

# Konjunkturanalysen mit DIWAX

Von Stefan Kooths\*

**Zusammenfassung:** DIWAX ist ein agentenbasiertes Softwarepaket für VGR-gestützte Konjunkturprognosen. Die Hauptkomponenten des Systems bilden eine relationale Metadatenbank, mit der Zeitreihen nach primär ökonomischen Kriterien beschrieben werden können, ein Ressourcenkonzept, das die Verbindung zu beliebig vielen heterogenen Datenquellen ermöglicht, sowie – als wesentlicher Teil des Systems – eine Analyseumgebung, die dem Anwender innerhalb eines frei gestaltbaren Ableitungsschemas bei der Prognoseerstellung unterstützt. Hilfsagenten übernehmen dabei Routineaufgaben, stellen die Aktualität der Datenbasis sicher, prüfen die Prognose auf Konsistenz und diagnostizieren mögliche Unregelmäßigkeiten. Der Beitrag beschreibt den konzeptionellen Aufbau der Software und stellt wesentliche Eigenschaften vor. Zudem wird ein Ausblick auf zukünftige Erweiterungen geworfen, zu denen vor allem die Integration makroökonomischer Modelle gehört.

**Summary:** DIWAX is an agent-based software package designed to support macroeconomic forecasting based on national accounting schemes. Its main building blocks comprise a relational metadatabank for organizing data along standard economic and statistical classifications, a resource-based project management for downloading and importing data from many different sources (including output from macroeconomic models) and, most importantly, an analytic interface for analyzing data and designing forecasts within a user-defined framework. Software agents assist the forecaster's work in terms of performing routine work and checking the results for inconsistencies, irregularities or outdatedness. This paper describes the key features and underlying concepts for the development of DIWAX and delivers a preview on future extensions.

JEL Klassifikationen: C63, E01, E27

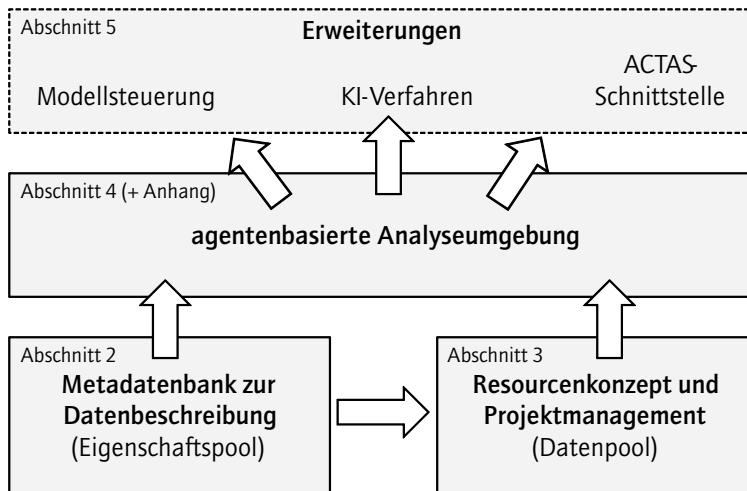
Keywords: Business cycle analysis, national accounting, forecasting, database software

## 1 Einleitung und Überblick

Die empirische Konjunkturforschung als Beobachtung, Diagnose und Prognose der gesamtwirtschaftlichen Entwicklung stellt hohe Anforderungen an die Informationsinfrastruktur. Dies gilt insbesondere für detaillierte Konjunkturanalysen, die vollständig in die Systematik der Volkswirtschaftlichen Gesamtrechnungen (VGR) eingepasst werden. Kennzeichnend für diese Anforderungen ist ein umfangreiches und multifrequentes Datenaufkommen, die Notwendigkeit zur Zusammenführung zahlreicher, zum Teil sehr heterogener Datenquellen, die Variabilität von Vergangenheitsdaten aufgrund häufiger Revisionen, eine schnelle Bereitstellung verschiedener Sichten auf den konsolidierten Datenbestand sowie die Erfüllung mehrdimensionaler Konsistenzbedingungen, die im Zuge so genannter Rundrechnungen zu beachten sind. Darüber hinaus ist im prognostischen Bereich die Verbindung von statistisch-ökonomischen Verfahren mit menschlichem Expertenwissen zu berücksichtigen.

\* DIW Berlin, E-Mail: [skooths@diw.de](mailto:skooths@diw.de)

Abbildung 1

**Überblick**

Quelle: Eigene Darstellung.

Diesem speziellen Anforderungsmix können Standardprogramme (Datenbanksoftware, Tabellenkalkulationsprogramme oder Statistikpakete) typischerweise nicht gerecht werden, da sie ein allgemeineres Anwendungsspektrum abdecken müssen. Vor diesem Hintergrund hat das DIW Berlin vor zwei Jahren entschieden, ein auf die Bedürfnisse der Konjunkturforschung zugeschnittenes Softwarewerkzeug zu entwickeln, das diesen Anforderungen gerecht wird. Dieses liegt nunmehr unter der Bezeichnung DIWAX in einer ersten Version vor, die im Folgenden näher erläutert wird. Hierbei stehen die konzeptionellen Konstruktionsprinzipien im Vordergrund, nicht aber die praktische Bedienung der Software selbst. Einen Überblick über die wichtigsten Programmkomponenten bzw. -konzepte und deren Behandlung in diesem Beitrag zeigt Abbildung 1.

Die Grundlage der DIWAX-Konzeption bildet die im zweiten Abschnitt behandelte Metadatenbank zur universellen Beschreibung der ökonomischen und statistisch-technischen Dateneigenschaften. Diese wird erweitert um ein Ressourcenkonzept zur Einbindung externer Datenquellen, die das Management von umfangreichen Prognoseprojekten unter Verwendung der in der Metadatenbank abgelegten Dateneigenschaften ermöglicht (Abschnitt 3). Beide Blöcke bilden die Grundlage, auf der die agentenbasierte Analyseumgebung für (insbesondere VGR-gestützte) Konjunkturanalysen aufbaut (Abschnitt 4): Ähnlich einem CAD-System, das den Ingenieur bei technischen Konstruktionsaufgaben unterstützt, stellt diese Analyseumgebung Hilfsmittel zur Verfügung, die beim Design von Konjunkturanalysen Routineaufgaben übernehmen und die Konsistenz des Ergebnisses überwachen. Die hierfür bereitstehenden Softwareassistenten sind als teilautonome Agenten gestaltet, deren Grundkonzept im Anhang näher erläutert wird. Dieser kann von weniger technikinteressierten Lesern übergangen werden, da für die Bedienung von DIWAX keinerlei Programmierkenntnisse erforderlich sind. Der Schlussabschnitt bietet einen Ausblick über geplante Erweiterungen.

