

Lutz Pietzka

## Techniken und Methoden wissensbasierter Systeme: Expertensysteme in Bibliothek, Information und Dokumentation

*Expertensysteme haben sich zu einem bedeutenden Wirtschaftsfaktor der Künstlichen Intelligenzforschung entwickelt, ebenso wie zu einem echten bibliothekswissenschaftlichen Arbeitsmittel. Verschiedene Ansätze in der Forschung und im praktischen Gebrauch beweisen die fortgeschrittene Anwendbarkeit dieser modernen Technologie im zeitgemäßen Informationsmanagement. Klassische Tätigkeitsfelder, wie es sie in Bibliotheken und Dokumentationsstellen immer schon gegeben hat, verändern unter dem Einfluß der KI das Tätigkeitsprofil wissenschaftlicher Bibliothekare zunehmend.*

Techniques and methods of systems based on knowledge: Expert based systems in library, information and documentation

*Expert based systems have become a significant economic factor of artificial intelligence research as well as a genuine librarian tool. A wide range of attempts in research and practice have proven the usability of this innovative technology in approach to a contemporary information management. Classic activities, like those which have always been in libraries and documentation centers, are changing under the impact of AI the profile of scientific librarians increasingly.*

Techniques et méthodes des systèmes basés sur le savoir: Des systèmes d'expert dans la bibliothèque, l'information et la documentation

*Des systèmes d'expert sont devenues un significant facteur économique de la recherche d'intelligence artificielle aussi bien que un vrai expédient de travail de la bibliothéconomie. Des différentes bases dans la recherche et dans l'usage pratique manifestent l'applicabilité avancée de cette technologie dans le management d'information moderne. Des activités classiques dans les bibliothèques et les service de la documentation changent de plus en plus sous l'influence de IA le profil d'activité des bibliothécaires de recherche.*

### 1 Einführung

Wer von Expertensystemen spricht, denkt im allgemeinen an phantastische Bilder einer schönen neuen Welt. Der Laie verbindet diesen Gedanken jedoch häufig mit der festen Überzeugung, bereits präsenste Systeme seien hoffnungslose und unpraktische Produkte einer fast jenseitsbezogenen Forschung. Der in der deutschen Öffentlichkeit entstandene, nahezu literarische Eindruck von einer zwar faszinierenden, aber ansonsten nur zeitverschwendenden Technologie, kann gravierende volkswirtschaftliche Konsequenzen nach sich ziehen. So darf es als originäre Aufgabe der Universitäten und Hochschulen angesehen werden, daß der Industrie Erkenntnisse und Erfahrungen vermittelt werden, die diese selbst noch nicht zu erbringen imstande sind. Mit diesem Axiom soll nun keinesfalls die stets korrespondierende interdisziplinäre Standortdiskussion erneut entfacht werden. Für den bibliothekswissenschaftlichen Bereich seien in diesem Sinne nur die angelsächsischen Aktivitäten genannt, wie die an der Dalhousie University (Halifax) und an der University of London/School of Librarian Sciences.

Zukunftsträchtige Technologien aus dem Bereich der Künstlichen Intelligenz, vermittelt von engagierten Fachkräften, dürften so die beste Grundlage bieten, um auch im 21. Jahrhundert die Überlegenheit angelsächsischer Bibliothekswissenschaft zu zementieren. In diesem Zusammenhang soll es Ziel und Aufgabe der vorliegenden Schrift sein, ein Diskussionspapier vorzulegen, welches geeignet sein mag, das Interesse der Fachwissenschaft zu wecken und um Verständnis für notwendige Positionswechsel zu werben.

### 2 Ökonomische Voraussetzungen

Laut einer noch von der Reagan-Administration in Auftrag gegebenen Marktstudie sollten Softwarelösungen für Expertensysteme<sup>1</sup> 1990 ein Marktvolumen von über einer Milliarde US-Dollar erreichen.

Der Markt konnte diese zum Zeitpunkt ihrer Veröffentlichung belächelte Prognose nicht stützen; sie wurde weit übertroffen ...

In der vergangenen Dekade haben über 95% der führenden US-amerikanischen Großkonzerne Pilotprojekte gestartet, um mögliche Einsatzgebiete zu evaluieren.

Die überwiegend positiven Erfahrungen, insbesondere auch im öffentlichen Dienst, haben dann zu einer weitgehenden Übernahme dieser Pilotprojekte in das upper information management geführt. In diesem Sektor ist ein eindeutiger Schwerpunkt im Bereich der Management Informationssysteme (MIS) zu konstatieren gewesen, da hier eine überproportionale Übernahme der XPS vorgenommen wurde. Diese US-Unternehmen, die frühzeitig einen praktischen Einsatz von XPS evaluiert haben, werden bei der absehbaren Marktverdichtung von ihren Erfahrungen profitieren; gleiches trifft auf das nord-amerikanische Bibliothekswesen zu, welches schon seit Jahren über seine bibliothekswissenschaftlichen Fakultäten in dieser Disziplin engagiert ist. Die nicht nur durch Mary Dykstra renommierte School of Library and Information Studies der Dalhousie University mag mit ihren vielseitigen KI-Forschungs- und Entwicklungsprojekten nur als ein Beispiel für viele gelten.

1 Der gängigen Literatur folgend im Fortlauf XPS genannt.

Der mit dieser Entwicklung korrespondierende ökonomische Erfolg vieler US-amerikanischer und zunehmend auch japanischer Softwarehäuser wird auch auf dem europäischen Markt konsequenterweise zu einer Marktvorherrschaft führen. Die Übermächtigkeit kalifornischer Softwarehersteller im Bereich der Standardlösungen mag hier einen genauen Prospekt der zukünftigen Marktlage bieten. Wenn die Europäer weiterhin der praktischen Anwendung von Ergebnissen der Künstlichen Intelligenz-Forschung reserviert gegenüberstehen und die ökonomische Bedeutung dieser Technologie außerdem in Frage stellen, wird der überseeische Technologievorsprung nicht wieder wettzumachen sein.

„Expertensysteme werden im Kreislauf der Information eine ähnliche Wichtigkeit in einer Volkswirtschaft erlangen wie die Institutionen, denen die Steuerung der anderen Produktionsfaktoren wie Kapital, Ressourcen und Arbeit obliegt“<sup>2</sup>.

In Fernost kulminierten diese entsprechenden Anstrengungen im japanischen Fifth Generation Computer System (FGCS) Project.

Das MITI (Ministry of International Trade and Industry) ist ein Instrument der Regierung Kaiser Akihitos zur Förderung und Stimulierung der japanischen Außenwirtschaft. Nach eigenem Selbstverständnis ist der staatliche Eingriff in ökonomische Mechanismen erforderlich, um eine kontinuierliche Entwicklung gerade auf dem Forschungssektor zu gewährleisten.

Unter diesen Auspizien wurde das FGCS Project durchgeführt. Erarbeitet wurden die Grundlagen für den Bau von Hard- und Softwarelösungen für den Künstlichen Intelligenz (KI)-Bereich, dabei wurde ein besonderer Schwerpunkt auf den Bereich der Expertensysteme gelegt. Das Institute for New Generation Computer Technology (ICOT) hat ad hoc als zentrale staatliche Lenkinstanz funktionieren können.

Von besonderer Bedeutung ist dieses von 1982 bis 1991 mit 100 Milliarden ¥ (Yen) dotierte Projekt insofern, da es ein signifikantes Beispiel für die engagierte japanische Subventionspolitik darstellt. Mit diesem „vorrangigsten nationalen Grundlagenprojekt“ (MITI-Definition) steuerte das MITI den gesamten vorwirtschaftlichen KI-Wettbewerb der achtziger Jahre auf nahezu globalem Niveau.

Ebenso wie in Japan und den Vereinigten Staaten insbesondere, müssen auch in Europa die Institutionen der Lehre und der Forschung eine exponierte Rolle in der Markteinführung von KI-Produkten übernehmen, da ein korrespondierendes staatliches Engagement, insbesondere auch vor dem Hintergrund der GATT-Vereinbarungen, nicht zu erwarten ist.

Einer seriösen Marketing-Studie der Firma Brainware GmbH zufolge, soll allein in der Bundesrepublik ein Marktvolumen von ca. 260 Millionen DM erreicht worden sein. Nach der Brainware-Markterhebung läßt sich eine Marktverdoppelung turnusmäßig für alle zwei Jahre prognostizieren.

Nicht umsonst ist der überwiegende Teil der heute international erfolgreichen Programme nicht im Auftrag etablierter Anbieter erstellt worden, sondern sie stellen vielmehr die erfolgreiche Kommerzialisierung von Produkten dar, die als Forschungsprojekte an vielen Universitäten realisiert wurden.

### 3 Informatik Grundlagen

Der Terminus „Wissensbasiertes System“ ist ein Kollektivum für diverse Teilgebiete der KI.

Aufgrund der ausgeprägt interdisziplinären Struktur der KI-Forschung werden verschiedene Disziplinen unter diesem Begriff in der Fachliteratur zusammengefaßt. Die gegenwärtig relevanten Bereiche setzen sich zusammen aus:

gebräuchliche Übersetzung	englische Originalform
Automatischer Theorembeweiser	automatic theorem prover
Natürlichsprachliche Systeme	natural language system
Bildverarbeitende Systeme	pattern recognition (visual systems, Synonym)
Robotik	robotics
Expertensysteme (Wissensbasierte Systeme, Synonym)	expert (based) systems

Die Theorem-Beweiser befinden sich noch in der Grundlagenforschung, dennoch haben sie sich durch ihren engen Bezug zur mathematischen Logik bereits als eigenständige Disziplin einrichten können. Eine echte Kommerzialisierung ist jedoch noch nicht zu erwarten.

Die Natürlichsprachlichen Systeme bilden in Verbindung mit (bereits existierenden) Produkten zur Spracherkennung und -verarbeitung einen profitablen Bereich innerhalb der KI-Forschung und bilden ein potentielles Einsatzgebiet für Multi-Media-Plattformen.

Die KI-gesteuerte Computerlinguistik wird nicht mehr als reine Exploration des menschlichen Sprachvermögens betrachtet, sondern bietet insbesondere in der technischen Dokumentation diverse Kommerzialisierungsformen<sup>3</sup>.

Bildverarbeitende Systeme stellen mit ihrer Technologie den zweifelhaften *state of the art* der zeitgemäßen Militärtechnik dar. Diese sogenannten *smart weapons* haben ihre Effizienz im Golfkrieg in furioser Weise demonstriert ...

Die Robotik mit ihrer Fähigkeit zur sensorischen Wahrnehmung ist schon seit Jahren nicht mehr fortzudenken von den Taktstraßen maschineller Fabrikation, wo sie in dem ihr gesetzten Rahmen motorische Fähigkeiten des Menschen ersetzt.

Expertensysteme sind Programme, die nicht strukturiertes Expertenwissen mit der Fähigkeit eigener Schlußfolgerungen verbinden. Originär sind XPS-Programme, die dem Faktenwissen von Fachleuten nachempfunden wurden, wobei Fakten, bzw. Regelwissen, Heuristiken und sog. unscharfes Wissen vom System gewonnen werden.

Expertensysteme charakterisieren sich durch ihre Fähigkeit, komplexes Wissen eines Experten anderen Systemen oder (menschlichen) Endbenutzern zu ihren spezifischer Verfügung zu stellen. Im wissenschaftlichen Bereich, wie der Forschung, fungieren XPS weniger als Träger von bereits allgemein bekanntem Wissen, son-

2 Bullinger, H.-J.: Zur Informatisierung der Arbeit. In: FhG Berichte 1 (1986) S. 7-12.

3 Roesner, D.: Computerlinguistik auf dem Wege in die Praxis. In: Informationstechnik und technische Informatik (1994) S. 48-58.

dem abstrahieren das ihnen bekannte Faktenwissen auf neue Situationen. Die so selbständig von der inference engine gewonnenen Schlußfolgerungen, dem informatischen Zentralstück eines XPS, werden dann von den menschlichen Experten weiter ausgewertet.

