

Nicole Thees

Statistische Qualität zusammengesetzter Indikatoren am Beispiel der Wissensökonomie

1. Einleitung

Das von der Europäischen Kommission innerhalb des sechsten Rahmenprogramms geförderte Projekt KEI (Knowledge Economy Indicators, <http://www.kei.publicstatistics.net>) ist Teil eines Spezialprogramms politikorientierter Forschung zur Integration und Stärkung des europäischen Forschungsraumes. Projektstart war im September 2004. Es endet nach einer Laufzeit von 44 Monaten im Mai 2008. Ziel des Projektes ist es, Indikatoren für die Wissensgesellschaft (Knowledge Economy) zu verbessern und weiter zu entwickeln. Hinzu kommen die Analyse der regionalen und überregionalen Aussagefähigkeit sowie die Konstruktion und Nutzung leistungsfähiger zusammengesetzter Indikatoren. Zu den Indikatoren der Wissensgesellschaft zählen beispielsweise "Schulabschlussrate", "Ausstattung der Haushalte mit Computern" oder "Ausgaben für Forschung und Entwicklung". Indikatoren dieser Art spielen eine bedeutende Rolle bei zahlreichen politischen Entscheidungen, so z.B. bei der finanziellen Förderung durch die EU mit Blick auf unterschiedliche wirtschaftliche und technische Entwicklungsstandards oder bei der Erfüllung der Lissabon-Ziele.

Im Forschungsprojekt kooperieren sechs Institutionen, die Universität Trier, die Eberhard Karls Universität Tübingen, das Joint Research Centre der Europäischen Kommission in Ispra (Italien), die Katholische Universität Leuven (Belgien), die Universität Maastricht (Niederlande) sowie das Statistische Amt Finnlands. Die wissenschaftliche Leitung des Projektes obliegt Ralf Münnich (Universität Trier).

Ausgangspunkt des Projektes ist die Analyse der Kernaspekte der wissensbasierten Ökonomie und deren Definition, zu deren Abbildung ein Indikatorenset entwickelt wurde. Zu diesem schließlich aus 144 Indikatoren bestehenden Bündel sind entsprechende Daten aus 30 europäischen und sechs nicht-europäischen Ländern zusammenzutragen, wobei im Verlauf des Projektes auch mögliche Lücken und Mängel zu analysieren sind, um das zukünftige Verständnis der wissensbasierten Ökonomie zu verbessern.

Ein zweiter Schwerpunkt des Projektes liegt in der Entwicklung und Analyse von zusammengesetzten Indikatoren, um mittels einer Kenngröße einen ersten Eindruck über den Stand der wissensbasierten Ökonomie zu erhalten. In diesem Kontext werden dabei auch unterschiedliche Aggregations- und Gewichtungsansätze sowie Präsentationstechniken näher betrachtet.

Die skizzierten inhaltlichen und methodischen Aspekte des Projektes werden schließlich um den Datenqualitätsanalyse abgerundet. Die Qualitätsrelevanz begründet sich nicht nur in der Tatsache, dass der zusammengesetzte Indikator als Grundlage des politischen Entscheidungsprozesses dienen könnte und somit seine Genauigkeit und Relevanz unabdingbar ist. Vielmehr erfordert die Vielfalt der Datenquellen eine ebenso exakte Betrachtung der Qualität: so haben 24 der schlussendlich verwendeten Indikatoren administrative Quellen zur Grundlage, 79 stammen aus amtlichen als auch nicht-amtlichen Stichprobenerhebungen und insgesamt 21 stammen aus einer Kombination administrativer Quellen wie auch Stichproben oder Zensen. Diese Quellenvariabilität resultiert aus der Vielzahl der betrachteten Länder, die für identische Fragestellungen unterschiedliche Datenquellen heranziehen. Schließlich blieb bei einem Indikator die Quelle ungeklärt und zu 19 ursprünglich angedachten konnten keine adäquaten Daten gefunden werden. Gerade auch die Breite der nicht-amtlichen Datenquellen erforderte dabei auch eine methodische Absicherung vor deren Nutzung. Letztendlich sollte die Qualitätsanalyse auch als Mittel der Indikatorenselektion herangezogen werden.

Zur Realisierung dieser Qualitätsanalyse entstand schließlich die Überlegung, in Anlehnung an den zu entwickelnden zusammengesetzten Indikator der wissensbasierten Ökonomie und dessen Gewichtungsschema, einen zusammengesetzten Indikator der Qualität zu generieren, der es ermöglicht, auf einen Blick einen ersten Eindruck über die Qualität zu vermitteln.

2. Definition der Einzelkomponenten

Ein zusammengesetzter Indikator basiert auf Einzelindikatoren, deren Grundlage Daten sind. Somit erfordert die Analyse der Qualität eines zusammengesetzten Indikators zunächst die Analyse der Qualität seines Inputs, wobei hier zunächst eine genaue Abgrenzung vorzunehmen ist.

Ein Indikator bzw. eine Variable ist ein empirisch beobachtbarer Sachverhalt, der unter Anwendung einer Messvorschrift ein theoretisches nicht-beobachtbares Konstrukt abbilden soll (Schnell, Hill, Esser, 2005, S. 131ff.). Ziel ist es, die Diskrepanz zwischen theoretischem und statistisch beobachtbarem Konstrukt auf ein logisches bzw. begriffliches Minimum zu reduzieren, was eine genuine nicht-stochastische Erkenntnisaufgabe der wirtschafts- und sozialstatistischen Methodologie ist (Menges, 1972, S. 40). Diese Diskrepanzreduktion bzw. Annäherung wird als Adäquation bezeichnet. Im KEI Projekt soll beispielsweise die Produktion und Verbreitung von Informations- und Kommunikationstechnologie (theoretisches Konstrukt) analysiert werden. Messbar ist hingegen beispielsweise *business receiving orders over the internet*, die somit einen Indikator für das theoretische Konstrukt bilden würden. Dabei werden diese wiederum aus einer Vielzahl beobachtbarer Fakten, also Daten, abgeleitet. Der Indikator ist somit zumeist eine singuläre Größe, eine Rate oder ein Durchschnitt, basierend auf einer Vielzahl von Daten (De Vries, 2001, S. 315) und die Frage der Adäquation ist dann u.a. ein Qualitätsaspekt eines Indikators neben der Frage der Qualität der Daten.