Die Software ist für den internationalen Einsatz konzipiert ist. Daher sind sowohl der Quellcode als auch die Bedienungsoberfläche in englischer Sprache verfasst (alle ökonomischen Fachbegriffe können jedoch auch in beliebig vielen anderen Sprachen hinterlegt werden). Aus diesem Grunde tragen alle im Folgenden vorgestellten Fachkonzepte englische Bezeichnungen.

## 2 DIWAX als Metadatenbank zur Zeitreihenorganisation

### *Zeitreihen und ökonomische Aktivität*

Datensammlungen zur Konjunkturanalyse dienen der Abbildung ökonomischer Aktivität (inklusive der relevanten Begleitumstände) im Zeitablauf. Die Beobachtung, Erklärung und Prognose der verschiedenen, sich typischerweise im Zeitablauf ändernden Aktivitätszustände steht im Mittelpunkt der Betrachtung. Hingegen ist eine einzelne Zeitreihe lediglich das technische Vehikel zur Messung einer speziellen Facette einer bestimmten Aktivität, die typischerweise erst durch eine Vielzahl von miteinander zusammenhängenden Zeitreihen sinnvoll erfasst und interpretiert werden kann. So gehört zum Beispiel zur Abbildung der privaten Konsumaktivität grundsätzlich eine nominale und eine preisbereinigte Zeitreihe nebst zugehörigem Deflator. Werden alternative Preisbereinigungsverfahren verwendet, so erhöht sich die Anzahl an Volumen- und Deflatorenzeitreihen entsprechend. Ferner können alle diese Zeitreihen in unterschiedlichen Frequenzen auftreten, nach verschiedenen Verfahren kalender- und ggfs. saisonbereinigt sein, als Absolutwerte, Indizes oder Veränderungsraten auftreten usw. Leicht können so Zeitreihenmengen im drei- und vierstelligen Bereich entstehen. Die inhaltliche Klammer, die den Zusammenhang zwischen den verschiedenen Zeitreihen herstellt, ist die ökonomische Aktivität selbst. Diese steht daher im Vordergrund der Datenorganisation mit DIWAX. Hingegen treten die einzelnen Zeitreihen – anders als bei der Datenverwaltung mit Standarddatenbanksystemen oder Tabellenkalkulationsprogrammen – für den Anwender zunächst in den Hintergrund. Diese werden primär nach ökonomischen und sekundär nach technischen Kriterien so strukturiert, dass sie jeweils nach Bedarf und Analysekontext verfügbar sind; sie stehen damit nicht einfach seriell nebeneinander, sondern sie sind auch für die Software selbst in ihrer jeweiligen Erscheinungsform identifizierbar, was die Übertragung vieler Routineaufgaben (z. B. Preis- oder Saisonbereinigung, Ver- und Entkettung von Volumengrößen auf Vorjahrespreisbasis, Berechnung von Wachstumsbeiträgen etc.) auf den Rechner vereinfacht.<sup>1</sup> Da das System „weiß“, mit welcher Datenkategorie es gerade zu tun hat, kann es auch selbständig entscheiden, welche Operationen in diesem Zusammenhang mathematisch erlaubt bzw. ökonomisch sinnvoll sind und welche nicht.

### *Semantische Datenbeschreibung*

Kennzeichnend für die Funktionsweise von DIWAX ist daher die strikte Trennung konkreter Zeitreihen von ihrer – unabhängig von konkreten Zahlenwerten gültigen – inhaltli-

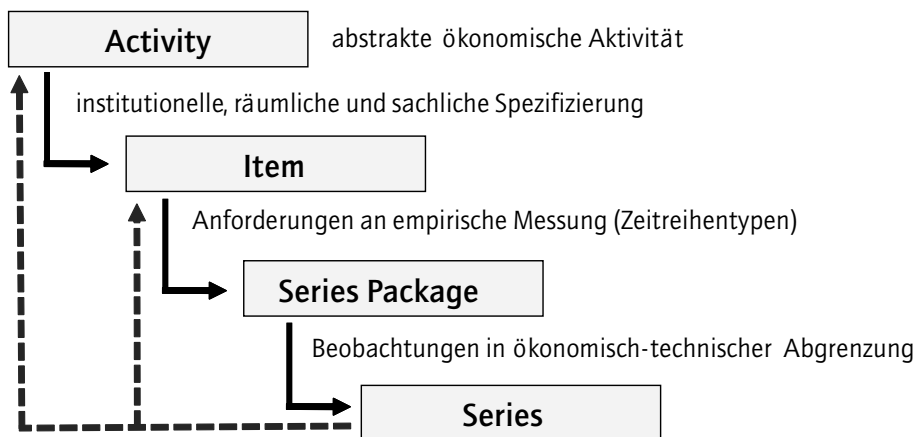
<sup>1</sup> Die Dimension der relevanten Datenbestände zur Konjunkturbeobachtung und die zahlreichen Sichten, die zur adäquaten Analyse auf diese Daten erforderlich sind, lassen tabellenkalkulationsbasierte Lösungen schnell an ihre praktischen Grenzen stoßen. Insbesondere erweist sich das System der relativen Zellbezüge zur Automatisierung von Standardaufgaben als wenig flexibel (z.B. keine Umkehrung der Flussrichtung) und die Komplexität des über verschiedene Arbeitsblätter und Dateien verstreuten Beziehungsgeflechts als äußerst fragil.

chen Kennzeichnung nach Kategorien, die für die ökonomische Auswertung relevant sind. Diese vollzieht sich auf mehreren Ebenen und nach verschiedenen Gesichtspunkten, wobei die jeweils allgemeineren Merkmale auf einer höheren und damit zugleich abstrakteren Ebene festgelegt werden. Alle für die Datenbeschreibung erforderlichen Merkmalsausprägungen sind in einer Metadatenbank zusammengefasst, die – wo immer dies möglich ist – auf standardisierte Klassifikationssysteme offizieller Organisationen (z. B. NACE Rev.1 zur Klassifikation von Wirtschaftszweigen) zurückgreift. Parallel können auch alternative Klassifikationssysteme hinterlegt werden, um sich den jeweiligen Besonderheiten eines Landes oder Wirtschaftsraumes anzupassen.

