DAS NATÜRLICHSPRACHLICHE INTERFACE NLI/AIDOS

H. Helbig; F. Zänker
Robotron-Projekt GmbH Dresden

Die Formulierung des Informationsbedarfs an ein Recherchesystem mit den Mitteln der natürlichen Sprache ist für viele Nutzer die adäquate und letztlich einzige Möglichkeit der Kommunikation mit dem Rechner.
Für das Informationsrecherchesystem AIDOS wurde ein Interface NLI/AIDOS geschaffen, das eine nutzerfreundliche Recherche nach Fakten aus dem Datenbestand des Systems AIDOS gestattet. Der Aufbau und das Zusammenwirken von NLI und AIDOS sind im Bild 1 dargestellt. Auf der linken Seite des Bildes sind die sprachlichen Schnittstellen des Systems widergegeben, in der Mitte stehen die für die automatische Verarbeitung der natürlichsprachigen Anfragen relevanten Moduln, und auf der rechten Seite ist das benötigte Hindergrundwissen repräsentiert.

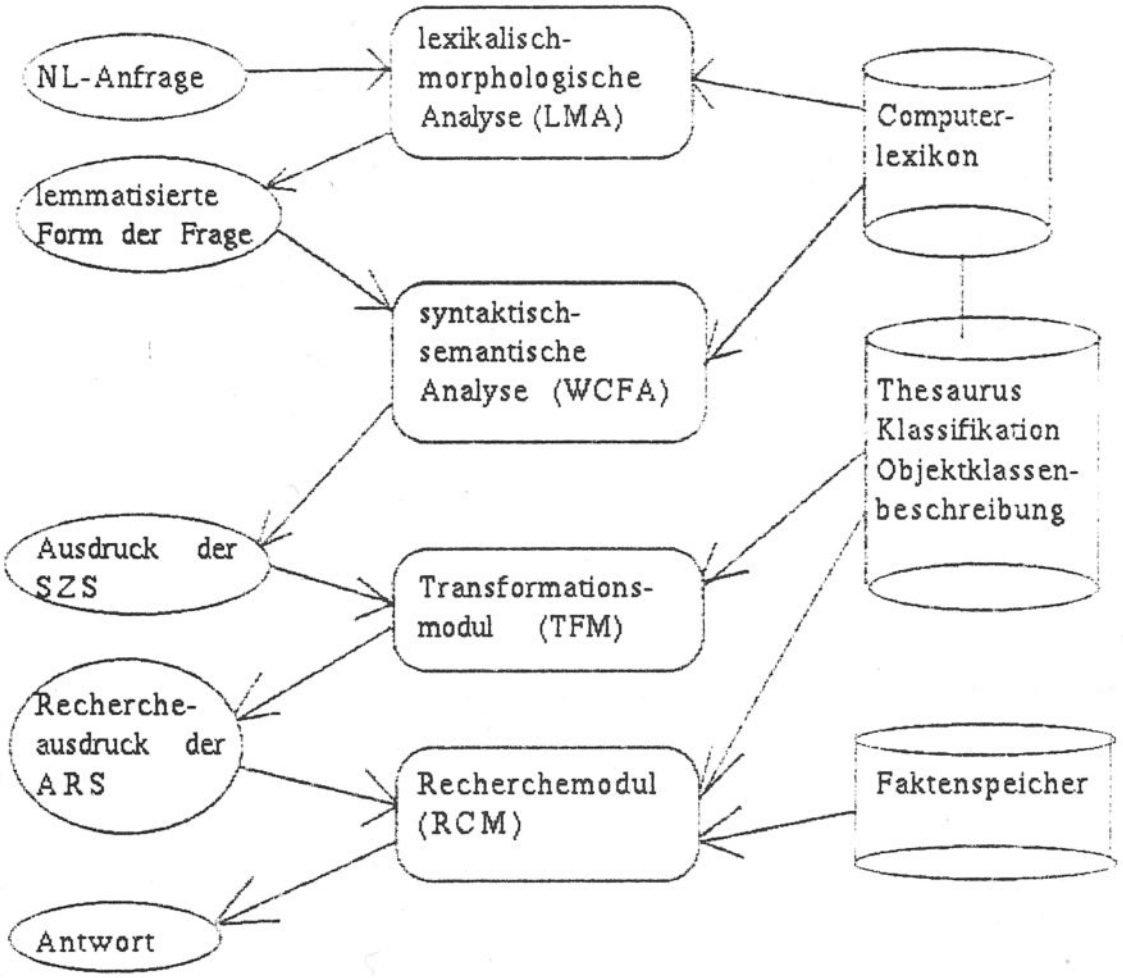

Bild 1: Struktur und Wirkungsweise des Gesamtsystems NLI plus AIDOS

Zur übersetzung der natürlichsprachigen Anfragen (NL-Anfragen) in Ausdrücke der AIDOS - Recherchesprache (ARS) benötigt man ein Computerlexikon, das die wichtigsten grammatischen Eigenschaften der Wörter enthält. Aber auch das im Datenbanksystem selbst gespeicherte Wissen ist für den übersetzungsprozeß relevant. Es umfaßt die Objektklassenbeschreibung, ein fakultativ einsetzbares Klassifikationssystem (mit dem das Anwendungsgebiet strukturiert werden kann) und den Thesaurus (s. Bild2). Aber auch das Faktenwissen des Systems, im gewählten Einsatzgebiet ist das die Verwaltung von Literatur zur künstliche Intelligenz (abgekürzt: KI), ist für die natürlichsprachliche Verarbeitung interessant. So gibt z.B. die Tatsache, daß ein bestimmter Eigenname dem Begriff VERFASSER zuzuordnen ist, einen Hinweis auf die Wortklassenzugehörigkeit (Substantiv).

