

Die Evaluation von Artefakten: Eine zentrale Herausforderung der Wirtschaftsinformatik

Ulrich Frank

Universität Koblenz-Landau, Institut für Wirtschaftsinformatik
Rheinau 1, D-56075 Koblenz

„There are potentially at least as many ways of dividing up the world into object systems as there are scientists to undertake the task.“

Clyde H. Coombs, Howard Raiffa, Robert M. Thrall

"Wir wollen in unserem Wissen vom Gebrauch der Sprache eine Ordnung herstellen: eine Ordnung zu einem bestimmten Zweck; eine von vielen möglichen Ordnungen, nicht die Ordnung."

Ludwig Wittgenstein

Inhalt

- 1 Evaluation: Ein vielschichtiges Thema
- 2 Artefakte in der Wirtschaftsinformatik
- 3 Evaluation als inhärenter Bestandteil wissenschaftlicher Forschung
- 4 Evaluation im Kontext der konzeptionellen Modellierung
- 5 Zum Stand der einschlägigen Evaluationsforschung
- 6 Abschließende Bemerkungen

Zusammenfassung

Der Gegenstand der Wirtschaftsinformatik ist wesentlich durch Artefakte geprägt. Das gilt einerseits für den Anwendungsbereich der Disziplin, andererseits ist die Forschung selbst auf die Konstruktion und Beurteilung von Artefakten, wie etwa von Modellen und Prototypen, gerichtet. Der vorliegende Beitrag untersucht die spezifischen Schwierigkeiten, die mit einer wissenschaftlichen Beurteilung informationstechnologischer Artefakte verbunden sind. Ausgehend von einem einführenden Überblick über das Phänomen Evaluation und die Bedeutung von Artefakten in der Wirtschaftsinformatik wird die Bewertung von konzeptionellen Modellen und zugehörigen Sprachen näher betrachtet. Dabei zeigt sich, daß Maßstäbe, wie sie von der Wissenschaftstheorie zur Beurteilung von Forschungsergebnissen vorgeschlagen werden, allenfalls eingeschränkt auf die Evaluation von Modellen übertragen werden können. Vor diesem Hinter-

grund gibt der Beitrag einen Überblick über die einschlägige Evaluationsforschung und skizziert zukünftige Herausforderungen.

Abstract

Artefacts are of pivotal importance for the information systems discipline. On the one hand, they constitute a major part of the domain that is subject of information systems research. On the other hand, research is producing artefacts, such as models and prototypes, itself. This paper provides an investigation of the specific problems that are related to a rational evaluation of information technology artefacts. After a general overview of evaluation and the role of artefacts in information systems research, we will focus on conceptual models and corresponding modelling languages. It will be shown that general epistemological criteria to evaluate research results are of limited help only when it comes to evaluate models. Against this background, the paper outlines the current state of the art in evaluating conceptual models and modelling languages as well as future challenges.

1 Evaluation - ein vielschichtiges Phänomen

Evaluation wird gemeinhin verstanden als die gezielte Bewertung von materiellen oder immateriellen Gegenständen unter Rückgriff auf Kriterien und Verfahren, deren Angemessenheit erläutert oder begründet werden sollte (vgl. [Hou93], S. 1). In dieser ersten Begriffsabgrenzung deuten sich schon die Schwierigkeiten an, die mit der Durchführung von Evaluationen verbunden sind. Gleichzeitig ist es offensichtlich, daß Evaluationen - auch wenn der Begriff eher selten verwendet wird - in vielfältiger Form nahezu alle Bereiche moderner Industriegesellschaften durchdringen. Sie sind typischerweise auf facettenreiche Sachverhalte gerichtet, also solche, die sich gegen eine einfache Erfassung wesentlicher Merkmale sperren. Sie reduzieren damit, jedenfalls vordergründig, Komplexität. Evaluationen dienen der Vorbereitung und nicht zuletzt der Legitimation von Entscheidungen. Es ist deshalb wenig überraschend, daß sie auch zur Instrumentalisierung von Interessen verwendet werden.

Evaluation ist - jedenfalls dem Anschein nach - durch das Bemühen um *Objektivität* gekennzeichnet. Schon bei oberflächlicher Betrachtung liegt es auf der Hand, daß die Chance, Evaluationen durchzuführen, die gemeinhin als objektiv bezeichnet würden, wesentlich von der Art des zu bewertenden Sachverhalts abhängen: Offensichtlich ist dies für die Kulturpolitik eines Landes schwieriger als für ein Medikament zur Behandlung von Schnupfen. Evaluation setzt Wahrnehmung und Urteilsvermögen voraus. Seit den Anfangszeiten abendländischer Wissenschaften wurden immer wieder die Möglichkeiten und Beschränkungen dieser menschlichen Fähigkeiten thematisiert. Dabei ist einerseits an erkennt-

nistheoretische Fragen nach den grundlegenden Grenzen unseres Urteilsvermögens zu denken. Beispielhaft dafür sind die Arbeiten Kants zur "Kritik der reinen Vernunft" sowie zur "Kritik der Urteilskraft". Daneben sind Arbeiten in der kognitiven Psychologie zu nennen, die auf die Erfassung von Verzerrungen und Anomalien der Urteilsfähigkeit gerichtet sind. Hier ist u.a. an umfangreiche Untersuchungen zu denken, die das Entscheidungsverhalten unter Unsicherheit zum Gegenstand hatten. Dabei hat sich u.a. die verbreitete Disposition gezeigt, in komplexen Entscheidungssituationen zu selektiver Wahrnehmung zu neigen, wobei vor allem die leicht verfügbaren Informationen berücksichtigt werden ("availability", [TvKa82], [RoSi82]).

Wenn Gegenstand und Zielsetzung einer Evaluation festgelegt sind, ist damit nicht notwendigerweise gesagt, welche Merkmale bzw. Indikatoren betrachtet und wie ihre Ausprägungen erfaßt werden sollen. Diese Anforderung an die für eine Evaluation ausgewählten Indikatoren ist vergleichbar mit gängigen Anforderungen an die Erhebung von Merkmalen im Rahmen empirischer Forschung - wie Reliabilität und Validität. So läßt sich etwa eine Vielzahl von Merkmalen denken, um die Qualität von Forschung und Lehre an einer Universitätseinrichtung zu beurteilen. Dieser Umstand hat subtile Konsequenzen. Mit der Auswahl von Merkmalen ist der Anspruch verbunden, eine angemessene Bewertung des betrachteten Sachverhalts zu realisieren. Die Merkmale haben danach einen instrumentellen Charakter. Tatsächlich hat aber die Auswahl der Merkmale einen erheblichen, wenn auch latenten, Einfluß auf das Ergebnis der Evaluation. Überzeichnend formuliert: Nicht das Ziel der Evaluation bestimmt die auszuwählenden Merkmale, sondern die ausgewählten Merkmale legen das Ziel der Evaluation fest - und mögen damit letztlich den zu beurteilenden Gegenstand beeinflussen, indem die Gestaltung des Gegenstands an den Indikatoren ausgerichtet wird. Die damit verbundenen, mehr oder weniger deutlichen Verzerrungen des eigentlichen Gegenstands der Evaluation erinnern an eine traditionsreiche Diskussion in der Ökonomie: die Frage nach der Wertneutralität von Mitteln bei gegebenen Zielen. Schon früh hat Myrdal darauf verwiesen, daß eine solche Trennung nicht möglich ist: "Die Wertsetzung bezieht sich jeweils auf einen ganzen Verlauf ...", ([Myr33], S. 310). In diesem Zusammenhang ist auch der "instrumentalistische Trugschluß" ([Alb60], S. 217) zu nennen.

In angelsächsischen Ländern gibt es seit langem eine rege Evaluationsforschung. Sie ist vor allem auf Evaluationsverfahren gerichtet, die im weitesten Sinne der Politikberatung dienen. Eine besondere Rolle spielen dabei das Bildungs- und das Gesundheitswesen (ein Überblick findet sich in [Hou93]). Die erhebliche Bedeutung, die derartige Evaluationen in den USA erlangt haben, spiegelt sich u.a. in dem Umstand, daß mit der "American Evaluation Association" eine Art Standesorganisation für diejenigen entstanden ist, die professionell Evaluationen durchführen. Die Ziele, denen sich dieser Verband mit mehreren tausend Mitgliedern ([Hou93], S. vii) verschrieben hat, sind bemerkenswert. So wird u.a. angestrebt "evaluation as a profession" zu etablieren und einen Beitrag zur "generation of theory and knowledge about effective human

action" (<http://www.eval.org/>) zu leisten. Ein weiteres Indiz für die Professionalisierung von Evaluationen ist in der Existenz einer dedizierten Fachzeitschrift (American Journal of Evaluation) und einer beachtlichen Zahl einschlägiger Monographien (z.B. [Hou93], [Paw97], [Scr80]) zu sehen.

Auch im deutschsprachigen Raum spielen aufwendige Evaluationsverfahren eine wichtige Rolle - nicht zuletzt in der Politikberatung. Dabei wird allerdings zumeist der Begriff "Gutachten" verwendet. Der Terminus Evaluationsforschung ist vor allem mit einem Teilgebiet der Psychologie verbunden (allgemein: Gesundheitswesen - ein Überblick findet sich in [Koc90]). Daneben gibt es in den Sozial- und Wirtschaftswissenschaften eine Fülle ähnlich ausgerichteter Forschung. Dazu zählen u.a. Untersuchungen der vielfältigen Folgewirkungen des Technologie-Einsatzes wie auch solche, die auf eine Bewertung wirtschaftspolitischer Maßnahmen zielen. In jüngerer Zeit hat auch im deutschsprachigen Raum die explizite Verwendung des Begriffs Evaluation deutlich zugenommen. Dabei ist vor allem an die Beurteilung der Forschung und Lehre an Hochschulen bzw. an die Beurteilung von Hochschullehrern und Hochschulen zu denken.