Im gängigen Gebrauch wird diese inference engine exakt von der Wissensbasis (knowledge base) unterschieden. Hierzu gehören meist diverse Hilfsprogramme und Benutzerschnittstellen, welche die Entwicklung eines XPS spürbar vereinfachen im Vergleich mit der Verwendung einer Programmiersprache.

Diese Fähigkeiten grenzen die XPS aus dem Bereich der konventionellen Datenverarbeitung aus, denn die wichtigsten Charakteristika, wie Deklaration, Traceing, Heuristiken und Flexibilität, sind der EDV in der vorliegenden Form unbekannt.

Beim Programmieren eines klassischen Programmes muß mittels Algorithmen vorgegeben werden, wie ein Problem gelöst wird; das XPS verlangt nur nach der Angabe, welches Problem gelöst werden soll. Datenbanken setzen die Eingabe von strukturierten Daten (Zeichen, Zahlen) voraus. Damit sind sie im Gegensatz zu den disparitativen Formen der KI-Wissensrepräsentation streng numerisch formalisiert.

Konträr zur herkömmlichen EDV ist ein XPS fähig, heuristisches Wissen zu verarbeiten. Vages, nicht näher eingrenzbare Wissen wird gespeichert und ausgewertet.

Besonders benutzerfreundlich ist die flexible Struktur des XPS durch die Trennung von Wissensbasis, Datenbasis und inference engine. Dadurch können stets neue Regeln implementiert werden, ohne das die Notwendigkeit besteht, aufwendige Änderungen an der Programmstruktur vorzunehmen.

Der Benutzer kommuniziert mit dem XPS über eine der Benutzerschnittstellen, dem sogenannten *user interface*. Diese Schnittstelle besitzt zwei Funktionen, die für alle genuinen Expertensysteme gültig sind und diese gleichzeitig charakterisieren.

Die Methodik der Wissensakquisition läßt sich in drei unabhängige Blöcke unterteilen:

1. Befragung und/oder Beobachtung des Experten durch den knowledge engineer;
2. (semi)-automatischer Erwerb, welcher nur noch eine input-Leistung durch nicht weiter qualifizierte Kräfte vorsieht;
3. der vollständig automatisierte Wissenserwerb, welcher einen Prozeß ohne menschliche Beteiligung darstellt (beachte hierzu besonders das Jeng Projekt sowie das Compulex-Verfahren).

Bei der Differenzierung ist der Umstand von besonderer Bedeutung, daß eine neutrale Wissenserhebung im erstgenannten Verfahren nicht möglich ist, da das betroffene Fachwissen vom KE manipuliert (Transformierung in die vorgegebene Regelsprache) werden muß.

#### 4 Wissensrepräsentation

Das wohl gravierendste Problem der Wissensrepräsentation liegt in der Beziehung zwischen den unterschiedlichen Arten, wie ein Problem formuliert werden kann und wie effizient seine maschinelle Problemlösung dann sein wird. Das Verständnis für diese Problematik ist die

Grundvoraussetzung, um die dem betreffenden Problem adäquaten Prozeduren entwickeln zu können.

Das komplexe menschliche Wissen entzieht sich wegen seines vielschichtigen Aufbaus einer allgemeingültigen Definition, wird in der Informatik jedoch zumeist auf die Begriffe *Daten* und *Informationen* subsumiert. Werden interaktive Beziehungen zwischen diesen beiden Gruppen hergestellt, so entsteht „Wissen“, zumindest in dieser Kausalität.

Die Wissensrepräsentation bezeichnet die Beschreibung dieses Wissens in einer manipulierten Form, um sie in einem System speichern und verarbeiten zu können. Das vom *knowledge engineer* rekonstruierte menschliche Fachwissen wird in eine symbolische oder numerische Computersprache eingebunden.

Bei der imperativen numerischen Sprache muß der Lösungsweg vom Programmierer als Prozedur vorgegeben werden, der dann befehlsorientiert abgearbeitet wird.

Diese Prozeduralen Programmiersprachen unterscheiden sich auch in der Wissensrepräsentation von den regelbasierten Sprachen der KI.

In einer algorithmischen Sprache der 3. Generation wird das Wissen durch eine festgefügte Programmstruktur mit ihren Routinen und Subroutinen repräsentiert.

In den Regelsprachen der 5. Generation hingegen wird das Wissen getrennt von der Programmstruktur in der Wissensbasis deklariert und verwaltet. Somit ist das Wissen demnach nicht wie in der herkömmlichen Form an den Programmcode gebunden. Das eigentliche Programm bei einem XPS ist in diesem Sinne die *inference engine*, da sie in ihrer Struktur nicht manipuliert werden kann.

Diese Form der Wissensrepräsentation gehört zu den zentralen Aspekten der KI-Forschung. Zur Problemlösung benötigt der Rechner ein Modell der realen Welt, das ihm in der Wissensrepräsentation geliefert wird. Das reale Wissen wird also in einer modellierten Form der Maschine zugänglich gemacht. Diverse Werkzeuge wurden entwickelt, um unter diesen Paradigmen eine Nachbildung des interessanten Domänenwissens vornehmen zu können.

Die erfolgreichsten Methoden finden sich in der symbol-, respektive objektorientierten Wissensrepräsentation, der Prädikantenlogik, den Semantischen Netzen, den Produktionsregeln und Frames. In der modernen Anwendung finden sich heute nur noch stark vereinzelt XPS, die lediglich auf eine dieser Repräsentationsformen beruhen. Der Einsatz von sogenannten hybriden Werkzeugen, die über mehrere Darstellungsformen verfügen und über interne Schnittstellen miteinander kommunizieren, hat sich als besonders effiziente, weil wirtschaftliche Lösung bewährt. Ein markantes Beispiel für ein hybrides System stellt die Sprache PROLOG dar, mit der Wissen in Frames, Produktionsregeln und Semantischen Netzen eingebunden werden kann.

#### 5 Logik

In der mathematischen Logik wird eine Sprache formalisiert, in der Aussagen ausgedrückt und zusammen mit anderen Aussagen zu Komplexen verbunden werden können.

Die Logik ist somit das Forschungsgebiet, welches sich

mit dem Wahrheitsgehalt von Aussagen in Ja/Nein-Determinanten beschäftigt. Andererseits ist die Logik in der KI aber auch eine Sprache, mit der aus expliziten Meldungen implizite Informationen gewonnen werden können.

Die KI differenziert die Logik hauptsächlich in drei Sektoren, in die der Aussagenlogik, der Prädikantenlogik und der Fuzzylogik. In dieser Reihenfolge nimmt die Fuzzylogik jedoch eine Sonderposition ein.

## 6 Aussagenlogik

Die Aussagenlogik ist die Darstellung von Aussagen, in denen Variablen, wie X und Y stellvertretend für Aussagen wie „die RAK sind ein Kodex“ oder „der Kodex ist veraltet“ stehen. Aussagen dieser Art werden als Theoreme bezeichnet, für die folgende Charakteristika gelten:

- Vollständigkeit,
- Widerspruchsfreiheit und
- Ableitbarkeit.

In der Wissensrepräsentation dominieren zwar die Frametechniken und die Produktionsregeln, darüber hinaus haben sich jedoch auch die Techniken der mit Booleschen Operanten arbeitenden Aussagenlogik durchgesetzt. Aus der Aussagenlogik hat sich die mit ihr stark verwandte Prädikantenlogik entwickelt.

## 7 Prädikantenlogik

Die Prädikantenlogik ist eine zweiwertige Logik, die aus den Aussagen „wahr“ und „falsch“ besteht.

Ein Prädikat stellt dabei die Eigenschaft der Argumente oder deren Beziehungen untereinander dar.

LISP bedient sich hauptsächlich folgender Prädikate:

EQUAL    GREATERP    LESS  
ZEROP    ATOM    NUMBERP

EQUAL prüft, ob mindestens zwei Argumente identisch sind; GREATERP prüft, welches der von EQUAL geprüften Argumente das größte ist; LESS prüft, welches der von EQUAL geprüften Argumente das kleinste ist; ZEROP prüft, ob ein Argument den numerischen Wert von Null erreicht; ATOM prüft, ob ein Argument die kleinstmögliche Größe besitzt; NUMBERP prüft, ob ein Argument ein numerischer Wert ist.

In vielen Bereichen jedoch reichen die Techniken der Aussagen- und Prädikantenlogik nicht aus, um Wissen in seiner weiteren Komplexität zu repräsentieren; hier müssen andere Methoden der KI berücksichtigt werden. Besondere Relevanz erlangt diese Logikform durch ihre Implementation in PROLOG, korrespondierend mit dessen ökonomischem Erfolg.

Auch LISP oder Small Talk sind prädestiniert zur Darstellung der Prädikantenlogik; ältere Sprachen der konventionellen Programmierung, wie Pascal, eignen sich hingegen durch ihre prozeduralen Funktionen wesentlich weniger.

## 8 Fuzzy-Logik

Die Fuzzy-Logik, auch Unschärfe Logik genannt, grenzt sich durch eine eigenständige Fähigkeit von den oben beschriebenen klassischeren Logikarten ab:

*die Verwendung von unscharfer Logik.*

Fuzzy arbeitet nicht mit Probalistiken (Unsicherheitswerten), sondern kann unsicheres Wissen in eine symbolische Form bringen. Mit dieser Fähigkeit ist die Fuzzylogik eher als andere Logikformen geeignet, kognitives Verhalten zu simulieren.

Fuzzy ordnet einem Ausdruck keine Ja/Nein-Wahrheitswerte zu, sondern gebraucht eine beliebige Anzahl von Werten. Es entsteht eine Skala, die als Endpunkte A und B wohl wieder die bekannten Wahrheitswerte besitzt, dazwischen jedoch mit Unsicherheitsfaktoren rechnet. So können Begriffe semantisch klassifiziert und interpretiert werden, die weder A noch B entsprechen; unscharfes Wissen wird repräsentiert.

Fuzzy stellt danach eine Generalisierung jenes KI-Axioms dar, nach dem ein Element nur partiell zu einer Gruppe gehören kann. Mit dieser Aussage befindet sich die Fuzzy-Logik in konträrer Beziehung zur Booleschen Logik, die eine derartige Möglichkeit, zumindest in diesem Kontext, nicht vorsieht.

Erst durch die Implementierung dieser Form der Logik wird ein XPS befähigt, die Denkweise eines Experten wirklich effizient zu simulieren.

## 9 KI-Sprachen

In den vergangenen Jahrzehnten sind eine Vielzahl von Computerdialekten entwickelt worden, eine babylonische Sprachverwirrung war die Folge, unter der die KI jedoch weniger leidet.

Dieses Spektrum wird, zumindest auf internationalem Terrain, lediglich von zwei Sprachen abgedeckt.

Diese beiden Programmiersprachen der KI sind das Ergebnis eines evolutionären Prozesses, der bereits in den vierziger Jahren begann. In diesem Sinne wird die chronologische Aufeinanderfolge moderner und häufig rasch wieder veraltender Sprachen in verschiedene Generationen eingeteilt.

Dieser Staffelung nach gehören die aktuellen Sprachen bereits der 5. Generation an.

Da die reine Maschinensprache nicht als Programmiersprache anerkannt werden kann, beginnt das Zeitalter der Computerprogrammierung mit dem ersten Computerdialekt Assembler, obwohl er noch starke strukturelle Bezüge zur Maschinensprache aufweist.

