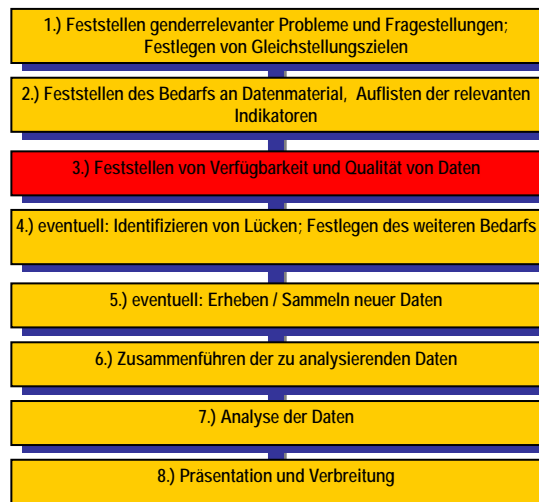


V.) VERFÜGBARKEIT UND QUALITÄT VON DATEN



Im letzten Kapitel haben wir uns damit beschäftigt, wie man die Indikatoren für die Themen, die man untersucht, festlegt. Nachdem diese aufgelistet sind, geht es im nächsten Schritt darum, unterschiedlichste Datenquellen danach zu durchforsten, **inwieweit die benötigten Daten verfügbar sind**, und deren **Qualität einzuschätzen**: Sind Indikatoren so definiert, dass sie auch wirklich die Realität umfassend abbilden, oder bleibt manches unberücksichtigt? Wurden geeignete Messmethoden angewandt? Liegen die Daten auch für die Region oder den Zeitraum vor, die man benötigt? Häufig werden dabei Lücken sowie Verbesserungsbedarf zum Vorschein kommen.

Je nach Fragestellung geht es dabei um **externe und / oder interne** Daten. Interne Daten sind solche, die in der eigenen Institution erhoben wurden oder erhoben werden könnten, beispielsweise Daten über die MitarbeiterInnen, über Kundinnen und Kunden, über das Budget oder über Einstellungen, die mittels Fragebogen erhoben werden. Die im Folgenden beschriebenen Kapitel beziehen sich prinzipiell auf interne und externe Daten: Bei beiden gilt es erst einmal festzustellen, ob sie verfügbar sind, und im Weiteren, inwieweit sie für das gerade anstehende Thema zu gebrauchen sind. Manche Fragen stellen sich freilich für

interne Daten anders, deshalb ist ihnen ein eigenes Zusatzkapitel gewidmet (siehe „...“, S. ...).

A) Verfügbarkeit

Welche der gesuchten Daten existieren überhaupt, und welche sind zu bekommen?

Als erstes geht es darum, festzustellen, inwieweit die Daten, nach denen man fahndet, überhaupt verfügbar sind. Ein Teil davon existiert möglicherweise gar nicht oder aber ist nicht (leicht) zugänglich. Prinzipiell gibt es hinsichtlich der Verfügbarkeit folgende Möglichkeiten¹:

Möglichkeiten hinsichtlich Verfügbarkeit:

Externe Daten:

- Die Daten wurden **erhoben**, nach den **gewünschten Kategorien tabellarisiert** (etwa nach den Alterskategorien, die man gerade benötigt – z. B. in 10-Jahres-Schritten) und in unterschiedlichen Formen **veröffentlicht** (in Jahrbüchern, Broschüren, auf CD-ROM, im Internet...).
- Sie wurden **erhoben und veröffentlicht**, allerdings **nicht nach den Kategorien tabellarisiert**, die man gerade benötigt.
- Die Daten wurden **erhoben und tabellarisiert**, aber **nicht veröffentlicht**. Sie sind daher nur auf Anfrage verfügbar.
- Die Daten wurden **erhoben, aber nicht tabellarisiert bzw. anderweitig systematisch erfasst**. Daher existieren sie **nur als Originaldokumente** auf Rechnern oder auf Papier, zum Beispiel in Form von Fragebögen. Wieder sind sie nur auf spezifische Anfrage verfügbar.
- Die gesuchten Daten wurden **gar nicht erhoben** – oder aber **nicht nach dem Geschlecht** erfasst.

Interne Daten

Vergleichbares gilt für **Daten aus der eigenen Organisation** bzw. dem **eigenen Unternehmen**. Auch hier stellt sich die Frage, inwieweit Daten überhaupt erhoben wurden und weiter, ob sie nur irgendwo lose auf Zetteln herumfliegen oder aber in geordneter Form erfasst sind. Schließlich stellt sich auch hier die Frage, ob sie zu haben sind – es ist ja nicht

¹ Hedman / Perucci / Sundström (1996): 57.

gesagt, dass Daten, die etwa der CheflInnenetage zur Verfügung stehen, auch allen MitarbeiterInnen so einfach zugänglich sind.

Verstreutes Material

Eine Herausforderung beim Aufstöbern von Daten liegt darin, dass das gesuchte Material oft **auf die unterschiedlichsten statistischen Quellen, Themenfelder und Publikationen verstreut** ist. Es bedarf also oft einiger Recherchearbeit, bis alles zusammengesucht ist.

Warum Daten fehlen

Ein wichtiger Grund dafür, dass Daten nicht erhoben werden, liegt darin, dass viele Datenbestände nicht oder zumindest nicht vorrangig für sozialwissenschaftliche, sondern vielmehr **für Verwaltungszwecke geschaffen** wurden, denkt man etwa an die Daten rund um Beschäftigung und Versicherung, die der Hauptverband der Sozialversicherungsträger sammelt, oder an die Arbeitslosenstatistik des Arbeitsmarktservice (AMS): Ihr Zweck liegt in den Verwaltungstätigkeiten rund um das Thema Arbeitslosigkeit, das heißt, es sind vorrangig Personen registriert, die Anspruch auf finanzielle Unterstützung oder auf Qualifizierungsmaßnahmen haben. Menschen ohne solche Ansprüche sind damit unterrepräsentiert. Daten zu Arbeitslosigkeit des AMS weichen daher von jenen ab, die etwa über den Mikrozensus erfasst werden und auf direkten Befragungen der Betroffenen beruhen².

Eine weitere Ursache für das Fehlen von Daten kann darin liegen, dass es sich um **Themen** handelt, die **neu** oder jedenfalls noch **wenig im öffentlichen Bewusstsein** sind. Eventuell sind die für Erhebungen Zuständigen für bestimmte Themen bzw. für die Gender-Thematik **nicht sensibilisiert** sind oder geben bei **knappen Ressourcen** anderen Themen den Vorrang. Mit gesellschaftlichen Veränderungen tauchen immer wieder neue Fragestellungen auf – die Konzeption von Erhebungen und Analysen muss also immer wieder daran angepasst werden.

Schließlich können die Gründe darin bestehen, dass die fragliche Angelegenheit **schwer zu erfassen und zu messen** ist oder dass sie als **zu sensibel** für Abfragen gilt (zB Gewalt, Sexualität). In all diesen Fällen entdecken wir Datenlücken.³

² Wroblewski / Leitner / Steiner (2005): 11

³ Hedman / Perucci / Sundström (1996): 57; Wroblewski / Leitner / Steiner (2005): 11