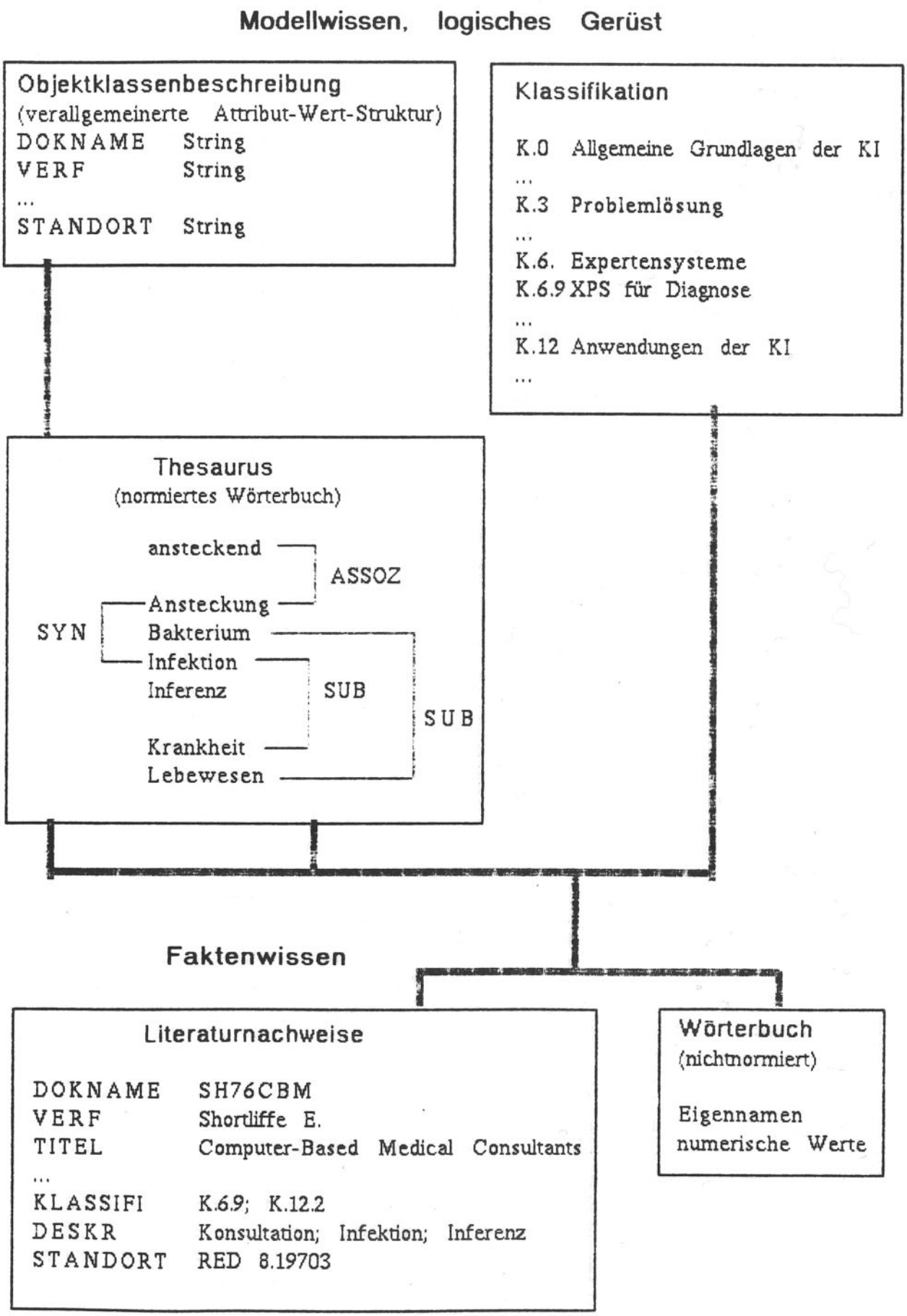

Bild 2: Wissensbasis des Informationsrecherchesystems AIDOS

Die Arbeitsweise des NLI/AIDOS

Die Arbeit mit dem Informationsrecherchesystem AIDOS bei natürlichsprachiger Anfrage soll anhand des Diskursbereiches "KI-Literatur" erläutert werden. Die zugrunde gelegte Faktenbasis enthält Beschreibungen von Büchern, Zeitschriftenartikeln und anderen Publikationen, die durch nachstehende Merkmale charakterisiert sind:

DOKNAME	-	Schlüsselwort (eindeutige Bezeichnung des Dokuments)
DOKTYP	-	Typ des Dokuments (Buch, Zeitschriftenartikel usw.)
DOKFORM	-	Form des Dokuments (Mikrofiche, gebundenes Buch usw.)
VERF	-	Verfasser
TITEL	-	Titel der Publikation
KLASSIFI	-	Nummer in einer vorgegebenen Dezimalklassifikation
DESKR	-	Deskriptoren zur Beschreibung des Inhalts
JAHR	-	Erscheinungsjahr
VERLAG	-	Verlag und Ort
STANDORT	-	Standort/Bibliothek, wo die Publikation vorhanden ist

Der Ablauf der Recherche geht wie folgt vonstatten. Nachdem der Modus für die natürlichsprachige Kommunikation mit Funktionstaste ausgewählt wurde, erscheint eine Aufforderung des Systems zur Eingabe. Wenn sich der Nutzer für Bücher über Künstliche Intelligenz interessiert, könnte der Schirm nach entsprechender Reaktion des Nutzers folgendes Aussehen haben:

Ihre Frage:
Welche Bücher über KI kennst du?

Man erhält als Ergebnis der linguistischen Verarbeitung (Schnittstelle SZS) den Ausdruck:

((ERG NIL) (G-1)((SETOF G-1 G-2)(SUB G-2 BUCH)(THM G-2 G-3)
(SUB G-3 KI)))