Mit den ersten der sogenannten „höheren Sprachen“ wurde bereits zu Beginn der sechziger Jahre erfolgreich gearbeitet; Fortran war die erste Sprache dieser mittlerweile vierten Generation. Gleichzeitig begannen jedoch auch unter der Leitung von John McCarthy die Arbeiten an einer symbolorientierten Sprache: LISP. Dieser Dialekt wurde aufgrund seiner neuartigen Leistungsmerkmale einer nächsthöheren Generation zugeordnet, der fünften. Im Gegensatz zu Fortran war intendiert, mit LISP keine Zahlen mittels (mathematischer) Prozeduren zu formen, sondern vergleichbare Muster zu suchen sowie zu identifizieren.

Aufgrund der mit LISP gesammelten Erfahrungswerte konnte sich diese Sprache zu einem internationalen Standard entwickeln. Die ersten und bedeutenden Expertensysteme wie Mycin, Dendral oder Prospector wurden folgerichtigerweise ausschließlich in LISP geschrieben.

Erst Anfang der siebziger Jahre wurde in Europa eine weitere KI-Sprache geschrieben, die sich auf dem inter-

nationalen Markt neben LISP durchsetzen konnte: PROLOG. Diese Sprache wurde von Professor Colmerauer an der Universität Marseille entwickelt. Anders als LISP, zu dem es gravierende strukturelle Unterschiede gibt, basiert PROLOG auf der Prädikantenlogik. Insbesondere in Europa und Japan konnte sich PROLOG noch vor LISP etablieren. Ein essentieller Grund für den Erfolg PROLOGs ist in den wesentlich geringeren Hardware-Anforderungen zu sehen. Während PROLOG ohne Einschränkungen auf Microcomputer läuft, die auf Intel (resp. AMD oder Cyrix), Motorola oder Risc Prozessoren basieren, mußte bisher für eine LISP-Applikation im generellen eine eigene Hardware, die sogenannte LISP-Machine, unter hohem Aufwand gestellt werden. Gegenwärtig wird diese Problemsituation gemildert durch die Einführung einer Plattform unabhängigen Version, die ohne Einschränkungen auch auf den IBM kompatiblen Rechnern eingesetzt werden kann: Common Lisp. Korrespondierend mit der technologischen Weiterentwicklung auf dem Mikroprozessormarkt, insbesondere mit der Verfügbarkeit der 32 Bit Rechner der Intel xxx-86 Baureihe, bzw. der 64 Bit Pentium Generation und dem Preisverfall der Risc-Prozessoren, wird dieses Manko in absehbarer Zeit aufgehoben sein.

## 10 Shells

Generell gib es zwei grundsätzlich verschiedene Möglichkeiten, mit einem Expertensystem zu arbeiten:

1. die Erstellung eines Systems mit einer geeigneten KI-Sprache; vorzugsweise LISP oder PROLOG (bzw. Small Talk als besondere Hochsprache des *object oriented programming*),
2. die Inanspruchnahme eines bereits fertigen Systems, welches jedoch noch nicht mit Daten/Fakten gespeist (im KI-Jargon „feuern“ genannt) ist.

Solche leeren Entwicklungswerkzeuge stellen gegenwärtig auf dem kommerziellen Markt das Produkt der Künstlichen Intelligenzforschung *par excellence* dar. Der große Vorteil dieser Shells ist der Umstand, daß sie auch ohne große Programmierkenntnisse nutzbringend angewendet werden können, im Gegensatz zu den KI-Implementaten. Der KE nutzt die leichte Bedienbarkeit der Shell zum *rapid prototyping*, kleine Beta-Versionen des zu entwickelnden Expertensystems, um dann in Kooperation mit dem eigentlichen Experten das System zu errichten.

Zwischen den Shells werden wiederum zwei Versionen gehandhabt, die deduktiven und induktiven Shells. Bei deduktiven Shells wird das Wissen in Form von herkömmlichen WENN-DANN-Regeln verarbeitet. Die auf dem Markt ungleich erfolgreicherer induktiven Shells arbeiten mit Regeln, welche auch unscharfes Wissen verarbeiten.

Shells hybride Tools, welche in der Lage sind, so verschiedene Wissensformen, wie wissenschaftliche Formeln, graphische Darstellungen, *real language*, importierte Fremdformate etc. zu bearbeiten, erlangen eine stetig steigende Marktakzeptanz.

## 11 KI im bibliothekarischen Informations-Management

Wie erläutert, gehören KI-Produkte und XPS im Besonderen, in der EDV-Welt zu einem Marktsegment, dem

die größten Potentiale in Bezug auf Prosperität und Marktakzeptanz zu attestieren sind.

Waren die ersten Generationen der XPS vorwiegend noch der medizinischen und naturwissenschaftlichen Forschung vorbehalten, so wurden in der zweiten Generation Lösungen hauptsächlich für die Wirtschaft erstellt. Erst zu Beginn der jetzigen Dekade entwickelte sich ein Trend zu allgemeineren Branchenlösungen hin. Mit dem zwischenzeitlich erreichten *know how* war es nun möglich, in den unterschiedlichsten Bereichen des öffentlichen Lebens Expertenwissen zu synthetisieren. Besonders in Berufszweigen, die akademisches Fachwissen mit praktischer Anwendung kombinieren, bietet sich der Einsatz von XPS an. Der gesamte Bereich des Informationsmanagements stellt somit ein konsequentes Anwendungsfeld der KI dar.

XPS im Bibliotheks- und Dokumentationswesen bieten die potentielle Möglichkeit, die Qualität in der Informationsversorgung zu angemessenen Kosten oder sogar kostenneutral zu maximieren. Zur Qualitätsverbesserung tragen diverse Komponenten bei:

1. schnellere Bereitstellung von aktuellem Wissen,
2. zuverlässige Aufbereitung einzelner Informationen zur Expertise im Auskunftsdienst,
3. Herstellung eines flächendeckend gleichmäßigen Informationsangebotes durch den Abbau regionaler Unterschiede, indem zunehmend auch bibliothekarische Zweigstellen durch XPS im Auskunftsdienst aufgewertet werden.

Wissenschaftliche Bibliothekare/Dokumentare und Informationswissenschaftler können unter Involvierung dieser zukunftsorientierten Technologie ihren eigenen Tätigkeitsbereich bedeutend aufwerten. Ein *job enrichment* ist insofern zu erreichen, da z.B. die Rationalisierung von Regelarbeiten in der Sachkatalogisierung oder in der Inhaltserschließung zur Verbesserung der Arbeitsplatzqualität führt. Ein Gedanke, der zunehmend auch von Arbeitswissenschaftlern geteilt wird.

Das Interesse des bibliothekarischen Fachpublikums, soviel kann konstatiert werden, ist erfreulicherweise in den letzten Berichtszeiträumen der Fachpublikationen gewachsen, wenn die sich häufenden Artikel zum Thema überhaupt einen Indikator darstellen können. Eine entsprechende Untersuchung der angelsächsischen Fachliteratur wurde von Cynthia C. Hsieh und Wendy Hall bereits 1989 veröffentlicht.

Bei der inhaltlichen Sichtung der vorliegenden Artikel und Referate dieser Untersuchung wird evident, daß besonders zwei Disziplinen der Bibliotheks- und Informationswissenschaft besonders intensiv mit dem Aufbau von XPS beschäftigt sind. Im Bibliothekswesen wird engagiert nach Techniken geforscht, die einen XPS Einsatz im Bereich der Formal- und Inhaltserschließung ermöglichen. Das Informationswesen stellt hingegen bereits heute leistungsfähige Lösungen für online-Dienste bereit.

Nach einer Studie der University of Loughborough/UK lassen sich im KI-Bereich sieben Einsatzgebiete lokalisieren, in denen die Fachkenntnisse von Bibliothekaren diese für eine Tätigkeit prädestinieren:

1. Bibliotheks- und Informationswissenschaftler als Spezialisten in der Aufbereitung von Informationen. In dieser Position stehen sie mit ihrem Expertenwissen während der Wissensakquisition dem KE zur Verfügung.

2. Der Einsatz eines Informationsspezialisten selbst als KE. Er verbindet sein bibliothekarisches Fachwissen z.B. aus den Bereichen der Inhaltserschließung (Mary Dykstra) oder der Formalerschließung (Ling Hwey Jeng) mit einer ausreichenden EDV-Kenntnis, um unabhängig von externen Fachkräften XPS zu erstellen.
3. Erfahrungen aus der Benutzerforschung,
4. Bibliothekare in der Rolle eines Information Brokers,
5. Pflege und Ausbau von bereits existierenden XPS,
6. Bibliothekare als Vermittler zwischen EDV und Informationswissenschaft und
7. Entwicklung eigener XPS für bestimmte Benutzergruppen.

In den folgenden Abschnitten wird dargelegt, inwieweit die zitierte Studie der University of Loughborough sich an den Realitäten des Informationsmanagements und der bibliothekarischen Tätigkeit orientiert.

## 12 Bibliothekare in der Wissensakquisition

Wie bereits ausgeführt wurde, ist die Leistungsfähigkeit eines XPS nicht nur von der Qualität der Software abhängig, sondern in besonderer Weise auch von den Fähigkeiten des KE, Sachverhalte in Regeln oder Frames umzusetzen. Im weiten Bereich der Informationsvermittlung sollten aufgrund ihrer Qualifikationen Bibliothekare besonders für diese Aufgabe prädestiniert sein. Anne Morris registriert bei ihrer Untersuchung von 75 nordamerikanischen Library and Information Schools, daß bei der Erstellung von Systemen nur selten auf bibliothekarisches Fachpersonal zurückgegriffen wird. Desweiteren stellt sie einen Kriterienkatalog auf, mit denen der Vorlesungsplan die KI-Studien an die Bedürfnisse der Wirtschaft anpassen kann. [Morris, 1991]

Während Bibliothekare in ihren Bibliotheksinformationssystemen (BIS) eine allgemein anerkannte Tätigkeit verrichten, kann festgestellt werden, daß die interessierten Träger Bibliothekare im generellen nicht konsultieren. Einschränkend muß vermerkt werden, daß diese Situation in Nordamerika eine grundlegend andere ist. Der *scientific librarian* stellt dort einen häufig konsultierten Ansprechpartner bei der Systemerstellung dar. Ein entscheidender Grund für dieses Mißverhältnis dürfte in der nicht ausreichenden Qualifikation hiesiger Fachvertreter zu finden sein.

Wohl ist die Erforschung von KI-Methoden noch nicht direkt in die *curricula* der bibliothekswissenschaftlichen Fakultäten aufgenommen, doch genügt bereits ein Blick in die entsprechenden Fachbibliographien um festzustellen, daß diese Lehrinrichtungen einen nicht unwesentlichen Anteil an der KI Forschung besitzen [siehe auch Hsieh, 1989].