Ein zusammengesetzter Indikator ist schließlich die aggregierte und gewichtete Zusammenfassung der Einzelkomponenten, wobei folgende zehn Schritte zu durchlaufen bzw. berücksichtigen sind (Nardo, Saisana, Saltelli, Tarantola, Hoffman, Giovannini, 2005):

- | | | |
|-------------------------------|--|-------------------------------------|
| 1. Modellentwicklung | 5. Datenstandardisierung | 8. Verlinkung mit anderen Variablen |
| 2. Datenauswahl | 6. Gewichtung & Aggregation | 9. Dekomposition |
| 3. Multivariate Analyse | 7. Robustheits- & Sensitivitätsanalyse | 10. Präsentation & Veröffentlichung |
| 4. Imputation fehlender Werte | | |

Die Betrachtung der Qualität des zusammengesetzten Indikators erfordert folglich die Analyse der Qualität in den einzelnen Schritten. Somit deutet die Definition der Einzelkomponenten des zusammengesetzten Indikators bereits auf die Komplexität der Qualitätsanalyse hin und verdeutlicht, dass es bisher an einem Gesamtkonzept der Qualität eines zusammengesetzten Indikators mangelt. Denn nicht lediglich die originäre Datenqualität ist von Relevanz, sondern wie bereits erwähnt, auch der Prozess der Adäquation aller Einzelindikatoren wie auch die Entwicklungsstufen des zusammengesetzten Indikators haben entscheidenden Einfluss auf die Qualität des Gesamtkonstruktes. Im Folgenden werden daher zunächst die Qualitätsaspekte der Einzelkomponenten näher betrachtet und existente Konzepte präsentiert.

3. Qualitätskonzepte der Einzelkomponenten

Die allgemeine Definition der Qualität deutet bereits auf die Komplexität der Fragestellung hin: laut DIN EN ISO 9000:2000 ist sie das "Vermögen einer Gesamtheit inhärenter Merkmale eines Produktes, eines Systems oder eines Prozesses zur Erfüllung von Forderungen von Kunden und anderen interessierten Parteien" zumeist mit der Nebenbedingung der Belastungs- und Kostenreduktion. Diese Komplexität hat sich in den letzten Jahren auch in Bezug auf die Datenqualität herausgestellt und der ursprüngliche Ansatz, sich lediglich auf die Genauigkeit zu beziehen, wurde deutlich erweitert. So existiert derzeit bereits eine Vielzahl an Datenqualitätskonzepten internationaler Institutionen (OECD, IMF, Eurostat), zumeist basierend auf dem Total-Quality-Management-Gedanken. Ihre Ausarbeitung und Umsetzung ist bereits sehr weit vorangeschritten. So wurde von Eurostat im Jahr 2005 ein Verhaltenskodex der Europäischen Statistik vorgestellt, der die Erfüllung der im Folgenden aufgeführten Qualitätsdimensionen differenziert nach drei Ebenen fordert.

- Institutioneller Rahmen: fachliche Unabhängigkeit, Datenerhebungsauftrag, angemessene Ressourcen, Qualitätsverpflichtung, statistische Geheimhaltung, Unparteilichkeit und Objektivität.
- Statistischer Produktionsprozess: solide Methodik, geeignete statistische Verfahren, Vermeidung übermäßiger Belastung der Auskunftspflichtigen, Wirtschaftlichkeit.

→ Statistischer Output: Relevanz, Genauigkeit und Zuverlässigkeit, Aktualität und Pünktlichkeit, Kohärenz und Vergleichbarkeit, Zugänglichkeit und Klarheit.

Letztgenannte Ebene bezieht sich dabei auf die Datenqualität. *Relevanz* wird dabei definiert als die Fähigkeit, aktuellen und zukünftigen Nutzeranforderungen gerecht zu werden. Sie bezieht sich darauf, ob alle notwendigen Daten erhoben werden und zu welchem Grad die verwendeten Konzepte (Definitionen, Klassifikationen etc.) die Nutzerbedürfnisse widerspiegeln. *Genauigkeit* im klassischen Sinn der Statistik ist die Nähe der Schätzung zum wahren Wert und somit die Höhe des statistischen Fehlers, der wiederum zu differenzieren ist in einen Stichprobenzufallsfehler und einen systematischen Fehler. Die dritte Dimension, die *Pünktlichkeit*, bezieht sich auf die zeitliche Diskrepanz zwischen der Veröffentlichung der Daten und dem Zeitpunkt der geplanten Veröffentlichung, der beispielsweise durch Veröffentlichungskalender oder Verordnungen vorgegeben ist. Die *Aktualität* hingegen beschreibt die Diskrepanz zwischen der Verfügbarkeit der Daten und ihrem Bezugszeitpunkt. Die Möglichkeiten, Bedingungen und der Umfang des Nutzers, Zugriff auf die Daten zu erlangen, werden als Qualitätskriterium des *Zugangs* definiert. Die *Klarheit* umfasst hingegen das zusätzliche Informationsangebot, das in Ergänzung zu den Daten bereitgestellt wird, um sie verstehen und beurteilen zu können (Eurostat, 2003a). Die vorletzte Qualitätsdimension, die *Vergleichbarkeit*, soll die Möglichkeit bieten, zeitliche und räumliche Veränderungen, die aus Unterschieden in den angewendeten statistischen Konzepten und Messmethoden resultieren, zu bewerten (Eurostat 2003b). Letztlich soll die Dimension der *Kohärenz* Aufschluss darüber geben, in wieweit die Möglichkeit zur reliablen Verknüpfung der Daten in unterschiedlicher Art und Weise und für unterschiedliche Bedürfnisse besteht. Insbesondere soll wiedergegeben werden, in wieweit eine gemeinsame Nutzung von Daten unterschiedlicher Datenquellen und unterschiedlichem primärstatistischen Auftrag möglich ist (Eurostat 2003a).