Jenseits dedizierter Forschungsprojekte und institutioneller Rahmenbedingungen ist Evaluation gleichsam inhärenter Bestandteil wissenschaftlicher Forschung: Forschung impliziert nach gängigem Verständnis das Bemühen um Erkenntnisfortschritt. Das wiederum erfordert die vergleichende Bewertung alternativer Angebote zur Beschreibung bzw. Erklärung der je interessierenden Phänomene. In der Wirtschaftsinformatik spielt Evaluation neben dieser grundsätzlichen Bedeutung auch im Anwendungsbereich eine wichtige Rolle: Entscheidungen im Informationsmanagement erfolgen in einem Umfeld von Systemen und Verfahrensweisen, deren Beurteilung in technischer wie auch in wirtschaftlicher Hinsicht erhebliche Herausforderungen mit sich bringt. Die Wirtschaftsinformatik erbt dabei gleichsam die spezifischen Evaluationsprobleme der Informatik und der Ökonomie (Aufwand, Nutzen!). Daneben sieht sich die Forschung in der Wirtschaftsinformatik einigen eigentümlichen Evaluationsproblemen gegenüber, die sich einerseits aus ihrem interdisziplinären Standort zwischen den Wirtschaftswissenschaften und der Informatik ergeben, andererseits eine Folge der Besonderheiten des Fortschritts im Bereich der Informationstechnologie darstellen.

Der vorliegende Beitrag ist darauf gerichtet, ein spezifisches Evaluationsproblem der Wirtschaftsinformatik zu analysieren: die Beurteilung von informations- bzw. sprachbasierten Artefakten. Diese Analyse erfolgt am Beispiel von konzeptionellen Modellen und zugehörigen Modellierungssprachen, da diese Artefakte eine herausragende Bedeutung in der Forschung haben und ihre Beurteilung mit erheblichen Problemen verbunden ist. An dieser Stelle sei dazu nur so viel angedeutet: Die Schwierigkeiten sind einerseits verursacht durch das eigentümliche Verhältnis von Modellen zur Realität, andererseits durch unsere beschränkte Fähigkeit, Sprachen unvoreingenommen zu beurteilen.

2 Artefakte in der Wirtschaftsinformatik: Anmerkungen zur gegenwärtigen Situation

Der Gegenstand der Wirtschaftsinformatik ist wesentlich durch Artefakte geprägt: Informationsbestände in unterschiedlichen Ausprägungen, Anwendungssysteme, Architekturen, Software-Werkzeuge, Standards, Informationsmodelle usw. Auch andere Disziplinen, wie etwa die Betriebswirtschaftslehre, die Rechtswissenschaften und vor allem die Ingenieurwissenschaften sind zu einem erheblichen Teil durch die Auseinandersetzung mit Artefakten geprägt. Die Artefakte, denen sich die Wirtschaftsinformatik gegenüberstellt, sind allerdings durch einige Besonderheiten gekennzeichnet. So wird die Erfassung existierender Artefakte durch die häufig erhebliche Komplexität informationstechnologischer Systeme bzw. Konzepte erschwert. Ein Problem, das durch den raschen technologischen Wandel noch verschärft wird. Dabei ist auch zu berücksichtigen, daß marktgängige Technologien häufig eine spezielle Begrifflichkeit prägen, die einerseits den Nachvollzug der wesentlichen Zusammenhänge behindert und andererseits zu der verwirrenden Begriffsvielfalt in der Informationstechnologie beiträgt. Aus der Sicht der Unternehmen führt die skizzierte Komplexität zu einem hohen Maß an Unsicherheit und damit zu einem erheblichen Beratungsbedarf. Die seriöse Beschreibung einzelner Produkte mag deshalb durchaus eine sinnvolle Aufgabe von Wissenschaft sein. Gleichzeitig ist allerdings zu berücksichtigen, daß die Auseinandersetzung mit solch singulären, allerdings eben ausgesprochen komplexen Facetten des Gegenstandsbereichs geeignet ist, den Blick für wesentlichere Zusammenhänge zu verstellen. Beispiele für diese Gefahr finden sich zuhauf: objektorientierte Technologien, betriebswirtschaftliche "Standardanwendungen" oder Internet-Technologien. Die Wirtschaftsinformatik als anwendungsorientierte Disziplin befindet sich hier in einer schwierigen Situation. Einerseits kann sie sich wohl nicht auf ein Abstraktionsniveau zurückziehen, wie es in weiten Teilen der Informatik üblich ist, da sie sich sonst zu weit von ihrer Praxis entfernt. Andererseits opfert sie den Anspruch auf Wissenschaftlichkeit, der ja wesentlich durch das Bemühen um Abstraktion und eine sorgsam entwickelte Terminologie gekennzeichnet ist, wenn sie sich allein darauf beschränkt, faktische Erscheinungsformen von Informationstechnologie zu untersuchen. Das skizzierte Problem erhält dadurch zusätzliches Gewicht, daß seit einiger Zeit die Durchführung von Projekten mit Beratungscharakter durch politische Vorgaben gefördert wird, die Drittmittelerwerb zu einem Wert *eo ipso* erheben. In solchen Projekten wird häufig, mit gutem Grund, die Praxis die Randbedingungen der Kooperation wesentlich prägen - also etwa die zu verwendenden Werkzeuge und damit auch: die Begrifflichkeit, die diesen Werkzeugen zugrunde liegen.

Doch es sind nicht nur Produkte oder am Markt vorherrschende Konzepte, die den Umgang mit Artefakten prägen. Daneben ist die Forschung in der Wirtschaftsinformatik durch eine Fülle selbst geschaffener Artefakte, wie Modelle und Prototypen gekennzeichnet. Die Konstruktion solcher Artefakte ist m.E. ein

zentraler Bestandteil der Forschung in unserem Fach: Die Erkenntnisse, die dabei gewonnen werden können, fördern ein tieferes Verständnis von Informationssystemen und liefern gleichzeitig einen Beitrag zu einer elaborierten Entwurfslehre. Tatsächlich ist der Entwurf von Artefakten in der Wirtschaftsinformatik (ähnliches gilt für die angewandte Informatik) auch von weniger erfreulichen Erscheinungen begleitet. So sind viele wissenschaftliche Arbeiten und auch Lehrbücher auf die - zumeist nur rudimentäre - Darstellung solcher Entwürfe gerichtet. Eine kritische Auseinandersetzung oder gar ein detaillierter Vergleich einzelner Arbeiten bleibt allerdings i.d.R. aus. Besonders augenfällig wird dieser Umstand auf einschlägigen Konferenzen: Nacheinander werden grobe Zusammenfassungen einzelner Modelle oder Prototypen präsentiert. Auf diese Weise entsteht mitunter der Eindruck – und dies ist die (selbst-) kritische Feststellung eines Betroffenen, daß die Forschung hier einer Spielwiese für die Schaffung von Artefakten gleicht, die vor allem einem dient: der Befriedigung der Konstrukteure.

Sowohl die Untersuchung informationstechnologischer Artefakte in der Praxis wie auch der Konstruktionen, die im Wissenschaftsbetrieb selbst entstehen, erfordert gehaltvolle Kriterien zu ihrer Evaluation. Seit einiger Zeit wird in der Forschungsförderung häufig auf "best practice" als Bewertungsgrundlage hingewiesen. Auch in der Wirtschaftsinformatik wird dieses Kriterium in Erwägung gezogen ([Buh97], [ReWi97], [Mer97]). Es mag durchaus interessant sein, zu untersuchen, welche Technologien unter den Randbedingungen der Praxis zu herausragenden Resultaten führen. Unabhängig davon, daß die Verwendung des "best practice" Kriteriums erhebliche methodische Schwierigkeiten aufwirft, ist es jedoch nicht geeignet eine wissenschaftlich begründete Evaluation zu ersetzen - allenfalls um den Preis, den Anspruch auf Wissenschaftlichkeit zu opfern.

3 Evaluation als inhärenter Bestandteil wissenschaftlicher Forschung

Die im letzten Kapitel geäußerte Forderung nach wissenschaftlich akzeptablen Kriterien zur Evaluation von Artefakten führt uns zu der Frage, wie eine solche Forderung umzusetzen ist. Die Bewertung der Ergebnisse wissenschaftlicher Forschung ist zentraler Bestandteil der Wissenschaftstheorie. Bevor wir auf die spezifischen Herausforderungen eingehen, die sich durch die Evaluation von Modellen ergeben, werden wir deshalb zunächst Orientierungen betrachten, die von der Wissenschaftstheorie zur Beurteilung wissenschaftlicher Aussagen bzw. wissenschaftlicher Begründungsverfahren angeboten werden.