Die primäre (ökonomische) und sekundäre (technische) Klassifikation der Datenbestände vollzieht sich von der abstrakten ökonomischen Aktivität hin zur konkreten Zeitreihe über 4 Stufen (Abbildung 2), die im Folgenden näher erläutert werden.

Abbildung 2

### Datenbeschreibungsebenen



Quelle: Eigene Darstellung.

#### *Von der Activity zum Item*

Den Ausgangspunkt bildet die abstrakte Activity-Ebene, die eine ökonomische Aktivität allgemein inhaltlich beschreibt, ohne sie bereits genauer zu verorten. Der Begriff der ökonomischen Aktivität ist dabei weit zu interpretieren. Er umfasst insbesondere die Abwicklung von Transaktionen, das Halten/Vorhandensein von Beständen und die Bewertung von Gütern, Leistungen oder Vermögenstiteln. Beispiele hierfür sind „Warenexporte“, „Arbeitnehmer“, „Verbraucherpreise“ oder „Auftragseingänge“. Auf dieser Ebene ist bereits festgelegt, ob es sich um eine Strom- oder Bestandsaktivität handelt (StockFlowType) und welcher Art die Einheit ist (UnitType), in der diese Aktivität grundsätzlich gemessen werden kann (z.B. Währung, Zeit oder Personen). Mittels der ActivityType-Eigenschaft wird darüber hinaus eine grobe Zuordnung getroffen, die bei der späteren Bereitstellung konkreter Zeitreihen hilfreich ist, um redundante Klassifizierungen zu vermeiden (Tabelle 1).

Tabelle 1

**Ausprägungen des Merkmals ActivityType**

| <b>Ausprägung (Code)</b> | <b>Beschreibung</b>        |
|--------------------------|----------------------------|
| Value                    | Monetäre Wertgröße         |
| Quantity                 | Nicht monetäre Mengengröße |
| Price                    | Bewertung, Preismessung    |
| Ratio                    | Abgeleitete Kennzahl       |

Quelle: Eigene Darstellung.

Durch eine räumliche, institutionelle und sachliche Spezifizierung wird aus einer allgemeinen Activity ein konkretes Item, das bei der Arbeit mit DIWAX eine zentrale Rolle einnimmt. So wird zum Beispiel aus der allgemeinen Activity „Konsum“ das konkrete Item „Konsum [Activity] der Gebietskörperschaften [Sector] in Deutschland [Region]“, das sich nur durch die institutionelle und räumliche Spezifizierung vom „Konsum der privaten Haushalte in Bayern“ unterscheidet.<sup>2</sup> Neben der sektoralen und regionalen Kennzeichnung kann zusätzlich noch der beteiligte Wirtschaftszweig (Industry) und die Gütergruppe (Product) näher spezifiziert werden (etwa „Konsum [Activity] von Dienstleistungen [Product] der Nachrichtenübermittlung [Industry] durch die privaten Haushalte [Sector] in der Eurozone [Region]“). Wird eine bestimmte Eigenschaft nicht explizit belegt, so gilt jeweils die weitestmögliche Abgrenzung (alle Wirtschaftszweige, alle Gütergruppen etc.). Die Merkmale Industry, Sector und Region können darüber hinaus nach Leistungs- und Empfängerseite (Aufkommen vs. Verwendung) bzw. nach Gläubiger- und Schuldnerposition (Aktiva und Passiva) differenziert werden.<sup>3</sup> Schließlich wird mit dem Übergang von der Activity zum Item auch die konkrete Einheit (Unit) spezifiziert (z.B. Stunden als Zeit- oder Euro als Währungsmaß).

*SeriesPackage und SeriesTypes*

Innerhalb eines Items bilden alle zugehörigen Daten das SeriesPackage, innerhalb dessen die verschiedenen Zeitreihen nach den folgenden ökonomisch-technischen Kriterien hierarchisch über die als Blockköpfe dargestellten Eigenschaften in Abbildung 3 geordnet sind.

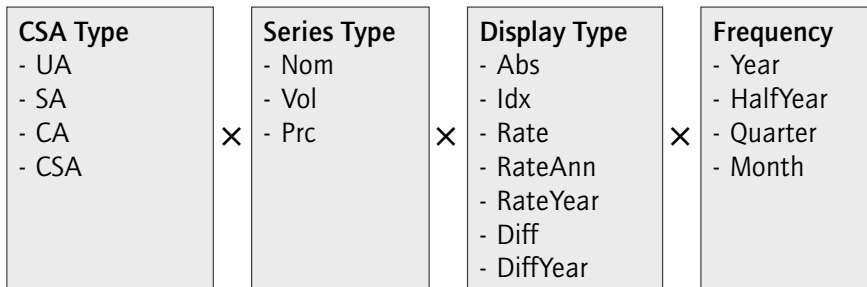
Dies ermöglicht einen baumartig strukturierten Zugriff auf sämtliche Zeitreihen eines Items. So gelangt man beispielsweise zur saison- und kalenderbereinigten annualisierten Quartalswachstumsrate des realen privaten Konsums über den im SeriesPackage des entsprechenden Items abgebildeten Zeitreihenbaum: [Item „Privater Konsum ...“]. CSA.Vol.RateAnn.Quarter.