Tabelle 1: Datenqualität

Dimensionen	Indikatoren	Typ
Relevanz	R1. Nutzerzufriedenheitsindex	3
	R2. Rate der Verfügbarkeit von Statistiken	1
Genauigkeit (und Zuverlässigkeit)	A1. Variationskoeffizient	1
	A2. Unit-response Rate (ungewichtet/gewichtet)	2
	A3. Item-response Rate (ungewichtet/gewichtet)	2
	A4. Imputationsrate und -verhältnis	2
	A5. Überdeckung und Fehlklassifikationsrate	2
	A6. Räumliche Unterdeckungsrate	1
	A7. Durchschnittliche Revisionsrate	1
Aktualität & Pünktlichkeit	T1. Pünktlichkeit der effektiven Veröffentlichung	1
	T2. Zeit zwischen Referenzzeitpunkt und den 1. Ergebnissen	1
	T3. Zeit zwischen Referenzzeitpunkt und endgültigen Ergebnissen	1
Zugang & Klarheit	AC1. Anzahl der veröffentlichten Publikationen	1
	AC2. Anzahl der Zugänge zur Datenbank	1
	AC3. Vollständigkeitsrate der Metadaten veröffentlichter Statistiken	3
Vergleichbarkeit	C1. Länge vergleichbarer Zeitreihen	1
	C2. Anzahl vergleichbarer Zeitreihen	1
	C3. Abweichungsrate zu europäischen Konzepten und Methoden	3
	C4. Asymmetrien bei Bilanzstatistiken	1
Kohärenz	CH1. Anforderungserfüllungsrate der Statistiken zur Sekundärnutzung	3

Die Eurostat Arbeitsgruppe Assessment of Quality in Statistics hat schließlich zur Messung dieser Qualitätsdimensionen eine Liste an Standardqualitätsindikatoren ausgearbeitet (vgl. Tabelle 1), die die Möglichkeit zur quantifizierbaren Messung der Qualitätserfüllung primär von Eurostat publizierter Daten bieten soll. Es werden dabei drei Indikatorentypen unterschieden. Sogenannte Schlüsselindikatoren (1), sie sind repräsentativ für die jeweilige Qualitätsdimension, ihre Methodik ist bereits ausgereift und sie sind leicht zu interpretieren. Die unterstützenden Indikatoren (2) sind bedeutende indirekte Messgrößen der jeweiligen Qualitätsdimension und schließlich gibt es Größen, die noch der weiteren Ausarbeitung bedürfen (3) (Ehling, Körner 2007).

Es wäre nun denkbar, in Anlehnung an den zu entwickelnden zusammengesetzten Indikator der wissensbasierten Ökonomie einen zusammengesetzten Indikator der Datenqualität basierend auf den soeben aufgelisteten Einzelindikatoren der Qualität zu entwickeln.

Im Gegensatz zur Datenqualität ist die Qualität von Indikatoren insbesondere mit Blick auf das Adäquationsproblem noch weitaus weniger diskutiert. De Vries (2001) diskutiert im Zusammenhang mit Politikindikatoren insgesamt vier primäre Qualitätsdimensionen: die technische Einwandfreiheit, die Verständlichkeit, Relevanz und Kosteneffizienz. Mit Bezug auf die erste Dimension soll die Frage nach der Genauigkeit, Validität, Reliabilität, Sensitivität und Messbarkeit beantwortet werden, somit streng genommen auch Aufschluss über die Frage der Adäquation gegeben werden. Der zweite Punkt zielt insbesondere auf den Aspekt Nutzerfreundlichkeit ab. Spezifische Vorschläge zur Messung der Qualitätsdimensionen macht de Vries indes nicht, ebenso fehlt eine einzelne aussagekräftige Maßzahl, die die Information zur Indikatorenqualität widerspiegeln könnte.

Es ist nunmehr zu überlegen, in wieweit die Datenqualitätsdimensionen Einfluss auf die der Indikatorenqualität haben (vgl. Tabelle 2). Die Datenrelevanz steht in direktem Zusammenhang zur Indikatorenrelevanz; haben Daten beispielsweise keine Relevanz für die Politik, können sie diese auch nicht für einen Politikindikator haben. Würden indes doch Daten verwendet, die nicht relevant sind, hätte dies Konsequenzen für die Genauigkeit ebenso wie für die Validität des Indikators. Auch würden nicht die Veränderungen gemessen, die gemessen werden sollen, demnach hätte dieser Aspekt auch Konsequenzen für die Sensitivität. Vergleichbare Schlüsse bezüglich des Einflusses auf die Indikatorenqualität können für ungenaue oder inkohärente Daten gezogen werden, jedoch wird zusätzlich auch die Reliabilität des Indikators beeinflusst. Denn erfolgen unterschiedliche systematische Fehler bei der Datenerhebung, ist es fraglich, ob auch der Indikator reliabel bleibt. Die Zusammenhänge zwischen Aktualität und Pünktlichkeit, Zugang und Klarheit wie auch Vergleichbarkeit sind offenkundig.