3.1 Zur Beurteilung wissenschaftlicher Aussagen

Die Anwendung angemessener Kriterien zur Beurteilung von Forschungsergebnissen ist ein wesentliches, wenn auch mitunter implizites Kennzeichen von Wissenschaft. Das gilt in zweifacher Hinsicht. So sind einerseits Abgrenzungskriterien erforderlich: Was sind wesentliche Merkmale wissenschaftlicher Erkenntnis, was unterscheidet sie von den "Träumereien eines Geistersehers" (Kant). Daneben - und durchaus damit zusammenhängend - impliziert das Bemühen um Erkenntnisfortschritt die vergleichende Beurteilung konkurrierender Erkenntnisangebote. Bei der Untersuchung dieser Frage sind einige miteinander verwobene Aspekte zu berücksichtigen ([Ham82]). So ist zu klären, welche Sachverhalte überhaupt als Gegenstand wissenschaftlicher Untersuchungen zu akzeptieren sind. Anders gewendet: Welche Sachverhalte sind prinzipiell einer Beurteilung durch wissenschaftliche Verfahren zugänglich, welche nicht? Wodurch wird subjektives Wissen zu wissenschaftlicher Erkenntnis? Welches ist die angemessene sprachliche Form der Darstellung und Vermittlung wissenschaftlicher Erkenntnisse?

Jenseits der in der Vergangenheit mitunter heftig geführten Grabenkämpfe zwischen den verschiedenen wissenschaftstheoretischen Schulen¹, lassen sich m.E. durchaus deutliche Gemeinsamkeiten im Hinblick auf grundsätzliche Anforderungen an realwissenschaftliche Aussagen rekonstruieren² - die Unterschiede zeichnen sich eher in der Umsetzung ab. Danach sollte eine sprachliche Darstellung von Erkenntnissen in den Realwissenschaften folgende Merkmale aufweisen:

- Es sollten generelle Aussagen, also Aussagen über *Klassen* von realen Sachverhalten, enthalten sein.
- Die beschriebenen Merkmale bzw. Eigenschaften dieser Sachverhalte sollten in ihrem Verhältnis zu Zeit und räumlicher Verteilung möglichst stabil, also *invariant* sein.
- Eine - wie auch immer geartete - *intersubjektive Überprüfung* der Aussagen sollte möglich sein.
- Es sollte neues Wissen und möglichst *informatives Wissen* enthalten sein - wobei die Vorstellungen darüber, was informativ bedeutet, divergieren. Nach der vom Kritischen Rationalismus vertretenen Ansicht nimmt der Informationsgehalt einer Aussage mit der Anzahl der durch sie ausgeschlossenen denkmöglichen Konstellationen zu.

¹ Hier ist im deutschsprachigen Raum vor allem an den Kritischen Rationalismus, die Kritische Theorie und den Konstruktivismus zu denken.

² Diese Annahme beruht darauf, daß meines Wissens die im folgenden dargestellten Anforderungen in den wissenschaftstheoretischen Auseinandersetzungen im wesentlichen von keiner Partei in Frage gestellt wurde.

- Die Aussagen sollten *bewährt* sein, d.h. eine Reihe von Überprüfungsversuchen erfolgreich überstanden haben.

Im Hinblick auf die Überprüfung wissenschaftlicher Aussagen gibt es eine Reihe unterschiedlicher Auffassungen. Häufig wird auf den Ansatz des Kritischen Rationalismus verwiesen, wonach eine Überprüfung durch die Konfrontation mit der Realität als objektiver Instanz zu erfolgen hat ("Fallibilismus"). Das setzt geeignete Abbildungs- bzw. Meßvorschriften voraus, um die Korrespondenz der Hypothesen mit der Wirklichkeit wiederum in einem intersubjektiv nachvollziehbaren Verfahren feststellen zu können. Daneben wird in den Wirtschafts- und Sozialwissenschaften mitunter auf diskursive Überprüfungs- bzw. Begründungsverfahren, auf die wir im nächsten Abschnitt kurz eingehen, verwiesen.

Die skizzierten Anforderungen sind hilfreich für die Abgrenzung wissenschaftlicher Aussagen. Zudem liefern sie eine Orientierung für deren Bewertung, denn die einzelnen Kriterien sind durchaus in unterschiedlichen Intensitäten denkbar. Gleichzeitig ist ihre Anwendbarkeit eingeschränkt. So ist die vergleichende Beurteilung des Informationsgehalts oder der intersubjektiven Überprüfbarkeit von Aussagen in den Wirtschafts- und Sozialwissenschaften ausgesprochen problematisch. Darüber hinaus begrenzt sie den Zuständigkeitsbereich von Wissenschaft auf solche Urteile, die auf eine Beschreibung faktischer Realität zielen. Es wird sich zeigen, daß gerade in der Wirtschaftsinformatik eine solche Einschränkung häufig nicht praktikabel ist.

3.2 Diskursive Urteilsfindung

Im Unterschied zur oben skizzierten Wissenschaftsauffassung des Kritischen Rationalismus gehen die Vertreter hermeneutischer Ansätze davon aus, daß sich soziale Realität allenfalls oberflächlich durch Messungen im Sinne der Naturwissenschaften beschreiben läßt. Statt dessen sei es erforderlich, sich um angemessene Interpretationen der je betrachteten Sachverhalte zu bemühen. Angemessenheit wird dabei in einem diskursiven Prozeß sachkundiger Wissenschaftler beurteilt. Um solche Prozesse dem Ideal *rationaler* Urteilsfindung anzunähern, nennen sowohl die Kritische Theorie ("herrschaftsfreie Kommunikation") wie auch der Konstruktivismus ("vernünftige Beratung") eine Reihe von Bedingungen ([Hab81], ([Lor74]):

- Die Teilnehmer sollten bemüht sein, in gemeinsamer Beratung einen Konsens zu erzielen.
- Es sollten keine Aussagen wider besseren Wissens gemacht werden (sie sollten nicht persuasiv sein).
- Die Teilnehmer sollten bemüht sein, eigene Interessen zu transzendieren und die Sichtweisen anderer Teilnehmer nachzuvollziehen.

- Die Teilnehmer sollten sachkundig sein.
- Die Diskurse sollten herrschaftsfrei sein. Lediglich der "eigentümlich zwanglosen Zwang des besseren Arguments" ([Hab81], S. 52 f.) ist zugelassen.

Im Unterschied zu den vom Kritischen Rationalismus vorgeschlagenen Prüfungsverfahren ist der Gegenstand wissenschaftlicher Urteilsfindung nicht auf falsifizierbare Aussagen über die Wirklichkeit eingeschränkt. Prinzipiell können danach auch Werturteile vernünftig begründet werden. Damit sind solche Ansätze diskursiver Begründung grundsätzlich gut geeignet, eine - wenn auch abstrakte - Methode für Evaluationen zu liefern. Gleichzeitig sind ihre Unzulänglichkeiten kaum zu übersehen. Im wesentlichen münden sie in die Frage, wie entschieden werden soll, daß die Teilnehmer die geforderten Voraussetzungen erfüllen. Es wird sich zeigen, daß diese Frage für unsere weitere Betrachtung von erheblicher Bedeutung ist.

4 Evaluation im Kontext der konzeptionellen Modellierung

Betriebliche Informationssysteme sind der zentrale Gegenstand der Wirtschaftsinformatik. Einschlägige Arbeiten sind u.a. darauf gerichtet, den Erfolgsbeitrag von Informationssystemen zu untersuchen, ihre Einbettung in den organisatorischen Kontext wie auch ihre Verwaltung und Pflege. Daneben sind die vergleichende Bewertung wie auch der Entwurf von Informationssystemen ein wichtiges Thema. Im Unterschied zur Informatik stehen dabei vor allem höhere Abstraktionsebenen im Vordergrund, wie etwa Architekturen oder - allgemein - Modelle. Die Evaluation konzeptioneller Modelle ist für die Wirtschaftsinformatik in zweifacher Hinsicht von Bedeutung. Zum einen gibt es in Unternehmen einen erheblichen Bedarf an einer gehaltvollen Beurteilung von Modellen, da ihre Erstellung einerseits mit erheblichen Kosten verbunden ist und sie andererseits einen nachhaltigen Einfluß auf den Nutzen der auf ihnen basierenden Informationssysteme haben. Zum anderen ist der Entwurf von konzeptionellen Modellen - zumeist wird es sich dabei um Referenzmodelle handeln - ein einschlägiges Forschungsziel, dem mitunter gar ein konstituierender Charakter für das Profil der Disziplin zugesprochen wird.¹ Ein solches Ziel empfiehlt eine überzeugende Beurteilung von Entwurfsalternativen. Die mit einer solchen Evaluation verbundenen Schwierigkeiten werden wir im folgenden auf zwei Ebenen betrachten. Zunächst werden die Anforderungen diskutiert, die eine Beurteilung konzeptioneller Modelle mit sich bringt. Im Anschluß daran begeben wir uns auf eine Meta-Ebene, da - wie sich zeigen wird - die Evaluation von Modellierungssprachen nicht vernachlässigt werden darf.

¹ beispielhaft dafür Weber, der die konzeptionelle Modellierung für den "core of the information systems discipline" ([Web97], S. 69 ff.) hält.