<sup>2</sup> Klassifikationssysteme bieten typischerweise eine hierarchisierte Struktur der verschiedenen Merkmale. Dies erlaubt es der Software, im vorliegenden Beispiel selbständig zu erkennen, dass Bayern ein Teil von Deutschland ist, was u. a. für die automatisierte Ermittlung bestimmter Anteilswerte oder einen Abgleich zwischen Bundesländerdaten und nationalen Daten genutzt werden kann.

<sup>3</sup> Die Unterscheidung in Aufkommens- und Verwendungsseite kann analog auch auf den Faktor Arbeit übertragen werden, um zwischen Arbeitnehmern nach Inlands- und Inländerkonzept zu differenzieren.

Abbildung 3

### Aufbau eines SeriesPackage



Quelle: Eigene Darstellung.

Eine zentrale Rolle nimmt dabei die Eigenschaft `SeriesType` mit den drei Basisausprägungen `Nom` (Wertgröße in laufenden Preisen), `Vol` (preisbereinigte Volumengröße) und `Prc` (Deflator) ein, deren Vorkommen von der `ActivityType`-Ausprägung der zugrunde liegenden `Activity` abhängt: Für `Value`-Items mit generischem Deflator treten sie gemeinsam als Trio auf, für alle übrigen kommen sie in einem `SeriesPackage` nur exklusiv vor. `Quantity`-Items (z.B. Arbeitnehmer, Auftragsreichweiten oder Kfz-Zulassungen) benötigen nur die Mengenkomponeute `Vol`. Bei `Value`-Items mit generischem Deflator dienen die drei Basistypen dazu, die Nominalgröße `Nom` in den generischen Deflator `Prc` und die preisbereinigte Volumenkomponente `Vol` aufzuspalten. `Value`-Items ohne eigenen Deflator (z.B. Staatseinnahmen) weisen nur `Nom`-Zeitreihen auf, während reine Preis-Items (z.B. Verbraucherpreise, Aktienkurse, Wechselkurse oder Zinsen) außer den `Prc`-Komponenten keine anderen Basistypen benötigen. Die zur Abbildung von Kennzahlen und Quoten verbleibenden `Ratio`-Items (z.B. Arbeitsproduktivität, Sparquote) verwenden definitionsgemäß den Basistyp `Vol`, um ihre Informationen abzubilden. Dies gilt konsequenterweise auch für solche Items, die aus einer nicht-generischen Preisbereinigung hervorgehen (z.B. Reallöhne).

Auf der Grundlage der `ActivityType`-Eigenschaft (und ggfs. anhand der Existenz eines generischen Deflators, die ebenfalls auf Item-Ebene kodiert wird), kann somit jedes Item entscheiden, welche `SeriesTypes` es innerhalb seines `SeriesPackage` zur Verfügung stellen muss.

Neben der programminternen Nutzung der semantischen Datenbeschreibung, die in Abschnitt 4 näher dargestellt wird, ermöglicht die systematische Klassifizierung der Datenbestände auch einen standardisierten Datenaustausch mit anderen Anwendungen.

Die Metadatenbank kann lokal von individuellen Nutzern weitergeführt und auf Wunsch automatisch mit der offiziellen DIW-Version, die über einen Internetserver bereitgestellt wird, abgeglichen werden. Dies stellt die Konsistenz der Datenbank sicher und ermöglicht eine dezentrale Erweiterung des Item-Vorrats. DIWAX stellt sämtliche Hilfsmittel bereit, um innerhalb der Programmumgebung die Metadatenbank zu verwalten.

Tabelle 2

**ActivityType und SeriesTypes**

| <b>ActivityType</b> | <b>SeriesType(s)</b>  |
|---------------------|---|
| Value               | Generischer Deflator: NOM, VOL, PRC<br>reine Wertgröße: NOM |
| Quantity            | VOL   |
| Price               | PRC   |
| Ratio               | VOL   |

Quelle: Eigene Darstellung.

### 3 Projektverwaltung, Ressourcen und Datenaktualisierung

Während der Aufbau bzw. die Erweiterung der Metadatenbank einmalig bzw. unregelmäßig erfolgt, findet die laufende Arbeit mit den definierten Items innerhalb von sogenannten Projekten (Projects) statt. Ein Projekt bezieht sich typischerweise auf einen bestimmten Prognosezeitpunkt und legt neben administrativen Angaben einheitliche Einstellungen für alle Items fest, die für dieses Projekt zum Einsatz kommen. Hierzu gehören das Projektverzeichnis, das Preisbereinigungsverfahren, das Saison- und Kalenderbereinigungsverfahren, der Analysezeitraum (getrennt nach Ex-post- und Ex-ante-Bereich), ein Referenzjahr sowie die Ressourcenliste, die im Folgenden kurz erläutert werden.

Obwohl sämtliche Projektdaten in einer singulären Excel-kompatiblen Datei gespeichert werden, sollte für jedes Projekt ein eigenes Verzeichnis bestimmt werden. DIWAX legt dann in diesem Projektverzeichnis automatisch ein Input- und ein Outputunterverzeichnis an, die für den Datenaustausch mit anderen Programmen bzw. für die Ablage automatisierter Aufbereitungen (z.B. VGR-Tableaus oder Verlaufsgrafiken) verwendet werden können. Die Projektdatei enthält neben einer Metadata-Tabelle, in der sämtliche Projekteinstellungen gespeichert werden, für jedes Item ein gesondertes Tabellenblatt. Im Tabellenkopf aller Zeitreihenblätter ist die jeweilige Zeitreihe über ihre individuellen Merkmale eindeutig beschrieben. Durch diese offene, explizit dokumentierte Form der Datenhaltung sind DIWAX-Projektdateien keine Insellösungen, sondern können problemlos von menschlichen Betrachtern und anderen Anwendungen dekodiert werden.

Im Zentrum der inhaltlichen Einstellungen für ein Projekt steht der Projektzeitraum, der über den Prognosezeitpunkt in einen Ex-post- und einen Ex-ante-Bereich aufgeteilt wird, wobei jede Periode, die in den Ex-ante-Bereich hineinragt (beispielsweise Jahreswert 2007 und Prognosezeitpunkt 1.10.2007) automatisch dem Ex-ante-Bereich zugerechnet wird.