Tabelle 2: Dimensionen der Daten- und Indikatorenqualität

Datenqualität - Dimensionen	Indikatorenqualität - Dimensionen							
	Genauigkeit	Validität	Reliabilität	Sensitivität	Messbarkeit	Nutzerfreundlichkeit	Relevanz	Kosteneffizienz
Relevanz	X	X		X	X		X	
Genauigkeit	X	X	X	X	X		X	
Aktualität & Pünktlichkeit						X		X
Zugang & Klarheit						X		X
Vergleichbarkeit			X		X		X	
Kohärenz	X	X	X	X	X		X	

Offen bleibt indes die Frage, wie diese Bereiche der Qualität zu messen sind. Ein Indikatorensystem wird von de Vries nicht vorgeschlagen, jedoch ist es denkbar, in Anlehnung an das europäische Indikatorensystem für Daten eines für Indikatoren zu entwickeln. Dies würde wiederum bedeuten, dass ein zusammengesetzter Qualitätsindikator der Einzelindikatoren ermittelt werden könnte.

Schließlich muss die Qualität des zusammengesetzten Indikators betrachtet werden, wobei hier dessen umfassende Konstruktionsschritte zu berücksichtigen sind. Hier greifen Nardo et al. (2005) in ihrem *Handbook on Constructing Composite Indicators* weitestgehend auf die bereits im Zusammenhang mit der Datenqualität bekannten Dimensionen zurück. Lediglich die Dimensionen *Interpretierbarkeit* und *Glaubwürdigkeit* sind zusätzlich aufgenommen bzw. die Dimension *Vergleichbarkeit* wurde bei ihnen nicht explizit berücksichtigt und wird daher in Verbindung mit der Kohärenz aufgeführt (vgl. Tabelle 3). Sie zeigen überdies auf, wie die einzelnen Schritte der Konstruktion Einfluss auf die Qualität des zusammengesetzten Indikators nehmen. Insbesondere ist zu beachten, dass die bereits erörterte Qualität des Inputs, die der Daten und der Einzelindikatoren, zu berücksichtigen ist. Sie fließen im zweiten Schritt, dem der *Datenauswahl* oder korrekter der *Indikatorauswahl*, in die Qualitätserörterung ein und beeinflussen die Relevanz, Genauigkeit, Pünktlichkeit und auch Glaubwürdigkeit des zusammengesetzten Indikators.

Tabelle3: Indikatorenqualität

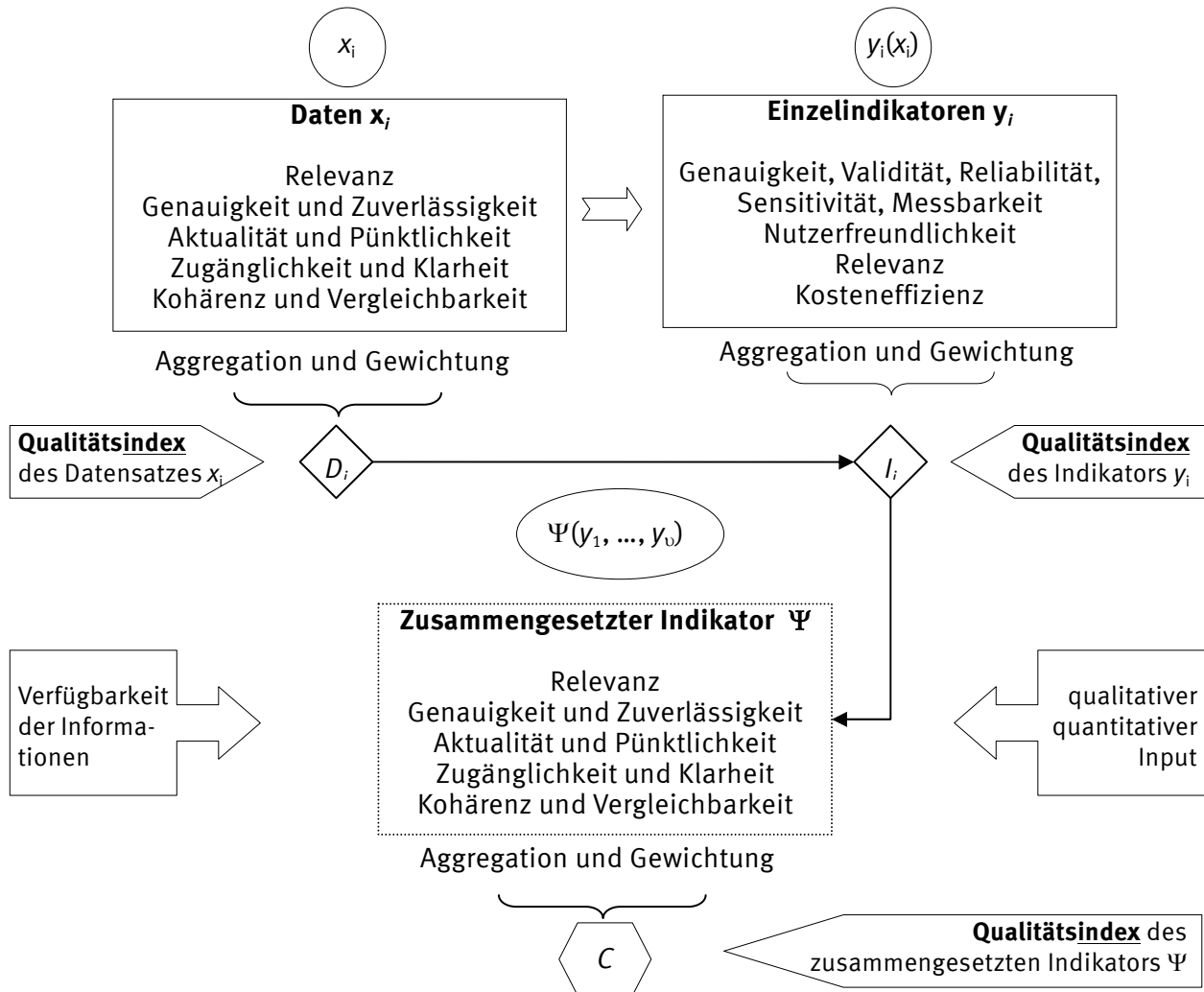
Entwicklungsschritte	Dimensionen						
	Relevanz	Genauigkeit (Zuverlässigkeit)	Pünktlichkeit (Aktualität)	Zugänglichkeit (Klarheit)	Koheränz (Vergleichbarkeit)	Inter- pretierbarkeit	Glaubwürdigkeit
Modellentwicklung	X					X	X
Datenauswahl	X	X	X				X
Multivariate Analyse		X			X	X	X
Imputation fehlender Werte	X	X	X				X
Standardisierung		X			X	X	X
Gewichtung & Aggregation	X	X			X	X	X
Robustheits- & Sensitivitätsanalyse		X				X	X
Verlinkung mit anderen Variablen	X				X	X	X
Dekomposition	X			X		X	X
Visualisierung & Veröffentlichung	X					X	X