Fast schon traditionsmäßig werden an der School of Library and Information Studies in Projekten Lösungen erarbeitet, die sehr schnell externe Bedeutung erlangen. Ein frühes Beispiel dieser bemerkenswerten Erfolgsserie stellt „Access to film information“ dar. Ein Indexierungs- und Retrieval-System für AV-Medien, welches auf der Grundlage von PRECIS<sup>4</sup> operiert. Unmittelbar nach der Publizierung wurde dieses ursprüngliche Studienprojekt von dem National Film Board of Canada übernommen und gehört heute weltweit zu einem probaten Instrument vieler AV-Archive. Diese fundierten wissen-

schaftlichen Studien haben auch im KI-Bereich ihre Wirkung nicht verfehlen können. Analog zum National Film Board of Canada hat eine andere Regierungsbehörde, das National Research Council of Canada, einen Entwicklungsauftrag an die School of Library and Information Studies abgegeben. An dieser Bibliotheksschule wurde kürzlich ein XPS erstellt, um als Interface zwischen Benutzer und dem National Building Code (NBC) zu fungieren. Wieder bildet PRECIS die Grundlage, nach der in diesem XPS eine syntaktische Analyse vorgenommen werden kann. Desweiteren beinhaltet dieses XPS eine Implementierung aller Suchbegriffe des NBC. Darüber hinaus wurde ein Modul erarbeitet, um die Suchbegriffe in eine PRECIS-gemäße Syntax zu transliterieren, verbunden mit einem umfangreichen Thesaurus. Das XPS verfügt ferner über einen Hypertexteditor, beherrscht Boolesche Suchstrategien und besitzt diverse Im- und Exportmöglichkeiten. Obwohl auch dieses Projekt unter der Ägide von Frau Professor Dykstra entstanden ist, sind im Wintersemester 1991/92 mehrere KI-Projekte von anderen Fakultätsmitgliedern angeboten worden. Im Auftrag des Canadian Council of Archives entwickelt ein Professor mit seinen Studenten ein XPS mit Hypertext zur Anwendung eines Regelwerks für archivalische Materialien. Ein weiterer Dozent betreibt Grundlagenforschung in der Entwicklung und Einbindung von Probalistiken in Hypertexte. Ein drittes Fakultätsmitglied führt Projekte durch, die alle KI-Komponente aufweisen. Bei der Erläuterung dieser Aktivitäten muß darauf hingewiesen werden, daß die School of Library and Information Studies mit ihren Unternehmungen durchaus repräsentativ ist, andere (nordamerikanische) bibliothekarischen Lehr- und Forschungseinrichtungen dürften ähnliche Aktivitäten vorweisen können.

Dieses Bild auf das deutsche Bibliothekswesen projiziert, kann nur einen bedenklichen Ausblick auf die Zukunft des eigenen Berufsstandes ergeben. Hier muß ein radikales Umdenken einsetzen, was den Einsatz zukunftsorientierter Technologien betrifft. Dieser Sachverhalt ist einer der hauptsächlichen Ursachen für die mangelnde Konsultation bibliothekarischer Kräfte von externen KE, die so nicht vermögen, die besonderen Fähigkeiten von Bibliothekaren und Dokumentaren adäquat einzuschätzen.

Im Rahmen einer von der Loughborough University erhobenen nicht-repräsentativen Umfrage wurde wiederholt ausgesagt, daß Experten, die auch auf konventionelle BIS nicht zugreifen, einer Kooperation mit Bibliothekaren ausgesprochen skeptisch gegenüberstehen, ungeachtet welche KI-Kompetenzen die Bibliothekare besitzen. Ein Mißstand, welcher jedoch durch ein zeitgemäßes Marketing abgebaut werden könnte, da die gleichen Interviewpartner nach Ablauf einer befristeten Zusammenarbeit die bibliothekarischen Dienste als wertvolle Ergänzung der eigenen Tätigkeit betrachteten.

## 13 Bibliothekare als *knowledge engineer*

Bibliothekare in der Rolle des KE stellen für das Berufsbild eine völlig neuartige Herausforderung dar, der je-

4 PREserved Context Index System.

doch durch Flexibilität und Technologieverständnis begegnet werden kann.

Den diversen Anforderungen, die an einen KE gestellt werden müssen, können weitgehend Bibliotheks- und Informationsspezialisten gerecht werden. Ihre Kenntnisse im online-Dienst, in der Erschließung von Sachzusammenhängen, die Fähigkeit, diffuse Problemstellungen zu konkretisieren und verbalisieren, verbunden mit der Fertigkeit, ein Benutzergespräch zu führen, bilden die beste Voraussetzungen für die Tätigkeit eines KE in der Funktion eines externen Beraters.

Im weitesten Sinne kann sogar festgestellt werden, daß die Tätigkeit eines KE nichts anderes ist, als die traditionelle bibliothekarische Aufgabe „Vermittlung von Wissen“ auf einer anderen Ebene.

Aus diesem Grunde sollte der Einsatz von Bibliothekaren besonders effizient sein, denn die Fähigkeiten eines KE besitzt der Bibliothekar bereits aufgrund seiner wissenschaftlich orientierten Ausbildung. Eine Fachkraft aus anderen Bereichen müßte sich dieses notwendige Fachwissen erst in einem eminent aufwendigen Lernverfahren aneignen. Anders als konventionelle Systemanalytiker, benötigen KE ein fundiertes Allgemeinwissen und eine gewisse Kommunikationsfähigkeit, um sich mit dem Experten auf seinem Niveau verständigen zu können. Kenntnisse der Indexierung, Katalogisierung, Klassifizierung, Erstellung von Thesauri, etc. bilden die Grundlage der Arbeit *im knowledge engineering*. Gleichzeitig stellen sie auch die Basis der hiesigen Bibliothekarsausbildung dar; somit verfügen Bibliothekare über das Expertenwissen eines KE, verbunden mit Fachkenntnisse des elektronischen Informations-Managements.

In summa bilden dies ausreichende Qualifikationen für wissenschaftliche Bibliothekare oder Dokumentare in diesem innovativen Tätigkeitsbereich.

#### 14 Konzeptionalisierung von Benutzerstudien

Eine der gegen XPS häufig, aber wenig fundiert artikulierten Kritikpunkte ist im allgemeinen eine mangelnde Benutzerakzeptanz einmal installierter Systeme. Auch wenn diese über Jahre hinweg unter hohem Kostenaufwand erstellt worden sind, kann ohne begleitende Rahmenmaßnahmen nicht erwartet werden, daß diese „künstlichen Kollegen“ ad hoc auf breite Zustimmung treffen.

Anders als bei herkömmlichen Programmen sind KI-Ingenieure, zumindest bei der Systempflege, auf ein gewisses *feed back* der Benutzer angewiesen. Ist dies nicht gegeben, kann dieses Manko gravierende Folgen für die Zuverlässigkeit des XPS haben<sup>5</sup>.

Bei der Wissensakquisition besteht die Gefahr, daß sich der KE weitgehend der Fachterminologie des Experten annähert; ein Umstand, der mit besonders hoher Frequenz bei ausgesprochen natürlich-sprachigen Systemen auftritt. Die Folge ist, das dem Benutzer die ihm oft unbekanntere Fachsprache als Hemmschwelle erscheint. Aufgrund diverser Benutzerbefragungen konnte die mangelnde Akzeptanz auf diese ausgesprochene Fehlleistung zurückgeführt werden. Der erfahrene Bibliothekar wird diesen Fehler zu vermeiden wissen. Ein entgegengesetzter Weg kann beschritten werden, wenn die bibliothekarische Fachkraft eine komplexe Fachtermino-

logie (z.B. eine Retrievalsprache) durch einen Thesaurus dem Benutzer erschließt.

In nicht-naturwissenschaftlichen Branchen werden XPS überwiegend zur Informationsaufbereitung eingesetzt. Installiert werden diese Systeme aber von Fachkräften aus der Fachrichtung Informatik/Systemanalyse, die keine detaillierteren Kenntnisse der Informationswissenschaft aufweisen können. Hier eröffnet sich schon in der Entwicklungsphase, spätestens aber bei der Wissensakquisition, für den Bibliothekar ein breites Einsatzfeld. Er kann sein Fachwissen dem KE zur Verfügung stellen, so daß einerseits eine möglichst effiziente Struktur der Wissensbasis gewährleistet bleibt, andererseits das System aber auf die sehr speziellen Bedürfnisse der Benutzer konfiguriert wird. Bedürfnisse, die oftmals nur durch eine in der praktischen Informationsvermittlung gewonnene Erfahrung erkannt werden können. Externe Fachkräfte werden hier Defizite aufweisen. Ebenfalls verweisend auf die gemeinsame Grundlage ist die Tatsache, daß in sehr vielen XPS leistungsfähige Thesauri implementiert sind, um dem Benutzer bei der Formulierung seines Problems entgegenkommen zu können.

So sind bibliothekarische Benutzerstudien auch für die Tätigkeit eines KE von Bedeutung. Obwohl beide Vertreter unterschiedliche Ausgangspositionen einnehmen, können sie sich mit ihren spezifischen Fähigkeiten zu beiderseitigem Nutzen ergänzen.

#### 15 Entwicklung eigener Systeme

Der zentrale Aspekt bleibt bei diesen Betrachtungen jedoch die Entwicklung von XPS für die vielschichtigen bibliothekarischen Anwendungsfelder.

Im Auftrag der British Library leitete Anne Morris eine Enquete, um die besonderen Qualifikationen für die XPS-Erstellung durch Bibliothekare beschreiben zu können.

Morris führte zu diesem Zwecke auch eine Umfrage unter MIS-Vertretern durch, aus der eine bezeichnende Antwort gewonnen werden konnte, die, ohne generalisieren zu wollen, durchaus als charakteristisch für die Einstellung dem bibliothekarischen Berufsstand gegenüber gelten kann. Auf die Frage nach der persönlichen Meinung über KI-Aktivitäten von Bibliothekaren antwortete ein befragter Informationsmanager:

*„I believe that LIS<sup>6</sup> researchers are only good at talking about expert systems but do not have much, if any, practical experience of actually using them“<sup>7</sup>.*

Aus dieser verbreiteten Meinung (die leider nicht vollkommen ohne Substanz ist ...) resultiert auch für die Fachwissenschaft die unabdingbare Notwendigkeit, eigene KI-Projekte zu starten.

Mit selbständigen Aktivitäten können dabei die BIS den Normen und Erfordernissen einer Kommunikationsgesellschaft angepaßt werden. Unter diesen Auspizien

5 Oder wie ein betroffener Systemanalytiker es weitaus prägnanter zu formulieren wußte: *„if nobody really uses them, who to hell is to tell us where they are going wrong?“* ...

6 Library Information System.

7 Morris, Anne: Expert systems and their implications for information science. In: Brittain, M: Curriculum development in Information Science to meet the needs of the information industries in the 1990's. BLRDD Report (1987).

wurden bereits weltweit für die bibliothekarischen Belange diverse XPS entwickelt, wobei eine explizites Paradigma beobachtet werden kann: XPS in der Katalogisierung.

Wie aus späteren Abschnitten hervorgeht, können KI-Techniken in der Klassifikation nur einen nachgeordneten Rang einnehmen, entsprechend fällt deren Bedeutung aus.

Im folgenden soll der erreichte Entwicklungsstand analysiert und die sich daraus ergebenden Konsequenzen für die bibliothekarische Tätigkeit kritisch diskutiert werden.

## 16 XPS in der Katalogisierung

Durch die Kortzfleisch-Analyse des bibliothekarischen Arbeitsablaufs ist bereits in den sechziger Jahren ermittelt worden, daß die Formalkatalogisierung eine der personalintensivsten Prozesse bildet. Mit der Einführung der elektronischen Datenverarbeitung auf breiter Basis wurde auf vielfältige Weise versucht, diesen Arbeitsschritt zu rationalisieren. Konventionelle Programme, wie das Braunschweiger Allegro, konnten jedoch nicht die gewünschten Konsequenzen im Sinne einer Personalminimierung erzielen. Eine entgegengesetzte Entwicklung war zu beobachten, da die EDV-gestützte Katalogisierung eine zusätzliche Qualifizierung bedingte. Um eine solche Software handhaben zu können, sind fundierte Systemkenntnisse erforderlich, die auch nicht auf die vollständige Beherrschung des betreffenden Regelwerks verzichten.

Spätestens bei einer Personalbedarfsanalyse wird offenkundig, daß rationalisierende Effekte in der Katalogisierung durch konventionelle Systeme nur bedingt, wenn überhaupt, erreichbar sind. Hier eröffnet der Einsatz von KI-Technologien neue Perspektiven, nach denen die gewünschten Vorgaben nun realisiert werden können.

