) - optional  
 Begriff + - ein- oder mehrmaliges Vorkommen möglich  
 Begriff \* - kein oder mehrmaliges Vorkommen möglich  
 Begriff1 | Begriff2 - alternatives Vorkommen  
 [Begriff] - konkretes Erscheinungsbild  
 "Begriff" - Vorkommen genau in dieser Form  
 Begriff1, Begriff2 | Begriff3 - Begriff1 und 2 oder Begriff3

etc.) und durch die Angabe eines möglichen Spektrums des Inhaltes dieser Strukturelemente (Datum kann sich zusammensetzen aus Teilen von 123..JanFebMärz.. etc.) gelingt es, die Dokumentenstruktur zu erfassen

**Tab. 8: DOKUMENTENSTRUKTUR UND -INHALT:  
ANALYSERESULTAT**

---

---

**Brief**

**(Rundschreiben**

**(Absender("DIRECTION FEDERALE DES FORETS  
3001 Berne  
Laupenstrasse 20  
Case postale 5662"),**

**Telefon("031 / 61 80 91"),**

**Referenz("267 AB/ef"),**

**Adresse**

**(Verteiler("Aux Inspections cantonales  
des forêts")),**

**Datierung("Berne, mai 1989"),**

**Betreff("Inventaire complet des vues  
aériennes des dégâts causés  
par les tempêtes"),**

**Anrede("Messieurs"),**

**....)).**

---

---

**5.1.2 Die automatische linguistische Analyse<sup>38</sup>**

Im Rahmen einer umfassenden Analyse von natürlichsprachlichen Dokumententeilen kommt an erster Stelle die morphologische Analyse. Diese ist bei einer flexionsreichen Sprache wie dem Französischen von grosser Bedeutung. Im Rahmen unseres NFP-23-Projektes ist ein Mor-

---

<sup>38</sup> Vgl. für die folgenden Ausführungen zur Morphologie: COCHARD 1992, für den syntaxo-semantischen Teil: HESS 1992.

phologieanalyseprogramm entwickelt worden, welches es erlaubt, mehr als 100'000 Wortformen zu verstehen.

Die Dokumententeile, die gewöhnlich die grösste Informationsdichte aufweisen, werden sprachlich analysiert. Viele Systeme beschränken sich dabei auf eine statistische Auswertung. Sie entfernen zuerst alle Funktionswörter und alle häufig auftauchenden Wörter. Die übrigen Begriffe werden auf ihre Stammform reduziert, wie z.B.

- (1) inventaire, inventorisation, inventorier --->  
inventaire

Die Fragen werden als Bool'sche Kombinationen von Begriffen formuliert. Die Frage etwa

- (2) Existe-t-il un inventaire des vues aériennes des dégâts causés par les tempêtes?

würde als Bool'sche Kombination der Begriffe

- (3) inventaire & vue\_aérienne & dégât & tempête.

formuliert. Ein solches Retrievalsystem würde als Antwort einerseits zwar richtigerweise auf Dokumente verweisen, deren Betreffangabe folgendermassen lautete:

- (4) Inventaire complet des vues aériennes de dégâts causés par une tempête.

Leider würden andererseits auch Dokumente angezeigt, deren Betreffzeile lauten könnte:

- (5) Inventaire complet et vues aériennes de dégâts causés par une tempête

Eine solche Antwort ist unbefriedigend wegen ihrer mangelnden Präzision, die sich, sind erst einmal Tausende elektronischer Dokumenten archiviert, verheerend auswirken würde, müssten doch schier endlose Referenzlisten von Hand bearbeitet werden, damit man die nützlichen von den unnützen Dokumenten trennen kann.

Um die Präzision der Retrievalresultate zu erhöhen, wird die linguistische Analyse um eine syntaktische und semantische Analyse ergänzt. Das stellt uns vor verschiedene Probleme:

◆ Die öffentliche Verwaltung beschäftigt sich mit fast allen Lebensbereichen, so dass wir mit Dokumenten konfrontiert sind, deren Sprache allgemein und von enormer Vielfalt ist.

◆ Die syntaktische Struktur einer Frage wird sich kaum in der gleichen Form in den Dokumenten wieder finden lassen, da alles sich in verschiedenen syntaktischen Konstruktionen formulieren lässt. Meist werden wohl die Fragen allgemeiner formuliert sein als die Betreffzeilen von Dokumenten.

◆ Selbst einfache Sätze sind häufig mehrdeutig, wenn man sie sorgfältig analysiert.