Auch in diesem Fall bleibt wiederum die Frage offen, wie die jeweilige Erfüllung der Qualität gemessen werden kann. Vorschläge zu einem einzigen, die Qualität des zusammengesetzten Indikators insgesamt widerspiegelnden Maß werden nicht gegeben, doch ist gerade hier zu überlegen, ob das bereits zur Analyse der Datenqualität vorgeschlagene Konzept auf das für einen zusammengesetzten Indikator übertragen werden kann.

4. Gesamtkonzept der Qualität

Werden nun die im Einzelnen diskutierten Qualitätsaspekte zu einem Gesamtkonzept zusammengeführt, würde sich die Ermittlung der Qualität eines zusammengesetzten Indikators wie folgt darstellen: für die Daten x_i werden die Indikatoren der Datenqualitätsdimensionen basierend auf einer Gewichtungsfunktion aggregiert und so ein Qualitätsindex D_i jedes einzelnen der i verwendeten Datensätze entwickelt. Im nächsten Schritt wird die Qualität jedes einzelnen Indikators y_i ermittelt, die sich wiederum aus den Einzeldimensionen der Qualität gewichtet aggregiert, wobei die Datenqualitätsaspekte mit einfließen. Ergebnis wäre ein Qualitätsindex für jeden der Einzelindikatoren I_i . Bei der Evaluation der Qualität des zusammengesetzten Indikators müssten diese wiederum unter Verwendung der Gewichtungsfunktion des eigentlichen zusammengesetzten Indikators Berücksichtigung bei der Erörterung der Datenauswahl finden. In Abbildung 1 ist der Gedankengang zusammenfassend veranschaulicht.

Abbildung 1: Qualität eines zusammengesetzten Indikators



Dies ist derzeit lediglich ein theoretisches Konzept, wie die Qualität eines zusammengesetzten Indikators abgebildet werden könnte. Es bleibt zu fragen, in wieweit dieses auch tatsächlich umgesetzt werden kann und auch Aussagekraft hat. Bei näherer Betrachtung zeigt sich nämlich, dass es derzeit bereits bei der Datenqualitätsmessung erhebliche empirische Schwierigkeiten gibt, denn gerade zwischen den Informationen, die den Datenproduzenten bzw. den Nutzern vorliegen, besteht aufgrund der Neuartigkeit der Diskussion eine große Diskrepanz.

Konkret haben sich Linden und Pellegrino in ihrem Beitrag (Eurostat 2006) bereits mit der Frage der Konstruktionsmöglichkeit eines zusammengesetzten Datenqualitätsindikators auseinandergesetzt. Dabei stellten sie eine theoretische Einfachheit der Konstruktion fest, insbesondere sofern die Datenqualitätsindikatoren quantitativ sind. Jedoch führen gerade konzeptionelle und praktische Gesichtspunkte zu Schwierigkeiten der Ermittlung: so bedingen Länderunterschiede z.B. rechtlicher Natur oder auch fehlende Standards für alle Bereiche des Erhebungsprozesses - wobei deren Definition nicht unumwundlich deren Anwendung impliziert - Probleme bei der Ermittlung eines geeigneten vergleichbaren Indikators. Hinzu kommt, dass davon auszugehen ist, dass wie bei jedem zusammengesetzten Indikator es auch in diesem Fall zu unterschiedlichen Forderungen bezüglich der Gewichtungsfunktion kommt. Auch ist die zeitliche Vergleichbarkeit fraglich, sofern es zu gravierenden Änderungen im Messprozess kommt.

Im Verlauf des Projektes wurden schließlich zu den bereits von Linden und Pellegrino angesprochenen Problemen der Datenqualitätsmessung noch weitere augenscheinlich: sofern dem Endnutzer denn Informationen zur Datenqualität vorliegen, haben diese zumeist qualitativen Charakter. Vor al-

lem aber fehlen solch immanenten Informationen für Survey-Statistiker wie Standardfehler oder Non-Response-Raten. Hinzu kommt, dass es bei den projektrelevanten Datensätzen lediglich auf allgemeiner EU-Ebene generelle Informationen zu den Datensätzen gibt. Länderspezifische Qualitätsangaben fehlen weitestgehend bzw. werden zumindest nicht zentral über Eurostat publiziert, so dass der Endnutzer kaum Zugangsmöglichkeiten hat (z.B. auch unter sprachlichen Gesichtspunkten). Schließlich fehlen auch weitestgehend die sogenannten Qualitätsberichte, die idealerweise auf den Qualitätsindikatoren fußen sollen. Einen Extrakt aus den potentiellen Qualitätsberichten stellen lediglich sogenannte Qualitätsprofile dar. Für insgesamt 19 Strukturindikatoren werden diese auf EU-Ebene publiziert, auf Länderebene fehlen sie indes.

Inhalt dieser Qualitätsprofile sind generelle Informationen zur allgemeinen Beurteilung der Genauigkeit und Vergleichbarkeit, Ziel und Relevanz der Daten, Details zur Datenverfügbarkeit, der generellen Genauigkeit, der Vergleichbarkeit zwischen Ländern und über die Zeit sowie den Entwicklungsperspektiven der Datenqualität. Sie beinhalten somit nicht Angaben zu den einzelnen Datenqualitätsindikatoren sondern lediglich einen qualitativen Überblick mit einem ersten Qualitätsrating.

