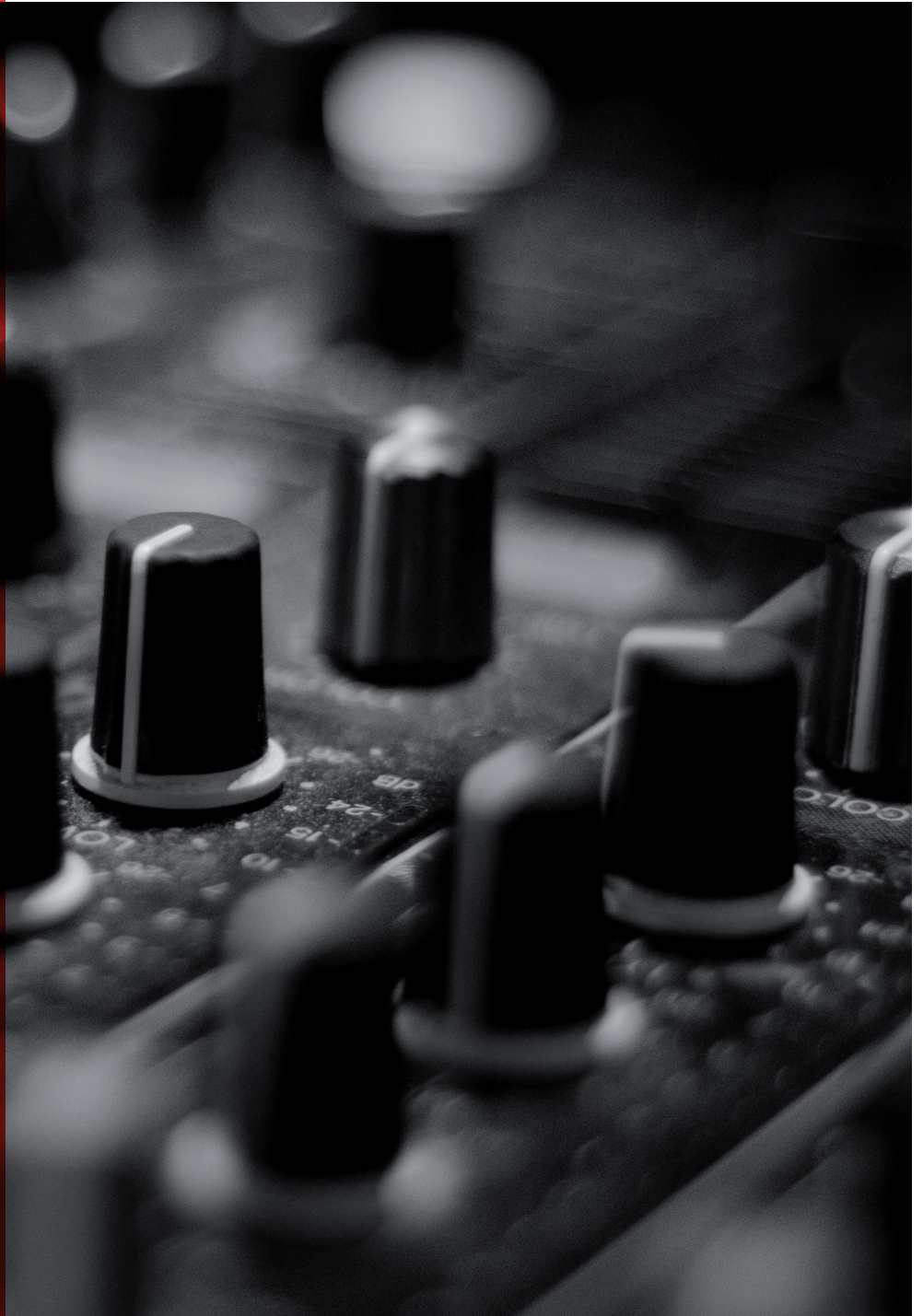


Swiss Insights News



#11

Dirty Data in Online Surveys

Dirty Data in Online Surveys: Wie Sie die Datenqualität vor und nach der Feldphase verbessern können.



Raffael Meier
Mitgründer und CTO
onlineumfragen.com



Nina Gwerder
Consultant
onlineumfragen.com

Immer mehr Daten werden immer schneller verarbeitet. Kundenbedürfnisse, Zeitdruck und ethische Aspekte treiben die Relevanz von Datenqualität in die Höhe. Wie optimieren wir Datenqualität schon vor der Feldphase? Was tun, wenn Daten schon erhoben sind?