Darüber hinaus wird innerhalb eines Projektes für alle Items, die eine generische Preisbereinigung zulassen, ein einheitliches Verfahren festgelegt. Ferner gilt innerhalb eines Projektes für alle Items ein einheitliches Saison- und Kalenderbereinigungsverfahren.

Innerhalb von DIWAX können mehrere Projekte gleichzeitig bearbeitet und miteinander verglichen werden. Diese können sich hinsichtlich der jeweiligen Einstellungen beliebig

voneinander unterscheiden. Eine typische Anwendung besteht z.B. in der nachträglichen Überprüfung von Prognosen oder in der schnellen Aufbereitung von Ergebnisabweichungen, die sich aus Revisionen der statistischen Datengrundlage ergeben.

Die konkreten Zeitreihenwerte für ein Projekt (sofern sie der Benutzer nicht selbst eingeben oder ändern möchte) werden über so genannte Ressourcen (Resources) eingebunden. Eine Ressource kennzeichnet eine externe Datendatei (typischerweise im Excel- oder CSV-Format), deren Inhalt und Struktur (Datenanordnung)<sup>4</sup> innerhalb von DIWAX hinterlegt werden kann. Für deutsche VGR-Daten stellen die als Excel-Dateien verfügbaren Veröffentlichungen der Fachserie 18 des Statistischen Bundesamtes typische Beispiele für Ressourcendateien dar. Solange die Struktur einer solchen Datendatei im Zeitablauf nicht verändert wird (sondern die Zeitreihen nur der bisherigen Systematik folgend fortgeschrieben werden), genügt es, einen Ressourcentyp einmalig zu definieren, um ihn dann nach jeder Aktualisierung neu einlesen zu können. Hierzu wird – ebenfalls einmalig – für jedes Item, das eine oder mehrere Zeitreihen aus einer Ressource bezieht, die Position innerhalb der Datendatei festgelegt. Sofern Zeitreihen innerhalb der Ressourcendatei eindeutig kodiert sind, kann auch alternativ der Zeitreihencode hinterlegt werden.<sup>5</sup> Ändert sich die Struktur einer Datei systematisch (z. B. werden vom Datenanbieter alle Zeitreihen um eine oder mehrere Zeilen oder Spalten verschoben), so genügt es, dies in der Ressourcendefinition zu vermerken. Alle übrigen Einstellungen können dann weiterhin genutzt werden.

Es lassen sich beliebig viele Ressourcen in DIWAX einbinden und über den ResourceManager verwalten. Für jeden Ressourcentyp wird im DIWAX-Datenverzeichnis ein eigenes Unterverzeichnis erstellt, in das fortwährend Aktualisierungen eingestellt werden können. Jede Ressourcendatei wird dabei anhand des Zeitraumcodes für die letzte Export-Periode (z.B. 2007-Q3 oder 2007-H1) identifiziert (übrige Namensbestandteile spielen keine Rolle und können daher frei gewählt werden, z.B. „2007-Q3 Fachserie 18 Reihe 1.2.xls“ oder „Staatskonto 2007-H1.xls“). Auf diese Weise werden alle externen Daten, die einer bestimmten Prognose zugrunde liegen, dauerhaft archiviert, so dass sie auch für spätere Analysen (etwa zur Bedeutung von Datenrevisionen für das Prognoseergebnis) genutzt werden können (automatische Versionierung der Daten).

Sofern Datenanbieter ihre Daten öffentlich über das Web zugänglich machen, können diese über DownloadAgents innerhalb von DIWAX direkt in die entsprechenden Ressourcenverzeichnisse transferiert werden. Hierzu ist nur einmalig der Standort und die jeweilige Bezeichnungsweise des Datenanbieters zu hinterlegen. Darüber hinaus besteht die Möglichkeit, den Zeitreihenserver der Deutschen Bundesbank direkt anzusprechen und entsprechende Zeitreihen abzurufen, die DIWAX dann in einer entsprechenden Ressourcendatei zusammenstellt. Ferner steht für das Datenangebot von Eurostat ein weiterer spezialisierter DownloadAgent bereit, der aus den komprimierten Datendateien im Euro-

<sup>4</sup> Hierzu zählt z. B. die Anordnung der Zeitreihen in Zeilen oder Spalten, aber auch komplexere Verteilungen von Zeitreihen über mehrere Arbeitsblätter oder die Abfolge unterschiedlicher Frequenzen innerhalb einer Zeile oder Spalte können hinterlegt werden.

<sup>5</sup> Die Deutsche Bundesbank macht die interne Kodierung ihrer Zeitreihenbestände zugänglich. Die inhaltliche Bedeutung des jeweiligen Zeitreihencodes ist unbedeutend, wichtig ist nur, dass jede Zeitreihe eindeutig benannt ist. Dies erhöht die Flexibilität im Umgang mit diesen Daten für die Nutzer erheblich. Es wäre wünschenswert, wenn auch andere offizielle Datenanbieter diesem Beispiel folgten.



stat-Textformat Excel-Dateien erzeugt, bei denen die Länderdaten spaltenweise auf jeweils einem gesonderten Datenblatt angeordnet werden. Weitere Spezialagenten für sonstige Datenanbieter können jederzeit ergänzt werden. Schließlich lässt sich der Datenabruf über das Hinterlegen von Update-Zyklen automatisieren, sodass der Datenbestand kontinuierlich auf dem aktuellen Stand gehalten wird.

Anhand des Ressourcenkonzeptes und mittels der DownloadAgents kann jeder DIWAX-Nutzer die für die eigenen Bedürfnisse erforderlichen Daten selbst zusammenstellen und jederzeit um weitere Quellen ergänzen. Diese Flexibilität besteht bei den meisten Datenbank Anbietern typischerweise nicht. Darüber hinaus lassen sich über das Ressourcenkonzept auch modellgestützte Prognosedaten reibungslos in ein Projekt einbeziehen, in dem die Modellergebnisse als selbständige Ex-ante-Ressource hinterlegt werden.