4.1 Der konstruktive Charakter von Informationsmodellen

Der Entwurf von Software erfordert eine angemessene Erfassung und Beschreibung ausgewählter Eigenschaften der jeweiligen Anwendungsdomäne. Da eine solche Beschreibung i.d.R. nicht allein im Kreis von Software-Entwicklern entsteht, sondern Anwender und Domänenexperten wesentlich an ihrer Entstehung zu beteiligen sind, sind Programmiersprachen oder auch formale Spezifikationsprachen zu ihrer Erstellung kaum geeignet. Sie sind deutlich an den Randbedingungen der Implementierung orientiert und sperren sich deshalb gegen Darstellungen, die von allen Beteiligten als verständlich und authentisch empfunden werden. Auf der anderen Seite ist auch die Beschränkung auf eine natürliche Sprache, sei es in Form der Umgangssprache und/oder einer dedizierten Fachsprache, unzureichend: Schließlich sollen Beschreibungen von Anwendungsdomänen Vorgaben für die Implementierung liefern, so daß daraus erwachsende Besonderheiten auch berücksichtigt werden müssen. Vor dem Hintergrund dieses Spannungsfelds ist die konzeptionelle Modellierung entstanden. Als ihr vorrangiges Anliegen wird gern die "natürliche" Abbildung des jeweils betrachteten Realitätsausschnitts oder gedanklichen Entwurfs genannt: "... descriptions of a world enterprise/slice of reality which correspond directly and naturally to our own conceptualizations of the object of these descriptions." ([MyLe84], S. 11) Eine solche Kennzeichnung ist jedoch allenfalls als didaktisch motivierte Überzeichnung zu akzeptieren: Eine natürliche Beschreibung in diesem Sinn dürfte i.d.R. mit einer natürlichen Sprache besser zu leisten sein. Gegenstand und Ziel der konzeptionellen Modellierung sind deshalb besser wie folgt zu charakterisieren: Die konzeptionelle Modellierung dient einer allen Beteiligten an einem Software-Entwicklungsprozeß verständlichen Darstellung einer Anwendungsdomäne. Die Darstellung beschränkt sich dabei auf die für die Erstellung des beabsichtigten Systems wesentlichen Aspekte dieser Domäne. Zusätzlich zu dieser Abstraktion muß ein konzeptionelles Modell auch den Randbedingungen Rechnung tragen, die durch die in einer späteren Phase zu verwendenden Implementierungssprachen entstehen.

Vor diesem Hintergrund gestaltet sich die Evaluation von konzeptionellen Modellen in zweifacher Hinsicht problematisch. So sind einerseits konkurrierende Ziele zu beachten, andererseits ist das eigentümliche Verhältnis konzeptioneller Modelle zur Wirklichkeit zu berücksichtigen. Das Bemühen um eine verständliche Modellierung zielt darauf, solche Abstraktionen zu verwenden, die den Wahrnehmungs- und Konzeptualisierungsmustern der Betrachter entsprechen. Während derartige Dispositionen der Betrachter grundsätzlich variieren mögen, ist die Wahrscheinlichkeit für Divergenzen dann besonders groß, wenn implementierungsnahe Sichten (also die von Programmierern) und anwendungsnahe Sichten aufeinandertreffen. Das in Abb. 1 dargestellte Beispiel verdeutlicht diesen Konflikt in ebenso vereinfachender wie krasser Weise: Während im Anwendungsbereich (hier: der Geometrie) ein Quadrat als eine Spezialisierung eines Rechtecks betrachtet wird, mag mancher Programmierer die umgekehrte Spezialisierungsrichtung bevorzugen, da sie im Hinblick auf Sprachkonzepte

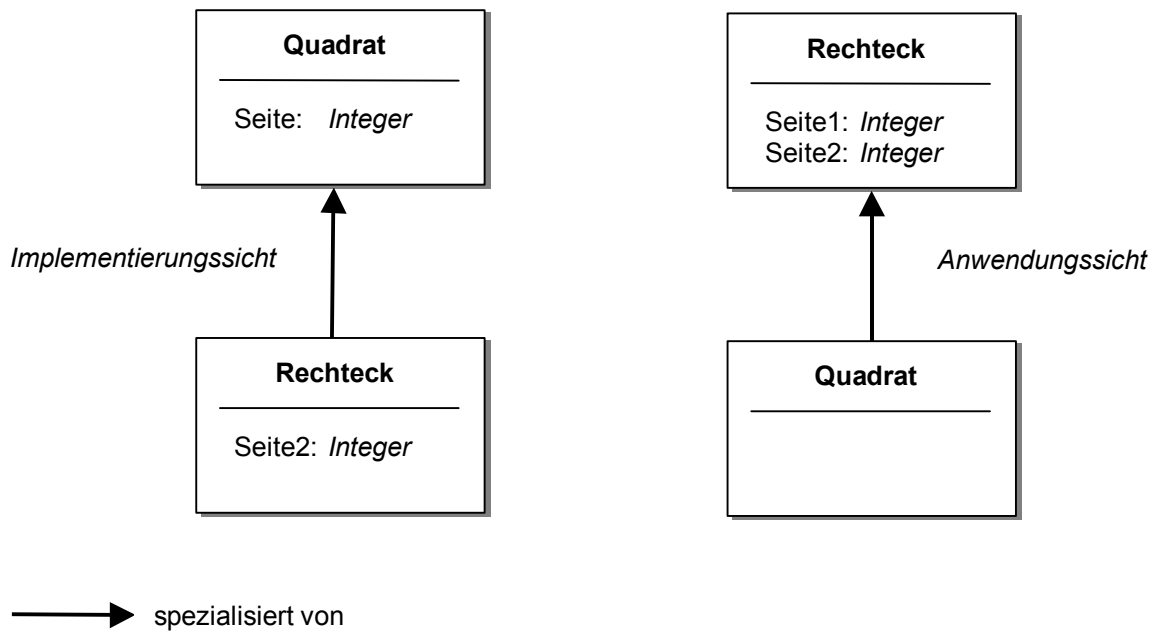


Abb. 1: Konkurrierende Kriterien für die Festlegung von Spezialisierungsbeziehungen

gängiger Programmiersprachen leichter zu implementieren ist. Man spricht denn auch im letzten Fall mitunter von "implementation inheritance" ([Gam95], S. 18 ff.). In der Praxis der Modellierung präsentieren sich derartige Konflikte zumeist komplexer und subtiler. Im Hinblick auf unsere Untersuchung können wir an dieser Stelle festhalten, daß eine einheitliche Bewertung solcher Fälle oft dadurch verhindert wird, daß die zu beteiligenden Personen unterschiedliche Präferenzen haben, für die jeweils durchaus gewichtige Argumente angeführt werden können

Im Unterschied zu gängigen wissenschaftlichen Aussagen sind konzeptionelle Modelle nicht unbedingt darauf gerichtet, den jeweils interessierenden Wirklichkeitsausschnitt authentisch zu beschreiben. Vielmehr wird man in aller Regel bewußt mehr oder weniger von der Realität abstrahieren, je nachdem, welchem *Zweck* das jeweilige Modell dienen soll. Die oben skizzierten Anforderungen an wissenschaftliche Aussagen über die Realität sind deshalb allenfalls eingeschränkt zu verwenden. Die Wahl von Abstraktionen kann an unterschiedlichen Kriterien orientiert sein. Sie ist abhängig von individuellen Wahrnehmungen und Präferenzen, oft aber auch Ausdruck einer gewissen Willkür. Das führt dazu, daß i.d.R. verschiedene Personen einen bestimmten Realitätsbereich in unterschiedlicher Weise modellieren. Die damit verbundene Varianz erschwert eine vergleichende Beurteilung komplexer Modelle - konzeptionelle Modelle sind häufig komplex - allein dadurch, daß der Bewertungsaufwand, jenseits des Problems geeignete Beurteilungskriterien zu finden, sehr hoch wird.

Doch damit nicht genug. Das Verhältnis konzeptioneller Modelle zur Wirklichkeit wird durch einen weiteren Umstand noch subtiler. So sind konzeptionelle Modelle häufig aus gutem Grund (vgl. dazu [Fra98a]) nicht darauf gerichtet, faktische Wirklichkeit abzubilden. Sie reflektieren statt dessen neue Formen der Organisation von Arbeit und neue Begrifflichkeiten - mit dem Ziel, auf diese Weise die Potentiale der Informationstechnologie weiter ausschöpfen zu können. Das äußert sich etwa in neu entworfenen Geschäftsprozessen und den mit

ihnen einhergehenden Informationsmodellen. Ein anschauliches Beispiel für Informationskonzepte, die auf diese Weise eher konstruiert werden als daß sie existierende Begriffe abbilden, sind Dokumentarchitekturen wie etwa ODA, SGML oder OpenDoc ([App91], [GoRu90], [Ope94]): Sie gehen weit über den traditionellen Dokumentbegriff hinaus - was vor allem für OpenDoc gilt.

Bei vordergründiger Betrachtung sind Modelle, die geplante Wirklichkeit abbilden, durch einen Vergleich mit faktischer Wirklichkeit kaum zu bewerten. Das ist allerdings nur zum Teil richtig, da es sich hier ja nicht um völlig willkürliche Szenarien handelt. Vielmehr ist mit entsprechenden Plänen ein Machbarkeitsanspruch verbunden, dessen Beurteilung zwangsläufig die Analyse gegebener Realität nicht vernachlässigen kann. Diese Überlegung führt uns zu einer zusätzlichen Schwierigkeit. Häufig ist in der Wirtschaftsinformatik der Entwurf konzeptioneller Modelle mit dem Anspruch verbunden, von den Besonderheiten einzelner Unternehmen zu abstrahieren. Es geht also eher um generelle Modelle mit Referenzcharakter. Solche Referenzmodelle sind aus ökonomischer Sicht ausgesprochen reizvoll: Sie versprechen ein hohes Maß an Integration innerhalb und zwischen betrieblichen Informationssystemen und fördern zudem eine komfortable und vielfältige Wiederverwendung. Gerade deshalb aber gestaltet sich die Bewertung von Referenzmodellen ausgesprochen tückisch: Wenn Referenzmodelle vor allem unter Hinweis auf wirtschaftliche Verheißungen eingeführt werden, können sie nur dann als erfolgreich bezeichnet werden, wenn sie wirksam werden, also als akzeptierte Ordnungen weite Verbreitung finden. Es ist allerdings unstrittig, daß die Verbreitung von Referenzmodellen letztlich über Märkte erfolgt. Dabei spielen wirtschaftliche Interessen und Machtverhältnisse eine Rolle - also Einflüsse, die jenseits der unmittelbaren Eigenschaften eines Modells liegen.