Um die Möglichkeiten heutiger Computerlinguistik nicht überzustrapazieren, werden also nur informationsdichte Dokumententeile einer syntakto-semantischen Analyse unterworfen, deren Ergebnis anschliessend in eine normalisierte Logik übertragen wird.

**Tab. 9: SYNTAKTISCHE VARIANTEN FÜR DIE FORMULIERUNG EINES SACHVERHALTES**

---



---

|    |  |
|----|--|
| 6a | Inventaire des vues aériennes de dégâts causés par une tempête             |
| 6b | Inventaire des vues aériennes de dégâts qui ont été causés par une tempête |
| 6c | Vues aériennes inventorisées de dégâts causés par une tempête              |
| 7a | Droit de douane sur l'importation de légumes du Marché commun en Suisse    |
| 7b | Droit de douane sur des légumes importées du Marché commun en Suisse       |

---



---

Die gemeinsame Bedeutung der Betreffangaben (7a-b) kann logisch in folgender Form dargestellt werden<sup>39</sup>:

- (8) droit(sk-1,sk-2),d1,p1,s1  
 douane(sk-2,sk-3),d1,p1,s1  
 importation(sk-3,sk-4,sk-5,sk-6),d1,p1,s1  
 légume(sk-4),d1,p1,s1  
 marché(sk-5),d1,p1,s1  
 commun(sk-5),d1,p1,s1  
 suisse(sk-6),d1,p1,s1

---

39 sk-n ist eine sogenannte Skolem-Konstante; jedem Begriff wird eine eigene solche fortlaufend nummerierte Konstante zugeordnet. Die Angaben d1, p1, s1 verweisen auf den Ort im Dokument, an welchem sich ein Begriff befindet (d1 = im 1. Dokument, p1 = im 1. Absatz, s1 = im 1. Satz etc.).

In einem Lexikon sind die entsprechenden Begriffe mit ihren obligatorischen (fett) und fakultativen (normal) Attributen definiert:

- (9) importation(de{1} qc [objet], de{2} [origine], en [but])
- importer(qn [agent], qc [objet], de [origine], en [but])
- importer(qc/qn [agent], qn [objet])
- douane(sur qc [objet])
- droit(de qc [objet])

Es ist weiter möglich, Variablen zu definieren, welche die Art und Weise, den Ort oder die Zeit wiedergeben. Mit diesen Angaben können die Betreffe 7a-b folgende Form annehmen:

- (10) objet(droit,sk-1,sk-2),d1,p1,s1
- objet(douane,sk-2,sk-3),d1,p1,s1
- action(importation,sk-3,sk-4,sk-5,sk-6),d1,p1,s1
- objet(légume,sk-4),d1,p1,s1
- propriété(marché,sk-5),d1,p1,s1
- propriété(commun,sk-5),d1,p1,s1
- propriété(suisse,sk-6),d1,p1,s1

Ein solches Vorgehen erlaubt eine sehr genaue Beantwortung von Fragen. Es ist sogar möglich, Antworten auf sehr allgemein formulierte Fragen zu finden, wie etwa auf diejenige über

- (11) le droit de douane protectionniste sur l'importation de légumes en Suisse,

welche logisch in nachstehender Form repräsentiert wird:

- (12) droit(Dr,D) n douane(D,I) n protectionniste(D) n importation(I,L,\_,S) n légume(L) n Suisse(S)

und durch den Verweis auf die Dokumente 7a-b beantwortet werden kann. Die Frage

- (13) Existe-t-il un inventaire des vues aériennes des dégâts causés par les tempêtes?

dagegen kann nicht durch den Verweis auf ein Dokument mit der Be-  
treffzeile

- (14) Inventaire des vues aériennes des forêts dévastées par une tempête

beantwortet werden. Da solche sprachlichen Inkongruenzen häufig sind, hatten wir einen Weg zu finden, der verhindert, dass negative Antworten die Regel werden. Dank der Möglichkeit einer graduellen Abstumpfung lässt sich der Präzisionsanspruch des Retrievalsystems schrittweise reduzieren (15-17). Die gesuchten Begriffe müssen sich

dann nicht mehr innerhalb eines einzelnen Satzes oder eines einzelnen Abschnittes befinden.

- (15) vue(V,D),d,p,s  
aérienne(V)d,p,s  
inventaire(I,V),d,p,s  
dégât(D),d,p,s  
tempête(T),d,p,s  
cause(T,D),d,p,s
- (16) vue(\_),d,p,\_  
aérienne(\_),d,p,\_  
inventaire(\_),d,p,\_  
dégât(\_),d,p,\_  
tempête(\_),d,p,\_  
cause(\_),d,p,\_
- (17) vue(\_),d,\_,\_  
aérienne(\_),d,\_,\_  
inventaire(\_),d,\_,\_  
dégât(\_),d,\_,\_  
tempête(\_),d,\_,\_  
cause(\_),d,\_,\_

Damit werden die herausgearbeiteten Beziehungen zwischen den Begriffen allerdings hinfällig, doch behält unser Prototyp immer noch die Präzision eines konventionellen Retrievalsystems.