Wie eingangs erwähnt, stellt die regelbasierende Katalogisierung ein natürliches Einsatzfeld für Bibliotheken dar.

Clarke und Cronin unterscheiden generell zwei Formen, in denen XPS den Bedürfnissen der Katalogisierung angepaßt werden können:

1. XPS, die ein menschliches Interface benötigen; einen Operator also, der überwiegend die Wissensakquisition vollzieht.
2. XPS, die vollständig autonom den Vorgang der Katalogisierung nachvollziehen und die fertige Titelaufnahmen generieren können.<sup>8</sup> Die Definition dieses Modells schließt eine menschliche Interaktion *expressis verbis* aus.

Ein nach den Vorgaben des ersten Modells entwickelte XPS ist das MITINET/marc System, welches 1986 in den USA konstruiert wurde. MITINET/marc gibt in Dialogform dem Katalogisierer Instruktionen, wie eine Titelaufnahme nach den AACR 2<sup>9</sup> vorgenommen werden muß. Anschließend transformiert MITINET/marc die Aufnahme in das maschinenlesbare MARC-Format, womit dieses XPS auch als Subsystem des OCLC interessant wird. Das wesentliche Merkmal dieses mittlerweile probaten XPS liegt weniger in der Katalogisierung, denn hier wird immer noch ein gut ausgebildeter Katalogisie-

rer vorausgesetzt, sondern im *handling* des Marc-Formates.

Die Arbeit mit OPACs erfordert vom Bibliothekar nicht nur die Beherrschung seiner eigenen Disziplin, vielmehr werden auch detaillierte Kenntnisse im Bereich der elektronischen Datenverarbeitung impliziert. Qualifikationen, die insbesondere ältere Bibliothekare nicht immer aufweisen können. Um deren wertvollen Erfahrungsschatz in der Katalogisierung zu nutzen, konvertiert MITINET/marc die in der Datenbasis gewonnenen (bibliographischen) Informationen zur Katalogisierung in die MARC-Terminologie. MITINET/marc kann als eines der wenigen bereits in der praktischen Anwendung stehenden XPS für Bibliotheken gelten, da es das Fachwissen einer Domäne (Terminologie, Kode und Prozeduren des OCLC) über eine Software weniger erfahrenen Benutzern erschließt. Dieses XPS wird somit vornehmlich zur Konvertierung, nicht aber zur eigentlichen Titelaufnahme eingesetzt. Analog zu den konventionellen Softwarelösungen, muß der Benutzer über ausreichendes Expertenwissen verfügen, um in MITINET/marc die jeweils gültigen Kriterien für die Haupttitelaufnahme selektieren zu können.

Der Vorteil dieses Systems bleibt jedoch offenkundig; die Arbeit einer online gestützten Katalogisierung wird auch den Bibliothekaren ermöglicht, die vorher konventionelle Systeme aufgrund ungenügender EDV-Kenntnisse nicht handhaben konnten.

Ein in seiner Struktur weitaus ausgereifteres XPS wurde an der University of California/Los Angeles (UCLA) entwickelt: MAPPER. Dieses System wurde für die von der herkömmlichen Katalogisierung deviiierenden Aufnahme von Landkarten erstellt. Wiederum bildeten die AACR die Grundlage für die Regelbasis MAPPERs. Anders als die RAK-WB, bieten die cataloging rules im Kapitel 21 auch Regeln zur Katalogisierung von Landkarten und vergleichbaren Medien an. Dieses betreffende Kapitel 21 wurde vollständig von den Programmautoren in MAPPER implementiert, wodurch dieses System über eine bemerkenswerte Flexibilität verfügt.

In der Konsultation werden vom Benutzer in mehreren Arbeitsschritten Informationen über die physikalische Beschaffenheit des Objektes eingegeben. Diese Befragung wird vom System dann beendet, wenn der *inferrence engine* in der Datenbasis ausreichend Informationen zur Verfügung stehen, um als Schlußfolgerung einen Katalogeintrag generieren zu können. Besonders interessant wird MAPPER darüber hinaus noch über die Tracekomponente, die dem Benutzer erlaubt, den vom System vorgenommenen Eintrag auf seine Authentizität hin zu kontrollieren.

Anders als bei MAPPER oder MITINET/marc war auch das sogenannte „Projekt ESSCAPE“<sup>10</sup> als reine Grundlagenstudie gedacht. Diese Forschungsstudie wurde von Roland Hjerpe und Birgitta Olander an der schwedischen Linköping Universität entwickelt. Die Linköping

8 Clarke, Ann: Expert systems and library information work/Ann Clarke; Blaise Cronin. In: Journal of Librarianship, 15 (1989), S. 289.

9 Anglo American Cataloging Rules, second edition. Angelsächsisches Pendant zu den deutschen RAK.

10 Expert Systems For Simple Choice Of Access Points For Entries.

Universität verfügt zusätzlich zur bibliothekswissenschaftlichen Fakultät noch über eine Einrichtung, die gänzlich der Forschung in dem Bereich der Informations- und Kommunikationstechnologie für Bibliotheken gewidmet ist: das renommierte LIBLAB, dem auch Hjerpe und Olander angehören.

Eine essentielle Erfahrung, die das LIBLAB durch seine Arbeit an ESSCAPE gewinnen konnte, wurde von Hjerpe folgendermaßen zusammengefaßt:

... *the interpretation and judgement ... played a much larger role in the cataloging process than had previously been assumed and acknowledged that an expert cataloging system would, ideally, need to be able to identify and interpret a title page, categorize the contents, interpret relationships between names and between bibliographic entities, and make judgements in determining uniform headings*<sup>11</sup>.

Dieses Projekt stellt zwar auch eine Repräsentation der AACR dar, ist aber wesentlich leistungsfähiger. Eine graphische Benutzeroberfläche erlaubt es, die Titelaufnahme mittels Menüs durchzuführen. In diesen Menüs brauchen die entsprechenden bibliographischen Angaben lediglich per Mausklick aktiviert zu werden.

Diese Arbeit war ausschließlich als Forschungsprojekt intendiert, so daß es zu keiner praktischen Anwendung gekommen ist. Dennoch konnten von den Projektleitern wertvolle Schlußfolgerungen für den bibliothekarischen Alltagseinsatz von regelbasierten Systemen gewonnen werden:

1. rules determining access points could be reinstated as production rules;
2. rules governing bibliographic description could be entered ... through the user of templates;
3. local rules ... could be accommodated;
4. menus and screen forms could be designed which would act as a user-friendly interface between a cataloguer and an expert system;
5. explanation facilities in terms of the systems attempt to decide which rules from AACR 2 apply could easily be incorporated<sup>12</sup>.

Nach Abschluß dieser gut dokumentierten Arbeit und deren Publikation konnte dieses Projekt insofern noch eine Bedeutung a posteriori erlangen, da es die Basis für die Arbeiten von Ling Hwey Jeng, einem Doktoranden der University of Texas/Austin, werden sollte.

Wie zu Beginn dieses Kapitels ausgeführt wurde, konkurrieren zwei konzeptionelle Vorstellungen miteinander, wie XPS in der Katalogisierung eingesetzt werden können. Sind die bisher erörterten XPS dem semi-autonomen Modell zuzuordnen, konzentriert sich die gegenwärtige Forschung hauptsächlich auf XPS, die ohne Involvierung menschlicher Experten fungieren.

Die grundlegenden Probleme bei der autonomen Titelerfassung und deren zufriedenstellende Lösungen mittels KI-Strategien wurden von Ling Hwey Jeng in einer aufwendigen Studie in den USA erarbeitet. Jeng selber faßt seine Arbeit wie folgt zusammen:

*The data from the title page are input into the computer and interpreted in terms of its physical appearance and its contents. In response, the system returns a decision as to the content of the title proper and its correct format*<sup>13</sup>.

Jeng stellt die Hypothese auf, daß ein auf der Titelseite aufgeführtes bibliographisches Element sich aus mehreren erkennbaren Merkmalen zusammensetzt und sich

gleichzeitig in seinen Ausprägungen von anderen Elementen unterscheiden läßt.

Die sorgfältige Analyse der formalen Merkmale eines bibliographischen Elements ist von besonderer Bedeutung für das spätere *knowledge engineering*. Mittels einfacher Produktionsregeln kann ein Element von der *inference engine* interpretiert werden. Zur Beweisführung hat Jeng ein Verfahren entwickelt, welches ihm ermöglicht, seine Schlußfolgerungen zu untermauern.

Für eine beispielgebende Versuchsserie mit dem Zweck einer automatischen Bibliographierung wurden aus dem Bestand der Perry Castaneda Bibliothek der University of Texas 203 Monographien ausgewählt.

Eine Einschränkung der Titelauswahl wurde insofern vorgenommen, da nur englischsprachige Monographien aus dem 20. Jahrhundert berücksichtigt werden konnten.

Klassifiziert waren die Titel nach der Library of Congress Classification (LCC) und entstammten deren Hauptklassen Literatur, Soziologie und Geschichte.

Korrespondierend zu den RAK, stellen die AACR 2 fest, daß der Titelseite eines Buches die zentrale Bedeutung bei der Katalogisierung zukommt. Regel 2.0 B 1 der AACR 2:

„The chief source of information for printed monographs is the title page or, if there is no title page, the source from within the publication that is used as a substitute for it“

Diese Regel fixiert demnach die Aufnahme gänzlich auf die Titelseite, andere Quellen bleiben dabei nachgeordnet.

Für den Einsatz eines XPS muß darum die Titelseite eines zu katalogisierenden Werkes besonders aufbereitet werden. Aus diesem Grund wird die Titelseite „maschinenlesbar“ gemacht; Kategorien werden entwickelt, um den Titelseiten logische Strukturen zu unterlegen, welche von einem XPS verwaltet werden können.

R. Hagler und P. Simmons nehmen zu diesem Zweck einer maschinengestützten Titelaufnahme folgende Unterteilung der Buchtitel in drei unabhängige Kategorien vor:

1. Daten, die ein besonderes Dokument identifizieren, um sie von anderen unterscheiden zu können (z. B. Teile eines enthaltenen Werkes);
2. Daten, die eine Beziehung von mehreren Dokumenten untereinander beschreiben und
3. Daten, die zur Inhaltserschließung geeignet sein könnten.

Von dieser Schematisierung abstrahierend, kann eine weiterführende Untersuchung der Titelseite vorgenommen werden, um diese für die besonderen Ansprüche eines XPS aufbereiten zu können.

Diese Vorbereitung erfolgt in zwei Stufen. Zuerst muß die Titelseite digitalisiert (d. h. maschinenlesbar ge-

11 Hjerpe, Roland: Project ESSCAPE: expert systems for simple choice of access points for entries; applications of artificial intelligence in cataloging. In: International Federation of Library Association General Conference 51. Chicago 1985. S. 46 ff.

12 Davies, Roy: Towards an expert system for cataloguing: some experiments based on AACR 2 / Roy Davies; Brian James. In: Program 18 (1984). S. 284-285.

13 Jeng, Ling Hwey: An expert system for determining title proper in descriptive cataloging: a conceptual model/Ling Hwey Jeng. In: Cataloging & Classification Quarterly 7/2 (1986). S. 67 ff.

macht) werden. Ein Vorgang, der von jedem handelsüblichen Scanner geleistet werden kann.

Für diese Zwecke einer innovativen Mustererkennung, wie sie bereits von handelsüblichen OCR-Programmen geleistet werden, kommen ausschließlich Neuronale Netze in betracht. Neuronale Netze werden auf absehbarer Zeit die günstigste Lösung in dem hier beschriebenen Problemfeld bleiben, allein aufgrund ihrer besonderen Effizienz im Bereich der *pattern recognition*.