Die 1986 kurz nach dem Start explodierte Challenger-Raumfähre der NASA mit sieben Astronauten^[1] und das wegen eines Fehlers in Google Maps versehentlich abgerissene Haus von Lindsey Diaz^[2] sind beides Konsequenzen mangelnder Datenqualität.

Als multifaktorieller, unscharfer Begriff ist Datenqualität kein «Messerwert», sondern wird multimodal erarbeitet: bei Online-Befragungen vor der Feldphase mittels elaboriertem Sampling, Fragebogenkonzeption und Pretesting. Aber auch danach in der Analyse und der Bereinigung der gewonnenen Daten. Hinzu kommen Meta-Aspekte der Datenqualität wie Passung der Daten zur Forschungsfrage, Prozesstransparenz, faire Datengewinnung oder proaktiver Datenschutz, die Empowerment für Kunden bewirken.

Datenqualität – was ist das?

57 Millionen Suchergebnisse bei google verdeutlichen die Relevanz des Begriffs «Data Quality». Im täglichen Sprachgebrauch wird der Begriff «Datenqualität» häufig mit der Richtigkeit (Präzision) von Daten gleichgesetzt. Um diese zu erhöhen, werden genauere Messinstrumente, rigorose Datenerhebungsverfahren und komplexe statistische Methoden zum Aggregieren von Daten entwickelt^[3]. Auch der Fokus der nachfolgend präsentierten Massnahmen zur Erhöhung der Datenqualität liegt auf der Verbesserung der klassischen Gütekriterien Validität und Reliabilität von

Umfrageergebnissen. Darüber hinaus differenzieren Konzepte wie das der Universität Greifswald Datenqualität weiter, hier in die Dimensionen Integrität, Komplettheit und Korrektheit (Konsistenz und Akkuratheit)^[4] – und es existieren in der Literatur zahlreiche überlappende Modelle^[5].

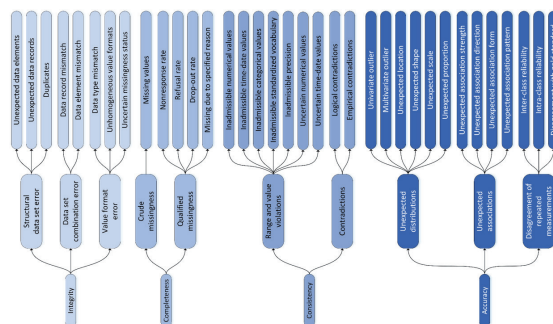


Bild 1: Datenqualitätsmodell der Universität Greifswald nach Schmidt et al. (2021)

Insbesondere im Bereich der Markt- und Meinungsforschung stehen am Ende der Datenverarbeitung Auftraggebende oder Forschende als Consumer der Daten. Daher ist es besonders wichtig, zu verstehen, dass Datenqualität aus Konsumentensicht oft über die Präzision von Werten hinausgeht.

Consumer als Zielgrösse

Welche Aspekte der Datenqualität für Consumer ausschlaggebend sind, untersuchten Richard Y. Wang und Diane M. Strong in einer zweistufigen empirischen Studie. Definiert wurde Datenqualität dabei über das verbreitete «fitness-for-use»-Konzept^[6], welches Datenqualität durch die Eignung der Daten zum von Consumern

angedachten Zweck bestimmt. Das Ergebnis der Studie war ein vierdimensionales Modell von Datenqualität, welches das bestehende Konzept empirisch operationalisiert^[6].

Die vier Dimensionen sind: intrinsic data quality, contextual data quality, representational data quality und accessibility data quality. Sie zeigen, dass für Consumer der Begriff der Datenqualität vielschichtig ist und nicht nur im Sinne der Richtigkeit verstanden werden sollte. Vielmehr sind auch Aspekte wie Glaubwürdigkeit, Relevanz, Vollständigkeit, Interpretierbarkeit, Konsistenz oder Zugänglichkeit der Daten für Consumer ausschlaggebend^[4].

Intrinsic Data Quality	Believability, Accuracy, Objectivity, Reputation
Contextual Data Quality	Value-added, Relevancy, Timeliness, Completeness, Appropriate amount of data
Representational Data Quality	Interpretability, Ease of understanding, Representational consistency, Concise representation
Accessibility Data Quality	Accessibility, Access security

Tabelle 1: Dimensionen der Datenqualität mit Ausprägungen nach Wang und Strong (1996).