## 5.2 Wissensbasierte Systeme (Expertensysteme)<sup>40</sup>

Expertensysteme sind wissensbasierte und regelgeleitete Systeme, welche versuchen, gemäss den ihnen zur Verfügung gestellten Informationen (Fakten und Regeln) eigenständig einen Weg durch einen Problemraum zu finden. Sie haben bereits mit Erfolg einen grossen Anwen-

---

40 Es war von Anfang an klar, dass eine sprachliche Analyse von einzelnen Dokumenten nicht einen sinnvollen selbständigen Zugang zu Archivmaterial schaffen kann, da die archivische Grundeinheit das Dossier ist. Deshalb sah das NFP 23-Projekt immer vor, neben den computerlinguistischen Teilen auch Expertensystemteile im geplanten Prototyp zu integrieren, welchen die Aufgabe zukommen würde, den kompetenz- und institutionenbezogenen Zugang zum Archivmaterial zu sichern.

Neben dem Expertensystemteil im NFP 23-Projekt hat das Bundesarchiv in Zusammenarbeit mit der ETHZ noch ein Projekt entworfen, welches im Rahmen des Schwerpunktprogrammes Informatik hätte realisiert werden sollen. Im folgenden werden die Ideen beider Projekte ohne weitere Unterscheidung zusammengefasst und vorgestellt.

dungskreis in der Praxis gefunden, und auch kommende Büroinformationssysteme werden sich ihre Vorteile zu Nutzen machen<sup>41</sup>. Sie können im Rahmen des Information Retrieval wertvolle Unterstützung bieten.

Wichtiger Bestandteil eines Expertensystems ist dessen Wissensbasis. Es ist klar, dass sich heute in einem Expertensystem nur spezifisches Fachwissen sinnvoll formalisieren lässt, nicht allgemeines Weltwissen. Da sich öffentliche Verwaltungen heute praktisch mit allen Lebensbereichen auseinandersetzen, ist es kaum möglich, alles benutzte Fachwissen in einem archivischen Expertensystem zu erfassen, um in einem intelligenten Retrievalsystem methodische und themenspezifische Unterstützung anbieten zu können. Es ist eine 'arbeitsteilige' Lösung anzustreben.

Einerseits werden die einzelnen Amtsstellen - das Bundesarchiv mit eingeschlossen - in Zukunft intelligente Büroinformationssysteme verwenden, welche die Erledigung der anfallenden Geschäfte sowohl verwaltungsseitig wie fachseitig unterstützen werden. Ihre Wissensbasen werden einerseits, wahrscheinlich für die ganze Bundesverwaltung weitgehend einheitlich, juristisch-administratives Wissen umfassen, andererseits höchst unterschiedliches fachlich-inhaltliches Wissen, das jede Amtsstelle für ihre spezifische Aufgabenerledigung benötigt<sup>42</sup>. In diese Systeme wird durch das integrierte Regelwerk ein Handlungskonzept eingebaut, was neue Probleme bezüglich ihrer Archivierung aufwirft; sie stellen zudem die Forschung vor neue (hermeneutische) Probleme, da Teile des Entscheidungsprozesses in Form von Programmcodes, nicht mehr in Dokumentenform vorliegen.

Ergänzend zu zukünftigen amtsstellenzentrierten Expertensystemen zur Unterstützung der Verwaltungsarbeit müsste das Archiv ein eigenes wissensbasiertes Retrievalsystem entwickeln, welches auf der Basis methodisch-archivkundlichen Wissens über erfolgversprechende Suchstrategien den BenutzerInnen einen möglichst anwendungsfreundlichen, selbständigen Zugang zum Archivgut ermöglichen sollte. Unter methodisch-archivkundlichem Wissen ist sowohl Wissen über die Beständestruktur, Kompetenzverteilung und Organisationsstruktur der aktenabliefernden Verwaltung zu verstehen wie auch Erfahrungen aus be-

---

41 Vgl. etwa SCHMIDT 1989.

42 MICHELSON 1991:13ff gibt einen Überblick über Anwendungsbeispiele in der amerikanischen Bundesverwaltung (Steuerverwaltung, Sozialversicherung, Finanzverwaltung).

reits gemachten Recherchen (heute in BAK-BV enthaltene Informationen), welche helfen könnten, nachfolgende Recherchen zu optimieren. Dieser konzeptionell-zukunftsorientierten Perspektive steht eine gegenwartsorientiert-pragmatische gegenüber: Es wäre theoretisch möglich, ausgehend von den in der EDIBAR-Datenbank enthaltenen Fakten, bereits heute ein flexibleres Retrievalinstrument zu entwickeln, welches die Erschließung bereits archivierter Bestände verbessern würde.