Erst in der zweiten Stufe kommen die Techniken der KI zur Anwendung: die Erkennung der eingescannten Muster.

Die Erkennung bibliographischer Bestandteile einer Titelseite erfolgt zuerst auf Basis der oben angegebenen „Drei-Stufen-Lösung“.

Bei näherer Betrachtung eines beliebigen aber typischen Titelbildes wird deutlich, daß bibliographische Daten oft in ihrer Form typographisch differenzieren. Der Titel z. B. wird häufig in einer größeren Schriftform wiedergegeben als der Name des Autors, der Hauptsachtitel wird oft von seinem Zusatz abgehoben etc.

Diese empirisch belegbaren Unterschiede müssen während des Prozesses der Systemerstellung in die Wissensbasis eingebunden werden.

Jeng stellt folglich Analogien in der Gestaltung der Titelseite fest, die generalisiert werden können, da sie stets wiederkehren:

„Thus we may say that the proper names in the middle of the page are the authors and not the publisher, even though we may not recognize them as personal names“.

Zusammenfassend kann gesagt werden, daß der erste Arbeitsschritt einer artifiziellen Titelaufnahme in der Erstellung eines leistungsfähigen Mustererkennungsverfahrens und einer Kategorisierung der Titelseite liegt.

Diese Kategorisierung ist, wie beschrieben, eine Aufteilung der Titelseite in einzelne Blöcke bibliographischer Daten, verbunden mit deren AACR 2-gemäßen Interpretation. Diese Zuordnung, die eigentliche Inferenz, kann auf zwei Arten erfolgen: durch einfache oder koordinierte Differenzierung der Blöcke.

Die einfache Differenzierung besteht aus den Merkmalen der jeweiligen Blöcke und ihrer Daten.

Dreizehn Charakteristika wurden definiert und in ihre Blöcke eingebunden:

01. die Position des Blockes auf der Titelseite
02. die Größe des Blockes auf der Titelseite
03. die Größe der Typologie
04. die Buchstaben (groß/klein/gemischt)
05. Typologie (Courier, Draft, Serif, etc.)
06. horizontale Stellung der Schrift
07. Konturenschärfe der Schrift
08. graphische Symbole (die Fähigkeit, Graphiken zu erkennen und ignorieren zu können)
09. eine zentrale Linie zur besonderen Hervorhebung einer Angabe (in der Grafik *ruler line* genannt)
10. der Endpunkt
11. die Raumgestaltung (der Zeilenabstand und die Größe der Typologie werden in einer Matrix berechnet und in eine Interaktion gesetzt)
12. vertikaler Stellung der Schrift
13. übergroße Blöcke (nur relevant, wenn mehr als eine halbe Seite beansprucht wird)

Die unterschiedlichen Merkmale setzen voraus, daß sie einzeln betrachtet werden können. Gleichzeitig bestäti-

gen sie, daß bibliographische Daten auf einer Titelseite unselbständig erscheinen können.

Diese Erscheinungsweise wird in die oben genannten Blöcke eingegliedert. Die unter Punkt 06 aufgelistete horizontale Stellung, verbunden mit einer typographischen Hervorhebung einer Eintragung auf der Titelseite, weist z.B. eben nicht auf eine Seriennummer, sondern mit größerer Wahrscheinlichkeit auf einen Sachtitel hin. Dabei wird der Schriftart eine weitere Bedeutung zugemessen, da mit ihrer Hilfe sich der Hauptsachtitel vom Zusatz zum Titel differenzieren läßt.

Mit diesem Wissen kann ein XPS bereits erste Schlussfolgerungen ziehen, wenngleich die Ergebnisse nicht jenen subtilen Grad erlangen können, wie bei der koordinierten Differenzierung.

Bei der weitergehenden koordinierten Differenzierung liegen hingegen Merkmale vor, welche die Beziehungen zwischen den Blöcken selbst analysieren.

Das Ziel dieser Segmentierung der Titelseite ist demgemäß die Identifizierung bibliographischer Merkmale der Blöcke.

Die ersten sieben Charakteristika sind in diesem Sinne nicht koordiniert, sondern einfach differenziert. Die Verbleibenden hingegen sind es.

Nach diesem Muster wurden die Daten der von den 203 Titelseiten gewonnenen Blöcke aufgenommen, wobei jeder Block eine bibliographische Angabe repräsentierte. Zwangsläufig ließen sich opportunerweise ausschließlich die Blöcke einer gemeinsamen Titelseite miteinander in Verbindung bringen.

In den wenigen Fällen, in denen ein Block mehrere bibliographische Angaben hatte, wurden alle Merkmale aufgenommen. So erhielten die Blöcke für jede ihrer Merkmale eine separate Aufnahme.

Im nächsten Arbeitsschritt wurden diese gewonnenen Daten nach der Häufigkeit ihres absoluten Vorkommens mit dem SAS-Statistikpaket von IBM analysiert. Die so erzielten Angaben bildeten dann die Grundlage für die in einem anderen Arbeitsprozeß stattfindende Bildung von Heuristiken.

Das Resultat dieser Datenanalyse soll die Lösung folgender Fragen sein, welche erst den Ansatz zur Formalschließung bilden.

1. Was sind bibliographische Elemente?
2. Wie sind diese Merkmale untereinander geordnet?
3. Was sind deren besondere Merkmale?
4. Wo auf der Titelseite erscheinen sie?

Eines der Hauptprobleme dieser Kategorisierung stellt der unbeschränkte Variationsreichtum an gestalterischen Möglichkeiten auf der Titelseite dar. Einige der von Jeng untersuchten Monographien wiesen auf ihrer Titelseite lediglich ein bis zwei Blöcke auf, andere wiederum bis zu elf.

Ein besonderes Problem ist in der Mehrfachnennung eines Elementes auf einer Titelseite zu sehen. Dennoch wird keine dieser Mehrfachnennungen später vom XPS während der Titelaufnahme berücksichtigt werden.

Es kann konstatiert werden, daß drei Elemente eindeutig die anderen dominieren; Titel, Autor und Herausgeber konnten bei fast allen der 203 untersuchten Titel maschinell isoliert werden.

Jeng folgert aus dieser nachgewiesenen Häufigkeit, daß diese Ausprägungen die Grundlage jeder maschinellen Titelaufnahme stellen werden.

Ein essentieller Anteil eines solchen Systems wird in der

Fähigkeit liegen, diese Heuristiken mit Wahrscheinlichkeitsfaktoren zu verbinden. Eine konsequente Schlußfolgerung aus der oben erläuterten Arbeitsweise ist, daß die visuelle Identifizierung bibliographischer Merkmale die Grundlage einer selbständigen Titelerfassung ist.

Damit wird eine deutliche Trennung zu diversen anderen Projekten vorgenommen, bei denen in der Titelaufnahme noch ein Mensch als Interface zu fungieren hat.

Aufgrund der diversen wiederkehrenden Darstellungstypen im Layout der Titelseiten lassen sich somit Heuristiken bilden, die zusammen mit dem z.B. in Produktionsregeln eingebundenen Expertenwissen (hier das Regelwerk) die Wissensbasis für die Inference Engine stellt.

Eine dieser von Jeng in seiner Reihenuntersuchung der 203 Titelseiten gewonnenen Heuristiken bezieht sich auf die vergleichsweise unproblematische Identifizierung des Autors.

Wenn ein Block, der durch seine näheren Angaben zum Autor (z.B. akademischer Titel) eindeutig identifizierbar ist, durch einen Zwischenraum von einem anderen Block getrennt wird, existiert eine hohe Wahrscheinlichkeit, daß in diesem Block der Name des Autors genannt wird. Liegt darüber hinaus noch eine typographische Abstufung vor, erhöht sich der die Heuristik bildende Wahrscheinlichkeitskoeffizient beträchtlich. Sollte nach diesem Autorenblock eine Grafik folgen, läßt sich aufgrund der erzielten Erkenntnisse schließen, daß im nächsten bibliographischen Merkmal der Herausgeber genannt wird.

Eine weitere dieser empirisch belegten Heuristiken exemplifiziert, daß ein Block, der ohne Zwischenraum einem identifizierten Autorenblock folgt, mit einer Wahrscheinlichkeit von über 80% Zusätze zum Verfassername enthält.

Jeng schlägt ein XPS vor, welches auf Frame-Techniken basiert, um den notwendig hohen Abstraktionsgrad für die Katalogisierung zu erzielen. Ausgehend von seinem Modell ergibt sich ein leistungsfähiges System, welches in der Wissensrepräsentation besonders intensiv mit Frames arbeitet.

Für die Darstellung in Frames wurden 12 der bibliographischen Angaben verwendet.

Diese Angaben wurden nach ihrer relativen und absoluten Häufigkeit ihres Vorkommens gestaffelt und mit vier Wahrscheinlichkeitswerten (95%, 50%, 20% und 1%) versehen.

Eine probate Methode ein Frame zu konstruieren, ist die Einbindung der bibliographischen Elemente in eine Frameumgebung. Die von Jeng isolierten zwölf Merkmale wurden nach der Häufigkeit ihre relativen und absoluten Verteilung in vier Ebenen gegliedert und einem Wahrscheinlichkeitskoeffizienten zugeordnet:

1. im ersten Bereich wird der Wahrscheinlichkeitskoeffizient mit 95% angegeben,
2. im zweiten Bereich wird der Wahrscheinlichkeitskoeffizient mit 50% angegeben
3. im dritten Bereich wird der Wahrscheinlichkeitskoeffizient mit 20% angegeben,
4. im vierten Bereich wird der Wahrscheinlichkeitskoeffizient mit 01% angegeben.

Ein Prozent ist dabei der Wert mit der geringsten relativen Häufigkeit.

Dieser Struktur zufolge, kann mit einer 95% Wahrscheinlichkeit davon ausgegangen werden, daß alle Angaben der ersten Ebene (Titel, Autor, Herausgeber) erkannt werden.

In diesem oberen Bereich werden die Werte/Merkmale aufgeführt, welche die höchste Wahrscheinlichkeitsrate erzielen konnten.

Bei den in der vorliegenden Felduntersuchung analysierten Werten, konnten, wie abgebildet, nur drei Blöcke durchgehend erkannt werden. Zusatzinformationen, wie Nebentitel, usw. konnten hingegen nur bei der Hälfte der Testobjekte gewonnen werden. Lediglich in zwei Fällen wurde die gesetzte Zielvorgabe einschränkungslos erreicht, sämtliche bibliographischen Angaben konnten von dem XPS erkannt und selbständig aufgenommen werden.

Die Framestruktur erlaubt es, ein ideales Titelbild zu herzustellen, einen Prototypen, von dem später abstrahiert wird. Dieses Modell wird konstruiert, indem die Titelbilder verschiedener Abstraktionsgrade in Frames eingebunden und zueinander in Relationen gesetzt werden.

Analog zum Tiefensuchverfahren werden die Frames nach dem optimalsten Wiedererkennungswert abgesehen:

Die unterste Erkennungsrate wird in dem Frame dargestellt, der erst geöffnet wird, nachdem die weniger abstrakten Frames ihre günstigeren nicht erzielen konnten. In diesem Frame ist das wichtigste Element der Titelblock, dem noch der Autor- und der Herausgeberblock beigefügt sind. Weitere Ordnungskriterien sind Größe, Kontrast und Typographie dieser Einträge. Der Block mit der Verlagsangabe folgt dem Autorenblock, wobei die Verlagsangabe durch einen breiten Raum (der oft von einer Grafik ausgefüllt wird) von den ersten Blöcken getrennt wird.