Vor der Befragung

Jede Onlinebefragung ist trotz Abwesenheit einer interviewenden Person und den damit wegfallenden primären Interviewer-Effekten (deshalb oft als «verzerrungsfrei» angepriesen) immer eine Situation, in der im Extremfall eine parasoziale Interaktion mit einem imaginären Interviewer stattfindet. Oder die mindestens irgendeine Art von kognitivem Stimulus-Response-Prozess bei Befragten auslöst. Dieser gliedert sich als Cognitive Aspects of Survey Methodology (CASM) vereinfacht in vier Schritte^[7-10]:

1. Lesen und Verstehen der Frage
2. Abrufen relevanter Informationen aus dem Gedächtnis
3. Beurteilung der abgerufenen Informationen bezüglich Vollständigkeit und Relevanz, kognitive Editierung der präferierten Antwort auf das gewünschte Antwortformat und Angemessenheit
4. Antwortabgabe

Die kognitive Verarbeitung folgt dabei den Zweiprozessmodellen^[11-12] und geschieht elaboriert (intensiv, zentral) oder peripher (oberflächlich, schnell, nebenbei, heuristisch). Eine

elaborierte Verarbeitung begünstigen unter anderem die gute Verständlichkeit der Frage, die kognitive Fähigkeit sowie die Motivation der Antwortenden, genügend Zeit und wenig Ablenkung. Eine periphere Verarbeitung öffnet die Tür für Verzerrungseffekte noch weiter.

Zehn Gebote für gute Fragebögen

Bei dieser Verarbeitung sind eine Menge latenter Störeinflüsse (Response Bias^[13]) im Spiel, die sich nur schwer kontrollieren lassen. Trotzdem kann im Design von Fragebögen mit zehn Geboten^[9], soweit möglich und ökonomisch sinnvoll, vorgesorgt werden, wobei diese weniger als «starre» Regeln, denn als Reflexionsbasis für das Finetuning von Fragebögen dienen:

1. einfache, unzweideutige Begriffe (werden von allen Befragten sehr ähnlich verstanden)
2. unklare Begriffe definieren
3. keine langen/komplexen Fragen
4. keine hypothetischen Fragen
5. keine Doppelstimuli/Doppelverneinung
6. keine Unterstellungen/Suggestivfragen
7. keine Fragen nach Informationen, die viele Befragte vermutlich nicht kennen
8. Fragen mit eindeutigem zeitlichen Bezug
9. Antwortkategorien sind erschöpfend und disjunkt (überschneidungsfrei)
10. Kontext einer Frage wirkt sich nicht auf deren Beantwortung aus

Zum Beispiel kann eine Frage wie «Wurden in Ihrer Schule bereits Projekte im Rahmen des LP21 umgesetzt?» vielleicht Lehrpersonen gestellt werden, aber nicht Eltern (Gebote 2 und 7). Und «Experten denken, dass CO2-Massnahmen zu langsam umgesetzt werden – halten Sie diese Ansicht für richtig oder für falsch?» verstösst gegen das sechste Gebot. Sogar in wissenschaftlich angesehenen Studien werden Fragen gestellt, die sich in erheblichem Masse der Erinnerbarkeit entziehen (Mobiltelefonnutzung in den letzten 6 Monaten): «detailed questions were asked about the initial pattern of use, including network operator and average number and duration of calls, and any subsequent changes in use patterns.

Questions were also asked about the proportion of time the phones were used in urban, suburban or rural settings»^[14] (Gebot 3, 7, 8). Entsprechend der Konversationsmaximen nach Grice^[15] gilt auch für Formulierungen in Fragebögen:

- Qualität (wahr und nicht suggestiv)
- Quantität (so ausführlich wie nötig, so kurz wie möglich)
- Relevanz (zielgruppengerecht, nur fragen und ansprechen, was zum Thema gehört)
- Modalität (klar, eindeutig, kompakt und geordnet)

Weiter sollten Übertragungseffekte minimiert werden: Steht vor der Frage, «Welche Partei wählen Sie am Wahlsonntag?» eine Frage zu Umweltkatastrophen oder eine zum Wirtschaftswachstum? Ist Fleisch 25% fett oder 75% mager?^[16] Übertragungs- und Framing-Effekte sollten mit möglichst objektiven Formulierungen und Kontexten verringert werden.