Ausgehend vom Recherche-Beispiel in Tabelle 2 lässt sich andeuten, was eine wissensbasierte Unterstützung der archivischen Suche bedeuten könnte: Bereits die Umsetzung einer Frage in ein Schlagwort oder eine Kompetenz könnte, in Verbindung mit dem oben vorgestellten computerlinguistischen Ansatz, wirkungsvoll unterstützt werden. Begriffsgruppen aus einer Frage könnten etwa als Suchbegriffe mit der Systematischen Rechtssammlung verglichen werden. Entsprechende Fundstellen (Gesetzestitel, Titel eines Gesetzesabschnittes, Titel eines Artikels, Verweis auf ein zweites Gesetz in einer Fussnote etc.) könnten vom Retrievalsystem entsprechend ihrer Stellung in einem hierarchisch strukturierten Text gewichtet werden. Sollte eine Suche zu keinen positiven Ergebnissen führen, so könnte das Retrievalsystem weitere Suchbereiche vorschlagen (Geschäftsberichte, Parlamentsprotokolle, Staatskalender etc.) oder die BenutzerInnen nach zusätzlichen Informationen befragen (vgl. Tabelle 4).

Nach ersten Entscheidungen könnten aufgrund gefundener Informationen die zuständigen Amtsstellen automatisch ermittelt werden.

Wie der Suchbaum zeigt, sind während der Recherche auf verschiedenen Ebenen Entscheidungen zu treffen: Sollen alle möglichen Kompetenzträger und alle Bestände in die Recherche eingeschlossen werden? Eine Auswahl einzelner Kompetenzträger bzw. die Priorisierung der Recherche hat inhaltliche Konsequenzen, denn die einzelnen Amtsstellen befassen sich unter je einem spezifischen Gesichtspunkt mit "Kokain", und entsprechend wird ihr Schriftgut andere Informationen zum Thema liefern, etwa gesundheitliche Aspekte in den Vordergrund rücken oder solche der Strafverfolgung. Durch die Berücksichtigung der geplanten Arbeitsintensität (Zeitungartikel - Dissertation) liesse sich die Recherche anders gestalten: Das System könnte vorschlagen, einen Teilaspekt auszuwählen, ev. je nach Quellenlage, Anzahl in Frage kom-

mender Dossiers oder nach thematischen Einschränkungen, oder es könnte ohne Rückfrage eine Übersicht über alle Fundstellen liefern.

Bei der manuellen Recherche erweisen sich ausserdem viele anfänglich offene Suchwege, manchmal bereits auf der Ebene der Kompetenzträger, als Sackgassen. Ein intelligentes System würde z.B. Regeln enthalten wie "wenn eine Amtsstelle noch keine Akten ans Archiv abgeliefert hat, dann zeige dies mit einem Vermerk an".

Beim Schritt von der Ebene der Kompetenzträger zur Übersicht über deren Schriftgut muss man beachten, dass die Akzession nur eine organisatorische Hilfsgrösse darstellt, welche vom logisch-thematischen Gesichtspunkt aus keine sinnvollen Informationen liefert. Ein intelligentes Retrievalsystem müsste demnach das Schriftgut einer Amtsstelle als Einheit behandeln. Die gestellte Frage der BenutzerInnen bedeutet die wesentliche Eingrenzung: es ist nicht alles Schriftgut einer Amtsstelle für die Recherche relevant. Die im EDIBAR-System grundsätzlich mögliche, allerdings manuell vorzunehmende Eingrenzung auf Akzessionen, welche Schriftgut zu einer bestimmten Registraturplanposition (hier: Betäubungsmittel) enthalten, könnte automatisiert werden; dies wird erst nach der Digitalisierung aller Detailfindmittel sinnvoll möglich sein (Betäubungsmittel und/ oder Kokain). Es könnte auch automatisch auf verschiedene Spezialfindmittel (z.B. Namenskarteien) aufmerksam gemacht werden, welche in Zukunft ebenfalls maschinenlesbar und integraler Bestandteil des archivischen Retrievalsystems sein werden.