Bevor wir uns den einschlägigen Forschungsansätzen zuwenden, die auf eine Bewertung konzeptioneller Modelle zielen, müssen wir den Fokus unserer Betrachtung zunächst auf ein weiteres, noch subtileres Problem ausweiten.

4.2 Die Beurteilung von Modellierungssprachen: Notwendiges Übel oder entbehrlicher Luxus

Konzeptionelle Modelle werden mit Hilfe künstlicher Modellierungssprachen erstellt. Solche Sprachen werden gewöhnlich durch eine mehr oder weniger präzise abstrakte Syntax und Semantik definiert. Außerdem beinhalten sie typischerweise eine graphische Notation, auch konkrete Syntax genannt ([FrPr97], S. 12 ff.). Es liegt auf der Hand, daß die Evaluation von konzeptionellen Modellen nicht unabhängig von der jeweils verwendeten Modellierungssprache erfolgen kann. Wenn beispielsweise eine Modellierungssprache verwendet wird, die bestimmte semantische Konstrukte - wie etwa Generalisierungsbeziehungen - nicht bereitstellt, können die zugehörigen Modelle eben keine Generalisierungen ausdrücken. Ähnliches gilt für den Formalisierungsgrad einer Spra-

che. Wenn die Semantik der Sprachkonzepte nicht eindeutig definiert ist, wird auch ein entsprechendes Modell einen Mangel an Eindeutigkeit aufweisen. Schließlich hat die graphische Notation einer Sprache einen wesentlichen Einfluß auf Anschaulichkeit und Verständlichkeit von Modellen. Die Evaluation von konzeptionellen Modellen setzt also die Beurteilung der jeweils verwendeten Modellierungssprache voraus. Aus pragmatischer Sicht könnte man hier einwenden, daß man durchaus von den spezifischen Qualitäten einer Modellierungssprache abstrahieren könne, da es für die konzeptionelle Informationsmodellierung mit dem Entity Relationship Model (ERM) eine vorherrschende Sprache gibt. Abgesehen davon, daß dies letztlich einen Stillstand der einschlägigen Forschung bedeuten würde, ist ein solcher Einwand auch aus anderen Gründen kaum zu akzeptieren. So gibt es nicht nur eine Reihe von Dialekten des ERM, darüber hinaus haben in den letzten Jahren objektorientierte Modellierungssprachen eine zunehmende Bedeutung erlangt. Zudem ist zu berücksichtigen, daß die konzeptionelle Modellierung nicht allein auf die Darstellung statischer Informationskonzepte beschränkt ist. Für andere Sichten, wie etwa Prozeßbeschreibungen, hat sich noch kein vorherrschendes Sprachparadigma etabliert.

Die Bewertung von Modellierungssprachen ist ein delikates Unterfangen. Das liegt vordergründig an den Sprachbeschreibungen selbst. Sie liegen i.d.R. in Form von Metamodellen vor. Solche Metamodelle sind nicht nur abstrakt, sondern mitunter auch ausgesprochen umfangreich. Gleichzeitig beschreiben sie - anders als Modelle auf der Objektebene - für den Betrachter zunächst unbekannte Artefakte, nämlich eine künstliche Sprache. Zusätzlich ist zu berücksichtigen, daß eine bestimmte Sprache mit verschiedenen Metamodellen beschrieben werden kann. Ein konkretes Metamodell reflektiert also auch zum Teil willkürlich wirkende Entwurfsentscheidungen. Die daraus resultierenden Schwierigkeiten werden ansatzweise in Abb. 2 verdeutlicht. Sie zeigt Ausschnitte aus der Spezifikation ähnlicher Sprachen in unterschiedlichen Metamodellen. In beiden Fällen handelt es sich dabei um kleine Teile des gesamten Metamodells.

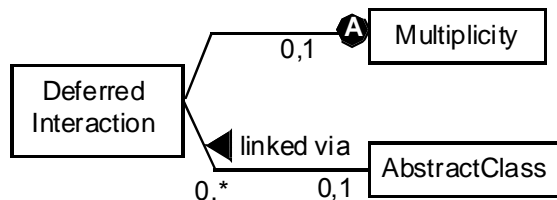
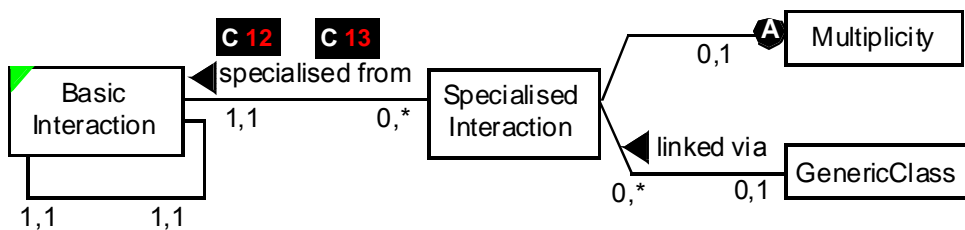
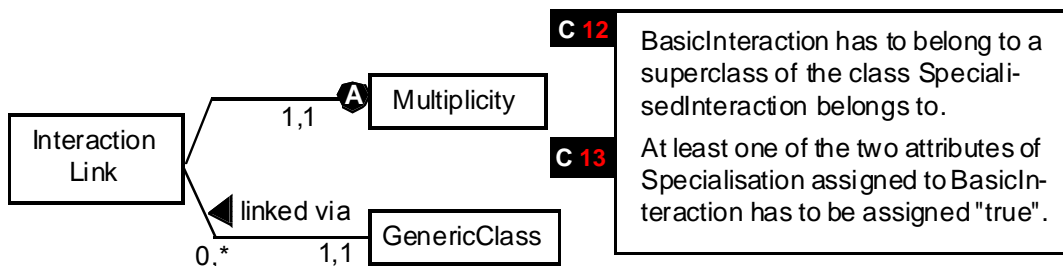
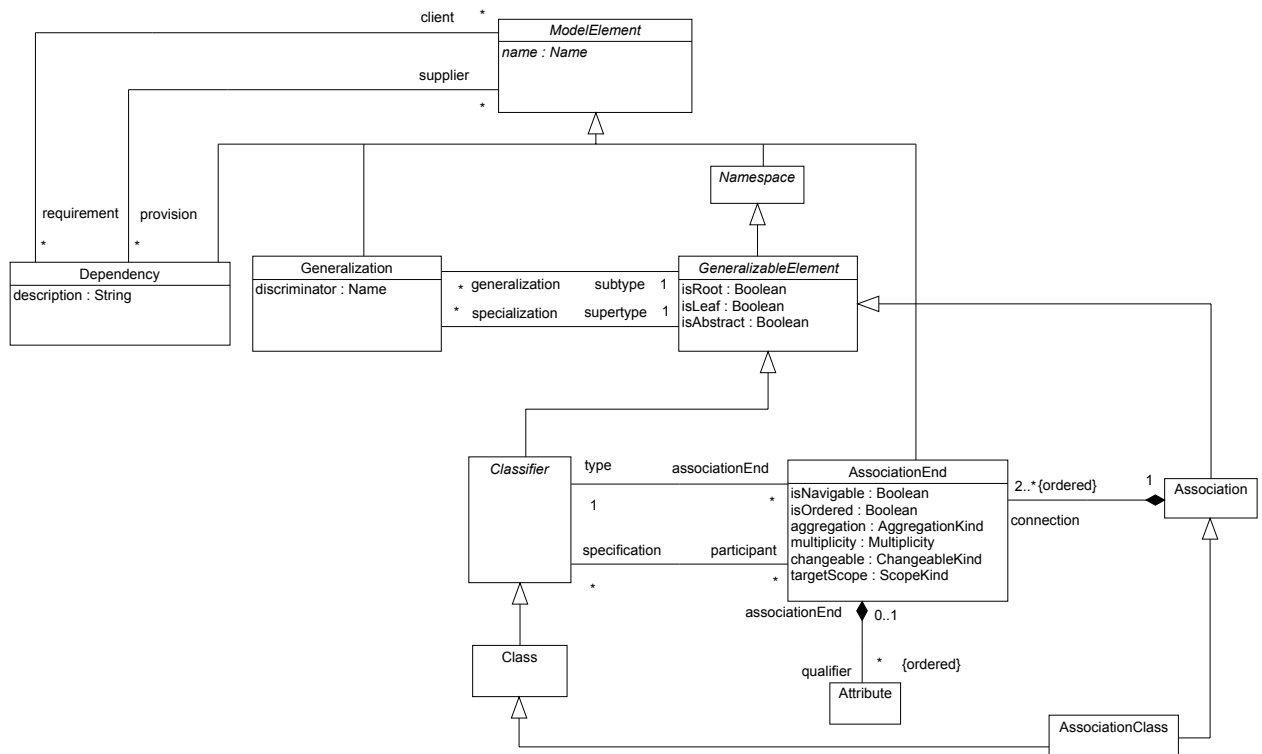


Abb.2 : Ausschnitt aus dem Metamodell von UML (*oben*) ([Rat97], S. 23) und aus dem Metamodell von MEMO-OML ([Fra98c], S. 21) (*unten*)

Zu diesen vordergründigen Problemen gesellen sich weitere, grundsätzliche Schwierigkeiten. Alle Wissenschaften, die Sprache und ihre Verwendung untersuchen, oder neue "Sprachspiele", also Sprachen und daran anknüpfende Handlungen entwerfen, sehen sich einer eigentümlichen Schwierigkeit gegenüber. Der Umstand, daß man als Forscher selbst in ein nicht vollständig auflösbares Geflecht von Sprache, Denken und Handeln eingebunden ist, markiert eine latente Befangenheit, die zu einem kaum zu überwindenden Paradoxon führt: Die Bewertung einer Sprache ist nicht möglich ohne sie zu verwenden. Gleichzeitig wird der Gebrauch einer Sprache ihre unvoreingenommene Beurteilung erschweren - oder, wie es der frühe Wittgenstein formuliert: "Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt ([Wit80], §5.6).