In diesem Frame wird folglich konzeptionalisiert, daß jede Titelseite mindestens die bibliographischen Angaben von Titel, Autor und Verlag aufweist. Gleichzeitig bestimmt dieser Frame (a) den untersten Abstraktionsgrad, den ein XPS erreichen muß.

Die vorgeordneten Frames (b) und besonders (c) geben weniger abstrakte Repräsentationsmodelle wieder. Der (b) Frame ergänzt dabei erforderliche Merkmale, die in (a) noch nicht enthalten sind, wie Zusätze zum Sachtitel und die Ortsangabe als Ergänzung zum Verlagsblock.

Im verbleibenden Frame (c) sind die von Jeng isolierten bibliographischen Merkmale in ihrer Gesamtheit enthalten. Diese Reduktionsfolge von (c) nach (a) gewährleistet eine günstige Streuung der Wiedererkennungswerte. Ausgehend davon, daß der sehr abstrakte Frame (a) das erste und der hoch differenzierte Frame (c) das letzte Muster bildet, nach dem das XPS inferenziert, wird gewährleistet, daß zumindest die konstitutiven Bestandteile eines Titelblatts erkannt werden.

Die Errichtung eines XPS für die Katalogisierung unter den von Jeng gemachten Erfahrungen, kann nur eine der vorliegenden Implikationen sein. Ebenso bedeutend sind die gewonnenen Erkenntnisse, mit denen die potentiellen Regelwerke nach einer möglichen Einbindung in regelbasierten Systemen untersucht werden können. XPS, wie das von Jeng gestaltete, zeigen aber auch die Grenzen der gegenwärtigen Technologien auf. Die herausragende Quintessenz der Arbeit Jengs liegt unzweifelhaft in der Tatsache, daß auch mit den noch begrenzten Werkzeugen der KI es bereits möglich ist, den Prozeß der Titelaufnahme vollständig zu automatisieren im Sinne einer optischen Erkennung der Schrift, verbunden mit deren bibliographischer Interpretation.

Die diskutierten XPS gewinnen erst durch eine erhebliche Steigerung ihres Leistungsumfanges an wirklichem praktischen Wert. Die gemessene Fehlerrate des Jeng-Systems verbietet zwar einen realitätsbezogenen Einsatz, bietet andererseits aber eine Entwicklungsumgebung für weitergehende Projekte.

Die Einführung der EDV im Bibliotheksmanagement muß konsequenterweise einen entsprechenden Impact auf die Bearbeitung bereits bestehender Daten haben. Die Transformierung bestehender Massendaten in konventionellem Aggregatzustand in die digitale Form stellt auch die KI in der retrospektiven Katalogisierung vor neue Herausforderungen.

Die sehr spezifischen Probleme der retrospektiven Katalogisierung (*human resources management, cost center, etc.*) sind bekannt.

Vor diesem Hintergrund setzt die Universitätsbibliothek Tübingen gegenwärtig in einem Pilotprojekt ein XPS zur Konversion von Katalogkarten nach den RAK-WB ein.

In Zusammenarbeit mit der Firma Compulex/Zürich konnten von diesem wissensbasierten System bereits über 100 000 Titelaufnahmen nach den RAK-WB im sog. „BiblioScan-Verfahren“ generiert werden. In BiblioScan werden im Prozeß der Katalogisierung die Schritte ERFASSUNG, KORREKTUR und FORMATIERUNG überwiegend vom System selbständig durchgeführt.<sup>14</sup>

Interimistisch muß festgestellt werden, daß die Katalogisierung als eine der arbeitsintensivsten bibliothekarischen Tätigkeiten durch XPS besonders gravierend betroffen sein wird.

So hat Michael Buckland wohl kein Verdikt über die Katalogisierer gesprochen, aber eine bereits begonnene Entwicklung erkannt, als er feststellte, daß „*the library catalog in its present state is fundamentally obsolescent in an online world*“<sup>15</sup>.

Eine amerikanische Bibliothekarin forderte in diesem Tenor schon vor über zwanzig Jahren mit der Katalogisierung beschäftigte Bibliothekare auf „*to rethink purposes and procedures to devise a system which will offer a totality, rather than a mere fragments of service.*“ Katalogisierer müssen heute wie ehemals ihre Ziele und Arbeitsmethoden neu überdenken, um Teil in dem Entwicklungsprozeß von der herkömmlichen Informationsverwaltung zum modernen Informations-Management zu werden. Um in diesem Vorgang bestehen zu können, ist es von unabdingbarer Notwendigkeit, daß die betroffenen Katalogisierer auf diese Herausforderung reagieren. Eine angemessene Antwort ist die Qualifikation als KE.

Nur so kann gewährleistet werden, daß dieses zentrale Objekt bibliothekarischer Arbeit auch zukünftig den Ansprüchen und Normen der Kommunikationsgesellschaft nachkommen kann.

## 17 XPS in der Klassifizierung

Weit größere Probleme verursacht die Erstellung von XPS für die Klassifizierung von Medien.

Die Möglichkeiten, ein entsprechendes System zu erstellen, wurden an der University of Strathclyde, Vereinigtes Königreich, in einem Projekt untersucht. Die Zielvorgabe dieses Unternehmens, an dem das British Library Research and Development Department paritätä-

tisch beteiligt waren, sah eine Marktsichtung von kommerziell vertriebener Shells vor, und deren Untersuchung auf einen möglichen Einsatz in bibliothekarischen Bereichen.

Nach Abschluß der Projekte blieben die erstellten XPS auch weiterhin für Schulungszwecke installiert.

Im Anschluß an dieser umfangreichen Sichtung wurden drei Shells ausgewählt und genauer examiniert in Hinblick auf ihre Fähigkeiten, Wissen zu repräsentieren, bzw. klassifiziertes Wissen zu manipulieren:

1. Expert-Ease  
(Intelligent Terminals, Ltd.)
2. ESP/Advisor  
(Expert Systems International)
3. Xi Plus  
(Expertech)

Der geplante Einsatz dieser Shells sollte in der Implementierung zweier Klassifikationssysteme liegen. Darüberhinaus waren die weiteren Möglichkeiten von XPS-Architekturen zur Verbesserung der zeitlichen Effizienz und die Zuverlässigkeit in der Vergabe von Notationen zu prüfen.

Dieses System sollte ebenso präzise wie der genuine Katalogisierer sein, software-ergonomischen Kriterien genügen und eine flexible Regelbasis besitzen, verbunden mit einer leistungsstarken Tracekomponente, um die getroffenen Entscheidungswege nachvollziehen und verifizieren zu können.

Zwei verschiedene Klassifikationsarten wurden in die Shells eingebunden:

1. mit der Dewey Decimal Classification der 19. Ausgabe (DDC 19) ein (enumeratives und) hierarchisches Klassifikationssystem sowie
2. eine Facettenklassifikation.

Die Wissensbasis wurde so strukturiert, daß auch Anfänger mittels einer durchgehenden Menüführung die XPS ohne hohem Lernaufwand benutzen können. Bei Xi Plus besteht diese Menüführung aus einer Liste mit vorgegebenen Fragen und Antworten, die vom Benutzer lediglich aktiviert zu werden brauchen.

Die Entwicklung der Regelbasen wurde auf verschiedene Arten vorgenommen. Der probate Schritt des *rapid prototyping* wurde auch möglich, weitere Systemstellen problemlos zu übernehmen, ohne Gefahr zu laufen, die Integrität der DDC Hierarchie zu verletzen.

Die eigentliche Konsultation erfolgt in Dialogform. Das System stellt Fragen, der Benutzer erstellt mit seinen Antworten (passiv) die Datenbasis. Bei dieser Dialogführung muß beachtet werden, daß keine Antworten gegeben werden dürfen, denen der Bezug zur Wissensbasis fehlt. Auf die Systemfrage hin: „Wollen Sie ein Buch klassifizieren?“, kann also auch nicht mit einem Unsicherheitswert wie „Ich bin mir nicht sicher“ geantwortet werden. Eine Einbindung von diesen Unsicherheitswerten, wie wünschenswert und notwendig sie auch sein mögen, sieht Xi Plus auch in der aktuellen Version nicht vor.

Problematisch bei der Regelbildung ist der Umstand,

14 Shah, Raymond: Compulex – Wissensbasierte Verfahren für kontextbezogene Textumsetzung (1993) *nicht veröffentlicht*.

15 Buckland, K. Michael: Bibliography, library records, and the redefinition of the library catalog. In: *Library Resources & Technical Services* 32 (1988). S. 299-311.

daß eine unreflektierte Übernahme der DDC Terminologie nicht möglich ist. Begriffe, die in einer Regel vorkommen, können schon in einer anderen Regel eine veränderte Semantik besitzen. Die unausbleibliche Folge ist, daß es zu unlogischen Notationen, bzw. zu Systemabstürzen kommen kann. Verletzungen der Systemintegrität sind damit vorprogrammiert ...

Dieses XPS repräsentiert das Wissen in einer enumerativen (oder nach Brian Buchanan auch „präkombiniert“ genannten) Weise, da es ausnahmslos alle relevanten Klassen aufführen kann.

Unter Beachtung dieser Merkmale und Einschränkungen konnte von der Projektgruppe also ein funktionales XPS auf Basis einer enumerativen Klassifikation vorgestellt werden.

Bei der Synthetisierung einer Facettenklassifikation wurde jedoch kein befriedigendes Ergebnis erzielt.

Die erstellten Prototypen versagten, da das XPS während einer Konsultation immer nur eine Notation für einen Begriff finden konnte. Andere mögliche Systemstellen wurden ignoriert.

Die praktische Arbeit mit Xi Plus gestaltete sich auch für Bibliothekare mit geringen EDV-Vorkenntnissen als möglich, besonders das Xi Plus eigene Tutorial (auf XPS-Basis) erleichtert die Wissensakquisition, und trägt somit dazu bei, die KI Akzeptanz unter den beteiligten Bibliothekaren zu verstärken.

Die Arbeitsgruppe kam abschließend zu der Erkenntnis, daß konventionell vertriebene Shells sehr wohl in vielen bibliothekarischen Arbeitsbereichen wertvolle Hilfe leisten könnten, keineswegs aber in der Klassifikation.

Besonders problematisch war das Repräsentieren des in der DDC formulierten hierarchischen Aufbaus. Die in Xi Plus enthaltene ansonsten leistungsfähige Regelsprache erwies sich als nicht ausreichend flexibel, um jeden Begriff einer Klasse auf einer gemeinsamen Ebene darzustellen, z.B.:

wenn die Hauptklasse ist x	(Ebene 1)
und die Unterklasse 1 ist y	(Ebene 2)
und die Unterklasse 2 ist z	(Ebene 3)
und ...	(Ebene n)
dann die Signatur ist ? <sup>16</sup>	

Die Anzahl der betroffenen Begriffe würde eine extrem hohe Redundanz der Regeln verursachen, was wiederum einen nicht tolerablen Abfall der Performance verursachen müßte.

Die entwickelten Prototypen wiesen darüber hinaus ein weiteres gravierendes Manko auf, welches in der kongruenten Menüführung von Xi Plus lag. Der Benutzer kann nur mit den angebotenen Auswahlmöglichkeiten arbeiten. Eine freie Eingabe von Begriffen, z. B. aus dem Titel, kann nicht verwirklicht werden, obwohl dieses ursprünglich so geplant war.

Ein Umstand, der erneut auf die ungenügende Zusammenarbeit zwischen Informatikern und Informationsmanagern hinweist.

Eine entsprechende Kooperation ist umso erforderlicher, da Shells aufgrund ihrer mangelnden Flexibilität für die Klassifikation eine unbefriedigende Lösung darstellen.