Bei einer allgemein gehaltenen Frage könnte ein wissensbasiertes Retrievalsystem die BenutzerInnen auch darauf aufmerksam machen, dass das Rechercheresultat aus dem EDIBAR-System in erster Linie auf Akten aus dem Gesetzesvollzug verweist. Ist man eher am Prozess der Rechtssetzung interessiert, so würde sich eine andere Vorgehensweise aufdrängen, welche im Rahmen eines erneuerten oder zukünftigen Retrievalsystems natürlich auch zu unterstützen wäre. Dies würde rückblickend ebenfalls eine Nacherfassung von Informationen voraussetzen, vorausblickend würde es die rechtzeitige Sicherstellung solcher Informationen im Rahmen des Verfahrens zur Archivierung elektronischer Sachakten voraussetzen (als Metadaten).

Im EDIBAR-System ebenfalls nicht direkt ersichtlich ist, dass man in vielen Fällen zusätzlich auch die Akten der Generalsekretariate oder die Handakten einzelner Magistraten oder Chefbeamter konsultieren müsste.

Sind die Detailfindmittel einmal digitalisiert, so würde die Recherche automatisch auch diese Bestände miterfassen. Die Bedeutung einzelner solcher Bestände könnte bewertet werden, so dass darauf nur hingewiesen wird, wenn sich nicht nur eine sprachliche Deckungsgleichheit feststellen liesse. Die Bewertung würde z.B. Bestände mit inhaltlich interessanten Dossiers von solchen unterscheiden helfen, die nur Doppel und Abschriften enthalten, materiell die Nachforschung also nicht weiterführen.

Ginge man davon aus, dass die bestehende hierarchische Organisation des Suchprozesses im EDIBAR-System nicht zwingend weiter erhalten bleiben muss, so könnte man eine Recherche sinnvollerweise bei der jeweils auf der Basis der Systematischen Rechtssammlung aktuell zu ermittelnden Kompetenzverteilung einsetzen lassen, von dort aus in einem nächsten Schritt die Kompetenzträger ermitteln und anschliessend - bevor man das Schriftgut selber in die Recherche miteinbezieht - anhand von Geschäftskontrolldaten herausuchen, welche Geschäfte für die Beantwortung einer Frage von Belang sind. Die Geschäftskontrolldaten könnten bereits erste Querbezüge zu anderen Amtsstellen anzeigen. Erst zuletzt würde dann eine möglichst klar eingegrenzte Teilmenge des Schriftgutes - Registraturplanpositionen - Dossiertitel (Abgabeverzeichnisse) - Dokumente - durch das Retrievalsystem analysiert. Solche Vorstellungen sind zwar zukunftsgerichtet, gehen sie doch von Daten aus, welche heute in dieser Form gar nicht zur Verfügung stehen. Gleichzeitig dienen sie aber auch dem Zweck, die archivischen Vorstellungen zu konkretisieren, welche Daten den ArchivarInnen nützlich sein könnten für die arbeitsökonomisch noch realisierbare Erschliessung der Bestände; mit Variationen liesse sich auf dem heutigen System aufbauen. Die Daten müssten neu geordnet werden, damit die Struktur der inhaltlich-erschliessenden Informationen nicht vermischt würde mit denjenigen, die sich aus den Bedürfnissen der physischen Verwaltung des Archivgutes ergibt. Mit verhältnismässig einfachen Veränderungen könnte das bestehende System verbessert werden, bevor KI-Applikationen zum Einsatz kämen. Später könnte dann ein archivisches Expertensystem, kombiniert mit einer automatischen computerlinguistischen Dokumentenanalyse, als Ausbauperspektive ins Auge gefasst werden für ein "Archiv 2000 Informationsversorgungssystem".

## 6 Schlussbemerkung

"Die Informationstechnologien werden die archivische Arbeit in den kommenden Jahren in einem heute noch kaum vorstellbaren Ausmass verändern"<sup>43</sup> und es wird praktisch jeder Bereich archivischer Tätigkeit von diesen Veränderungen betroffen sein. Es geht dabei nicht nur um die Entwicklung neuer Hilfsmittel, sondern die ganze Methodik steht zur Diskussion. Es geht um die informatikgerechte Gestaltung künftiger Archivarbeit.

Das Bundesarchiv befindet sich dabei in der komfortablen Lage, bereits über ein umfassendes Archiv-Informationssystem zu verfügen. Mit den gewandelten Vorstellungen und Bedürfnissen und den neuen technischen Möglichkeiten ergibt sich aber jetzt ein gewisser Ausbau- und Erneuerungsbedarf. Es ist deshalb an der Zeit, das Erreichte kritisch zu würdigen und darüber nachzudenken, wie zukunftsweisende Verbesserungen in Richtung auf ein vollelektronisches benutzerfreundliches Informationsversorgungszentrum realisiert werden könnten.