Die gängigste Informationsquelle des Endnutzers, auch im Projekt, um einen ersten Eindruck über die Daten, jedoch nicht detailliert zu deren Qualität, zu erhalten, sind die in Verbindung mit der Datenbank NewCronos von Eurostat veröffentlichten Metadaten. Sie stehen dabei in keinem Zusammenhang zu den in Tabelle 1 dargestellten Datenqualitätsindikatoren. Sofern vorhanden, findet sich lediglich unter der Rubrik Qualität ein Hinweis auf die Existenz möglicher soeben vorgestellter Qualitätsprofile. Die Metadaten sind untergliedert in eine Basisseite mit sehr allgemeinen Informationen und einen Überblick über die Methodik. Konkret finden sich folgende Informationen:

Tabelle4: Metadaten

Basisseite	Methodologiezusammenfassung
Daten: Erfassungsbereich, Periodizität und Aktualität	Konzepte, Definitionen & Klassifikationen
Zugang der Öffentlichkeit - Vorherige Bekanntgabe des Veröffentlichungskalenders - Gleichzeitige Freigabe für alle Nutzer	Erfassungsbereiche, zeitliche und räumliche Abdeckung
Integrität - Verbreitung der Regeln für die Erstellung amtlicher Statistiken - Interner Zugang zu den Daten vor ihrer Freigabe - Anmerkungen zur Veröffentlichung von Statistiken - Angaben zu Revisionen und Ankündigung wesentlicher methodologischer Änderungen	Buchungsregeln Angaben zu den zugrundeliegenden Daten Aufbereitungsmethoden (Datenverarbeitung)
Qualität - Veröffentlichung: Methodenunterlagen und Datenquellen - Veröffentlichung von Informationen in tieferer Gliederung, Abgleich mit verwandten Daten und Angabe des statistischen Rahmens zur Durchführung von Gegenproben und Plausibilitätskontrollen	Sonstige Aspekte

Werden nun Metadaten herangezogen, um einen ersten Anhaltspunkt zur Datenqualität zu erhalten, ist zu beachten, dass die Verfügbarkeit von Metadaten als solches lediglich einen der vielen ursprünglich vorgestellten Datenqualitätsindikatoren darstellt (vgl. AC3, Tabelle 1). Im Umkehrschluss folgt jedoch, dass mittels dieser Informationen auch lediglich ein rudimentäres Bild über die Qualität entstehen kann. Im Projekt blieb diese Verfügbarkeitsrate jedoch der einzige Anhaltspunkt, sich einen ersten Eindruck über die Qualität der Daten zu verschaffen, eine detaillierte Auswertung der qualitativen Informationen bei der Anzahl betrachteter Datensätze war nicht leistbar.

Um die Ergebnisse der Verfügbarkeitsanalyse besser einordnen zu können, sei im Folgenden zunächst kurz die Indikatorengliederung dargelegt (Tabelle 5). Insgesamt wurden die Daten in drei Oberkategorien gegliedert mit A: Maßgebende, charakteristische Aspekte, mit insgesamt 80 von ursprünglich 91 angestrebten Indikatoren, B: Outputindikatoren mit 26 betrachteten von ursprünglich 28 und schließlich der Gruppe C: Globalisierung, mit 19 von 25. Wie bereits erwähnt resultiert die

Diskrepanz der Indikatoren aus der Tatsache, dass nicht für alle zur Messung der wissensbasierten Ökonomie relevanten Indikatoren auch Daten vorlagen. Zusätzlich wurden die Daten, wie Tabelle 5 zu entnehmen ist, in weitere Unterkategorien untergliedert.

Tabelle 5: Indikatoren der wissensbasierten Ökonomie

→ <i>A1 - Produktion und Verbreitung von IKT (20)</i> - Economic impact of ICT sector (5) - Internet use by firms (7) - Internet use by individuals (5) - Government ICT (3) - Missings (2)	→ <i>A4 - Innovation, Unternehmertum und Ideenvernichtung (17)</i> - Entrepreneurship (3) - Demand for innovative products (7) - Financing of innovation (1) - Market innovation outputs (5) - Organizational indicators (1) - Missings (3)
→ <i>A2 - Arbeitskraft, deren Fähigkeiten und Kreativität (18)</i> - General education (7) - Human resource in science and technology education (5) - Skills (3) - Credibility (2) - Mobility (1) - Missings (2)	→ <i>B1 - Wirtschaftsindikatoren (7)</i> - Income (2) - Productivity (3) - Employment (2)
→ <i>A3 - Wissensbildung und -verbreitung (25)</i> - Research and experimental development family (12) - Patents (5) - Bibliometrics (2) - Knowledge flows (5) - Total investment in intangibles (1) - Missings (4)	→ <i>B2 - Soziale Entwicklung (19)</i> - Environmental (5) - Employment and economic welfare (7) - Quality of life indicators (7) - Missings (2)
	→ <i>C - Globalisierung (19)</i> - Trade (3) - Knowledge production and diffusion (7) - Economic structure (2) - Human resources (7)

Zunächst wurde nun bei der Verfügbarkeitsanalyse der durchschnittliche Metadatenanteil je Land eines Indikators betrachtet (Abbildung 2).