#### 4 Analyseumgebung für VGR-gestützte Prognosen

##### *ItemAgents*

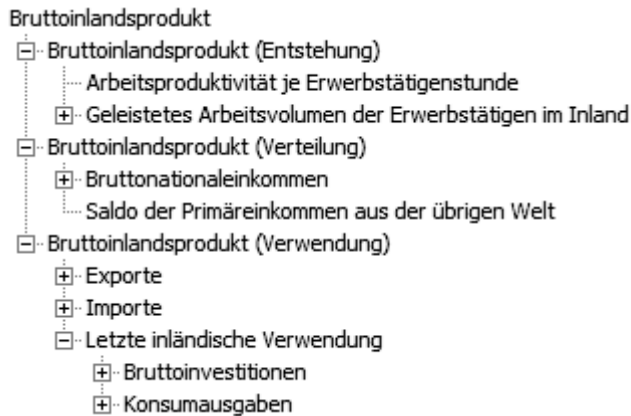
Mit der in der Metadatenbank hinterlegten ökonomischen und technischen Beschreibung der Items einerseits sowie der über das Ressourcenkonzept realisierten Anbindung eines Item-Series Package an konkrete Datenbestände andererseits sind die wesentlichen Bausteine für eine VGR-gestützte Analyseumgebung bestimmt. Innerhalb von DIWAX wird jedes Item einem eigenen ItemAgent übertragen, das die jeweilige ökonomische Aktivität mit konkreten Daten repräsentiert und die ihm übertragenen Aufgaben autonom oder im Verbund mit anderen ItemAgents ausführen kann. Ein ItemAgent „handelt“ somit für das von ihm repräsentierte Item. Hierfür spielt das Verhältnis der Items zueinander eine zentrale Rolle. Ein Item A kann als positives oder negatives SubItem einem oder mehreren anderen Items untergeordnet werden, die dadurch zum SuperItem für A werden. Ein SuperItem kann gegenüber seinen SubItems eines der drei Zustände Sum (additive Verknüpfung), Ratio (multiplikative Verknüpfung) oder Equivalence (Identität) einnehmen.<sup>6</sup> Hierdurch lässt sich eine Baumstruktur aufbauen, mit der die durch einen VGR-Kontenrahmen vorgegebenen Ableitungsschemata nachgebildet werden können.

Am Beispiel der Disaggregation eines Items sei das Zusammenspiel der hinter den Items stehenden ItemAgents kurz erläutert. Enthält ein Item mehrere SubItems, so lässt sich eines davon als aktives SubItem markieren; dieses kann dann vom jeweiligen SuperItem in Abhängigkeit dessen eigener Werte und ggfs. der Belegung der übrigen SubItems angepasst werden. Wie diese Anpassung erfolgt, kann ein SuperItemAgent aus der Datenbeschreibung der beteiligten Items lokal ermitteln. So reagiert im Beispiel aus Abbildung 4 der ItemAgent „Bruttoinlandsprodukt (Verwendung)“ auf den Disaggregationsbefehl, indem er die Werte des Items für die Inländische Verwendung durch Entkettung ermittelt (bei Vorjahrespreisbereinigung), da sämtliche beteiligten Items generische Deflatoren aufweisen. Hingegen wird der ItemAgent „Bruttoinlandsprodukt (Verteilung)“ bei demselben Befehl nur die nominale Komponente mit dem Saldo der Primäreinkommen aus der übr-

<sup>6</sup> Hat ein SuperItem X die beiden positiven SubItems A und B sowie zwei negative SubItems C und D, so ergibt sich X bei additiver Verknüpfung als  $X = A + B - C - D$  und bei multiplikativer Verknüpfung als  $X = (A \cdot B) / (C \cdot D)$ . Bei Identität gilt  $X = A = B = C = D$ .

Abbildung 4

### Ausschnitt aus einem Item-Baum



Quelle: Eigene Darstellung.

gen Welt verrechnen, um das lediglich als Nom-Type vorliegende Item für das Bruttonationaleinkommen zu aktualisieren. Schließlich berechnet der für das Ratio-Item „Bruttoinlandsprodukt (Entstehung)“ zuständige Agent die Arbeitsproduktivität durch Auflösung der hinterlegten Verhältnisformel. Das übergeordnete Identitäts-Item „Bruttoinlandsprodukt“ hingegen sorgt lediglich dafür, dass alle SubItems identische Werte aufweisen.

Aggregations-, Disaggregations- und Synchronisierungsaufgaben können entweder lokal (zwischen SuperItem und den direkten SubItems) oder über ganze Ableitungsbäume hinweg durchgeführt werden.

#### *Verwaltungsagenten*

Obwohl die Datenorganisation über inhaltlich abgegrenzte Items den Umfang der Zeitreihenbestände, mit denen der Nutzer konfrontiert wird, bereits drastisch reduziert, kann gleichwohl die Anzahl der für eine vollständige VGR-gestützte Analyse zu definierenden Items unübersichtlich werden. Nicht immer ist jedoch ein Analyst an allen Items gleichzeitig interessiert. Dies gilt insbesondere für die jeweiligen Bereichsspezialisten. So mag für den Europafachmann ein Blick auf die groben Verwendungsaggregate der verschiedenen Länder zusammen mit einigen zentralen Entstehungs- und Verteilungsgrößen relevant sein, nicht aber ein vollständiger Kontenrahmen für alle 27 EU-Länder mit jeweils über 200 Items. Andersherum mag ein Finanzexperte für Deutschland zwar an einer sehr detaillierten Darstellung des deutschen Staatskontos und dessen wichtigster Determinanten (Einkommen und Beschäftigung) interessiert sein, nicht aber an den Dienstleistungsexporten Schwedens. Um die Menge der Arbeits-Items auf den gerade interessierenden Analyseschwerpunkt zu begrenzen, können über die ThemeAgents beliebig viele Themen definiert (und miteinander kombiniert) werden, die die Grundgesamtheit aller Items nach bestimmten Kriterien und individuellen Setzungen filtern. Sämtliche Analysewerkzeuge des Programms passen sich dabei der jeweiligen Menge aktivierter Items an, insbesondere

werden die Item-Systeme, die aus der Hierarchie von Sub- und SuperItems entstehen, dynamisch angepasst. Die Klassifizierung der Items nach ökonomischen Kriterien erleichtert es dabei, ganze Gruppen von Items zu- oder abzuschalten.