In diesem Artikel haben wir uns fast ausschliesslich mit der Retrievalkomponente des EDIBAR-Systems befasst, der Komponente, welche, von einem methodischen Standpunkt aus betrachtet, zu den interessantesten und anspruchsvollsten Teilen eines Archivinformatiksystems gehört. Eine mögliche Ausweitung eines archivischen Suchinstrumentes in Richtung auf einen methodischen Pluralismus müsste logischerweise auch Konsequenzen haben auf die Art der Erschliessung von Archivmaterial. Heute ist das Bundesarchiv an verschiedenen Forschungsprojekten und Arbeitsgruppen beteiligt und kümmert sich aktiv um die Zukunftsgestaltung. Damit ist der Zeitpunkt gegeben, verschiedene neue Trends etwas zusammenzufassen und ihre Auswirkungen zu überdenken. Informatischem Optimismus muss dabei oft mit etwas Skepsis begegnet werden. Gleichzeitig darf man aber nicht verkennen, dass sich bei allen Schwierigkeiten auch konkrete positive Entwicklungen abzeichnen, wie wir sie hier skizziert haben. Erst das eigene Engagement hilft, solche Trends - auch in archivischem Sinn - zu stärken. In dieser offenen Situation werden die ArchivarInnen dauernd herausge-

---

43 DOLLAR 1992a:1.

fordert, und diese Herausforderungen verleihen unserem Beruf einen grossen Reiz.

## 7 Literaturverzeichnis

- APPELT 1990 Wolfgang, Dokumentenaustausch in Offenen Systemen. Einführung in die ISO-Norm 8613: Office Document Architecture (ODA) and Interchange Format, Berlin.
- ARCHIVIERUNG 1988 Die, elektronisch gespeicherter Daten, Referate gehalten an der Arbeitstagung der VSA in Bern, 25.3.1988, in: ARBIDO-R 3, S. 69-84.
- BEARMAN 1991 David, Archival Principles and the Electronic Office, Vortrag Marburg, Oktober 1991, Ms., Pittsburgh.
- BEARMAN 1992 David, Documenting Documentation, in: Archivaria 34 (im Druck).
- BUCHMANN 1988 Wolf, Anforderungen aus der Sicht des Archivars, Referat gehalten an der VSA-Arbeitstagung in Bern, 1988, in: ARBIDO-R 3, S. 70-75.
- BÜTIKOFER 1991a Niklaus, CADUFF, Hugo, GRAF, Christoph, Systematische Beständeübersicht, Bern.
- BÜTIKOFER 1991b Niklaus, Archivierung von statistischen Daten, in: Studien und Quellen 17, Bern, S. 263-294.
- CADUFF 1978 Hugo, RUFER, Kurt, Der Einsatz der maschinellen Datenverarbeitung im Bundesarchiv, in: Studien und Quellen 4., S. 129-170.
- CADUFF 1990 Hugo, FLÜCKIGER, Bernhard, GRAF, Christoph, Computer Science in the Service of Research and Administration. Electronic Data Processing in the Swiss Federal

- Archives, in: JANUS, Archival Review 2, S. 5-24.
- CADUFF 1989** Hugo, FLÜCKIGER, Bernhard, GRAF, Christoph, Informatik im Dienste von Forschung und Verwaltung, Studien und Quellen 15, S. 213-272.
- COCHARD 1992** Jean-Luc, Une technique efficace de traitement en Prolog de la morphologie flexionnelle du français, im Druck.
- COMMISSION FEDERALE POUR L'INFORMATION SCIENTIFIQUE 1988**, Versorgung der Schweiz mit wissenschaftlicher Information in Gegenwart und Zukunft, Zürich.
- DOLLAR 1992a** Charles, Die Auswirkungen der Informationstechnologien auf archivische Prinzipien und Methoden, übersetzt und hg. von Angelika Menne-Haritz, Veröffentlichungen der Archivschule Marburg 19, Marburg.
- DOLLAR 1992b** Charles, Archivists and Librarians. Our Common Ground in the Information Age, Referat, Ms., Washington.
- GRAF 1990** Christoph, Schweizerische Erfahrungen im Spannungsfeld zwischen Büroautomation und Archiven, in: Der Archivar 43, Sp. 102-106.
- GREEN 1991** Adam, The Development of Policies and Plans in Archival Automation, A RAMP Study with Guidelines, Paris.
- HESS 1992** Michael, An Incrementally Extensible Document Retrieval System Based on Linguistic and Logic Principles, in: Proceedings of SIGIR, Kopenhagen.
- INTERNATIONAL COUNCIL ON ARCHIVES 1990**, Statement of Principles Regarding Archival Description, adopted by the Ad Hoc Commission on Descriptive Standards, Höhr-Grenzhausen.
- MENNE-HARITZ 1988** Angelika, Büroautomation und Schriftgutverwaltung. Der Stand der Entwicklung und