Skalen sind Korsetts, die passen müssen

Auch *Rating-Skalen* an sich können Verzerrungen unterstützen, da Antworten von den Befragten «passend» gemacht werden. «Trash-Antworten» können reduziert werden, in dem eine Ausweichkategorie «keine Antwort» angeboten wird. Dadurch gehen leider einige gültige Antworten verloren, die versteckte Datenverschmutzung, wenn wegen Fehlen einer Ausweichkategorie einfach irgendwas angekreuzt wird^[17], nimmt jedoch ab. Eine mittlere Ausprägung wie «teils/teils» wird ebenfalls empfohlen^[18]. Rating-Skalen mit 5 bis 7 vollverbalisierten Antwortkategorien werden mehrheitlich als ratsam erachtet – zu wenige Kategorien differenzieren ungenügend, zu viele sind inhaltlich schwer zu unterscheiden, fehlende Verbalisierung ist schwerer zu interpretieren^[18]. Einige weitere Verzerrungseffekte, denen in Online-Befragungen Beachtung geschenkt werden sollte, sind Akquieszenz (generelle Zustimmungstendenz in Skalen), Non-Attitude (beliebige Antworten bei Desinteresse, Zeitknappheit), soziale Erwünschtheit (Erwartungskonformität),

Sponsorship-Effekte (Antworten dem Auftraggeber zuliebe), Situationseffekte (Kontext und Ort des Ausfüllens, Anwesenheit von Bekannten, Helfern), Tendenzen zur Mitte (Meiden von Extremen), Lageeffekte (Primacy/Recency-Effekt: Erste und letzte Antwortvorgaben werden besser erinnert und bevorzugt, weil salient) oder Ankereffekte (Übertragung, Beeinflussung durch Vorfragen, selektives Antwortuniversum, Priming). Schwarz^[19] hat zu Letzterem schon 1999 festgestellt, dass Antwortvorgaben implizite Anker setzen. So antworteten unter Verwendung der linken Skala (vgl. Bild 2) nur 16% der Befragten mit «mehr als 2 ½ Stunden». Mit der rechten Skala hingegen insgesamt 38%, obwohl dieselbe Frage gestellt worden war. Die Extremposition der Antwortvorgabe «mehr als 2 ½ Stunden» in der linken Skala lässt die Teilnehmenden unterreportieren, rechts ist der umgekehrte Effekt der Fall.

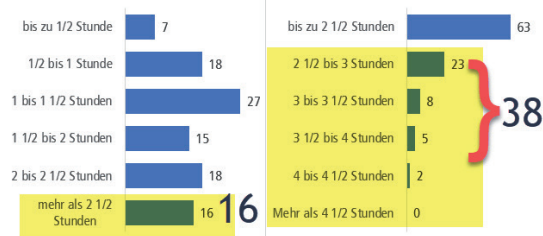


Bild 2: Ankereffekt der Antwortvorgaben bei Schwarz (1999)

Die Datenqualität ist auch beeinflusst durch die Frageformate: offene Textfelder versus Skalen versus halboffene Formate, Validierungen direkt im Fragebogen, Regie-Anweisungen, visuelle Darstellung von Skalen, Mobildarstellung, technische Einflüsse wie schwer bedienbare Schieberegler; und ebenso Herausforderungen bestimmter Zielgruppen wie Kinder oder digital unerfahrene Benutzer. Pretesting sowie Regie-Hinweise bezüglich der erwarteten Dateneingaben in punkto Umfang, Art und Format der Eingaben sind im Fragebogen fast immer vorteilhaft.

Einige Probleme der Datenqualität können somit durch intensive Reflexion, Einbinden von Fragebogen-Consultants, ausführliches Pretesting (z. B. auch mit Explorieren oder Split-Half-Setups verschiedener Frageformate) und Überarbeiten von bereits bestehenden Fragebögen durchaus adressiert werden.

Nach der Befragung

Auch im Nachgang können Massnahmen zur Verbesserung der Datenqualität getroffen werden. Diese beziehen sich hauptsächlich auf verzerrende Effekte, die Teilnehmende durch ihr Verhalten im Fragebogen erzeugen.

Während einige dieser Effekte, wie z. B. *Ausreisser*¹, *Speeder*² oder *inkomplette Fragebögen* unabhängig vom Befragungsmodus auftreten (z. B. auch in CATI, Face-to-Face), häufen sich gerade bei schriftlichen und Online-Befragungen sogenannte «Response-Sets» (Tendenzen von Teilnehmenden, eine Reihe von Fragen in einem bestimmten Muster zu beantworten)^[20]. Bei Online-Befragungen wird dies zusätzlich durch die erhöhte empfundene Anonymität verschärft, da mit steigender Anonymität die Tendenz steigt, Items weniger ernsthaft auszufüllen und falsche oder fiktive Antworten abzugeben^[20].