Solche Betrachtungen mögen als entbehrliche philosophische Spitzfindigkeiten erscheinen. Sie reflektieren allerdings in treffender Weise ein faktisches Dilemma der Evaluation von Modellierungssprachen. Die Beurteilung von Modellierungssprachen setzt eine erhebliche fachliche Kompetenz voraus, die man i.d.R. nicht zuletzt durch die Benutzung und evtl. den Entwurf solcher Sprachen erwirbt. Ähnlich wie bei Programmiersprachen ist es wohl unstrittig, daß der Umgang mit bestimmten Sprachen die Präferenzen und Konzeptualisierungen des Betrachters prägt. Jenseits solcher subtilen Beeinträchtigungen individueller Urteilsfähigkeit ist der Umstand zu berücksichtigen, daß eine Modellierungssprache eine Reihe, zum Teil konkurrierender Anforderungen zu genügen hat. So sollte sie Sprachkonzepte bieten, die verständliche Abbildungen auf verschiedenen Abstraktionsebenen erlauben. Im Hinblick auf die Software-Entwicklung sollte sie die Darstellung software-technisch relevanter Zusammenhänge erlauben. Zur Unterstützung der Implementierung sollten die entsprechenden Konzepte möglichst präzise, ggfs. formal spezifiziert sein.

Typischerweise reflektiert der Entwurf von Modellierungssprachen vor allem die Anforderungen der Sprachentwickler. Dabei mögen durchaus Annahmen über die Wahrnehmungsmuster der Sprachverwender einfließen, sie beruhen allerdings zumeist auf gewagten Generalisierungen eigener Vorstellungen. Beispielhaft dafür stellen etwa Jacobson et al. ohne weitere Begründung fest: „People regard their environment in terms of objects. Therefore it is simple to think in the same way when it comes to designing a model.“ ([JaCh92], S. 42)

Die enge Beziehung, die ein Sprachentwickler i.d.R. zu "seiner" Sprache aufbaut, sowie der hohe Aufwand einer Sprachspezifikation führen häufig zu der Tendenz, an den eigenen Entwürfen um ihrer selbst willen festzuhalten. Der Widerstand gegen die Übernahme konkurrierender Sprachen kann dabei deutlich wirksamer sein als derjenige, den Kuhn ([Kuh70]) für die Evolution wissenschaftlicher Theorien rekonstruiert hat: Im Unterschied zu wissenschaftlichen Theorien können Modellierungssprachen kaum "falsifiziert" werden.

5 Zum Stand der einschlägigen Evaluationsforschung

Die Schwierigkeiten, denen sich eine wissenschaftliche Bewertung von konzeptionellen Modellen und Modellierungssprachen gegenüberstellt, sind erheblich. Sie lassen sich allerdings relativieren: Auch in anderen Bereichen der Wirtschaftswissenschaften wie der Informatik sind Gegenstände bzw. Phänomene zu verzeichnen, die sich gegen eine letztlich überzeugende Evaluation sperren. Dennoch ist offenbar zielgerichtetes Handeln wie auch Forschen in derartigen Bereichen möglich. In der Tat wäre es verfehlt, aus den grundsätzlichen Schwierigkeiten den Schluß zu ziehen, daß eine sinnvolle Evaluation überhaupt nicht möglich ist.

5.1 Ansätze zur Evaluation konzeptioneller Modelle

Angesichts der zentralen Bedeutung, die konzeptionelle Modelle für die Wirtschaftsinformatik haben, ist es erfreulich, daß in den letzten Jahren eine Reihe von Arbeiten zu verzeichnen ist, die die Evaluation von Modellen zum Gegenstand haben. Der größere Teil der einschlägigen Arbeiten ist darauf gerichtet, Bezugsrahmen zur Bewertung anzubieten. Sie enthalten Anforderungen an Modelle und diskutieren deren Erfassung. Moody und Shanks ([MoSh94]) schlagen sechs Kriterien zur Evaluierung von Entity Relationship Modellen vor: *Simplicity, Understandability, Flexibility, Completeness, Integration, Implementability*. Batini et al. ([BaCe92]) behandeln in einem Lehrbuch über Datenmodellierung ebenfalls solche Kriterien. Ähnlich wie Moody und Shanks betonen Krogstie et al. die verschiedenen Perspektiven auf ein Modell und die mit ihnen einhergehenden Anforderungen. Sie unterscheiden sie u.a. "knowledge quality", "syntactic quality", "semantic quality" und "pragmatic quality" ([KrLi95]). Zur Anwendung des Bezugsrahmens wird dabei explizit auch auf die Einbeziehung der Betrachter eines Modells hingewiesen. So wird etwa "perceived semantic quality" als ein Kriterium genannt. In [Fra98b] werden u.a. modellinterne Kriterien, das Verhältnis zum Betrachter und zur Realität vorgeschlagen. Becker et al. ([BeRo95]) unternehmen den originellen Versuch, in Anlehnung an die "Grundsätze ordnungsmäßiger Buchführung" "Grundsätze ordnungsmäßiger Modellierung" zu entwerfen. Sie formulieren dazu sechs Prinzipien, die von Modellen zu erfüllen sind: *Richtigkeit, Relevanz, Wirtschaftlichkeit, Klarheit, Vergleichbarkeit, systematischer Aufbau*. Ähnlich wie [Fra98b] sehen Becker et al. ihren Ansatz nicht auf die Datenmodellierung beschränkt. Er soll vielmehr generell für konzeptionelle Modelle einsetzbar sein. Aufbauend auf diesem Ansatz entwickelt Schütte einen ähnlichen Kriterienkatalog, der durch eine erkenntnistheoretische Relativierung der ursprünglichen Kriterien gekennzeichnet ist. Das drückt sich etwa darin aus, daß das Kriterium "Richtigkeit" durch "Konstruktionsadäquanz" ersetzt wird ([Sch98]).

Eine zweite Gruppe von Arbeiten ist empirisch ausgerichtet. So untersucht Maier ([Mai96]) im Rahmen einer empirischen Erhebung nach den Randbedin-

gungen erfolgreicher Datenmodellierung in der Praxis. Dabei wird auch Erfolg daran gemessen, wie die betrachteten Modelle im Kreis der Befragten beurteilt wurden. Andere empirische Untersuchungen fokussieren auf die Wahrnehmung der Modellbetrachter. So haben Goldstein und Storey ([GoSt90]) wie auch Hitchman ([Hit95]) in - allerdings kaum repräsentativen - Erhebungen belegen können, daß viele Betrachter Entity Relationship Modelle für wenig verständlich halten.

Dedizierte Untersuchungen zur Beurteilung von Modellierungssprachen sind ausgesprochen selten. Zwar thematisieren einige Ansätze zur Evaluation von konzeptionellen Modellen auch Kriterien, die auf Modellierungssprachen bezogen sind - so führen Krogstie et al. das Kriterium "language quality" ein; eine differenzierte Untersuchung des Kriteriums bleibt allerdings aus. In der Informatik, hier ist vor allem an den Teilbereich Software-Technik zu denken, gibt es durchaus sehr deutliche und relativ gut überprüfbare Bewertungskriterien. Charakteristisch für solche Ansätze ist die Anforderungsliste von Süttenbach und Ebert [SüEb97], S. 2 f.). Danach sollte die Beschreibung einer Modellierungssprache mittels eines Metamodells ein *korrektes* Modell der Sprache liefern: Die Sprachbeschreibung sollte die Identifikation unzulässiger Modelle gewährleisten und gleichzeitig erlauben, die Menge aller zulässigen Modelle zu generieren. Die Sprachbeschreibung sollte zudem *vollständig* sein, d.h. alle in der Sprache verwendeten Konzepte sowie die Bedingungen ihrer Verwendung sollten definiert sein. Ein Metamodell sollte auch dem Prinzip der *Klarheit/Einheitlichkeit* genügen. Das umfaßt allgemein die Forderung nach einer verständlichen Darstellung des Metamodells, konkret die Forderung danach, daß ähnliche Konzepte auch in ähnlicher Weise im Metamodell repräsentiert werden. Schließlich sollte ein Metamodell so *einfach wie möglich* dargestellt werden, was u.a. die Vermeidung redundanter Informationen empfiehlt. So wichtig diese Kriterien aus formalwissenschaftlicher Sicht auch sind, so wenig reichen sie hin, um eine umfassende, also allen relevanten Anforderungen gerecht werdende Bewertung zu ermöglichen. So wird beispielsweise - nicht untypisch für den Blickwinkel der Informatik - weitgehend vom Modellierungszweck und den Perspektiven der verschiedenen Betrachtern abstrahiert.