- Überlegungen zu den Konsequenzen für die Archive, in: Der Archivar 41, Sp. 365-386.
- MENNE-HARITZ 1990** Angelika, Büroautomation und Schriftgutverwaltung. Stand der Entwicklung und Konsequenzen für die Archive, in: Der Archivar 43, Sp. 89-95.
- MENNE-HARITZ 1992** Angelika, Schlüsselbegriffe der Archivterminologie. Lehrmaterialien für das Fach Archivwissenschaft, Veröffentlichungen der Archivschule Marburg Nr. 20, Marburg.
- MICHELSON 1991** Avra, Expert Systems Technology and its Implication for Archives, National Archives Technical Information Paper No. 9.
- MICHELSON 1992** Avra, ROTHENBERG, Jeff, Scholarly Communication and Information Technology: Exploring the Implications of Changes in the Research Process for Archives, in: American Archivist (im Druck).
- NATIONAL ARCHIVES TECHNICAL INFORMATION PAPER NO. 8 1990:** A National Archives Strategy For the Development and Implementation of Standards for the Creation, Transfer, Access, and Long-Term Storage of Electronic Records of the Federal Government, National Archives Technical Information Paper No. 8, Washington.
- PAPRITZ 1964** Johannes, Zum Massenproblem der Archive, in: Der Archivar 17, Sp. 213-220.
- REALISIERUNGSKONZEPT 1989** für die Büroautomation in der schweizerischen Bundesverwaltung.
- SCHÄRER 1976** Erich, Die systematische Kompetenzenkartei des Bundesarchivs, in: Studien und Quellen 2, Bern, S. 113-153.
- SCHMIDT 1989** Stephanie, An Information Model for an Office Information System, in: Angewandte Informatik 10, S. 421-428.

## Résumé

Les Archives sont depuis un certain temps déjà confrontées à un flot toujours croissant de dossiers, que l'introduction récente de systèmes informatiques dans les services administratifs n'a pas réussi à diminuer. L'informatique offre bien cependant, là aussi, des possibilités de diriger dans une certaine mesure ce flux d'informations vers des canaux contrôlés.

Suite à ces considérations introductives, objet du premier chapitre de l'article, le deuxième chapitre distingue trois domaines d'utilisation de l'informatique aux Archives: la bureautique, l'informatique d'archives proprement dite, en tant qu'automatisation ou soutien électronique de fonctions archivistiques spécifiques, et l'archivage des documents électroniques. L'informatique d'archives sert, pour une part, à la rationalisation des procédures archivistiques, pour une autre part aussi, à l'amélioration qualitative de l'expédition des tâches traditionnelles; elle permet encore, parallèlement, d'offrir de nouveaux services.

Le troisième chapitre présente un aperçu du développement de l'archivage des données électroniques au plan fédéral, en portant une attention particulière aux efforts actuels d'amélioration de la procédure de versement pour permettre la sauvegarde, non seulement des données, mais aussi des documents électroniques eux-mêmes. Suit une appréciation critique du système informatique d'archives actuel, EDIBAR, complétée par la description des projets concrets de développement que les Archives fédérales ont l'intention de réaliser dans les années à venir. Enfin, des idées pour la mise sur pied d'un centre d'information pour un domaine spécifique sont proposées à la discussion.

Le quatrième chapitre analyse les stratégies archivistiques de questionnement reposant sur le principe de provenance et du respect des fonds et les possibilités de questionnement dans le système EDIBAR.

Dans la foulée de cette analyse et partant du principe que des documents électroniques seront archivés à l'avenir, le cinquième chapitre propose des élargissements possibles de la stratégie archivistique de questionnement. L'occasion de ces réflexions est offerte par l'achèvement d'un projet de recherche que les Archives fédérales ont pu réaliser conjointement avec l'Institut pour les études sémantiques et cognitives (ISSCO)