Der linguistische Prozessor des NLI hat also ermittelt, daß es sich bei der Anfrage um eine Ergänzungsfrage handelt, daß keine Referenz auf einen vorhergehenden Dialogschritt auftrat und daß schließlich der Fragefokus G-1 dem Begriff Buch (mittelbar über die Mengenrelation SETOF und die Subordinationsbeziehung SUB) untergeordnet ist, wobei der Inhalt (das Thema THM) von G-2 auf den Begriff KI verweist.
Durch den Transformationsmodul wird anschließend eine Folge von Suchtermen erzeugt, wobei von AIDOS sofort die zu jedem Suchterm in direkter Suche gefundene Anzahl relevanter Nachweise angezeigt wird:

AIDOS/M	Recherche	dir.Suche	Nachweisart:	Dokumente
1	DOFORM = BUCH			594
2	DOKTYP = BUCH			893
3	1 ODER 2			1174
4	KLASSIFI = K.0			256
5	KLASSIFI = K.12			281
6	4 ODER 5			524
7	3 UND 6			200

Da das natürlichsprachige Wort "Buch" sowohl den Dokumententyp als auch die Dokumentenform bezeichnen kann, müssen die Terme 1 und 2 disjunktiv verknüpft werden. Der Begriff KI weist in der Klassifikation auf die Rubrik K.0 (Allgemeine Grundlagen der KI) bzw. auf K.12 (Anwendungen der KI). Deshalb müssen die Terme 4 und 5 ebenfalls disjunktiv verknüpft sein. Dagegen müssen

die Terme 3 und 6 gleichzeitig auf die gesuchten Objekte zutreffen, weshalb für sie nur eine konjunktive Verknüpfung in Frage kommt.
Wenn dem Nutzer 200 relevante Nachweise für die Ausgabe zuviel sind, kann er durch Zusatzfragen den Suchraum weiter einengen. Durch Funktionstastendruck (F6) erhält er wieder die Initiative:

```
Ihre Frage:
Sind diese in der Robotron-Bibliothek vorhanden?
```