Ebenfalls der Übersichtlichkeit dienen die TabletAgents, mit denen sich individuell konfigurierbare Tableaus erstellen lassen. Diese fassen eine Auswahl an Items blockweise mit einheitlicher Datensicht (z.B. reale Wachstumsraten, absolute nominale Veränderung zur Vorjahresperiode etc.) für beliebige Frequenzkombinationen zusammen. Die in den DIW-Vierteljahresprognosen enthaltenen Tabellen „Die wichtigsten Daten der Volkswirtschaftlichen Gesamtrechnung für Deutschland“<sup>7</sup> sind ein Beispiel für ein Tablet, das sich in DIWAX mit wenigen Mausklicks zusammenstellen lässt, um den gegenwärtigen Stand der Analyse im Überblick darzustellen.

Für die Anbindung an externe Programme zur Datenweiterverarbeitung lassen sich sämtliche Größen über TransferAgents im Excel- oder CSV-Format exportieren.

### *Analyseagenten*

Zur Unterstützung der Analyse- und Prognosetätigkeit stehen JobAgents zur Verfügung, mit denen sich Arbeitsabläufe – ähnlich einer Makro-Funktion – automatisieren lassen. Über einen JobAgent können die von einem ItemAgent veröffentlichten Befehle angesprochen und – falls erforderlich – situativ parametrisiert werden. JobAgents unterstützen insbesondere unterschiedliche Rundrechnungsstile (Ableitungsabfolgen und quasiendogene Größenfestlegung).

Darüber hinaus stehen spezialisierte Sonderagenten zur Verfügung, deren Zahl zukünftig noch erweitert werden kann. Bislang zählen hierzu die SplitterAgents, mit denen sich niedrige in höhere Frequenzen aufspalten lassen, sowie die ShareSatellites, die für eine beliebige Zusammenstellung von Items Anteile und Wachstumsbeiträge berechnen. Weitere Sonderagenten zur Konsistenz- und Auffälligkeitsprüfung sind bereits in der Entwicklung.

## 5 Erweiterungen

### *Schnittstelle für Modellsimulationen (im Aufbau)*

Moderne makroökonomische Prognosen basieren auf den von ökonometrischen Modellen berechneten Vorlagen, die dann typischerweise um Expertenwissen ergänzt oder für Präsentationszwecke geglättet werden. Oft wird erst während einer intensiven VGR-Analyse deutlich, ob die Modellergebnisse plausibel sind oder ob für bestimmte Komponenten über Add-Faktoren weiteres Expertenwissen, das sich für das Modell nicht aus der beobachteten Vergangenheit erschließen konnte,<sup>8</sup> in die Prognose einfließen muss. Die enge, iterative Verbindung zwischen Modell- und Expertenprognosen legt es nahe, die ökonomischen

<sup>7</sup> Vgl. z.B. Wochenbericht des DIW Berlin, Nr. 42/2007, 617 ff.

<sup>8</sup> Die Abschätzung der mit der Mehrwertsteuererhöhung zum 1. Januar 2007 verbundenen Vorzieheffekte sind ein Beispiel für die Grenzen rein ökonometrischer Abschätzung relevanter wirtschaftspolitischer Eingriffe.

metrische Lösung direkt aus DIWAX heraus zu steuern; hierzu zählt neben der Vorgabe von Politikvariablen (z.B. Modellierung der durch die Fiskalpolitik vorgegebenen Ausgabepfade) auch die Konfiguration von Prognoseszenarien (Zu- und Abschaltung von Modellblöcken, Endo- oder Exogenisierung bestimmter Variabler), bei denen die Modellintegrität gewahrt bleibt. Derzeit wird prototypisch an der Integration des in Eviews implementierten DIW-Konjunkturmodells gearbeitet. Auch hierfür wird ein agentenbasiertes Verfahren (vgl. Anhang) entwickelt, das sich dann leicht auf beliebige andere Modelle übertragen lässt. Auf diese Weise wird die Durchrechnung von Alternativszenarien und Politiksimulationen erleichtert. Darüber hinaus reduziert die klare und systematische Trennung von Modellentwicklung und -anwendung mögliche Fehlerquellen und erleichtert sowohl die Datenvorgabe als auch die Ergebnisaufbereitung der in dieser Hinsicht meist weniger gut ausgestatteten Statistikpakete.

### *Sockel für Methoden der Künstlichen Intelligenz (in Vorbereitung)*

Eine strukturierte Datenumgebung ist auch für nicht-ökonometrische Verfahren zur Konjunkturanalyse aus dem Bereich der Künstlichen Intelligenz eine wichtige Grundlage. Hierzu zählen insbesondere Techniken des Soft Computing wie Neuronale Netze, Fuzzy Logik und Genetische Algorithmen.<sup>9</sup> Auch multikriterielle Entscheidungsverfahren (z.B. der Analytic Network Process)<sup>10</sup>, mit denen sich unscharfes, aber prognoserelevantes Expertenwissen in numerische Signale übersetzen lassen, stellen eine vielversprechende Bereicherung des konjunkturanalytischen Instrumentenkastens dar. Für diese Ansätze sind derzeit kaum standardisierte Softwareumgebungen verfügbar. Daher soll DIWAX künftig auch als Sockel genutzt werden, um diese Verfahren an die Bedürfnisse der Konjunkturforschung anzupassen, zu erproben und im Erfolgsfalle bedienerfreundlich zu implementieren.

### *Verbindung mit dem CAL-System ACTAS (in Planung)*

Als Beitrag für die computergestützte Lehre steht Interessenten, die sich in das Gebiet der Volkswirtschaftlichen Gesamtrechnung einarbeiten, das Softwaresystem ACTAS zur Verfügung,<sup>11</sup> in dem der vollständige ESVG Kontenrahmen sowie die Zahlungsbilanzsystematik als Buchungssysteme implementiert sind. Die Versorgung von ACTAS mit empirischen Daten soll zukünftig über die DIWAX-Systematik realisiert werden. Umgekehrt soll es in DIWAX demnächst möglich sein, alle Ergebnisse auch in Form einer dem ESVG entsprechenden Kontendarstellung aufzubereiten.