Abbildung 2: Durchschnittlicher Metadatenanteil je Land eines Indikators

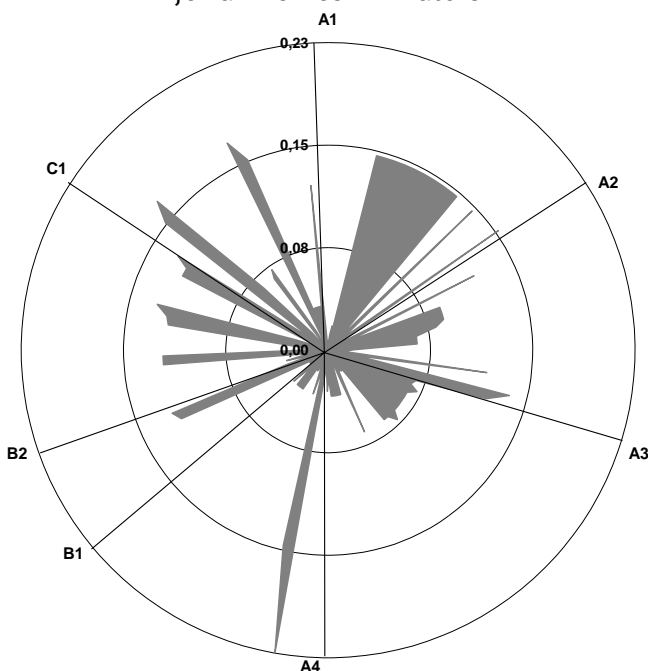
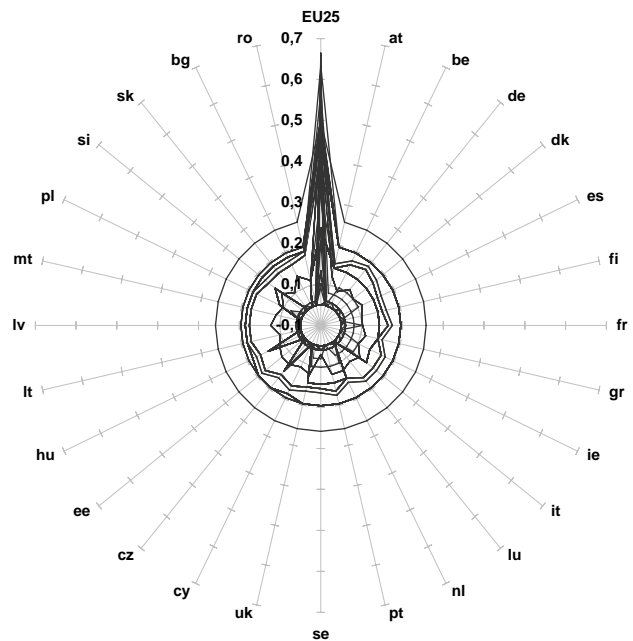


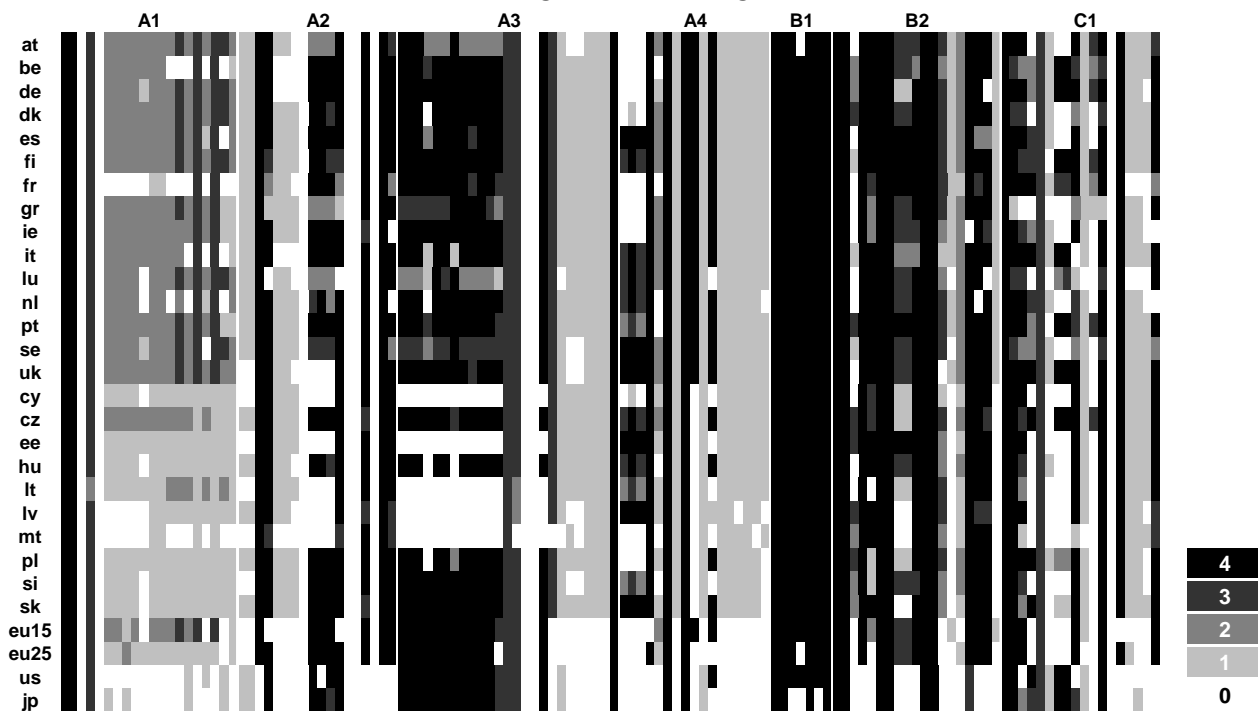
Abbildung 3: Metadatenanteil je Land und Indikator