Wegen der Referenz auf die vorher gefundene Antwort (ausgelöst durch das Demonstrativpronomen "diese") erzeugt das NLI keine neue, unabhängige Folge von Suchtermen, sondern setzt die vorhergehende Folge sinnvoll fort:

```
AIDOS/M          Recherche    dir.Suche          Nachweisart:  Dokumente
....
7    3 UND 6                                                     200
8    STANDORT = RED bis RED?                                     244
9     7 UND 8                                                     21
```

Es wurden aus den bis zum siebenten Term (vorhergehende Anfrage) gefundenen 200 Nachweisen nur die ausgewählt, die den Standort "Robotron-Bibliothek" (Bibliothekskürzel RED) besitzen. Da bei Bibliothekskürzeln sehr unterschiedliche Schreibweisen angetroffen werden, wurde vom NLI für das Merkmal STANDORT eine zusätzliche Suche mit Wildcard (Zeichen "?") vorgesehen, so daß alle Zusammensetzungen, die mit RED beginnen, ebenfalls noch als Benennungen für "Robotron-Bibliothek" akzeptiert werden.

Wenn auch die nunmehr ermittelten 21 Nachweise noch zuviel für eine Anzeige auf dem Bildschirm sind, oder wenn sich jemand nur für Literatur interessiert, die neueren Datums ist, kann das Spektrum der Antworten noch weiter eingeengt werden:

```
Ihre Frage:
Welche davon  sind nach 1980 erschienen ?
```

Durch diese Frage wird wiederum eine Fortsetzung der bisherigen Recherche erreicht (Referenzwort "davon") und die automatisch generierten Suchterme haben die Gestalt:

```
AIDOS/M          Recherche    dir.Suche          Nachweisart:  Dokumente
....
8    Standort = RED bis RED?                                     244
9    7 UND 8                                                      21
     IN 9
     1 JAHR > 1980
10   START                                                         7
```

Die Besonderheit der durch die letzte Frage ausgelösten Fortsetzung der Termfolge besteht darin, daß ein sequentielle Suche, aufbauend auf dem Ergebnis des Suchterms 9, ausgelöst wird. Das hat seine Ursache darin, daß beim Einrichten der AIDOS-Datenbank das Merkmal JAHR (für "Erscheinungsjahr") als ein Merkmal gekennzeichnet wurde, nach dessen Werten nicht invertiert wird. Es ist auch bemerkenswert, daß die zeitliche Nachfolgerelation (ausgedrückt durch die Präposition "nach") in der Frage richtig durch die Vergleichsoperation ">" im AIDOS-Suchterm wiedergegeben wird. Dies wird nicht einfach durch Oberflächenpräpositionen wie "nach", "ab" oder "seit" gesteuert, sondern durch eine echte semantische Deutung von syntaktischen Beziehungen des natürlichsprachigen Fragesatzes erreicht.
Die bis hierhin gefundenen sieben Nachweise kann man sich ohne weiteres auf dem Bildschirm

durch Einsatz einer besonderen Funktionstaste anzeigen lassen. Für den ersten der sieben Nachweise erhält man z.B. folgende Anzeige:

DOKNAME	NI82POA
DOKTYP	BUCH
DOKFORM	BUCH
VERF	Nilsson N.J.
TITEL	Principles of Artificial Intelligence
KLASSIFI	K.0, K.3
DESKR	KI-Methoden, automatische Problemlösung, Suchalgorithmen, Deduktion
JAHR	1982
VERLAG	Springer-Verlag, Berlin, Heidelberg, New York
STANDORT	RED/EB B.08.05.311

Literatur

1. Helbig, H.: Natural language access to the data base of the AIDOS/VS information retrieval system. Proc. AI Conference Smolenice 1984, North-Holland Publ. Comp., Amsterdam 1984

2. Helbig, H: Forschungsergebnisse auf dem Gebiet der natürlichsprachigen Kommunikation mit dem Rechner. Neue Technik im Büro, Berlin 29(1985)5 S. 140-142

3. Naumann, P.: Informacionno-poiskovaja sistema AIDOS/VS. Moskva, ICSTI, 1986