Im Bereich der objektorientierten Modellierung gibt es eine Vielzahl von Untersuchungen, die auf einen Vergleich verschiedener Modellierungsansätze gerichtet sind (z.B. [Big97], [DeFa92], [HoGo93], [Hsi92], [MoPu92]). Sie differenzieren allerdings zumeist nicht zwischen semantischen Eigenschaften, der Notation und sonstigen Eigenschaften einer Modellierungsmethode. Oft findet man Listen mit objektorientierten Sprachmitteln, wobei die Verfügbarkeit eines Sprachmittels gern per se als ein Vorteil angesehen wird. Es liegt auf der Hand, daß bei solchen Vergleichen diejenigen Sprachen am günstigsten abschneiden, die sehr umfangreich sind. In [FrPr97] findet sich der Versuch einer umfassenderen Beurteilung von Modellierungssprachen. Er beruht auf einem Bezugsrahmen, in den Anforderungen aus verschiedenen Blickwinkeln einfließen. Jedes Kriterium wird hinsichtlich seiner Erfäßbarkeit sowie der sich daraus erge-

benden Konsequenzen für eine Modellierungssprache diskutiert. Der Bezugsrahmen wird anschließend auf zwei neuere objektorientierte Modellierungssprachen ([FiHe96], [Rat97]) angewendet. Die Untersuchung hat die erheblichen Schwierigkeiten der Evaluation von Modellierungssprachen deutlich gezeigt. So konnten einige durchaus wichtige Aspekte - wie etwa die Verständlichkeit der Sprachkonzepte und die Anschaulichkeit der Notation - mangels fachlicher Kompetenz nur sehr eingeschränkt analysiert werden. Dessen ungeachtet war die Untersuchung mit einem erheblichen Aufwand verbunden. So mußten neben dem Entwurf des Bezugsrahmens Entwurfsdokumente im Umfang von mehr als tausend Seiten berücksichtigt werden. Ein solcher Aufwand ist vor allem dann wenig motivierend, wenn man es - wie im vorliegenden Fall - nur mit vorläufigen Sprachentwürfen zu tun hat.¹

5.2 Herausforderungen

Eine völlig objektive Beurteilung von konzeptionellen Modellen ist nicht zu erwarten. Diese Feststellung gilt ebenfalls, allerdings mit noch größerem Nachdruck, für Modellierungssprachen. Angesichts der großen Bedeutung solcher Artefakte für die Leistungsfähigkeit betrieblicher Informationssysteme und die Forschung in der Wirtschaftsinformatik ist es allerdings erforderlich, eine wissenschaftlich akzeptable Evaluation soweit wie möglich voranzutreiben. Die in jüngerer Zeit zu verzeichnenden einschlägigen Untersuchungen in der Wirtschaftsinformatik geben - allen Unzulänglichkeiten zum Trotz - Anlaß zur Hoffnung, denn sie sind geeignet, Problembewußtsein zu schaffen und liefern zudem Ansatzpunkte für weitere Forschungen. Dabei ist allerdings zu berücksichtigen, daß die Beurteilung von konzeptionellen Modellen wie auch die von Modellierungssprachen eine erhebliche Bandbreite von Anforderungen berücksichtigen muß. Die Wirtschaftsinformatik allein ist hier überfordert. Sie ist auch gewiß nicht allein zuständig: Die Tragweite, die Modellierungssprachen und den durch sie beschriebenen Artefakten zukommt, geht weit über den engeren Bereich betrieblicher Informationssysteme hinaus. Die Artefakte, die sich langfristig durchsetzen, stellen einen in zweifacher Hinsicht wesentlichen Beitrag zur Konstruktion zukünftiger Welten dar: Nicht nur, daß sie neue Formen der Organisation von Arbeit und des Austausches von Leistung reflektieren, sie beeinflussen auch die Ordnung von Sprache und bestimmen damit gesellschaftlich relevantes Wissen.²

Um den vielfältigen Facetten der Gestaltung und Nutzung von Informationsmodellen und Modellierungssprachen gerecht zu werden, ist sicherlich eine verstärkte Zusammenarbeit mit Disziplinen anzustreben, die im Hinblick auf die

¹ Gerade dieser Umstand bildete allerdings den Anlaß für den Vergleich, da einer der beiden Entwürfe als Vorschlag für einen Standard eingereicht worden war.

² "Wissen objektiviert die Welt durch Sprache und den ganzen Erkenntnisapparat, der auf Sprache beruht. Das heißt, es macht Objekte aus dieser Welt, auf daß sie als Wirklichkeit erfaßt werde." ([BeLu80], S. 71)

die eine oder andere Facette fachlich zuständig sind. Hier ist etwa an die kognitive Psychologie oder die Linguistik zu denken. Dabei muß man zwar berücksichtigen, daß sich solche - gern gestellten - Forderungen nach interdisziplinärer Forschung schwer umsetzen lassen. Ein erster Schritt ist allerdings getan, wenn man den eigenen Kompetenzbereich abgrenzt und andere Disziplinen auf ihre Zuständigkeit aufmerksam macht. Das öffnet die Chance für eine, wie Gutenberg es für die Betriebswirtschaftslehre formuliert hat, "Kommunikation auf der Basis größtmöglichen Sachverstandes" ([Gut89], S. 158).

Unabhängig davon, in welchem Ausmaß es gelingt, interdisziplinäre Kooperationen zu etablieren, muß man bei der Evaluation von Modellen und Modellierungssprachen von einem kontinuierlichen Prozeß der Erkenntnismehrung ausgehen: Auf generelle und gleichzeitig stabile Urteile darf man angesichts der Vielfalt von Modellen und möglicher Modellierungssprachen kaum hoffen. Zudem ist zu erwarten, daß alle an einem solchen Prozeß Beteiligten im Zeitverlauf ihre Ansichten ändern mögen. Daraus ergeben sich m.E. zwei Konsequenzen: Einerseits muß im wesentlichen auf Verfahren der diskursiven Urteilsfindung zurückgegriffen werden, weil fallibilistische Ansätze allenfalls eingeschränkt einsetzbar sind. Die mit diskursiven Ansätzen einhergehenden Probleme (wie etwa eine gewisse Beliebigkeit der Begründung) mahnen zu einem disziplinierten Umgang mit Sprache. Daneben ist es erstrebenswert, nicht nur die einschlägige Forschung zu forcieren, sondern auch die Ergebnisse in vergleichbarer Form aufzubereiten und zu dokumentieren - etwa in der Weise wie man Entwurfsmuster einsetzt, um Frameworks zu dokumentieren. Es muß nicht betont werden, daß eine solche Aufbereitung gerade einer Disziplin, zu deren Untersuchungsgegenständen das Informationsmanagement gehört, gut zu Gesicht steht. Auf diese Weise könnte im Zeitverlauf ein Verzeichnis von beispielhaften Modellen und deren Beurteilung aus verschiedenen Sichten entstehen¹ - und damit eine Grundlage für eine zukünftige Modellierungslehre.

6 Abschließende Bemerkungen

Die geradezu dominierende Rolle, die Artefakten im Gegenstandsbereich der Wirtschaftsinformatik zukommt, macht eine Evaluation konkurrierender Angebote unvermeidlich. Das gilt für Entscheidungssituationen in der Praxis wie auch für die Forschung in einer Disziplin, zu deren wesentlichen Gegenständen die Untersuchung von Gestaltungs- und Einsatzbedingungen solcher Artefakte zählt. Es wäre Ausdruck eines naiven Positivismus, würde man eine vollständige und dabei objektive Evaluation für möglich halten. Die Möglichkeiten der Wirtschaftsinformatik sind hier ausgesprochen bescheiden. Dennoch ist der Rekurs auf Bewährung in der Praxis ("best practice") oder die Befragung von O-

¹ In [Fra98b] findet sich ein Bezugsrahmen für solche multiperspektivischen, evolutionären Forschungsprozesse.

rakeln (Ergebnisse sog. Expertenbefragungen) m.E. nicht zu akzeptieren, da er zur Preisgabe eines wissenschaftlichen Anspruchs führt.

Die in diesem Beitrag skizzierten Überlegungen sind einerseits von grundsätzlicher Bedeutung für die Wirtschaftsinformatik, andererseits haben sie auch eine hohe Aktualität. Die gegenwärtige Standardisierung einer objektorientierten Modellierungssprache vollzieht sich ohne nennenswerte Beteiligung der Hochschulforschung, da die OMG eine kommerzielle Verankerung der eingereichten Beiträge fordert ([OMG96]). Dessen ungeachtet können wir davon ausgehen, daß dieser Standard Forschung und Lehre in der Wirtschaftsinformatik nachhaltig beeinflussen wird. Die Folgen dürften ambivalent sein: Einerseits erhöht die einheitliche Verwendung einer gemeinsamen Modellierungssprache die Vergleichbarkeit von Modellen, andererseits droht sie die Autonomie der einschlägigen Forschung auf subtile Weise einzuschränken. Das sollte uns zu denken geben.