**9** Vgl. z.B. S. Kooths und E. Ringhut (2003): Modelling Expectations with GENEFER – an Artificial Intelligence Approach. *Computational Economics*, 21, 173–194.

**10** Vgl. T.L. Saaty (2001): *The Analytic Network Process: Decision Making with Dependence and Feedback*. Pittsburgh.

**11** Zu beziehen über <http://actas.uni-muenster.de>; vgl. S. Kooths: Using MAKROMAT and ACTAS for Undergraduate Teaching in Economics; 79<sup>th</sup> Annual Conference of the Western Economic Association International, Jun./Jul. 2004, Vancouver.

## Anhang

### Softwarearchitektur (Konstruktionsprinzipien)

DIWAX besteht technisch aus verschiedenen Komponenten, die jeweils auf einer von vier aufeinander aufbauenden Schichten (Tiers) angeordnet sind (Abb. A1).

Abbildung A1

#### Schichtenmodell



Quelle: Eigene Darstellung.

Die Datenbankschicht enthält sämtliche Ausprägungen der listenweise definierten und für die ökonomische Analyse bedeutsamen Variablen. Die in der Datenbank abgelegten Werte sind vergleichbar mit dem Wortschatz einer Sprache. Durch die Auslagerung in eine separate Datenbank können die Listenwerte flexibel verwaltet und erweitert werden, ohne in den Programmcode eingreifen zu müssen. Die relational aufgebaute Datenbank kann technisch wahlweise als Excel-Workbook oder als SQL-Datenbank vorliegen. Die Verwaltung (Erweiterung, Modifikation, Löschung) und der Zugriff auf die in der Datenbank gespeicherten Listen erfolgt über die Proxy-Schicht. Auf diese Weise ist die Fachkonzeptebene, die alle für die inhaltliche Arbeit des Programms wesentlichen Komponenten enthält (die Grammatik der Sprache), völlig unabhängig von den technischen Aspekten der zugrundeliegenden Datenbank, da die ALT-Schicht nur über die Proxies als Makler auf die Datenbestände zugreift. Wie und woher die Proxies die Daten beziehen, ist den ALT-Komponenten gleichgültig. Schießlich setzt auf die Fachkonzeptebene die Benutzerschnittstelle (Programmoberfläche) auf. Benutzerführung und Reaktion auf Benutzeranweisungen sind somit voneinander getrennt. Grundsätzlich bauen alle Schichten sukzessive aufeinander auf. Eine Ausnahme davon bilden lediglich die Proxies, die in ihrer Zweitfunktion als Datenbankverwalter (neben ihrer Hauptaufgabe der Datenbereitstellung für die Fachkonzepte) ebenfalls Zugriff auf die GUI-Schicht haben.

Mit Ausnahme der relational aufgebauten Datenbank folgen alle Komponenten der übrigen Schichten dem objektorientierten Paradigma. Dieses zeichnet sich vor allem durch folgende Wesensmerkmale aus:

- Kapselung und Geheimnisprinzip
- Klassendefinition und Instanziierung
- Vererbung und Polymorphismus
- Ereignisorientierte Ablaufsteuerung

Ein Objekt ist ein informatisches Konstrukt, bei dem Daten mit den für ihre Bearbeitung erforderlichen Methoden zusammengefasst (gekapselt) werden. Dies ermöglicht eine dezentrale und kooperative Arbeitsweise: Nach außen, d.h. gegenüber anderen Objekten, braucht ein Objekt über seine Schnittstellen (Slots) nur offenzulegen, welche Aufgaben es übernehmen kann, nicht aber, wie es diese Aufgaben erledigt (Geheimnisprinzip). Ein Objekt kann man sich deshalb wie einen Agenten vorstellen, der mit einer bestimmten Aufgabe eigenverantwortlich beauftragt wird. Seine Aufgaben kann dieser entweder völlig selbständig lösen, oder er zieht andere Agenten hinzu. Da das Bild eines informierten, handelnden Agenten die Rolle der zentralen Objekte eines Programms besser zum Ausdruck zu bringen vermag als die eher passiv anmutende Bezeichnung eines Objekts, treten in DIWAX die Objekte der ALT-Schicht dem Nutzer gegenüber als Agenten in Erscheinung.

Der Gattungstyp eines Objektes wird „Klasse“ genannt. Von einer solchen Klasse können beliebig viele Instanzen erzeugt werden. Jede einzelne Instanz verfügt dann über die durch die Klassendefinition bestimmten Fähigkeiten. Die einzelnen Individuen einer Instanzpopulation (die Menge der Agenten gleicher Bauart) unterscheiden sich nur in der Belegung ihrer jeweiligen Slotbelegung. Bei der Definition einer Klasse kann auf einer bereits bestehenden Klasse (Vorfahrklasse) aufgebaut werden, dessen Funktionalität den Nachkommen vererbt wird. Diese können dann auf dieser Grundlage neue Funktionen hinzufügen oder bestehende neu interpretieren. Dies ermöglicht den Aufbau einer Klassenhierarchie, bei dem gemeinsame Funktionen weit oben im Stammbaum allgemeingültig formuliert werden, während die Nachkommen Spezialverhalten entsprechend ausdifferenzieren können. Auf diese Weise wird erreicht, dass verschiedene Nachkommen, auf formal identische Anweisungen ihrer jeweiligen spezifischen Ausprägung entsprechend unterschiedlich reagieren, ohne dass dies dem Auftraggeber (dies kann entweder ein anderes Objekt oder der Benutzer selbst sein) bekannt bzw. bewusst sein muss (Polymorphismus). Schließlich sorgt die Steuerung über Ereignisse dafür, dass ein Programm nicht schematisch nach vorgegebenen Abfolgen abläuft, sondern sich situativ nach dem Anfall bestimmter Aufgaben richtet. Hierzu tauschen Objekte untereinander Botschaften aus, die es ihnen ermöglichen, eine vorgegebene Problemstellung kooperativ einer Lösung zuzuführen.