Durch die eigene Programmierung eines XPS für die Klassifikation würde gewährleistet werden, daß eine Regelsprache entwickelt wird, welche die syntaktischen Charakteristika einer Klassifikation eher beherrschen könnte als ein vorgefertigtes Produkt.

Inwieweit jedoch eine solch immens aufwendige Programmierarbeit sich aus betriebswirtschaftlichen Gründen realisieren ließe, bleibt jedoch fraglich.

Die Summe dieser Einschränkungen läßt es als wenig opportun erscheinen, diesen Shell für die beschriebenen bibliothekarischen Aufgaben zu empfehlen.

Den besonderen Anforderungen in der Erforschung von KI-Techniken und deren Bearbeitung für Informationsmanagement Systeme (IMS) und BIS konnte sich Xi Plus hingegen als gewachsen zeigen.

Es zeigte sich bereits in einer frühen Phase dieses Projektes, daß Shells durchaus die Strukturen diverser Kodizes explizieren können. Ein Regelwerk für die Formalerschließung (wie die AACR oder auch die RAK) bildet ein in sich geschlossenes Objekt, welches relativ problemlos in ein regelbasiertes System eingebunden werden kann. Konträr verhält es sich in der Inhalterschließung.

Klassifikationssysteme (wie die DDC, resp. die DK) sind Instrumente zur universalen Darstellung des gesamten menschlichen Wissens. Die in diesen Klassifikationssystemen enthaltenen Regeln repräsentieren nicht nur einen in sich abgeschlossenen Komplex, sondern abstrahieren vielmehr eine unendliche Anzahl von Regeln (Daten, Fakten und Relationen): das menschliche Wissen. Der mit Klassifikationen vertraute Bibliothekar verfügt über ein breites Allgemeinwissen. Es versetzt ihn in die Lage, klassifizierende Begriffe zueinander in Beziehungen zu setzen. Erst dieses Geflecht stellt eine Klassifikation dar. Universales Wissen wird nur auf dieser Weise zu repräsentieren sein.

Ein XPS für die Klassifikation dagegen müßte erst eine Verbindung aller Begrifflichkeiten neben ihren Synonymen und Homonymen zu leisten haben.

Derartige Methoden für eine wirklich zufriedenstellende Artifizierung von Allgemeinwissen (im Gegensatz zum begrenzten Expertenwissen) kennt die KI noch nicht.

Vor dem Hintergrund dieser systembedingten Einschränkung konnte nicht nur Xi Plus den gestellten Anforderungen nicht genüge leisten, alle anderen getesteten Shells mußten konsequenterweise ebenfalls scheitern. Die Schlußfolgerung aus diesen Untersuchungen müssen in der Einsicht liegen, daß auf absehbarer Zeit auch die innovativsten Techniken und Methoden der KI eine autonome Inhalterschließung zumindest als fragwürdig erscheinen lassen.

## 18 Chancen?

Zusammenfassend muß konstatiert werden, daß der Einsatz dieser neuen Technologie auch einen sozialen Prozeß ausgelöst hat.

Nicht nur die forcierte Entwicklung der Informations- und Kommunikationstechnologien durch den KI-Einsatz muß eine entsprechende Bewertung finden. Diese Technik kann nicht losgelöst von den betroffenen Menschen betrachtet werden.

Die Notwendigkeit einer wissenschaftlich fundierten Technikfolgeabschätzung ist dementsprechend als evident anzusehen.

Die in den vorangegangenen Abschnitten dieser Schrift vorgenommene Hypothesenbildung kann nur durch eine

16 Man beachte die der englischen Grammatik entlehnte Syntax.

methodische Analyse erhärtet werden: Das Deutsche Bibliotheksinstitut ist aufgerufen, diese notwendige Erarbeitung einer Untersuchung von Handlungsoptionen vorzunehmen. Die Konstitutionalisierung einer Arbeitsgemeinschaft aus Wissenschaftlichen Bibliothekaren, Informationswissenschaftlern, Dokumentaren und Informatikern könnte den Kontext bilden, in dem eine Bewertung der Chancen und Risiken des XPS-Einsatzes möglich ist.

Die betroffenen Berufsverbände haben mit der Fragestellung konfrontiert zu werden, inwieweit die beschriebene Technikentwicklung und die beruflichen Strukturen einander angepaßt werden können, resp. wo sie sich gegenseitig ausschließen.

Durch die Erstellung möglichst realistischer Szenarien von XPS-Anwendungen in Bibliotheken, Informationszentren und Dokumentationsstellen könnten die sozialen Auswirkungen auf die betreffenden Berufsbilder antizipiert werden. Eine derartige Methode setzt jedoch das Vorhandensein interessenorientierter Bewertungskriterien als Ergebnis einer Technikfolgeabschätzung voraus.

Ausschließlich eine derartige Untersuchung könnte die Grundlage für die Sicherheitskontrolle von XPS schaffen. Der Einsatz von KI-Technologien ist auch im Informations-Management mit Unsicherheitsfaktoren belastet. Eine gewisse Uniformität in der verbleibenden bibliothekarischen Arbeit könnte ebenso auftreten. Die Beschäftigungsauswirkungen im Bibliothekswesen können bivalent betrachtet werden. Zum einen ergibt sich ein problematischer Rationalisierungseffekt im Bereich bibliothekarischer Regelarbeiten. Andererseits resultiert sich aus der Arbeit mit/von XPS die bereits angesprochene Höherqualifikation von Bibliothekaren bei der verbleibenden Tätigkeit. Der individuelle Bibliothekar wird einen persönlichen Vorteil erzielen können, da durch die Freisetzung kreativer Potentiale persönliche Freiräume in der bibliothekarischen Tätigkeit geschaffen werden können.

Die hier erarbeiteten Auswirkungsbilder müssen in Verbindung mit dem Stellenwert von XPS im Prozeß der Automatisierung der bibliothekarischen Arbeitswelt verstanden werden. Die Sozialverträglichkeit von XPS muß vor einem flächendeckenden Einsatz in der Fachwissenschaft diskutiert werden.

Besonders vor dem Hintergrund eines sich dramatisch wandelnden Berufsbildes (fort von der konventionellen bibliothekarischen Vermittlung von Printmedien, hin zu einer Erschließung der digitalen Wirtschaftsware „Information“ durch den *information specialist*) sollte dieser Umstand keineswegs unterbewertet werden.

### Auswahlbibliographie

- Aberico, Ralph: Expert Systems for reference and information retrieval. London 1990. (Supplements to computers in libraries; 10).
- Bailey, Carles: The intelligent Reference Information System. In: CD-ROM Librarian 5 (8) S. 10-19.
- Brittain, M.: The implications for LIS education of recent development in expert systems. In: Information Processing & Management 23 (1987) S. 139-152.
- Brewis, L.-E.: Application of artificial intelligence in information work. In: Mousaion 8 (1990) S. 65-81.
- Buckland, K. Michael: Bibliography, library records, and the redefi-

- inition of the library catalog. In: Library Resources & Technical Services 32 (1988) S. 299-311.
- Bullinger H.-J.: Zur Informatisierung der Arbeit. In: FhG Berichte 1 (1986) S. 7-12.
- Butkovich, Nancy J.: An expert system at the reference desk: impressions from users. In: Reference Librarian 23 (1989) S. 61-74.
- Caregnato, S.: Expert systems support for subject librarians or subject specialists in academic libraries. In: Information research news 4 (1994) S. 2-6.
- Clarke, Ann: Expert systems and library information work. In: Journal of Librarianship 15 (1989) S. 289.
- Davidson, Lloyd: Expert systems for library applications. In: Database 13 (1990) S. 80-83.
- Davies, Roy: Towards an expert system for cataloguing: some experiments based on AACR 2. In: Program 18 (1984) S. 284-285.
- Dykstra, Mary: Expert Systems to aid in retrieval from technical documentation. Halifax: Dalhousie University.
- Dykstra, Mary: The structured encoding of document content: a logolinguistik study of PRECIS as a possible model in the shift from indexing to automated text analysis. Sheffield 1986.
- Efthimiadis, Efthimis N.: A review of online searching aids. London 1990 (British Library Research Paper; 86).
- Gilroy, Joyce: An expert system for classification. In: Its News 22 (1990) S. 39-40.
- Hanfman, Deborah: AquaRef: an expert advisory system for reference support. In: Reference Librarian 23 (1989) S. 113-133.
- Heller, Martin: AI in Practice. In: BYTE 1 (1991) S. 267-278.
- Herfurth, M.: Von bibliographischen Datenbanken zu Wissensbanken. Frankfurt/M. 1990.
- Hjerpe, Roland: Project ESSCAPE: expert systems for simple choice of access points for entries; applications of artificial intelligence in cataloging. In: International Federation of Library Association General Conference 51. Chicago 1985. S. 46 ff.
- Hsieh, Cynthia C.: Survey of artificial intelligence and expert systems in library and information science literature. In: Information Technology and Libraries 8 (1989) S. 209-214.
- Jeng, Ling Hwey: An expert system for determining title proper in descriptive cataloging: a conceptual model. In: Cataloging & Classification Quarterly 7 (1986) S. 67.
- Jeng, Ling Hwey: Knowledge representation of the visual image of a title page. In: Journal of the American Society for Information Science 42 (1990) S. 99-109.
- Klee, Hans Werner: Zur Benutzergerechtigkeit von Expertensystemen. In: Nachrichten für Dokumentare 41 (1990) S. 299-305.
- Morris, Anne: Expert systems and their implications for information science. In: Brittain, M: Curriculum development in Information Science to meet the needs of the information industries in the 1990's. BLRDD Report 1987.
- Morris, Anne: Expert systems: Teaching and Research in North American Library and Information Science Schools. Loughborough 1990.
- Morris, Anne: Information professionals – roles in the design and development of expert systems? In: Information & Processing 24 (1988) S. 173-181.
- Morris, Anne: Shortage of expert system personnel in the UK – hype or reality? In: Expert Systems 6 (1989) S. 218-226.
- Nitecki, Danuta A.: Managing experts: creating links between librarians, users and systems. In: Journal of Library Administration 13 (1990) S. 59-79.
- Ohly, Hans Peter: Conceptual information retrieval by knowledge-based programming techniques. In: International Classification 18 (1991).
- Rasmus, Daniel W.: Putting the Experts to Work: Expert systems can capture and deploy the intricate thought processes of those who think for a living. In: BYTE 1 (1991) S. 281-287.
- Roesner D.: Computerlinguistik auf dem Weg in die Praxis. Das Beispiel multilinguale technische Dokumentation. In: Informationstechnik und technische Informatik 2 (1994) S. 48-58.
- Schuegraf Ernest J.: A survey of expert systems in library and information science. In: Canadian Journal of Information Science 15 (1990) S. 42-57.

Stein Richard Marlon: Real Artificial Life: How do algorithmic evolutionary mechanisms and knowledge representations combine to synthesize an alife organism? In: BYTE 1 (1991) S. 289-298.

Valauskas, E.J.: Using the Internet in libraries. In: IFLA journal 1 (20). S. 22-28.

Vickery, Alina: PLEXUS – the expert system for referral. In: Information Processing & Management 13 (1987) S. 99-117.

Walton, Kathryn: Manpower requirements and availability in the expert systems field. London 1989 (British Library Research Paper; 65).

Weckert, John: Artificial intelligence, expert systems and librarianship: a review of the literature. In: Australian Library Review 7 (1990) S. 281-299.

Weizenbaum, Joseph: Die Macht der Computer und die Ohnmacht der Vernunft. Frankfurt am Main. 1987.

**Anschrift des Autors:**

Lutz Pietzka  
BDO Deutsche Warentreuhand AG  
Wirtschaftsprüfungsgesellschaft  
Ferdinandstraße 59  
D-20095 Hamburg