Der Anteil bezieht sich dabei darauf, in wie vielen der Metadatenkategorien Informationen bereitgestellt werden. Dabei zeigt sich, dass die Rate insgesamt gesehen sehr gering ist, lediglich für einige Indikatoren sind höhere Anteile zu verzeichnen. Bei näherer Betrachtung zeigt sich, dass es sich zumeist um Indikatoren aus identischen Erhebungen handelt: in der Gruppe A1 sind es Daten des European ICT Surveys, die eine im Vergleich zu den übrigen Indikatoren umfassendere Metadatenbasis liefern. In der zweiten Gruppe sind es einerseits Zensus bzw. Mikrozensusquellen und andererseits Daten des Labour Force Surveys. In der folgenden Gruppe sind Daten aus der Gruppe der Main Science and Technology Indicators (MSTI) der OECD. Den höchsten Anteil hat ein Indikator in der Gruppe A4, ein Populationsindikator, der registerbasiert ist. Schließlich sind es in den verbleibenden Gruppen wiederum Daten des Labour Force Surveys bzw. aus EU-SILC, die sich von den sonst geringen Anteilen abheben. Werden die durchschnittlichen Anteile je Gruppe betrachtet, ist klar, dass in der Gruppe A1 der höchste durchschnittliche Metadatenanteil besteht und die Gruppen B2 und C1 auch besser abschneiden. Differenziert nach Ländern betrachtet zeigt sich (Abbildung 3), dass wie bereits erwähnt, lediglich auf EU-25 Ebene der Anteil recht hoch ist und einzig für den bereits erwähnten Populationsindikator der Anteil vorhandener Metadaten auch über die Länder hoch ist. Insgesamt ist der Anteil der Metadateninformationen jedoch sehr gering. Mit Blick darauf, dass er nur einen der in Tabelle 1 vorgestellten Qualitätsindikatoren repräsentiert, ist die Evaluation der Datenqualität mehr als unzufriedenstellend.

Es kann jedoch noch ein zweiter Aspekt der Datenqualitätsindikatoren analysiert werden, und zwar der der Datenverfügbarkeit (Abbildung 4). Hier zeigen sich noch deutlichere Lücken insbesondere bei der Verfügbarkeit über den insgesamt zu betrachtenden vierjährigen Zeitraum. Zu den zur Auswahl stehenden 144 Indikatoren konnten bekanntlich 125 Datensätze ermittelt werden, wobei für den im Projekt definierten Betrachtungshorizont sich diese Anzahl nochmals auf 116 verfügbare Datensätze reduzierte. Hier zeigen sich insbesondere bei den neueren Fragestellungen zur Produktion und Verbreitung von IKT (A1) wie auch in der Gruppe A4 eine sehr lückenhafte Datenverfügbarkeit. Lediglich die Daten der Gruppe B1 sind sehr umfassend, was damit zu begründen ist, dass es sich hierbei um klassische und etablierte Wirtschaftsindikatoren handelt.

Abbildung 4: Datenverfügbarkeit



Die soeben diskutierten Aspekte beziehen sich damit lediglich auf die Datenqualität. Zur Analyse der Frage der Indikatorenqualität oder gar des zusammengesetzten Indikators konnten keine Aussagen getroffen werden aus Mangel an Informationen.

5. Schlussfolgerungen

Theoretisch lässt sich die Qualität eines zusammengesetzten Indikators in einer Grundstruktur diskutieren, jedoch besteht hier noch weiterer Ausarbeitungsbedarf, da Indikatoren für die Messung der Indikatorenqualität noch zu entwickeln sind. Auch die Klärung der Frage, in wieweit ein mögliches Qualitätsmaß eine Aussagekraft besitzen könnte und als Wegweiser zur Verwendung dienen könnte, muss noch offen bleiben, da bisher noch nicht ansatzweise eine empirische Größe ermittelt werden kann.

Eine empirische Umsetzung des theoretischen Konzeptes ist derzeit noch nicht ansatzweise leistbar. Auch auf der ersten Stufe der Qualitätsmessung, der Datenqualität, ist eine standardisierte, vergleichende Evaluation zumindest für den Endnutzer derzeit nicht möglich, obwohl von Seiten der Ämter bereits umfassende Arbeit geleistet wurde. Hier liegt die Problematik insbesondere in den Zugangsmöglichkeiten des Endnutzers, um sich ein Bild über die Qualität zu verschaffen. Die für den Endnutzer frei zugänglichen Metadaten sind hier lediglich eine erste Basisinformation und insbesondere für den Einzeldatengebrauch geeignet. Verbesserungen sind jedoch auch hier wünschenswert: so sollten die Metadaten von zentraler Stelle (Eurostat) auch auf Länderebene bereitgestellt werden. In einem zweiten Schritt hin zu einer Veröffentlichung aller Datenqualitätsindikatoren sollten dann zunächst auf EU-Ebene für alle Indikatoren Qualitätsberichte bereitgestellt werden die dann auch in der Folge für jeden Indikator eines Landes auszuarbeiten wären.

Literatur

De Vries, Willern F.M. (2001): Meaningful Measures: Indicators on Progress, Progress on Indicators, *International Statistical Review*, 69, 313-331

Ehling, Manfred; Körner Thomas (2007): *Handbook on Data Quality Assessment Methods and Tools*.

Eurostat (2003a): Definition of quality in statistics (Eurostat/A4/Quality/03/General/Definition). http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47141301/DEFINITION_2.PDF.

Eurostat (2003b): Standard Quality Report (Eurostat/A4/Quality/03/General/Standard Report) http://epp.eurostat.ec.europa.eu/pls/portal/docs/PAGE/PGP_DS_QUALITY/TAB47143233/HOWTO_MAKEQUALITYREPORT_0.PDF.

Eurostat (2006): Quality reporting: reference metadata and quality (ESTAT/DDG-02/Quality/2006/5d)

Menges, Günter (1972): *Grundriß der Statistik: Teil1: Theorie*.

Nardo, Michela; Saisana, Michaela; Saltelli, Andrea; Tarantola, Stefano; Hoffman, Anders and Giovannini, Enrico (2005): *Handbook on Constructing Composite Indicators: Methodology and User Guide (STD/DOC(2005))*.

Schnell, Hill, Esser (2005): *Methoden der empirischen Sozialforschung*.