Literatur

- [Alb60] Albert, H.: Wissenschaft und Politik. Zum Problem der Anwendbarkeit einer wertfreien Sozialwissenschaft. In: Topitsch, E.v. (Hg.): Probleme der Wissenschaftstheorie. Wien 1960, S. 201-232
- [Ale91] Alexy, R. (Hg.): Rechts- und Sozialphilosophie in Deutschland heute. Beiträge zur Standortbestimmung. Stuttgart 1991
- [App91] Appelt, W.: Document Architecture in Open Systems. The ODA Standard. Berlin, Heidelberg u.a 1991
- [BaCe92] Batini, C.; Ceri, S.; Navathe, S.B.: Conceptual Database Design. Redwood City 1992
- [BeRo95] Becker, J.; Rosemann, B.; Schütte, R.: Grundsätze ordnungsmäßiger Modellierung. In: Wirtschaftsinformatik, Nr. 37, Heft 5, 1995, S. 435-445
- [BeLu80] Berger, P.L.; Luckmann, T. : Die gesellschaftliche Konstruktion der Wirklichkeit. Eine Theorie der Wissenssoziologie. Frankfurt/M. 1980
- [Big97] Biggs, P.: A Survey of Object-Oriented Methods. (<http://www.dur.ac.uk/~dcs3pjb/survey.html>)
- [Buc90] Buchwald, D.: Der Begriff der rationalen juristischen Begründung. Zur Theorie der juristischen Vernunft. Baden-Baden 1990
- [Buh97] Buhl, H.U.: Best practices vs. common practices bei der Softwareentwicklung. In: Wirtschaftsinformatik, Heft 6, 1997, S. 639 f.
- [DeFa92] De Champeaux, D.; Faure, P.: A comparative study of object-oriented analysis methods. In: Journal of Object-Oriented Programming, No. 1, Vol. 5, 1992, S. 21-33

- [FiHe96] Firesmith, D.; Henderson-Sellers, B.; Graham, I.; Page-Jones, M.: OPEN Modeling Language (OML). Reference Manual. Version 1.0. December 1996, (<http://www.csse.swin.edu.au/OPEN/comn.html>)
- [Fra98a] Frank, U.: Zur Verwendung formaler Sprachen in der Wirtschaftsinformatik: Notwendiges Merkmal eines wissenschaftlichen Anspruchs oder Ausdruck eines übertriebenen Szientismus? Erscheint in: Tagungsband der Tagung "Wirtschaftsinformatik und Betriebswirtschaftslehre", 1998, in Druck
- [Fra98b] Frank, U.: Evaluating Modelling Languages: Relevant Issues, Epistemological Challenges and a Preliminary Research Framework. Arbeitsberichte des Instituts für Wirtschaftsinformatik. Nr. 15, Koblenz 1998
- [Fra98c] Frank, U.: The MEMO Object Modelling Language (MEMO-OML). Arbeitsberichte des Instituts für Wirtschaftsinformatik. Nr. 10, Koblenz 1998
- [FrPr97] Frank, U.; Prasse, M.: Ein Bezugsrahmen zur Beurteilung objektorientierter Modellierungssprachen - veranschaulicht am Beispiel von OML und UML. Arbeitsberichte des Instituts für Wirtschaftsinformatik. Nr. 6, Koblenz 1997
- [GaHe95] Gamma, E.; Helm, R.; Johnson, R.; Vlissides, J.: Design Patterns. Elements of Reusable Object-Oriented Software. Reading/Mass. et al. 1995
- [GoRu90] Goldfarb, C.; Rubinsky, Y.: The SGML Handbook. Oxford 1990
- [Gut89] Gutenberg, E.: Zur Theorie der Unternehmung. Schriften und Reden Erich Gutenbergs aus dem Nachlaß, Hg. von H. Albach, Berlin 1989
- [Hab81] Habermas, J.: Theorie des kommunikativen Handelns. Bd. 1: Handlungsrationalität und gesellschaftliche Rationalisierung. Frankfurt/M. 1981
- [Ham82] Hamminga, B.: Neoclassical Theory Structure and Theory Development: The Ohlin Samuelson Programme in the Theory of International Trade, Stegmüller, W.; Balzer, W.; Spohn, W. (Hg.), Philosophy of Economics, Berlin, Heidelberg et al. 1982, S. 1-15
- [HoGo93] Hong, S.; Goor, G.: A Formal Approach to the Comparison of Object-Oriented Analysis and Design Methodologies. In: Nunamaker, J.F.; Sprague, R.H. (Hg.): Information Systems: Collaboration Technology, Organizational Systems, and Technology. Proceedings of the 26th International Hawaii International Conference on System Sciences. Los Alamitos 1993, S. 689-698
- [Hou93] House, E.R.: Professional Evaluation: Social Impact and Political Consequences. Newbury Park, Ca. 1993
- [Hsi92] Hsieh, D.: Survey of object-oriented analysis/design methodologies and future CASE frameworks. Menlo Park, Ca. 1992
- [JaCh92] Jacobson, I.; Christerson, M; Jonsson, P; Overgaard, G.: Object-Oriented Engineering. A Use Case Driven Approach. Reading, Mass. 1992
- [Koc90] Koch, U. (Hg.): Evaluationsforschung. Bewertungsgrundlage von Sozial- und Gesundheitsprogrammen. Berlin u.a. 1990

- [KrLi95] Krogstie, J.; Lindland, O.I.; Sindre, G.: Towards a Deeper Understanding of Quality in Requirements Engineering. In: Iivari, K.; Lyytinen, K.; Rossi, M. (Eds.): Proceedings of the 7th Conference on Advanced Information Systems Engineering (CaiSE '95). Berlin et al. 1995, pp. 82-95
- [Kuh70] Kuhn, T.S.: The structure of scientific revolutions. 2nd Ed., Chicago, Ill. 1970
- [Lor74] Lorenzen, P.: Konstruktive Wissenschaftstheorie. Frankfurt/M. 1974
- [Mai96] Maier, R.: Qualität von Datenmodellen. Wiesbaden 1996
- [McC85] McCloskey, D.N.: The Rethoric of Economics. Madison, Wis. 1985
- [Mer97] Mertens, P.: Best practice. In: Wirtschaftsinformatik, Heft 6, 1997, S. 640
- [MoPu92] Monarchi, D.E.; Puhr, G.: A Research Typology for Object-Oriented Analysis and Design. In: Communications of the ACM, Vol. 35, No. 9, 1992, S. 35-47
- [MoSh94] Moody, D.L.; Shanks, S.: What Makes a Good Data Model? Evaluating the Quality of Entity Relationship Models. In: Loucopoulos, P. (Hg.): Entity-Relationship Approach - ER'94. Business Modelling and Re-Engineering. 13th International Conference on the Entity-Relationship Approach. Berlin, Heidelberg etc. 1994, S. 94-111
- [MyLe85] Mylopoulos, J.; Levesque, H.J.: An Overview of Knowledge Representation. In: Brodie, M.L.; Mylopoulos, J.; Schmidt, J. (eds.): On Conceptual Modelling. Perspectives from Artificial Intelligence, Databases and Programming. Berlin/Heidelberg: Springer 1985, pp. 3-17
- [Myr33] Myrdal, G.: Das Zweck-Mittel-Denken in der Nationalökonomie. In: Zeitschrift für Nationalökonomie. 1933, S. 305-329
- [OMG96] Object Management Group: Object Analysis & Design RFP-1, ad/96-05-01, 1996, (<http://www.omg.org/library/public-doclist.html>)
- [Ope94] The OpenDoc Design Team: OpenDoc Technical Summary. Component Integration Laboratories, 1994 (<http://www.cilabs.org/>)
- [Paw97] Pawson, Ray, Tilley, N.: Realistic Evaluation. London 1997
- [Rat97] Rational: UML Semantics. Version 1.1 alpha R6, 07-21-97, 1997. (<http://www.rational.com>)
- [ReWi97] Reitwieser, N.; Will, A.: Best practice oder common practice: Eine Frage der Wirtschaftlichkeit. In: Wirtschaftsinformatik, Heft 6, 1997, S. 640 f.
- [Sch98] Schütte, R.: Grundsätze ordnungsmäßiger Referenzmodellierung. Konstruktion konfigurations- und anpassungsorientierter Modelle. Wiesbaden 1998
- [Scr80] Scriven, M.S.: The Logic of Evaluation. Inverness, Ca. 1980
- [SüEb97] Süttenbach, R.; Ebert, J.: A Booch Metamodel. Fachberichte Informatik 5/97, Universität Koblenz-Landau 1997
- [Tim95] Timmreck, T.C.: Planning, Program Development, and Evaluation. A Handbook for Health Promotion, Aging and Health Services. Boston u.a. 1995

- [TvKa82] Tversky, A.; Kahnemann, D.: Judgement under Uncertainty: Heuristics and Biases. In: Kahnemann, D.; Slovic, P.; Tversky, A. (Hg.): Judgement under Uncertainty: Heuristics and Biases. Cambridge 1982
- [Web97] Weber, R.: Ontological Foundations of Information Systems. Melbourne 1997
- [WeWi97] Wernick, P.; Winder, R.L.: Software Engineering as a Kuhnian Discipline. In: Winder, R.L.; Probert, S.K.; Beeson, I.A. (Eds.): Philosophical Aspects of Information Systems. London 1997, S. 117-129
- [Wit80] Wittgenstein, L.: Philosophische Untersuchungen. 2. Aufl., Frankfurt/M. 1980