de Genève et l'Istituto dalle Molle di studi sull'intelligenza artificiale (IDSIA) de Lugano, dans le cadre du Programme national de recherche 23 "Artificial Intelligence and Robotics". Dans ce projet de recherche, des méthodes de linguistique computationnelle ont été mises en relation avec celles des systèmes d'experts, pour développer un prototype de système de questionnement intelligent, capable de répondre aux besoins des Archives, en respectant les principes archivistiques mentionnés au chapitre 4. Ce prototype reconnaît la structure des documents électroniques, il isole les éléments porteurs de sens (indication de l'objet du document, etc.) et analyse ceux-ci de telle sorte qu'il puisse automatiquement générer, pour chaque document, une entrée d'index complexe prenant en considération non seulement des notions isolées mais aussi des informations syntaxiques, qui complètent la caractérisation du document. Il s'ensuit que la précision de la réponse peut être considérablement augmentée. Même si le prototype réalisé ne peut encore traiter automatiquement qu'un type de document (lettre), les résultats obtenus sont encourageants et appellent de nouvelles recherches.

## Compendio

Da parecchio tempo gli archivi sono confrontati con un flusso di documenti sempre crescente, che la recente introduzione dei sistemi informatici nell'amministrazione non è riuscita a ridurre. Tuttavia l'informatica permette di canalizzare e di tenere sotto controllo il flusso delle informazioni.

Dopo alcune considerazioni introduttive, nel secondo capitolo si distinguono tre campi d'applicazione dell'informatica nell'archivio: l'automazione degli uffici, l'informatica d'archivio in senso stretto (intesa cioè come automatizzazione o come supporto elettronico di funzioni archivistiche specifiche) e l'archiviazione di documenti informatizzati.

L'informatica per uso archivistico serve in parte a garantire l'esecuzione dei compiti tradizionali, in parte a migliorarla qualitativamente, per-

mettendo contemporaneamente di assumere nuovi compiti e di fornire nuove prestazioni.

Nel terzo capitolo è presentato per sommi capi lo sviluppo dell'archiviazione di dati informatizzati a livello federale, con particolare riguardo agli sforzi compiuti attualmente per lo sviluppo di una migliorata procedura di consegna per gli atti elaborati elettronicamente. Ne segue un apprezzamento critico dell'attuale sistema d'archivio informatico EDIBAR, corredata dalla presentazione di concreti progetti d'espansione che l'Archivio federale intende realizzare nei prossimi anni. Il capitolo si conclude presentando, affinché siano discusse, diverse idee riguardo alla creazione di un centro d'informazione specializzato.

Nel quarto capitolo sono analizzate le strategie archivistiche "retrieval" e le loro possibilità di applicazione, basate sul principio di provenienza e sul rispetto del fondo nel sistema EDIBAR.

In relazione con quest'analisi e partendo dal principio che in avvenire saranno archiviati altri documenti elettronici, il quinto capitolo propone possibili ampliamenti della strategia archivistica retrieval. Spunto concreto di discussione è la fine di un progetto di ricerca che l'Archivio federale ha realizzato congiuntamente all'Institut pour les études séman-tiques et cognitives (ISSCO) di Ginevra e all'Istituto dalle Molle di studi sull'intelligenza artificiale (IDSIA) di Lugano, nel quadro del Programma nazionale di ricerca 23 "Artificial Intelligence and Robotics". Nel progetto sono messi in relazione applicazioni di linguistica computazionale e sistemi di esperti per sviluppare il prototipo di un sistema retrieval intelligente, in grado di soddisfare le necessità ed i principi archivistici. Il prototipo riconosce le strutture dei documenti elettronici, isola gli elementi significanti (indicazione dell'oggetto del documento, ecc.) e li analizza in modo da potere generare automaticamente per ogni documento una complessa registrazione dell'indice, che non contenga solo singoli concetti bensì che tenga presente anche i loro rapporti sintattici. In tal modo la precisione delle risposte può essere sostanzialmente migliorata. Nonostante il prototipo realizzato possa trattare autonomamente un solo tipo di documenti (lettere), i risultati sono incoraggianti ed invitano al proseguire le ricerche.



## Archivisches Suchen und moderne Informatik

|         |                            |
|---------|----------------------------|
| In      | Studien und Quellen        |
| Dans    | Etudes et Sources          |
| In      | Studi e Fonti              |
| Jahr    | 1992                       |
| Année   |                            |
| Anno    |                            |
| Band    | 18                         |
| Volume  |                            |
| Volume  |                            |
| Autor   | Kellerhals-Maeder, Andreas |
| Auteur  |                            |
| Autore  |                            |
| Seite   | 65-100                     |
| Page    |                            |
| Pagina  |                            |
| Ref. No | 80 000 109                 |

Das Dokument wurde durch das Schweizerische Bundesarchiv digitalisiert.

Le document a été digitalisé par les Archives Fédérales Suisses.

Il documento è stato digitalizzato dell'Archivio federale svizzero.