Klassische Antwortmuster, die in schriftlichen Befragungen mit Rating-Skalen beobachtet werden können, sind: Response-Ranges (Verwendung nur eines bestimmten Bereichs der Skala, unabhängig vom Inhalt der Frage oder der Ausrichtung der Antwortvorgaben), Extreme Checking Style (auch Extreme Response Style^[21], ERS, genannt: abwechselndes Anwählen der linken und rechten Extrempunkte einer Skala), Muster-Ankreuzer (Ankreuzen von Mustern wie z. B. Diagonalen oder Pfeilen in Tabellenfragen) und Straightliner (Null-Varianz-Antwortverhalten, d. h. Auswahl eines bestimmten Skalenpunkts, unabhängig von der Skalenbreite, -ausrichtung und Frageformulierung). Gerade Letztere sind bei unmotivierten Teilnehmenden beliebt. Jandura identifiziert in einer offenen Online-Befragung zu Mediennutzungsverhalten von Jugendlichen bei einem Viertel der Befragten ein Null-Varianz-Antwortverhalten in mindestens

einer der neun beantworteten Fragen^[20]. Allerdings muss hier relativiert werden, dass Straightlining unter gewissen Umständen valide ist, beispielsweise, wenn eine Item-Batterie eine hohe interne Konsistenz aufweist und alle Items in dieselbe Richtung formuliert sind^[22]. Um valides Straightlining auszuschliessen, könnte ein Item pro Frageblock vor der Feldphase umgepolt werden.

Um die Qualität der Ergebnisse einer Befragung zu erhöhen, sollten Fälle mit den beschriebenen Effekten untersucht und allenfalls aus der Analyse exkludiert werden. Während dies bei einigen Effekten relativ einfach möglich ist (z. B. können Speeder anhand der Bearbeitungszeit und inkomplette Fragebögen anhand der fehlenden Antworten schnell identifiziert werden), ist für andere Effekte ein genaueres Hinsehen notwendig (siehe Tabelle 2). Der Prozess kann dabei digital mit multifaktoriellen, statistischen Prozeduren unterstützt werden. Die Entscheidung, ob ein spezifischer Fall nun als Quality Fail ausgeschlossen werden soll oder nicht, sollte jedoch durch geschulte Mitarbeitende nach der Sichtung des entsprechenden Fragebogens und nicht auf Basis einzelner Kriterien erfolgen.

Speeder	Unmöglich tiefe Bearbeitungszeit
Inkomplette Fragebögen	Fehlende Antworten, Abbruch der Befragung
Ausreisser	Scatterplot aller Datenpunkte, signifikante Abweichung vom Stichprobenmittel
Response-Range	Geringe Standardabweichung über alle Items
Extreme Checking Style	Quotient (<i>Anzahl Male, die eine Antwortmöglichkeit gewählt wurde</i> / <i>Anzahl Items in der entsprechenden Tabellenfrage</i>) ≈ 0.5 (exakt 0.5 bei Tabellen mit ungerader Anzahl Items nicht möglich)
Muster-Ankreuzer	Optisches Prüfen der Antworten (z. B. von Teilnehmern, die bei den entsprechenden Tabellenfragen vergleichsweise kurz verweilt sind)
Straightliner	Quotient (<i>Anzahl Male, die eine Antwortmöglichkeit gewählt wurde</i> / <i>Anzahl Items in der entsprechenden Tabellenfrage</i>) = 1 Wichtig: Straightlining kann u.U. valide sein!

Tabelle 2: Erkennen von verzerrenden Effekten im Datensatz, angelehnt an Jandura (2018) und Reuning und Plutzer (2020).

¹ Datenpunkt, der bedeutend von den restlichen Datenpunkten entfernt liegt.

² Teilnehmende, die den Fragebogen in einer Zeit komplettieren, die bei seriösem Durchlesen der Fragen und Antwortmöglichkeiten nicht plausibel ist.

Fazit

In einem kurzen Artikel kann das Thema Datenqualität lediglich anhand ausgewählter Aspekte angeschnitten werden. Zusammenfassend ergeben sich drei essenzielle Punkte:

- Eine Ausrichtung der Datenqualität auf Abnehmergruppen lohnt sich.
- Wesentliche Weichen für die Datenqualität werden schon vor der Feldphase gestellt.
- Eine Datenqualitätsanalyse und Datenbereinigung nach der Feldphase ist technisch möglich und nützlich.

Dies sind wichtige Erkenntnisse für validere Analysen, für die Kundenkommunikation, den Projektaufbau und das Branchen-Image.

Datenqualität geht zudem einher mit Themen wie Erhebungs- und Verwendungsethik, Zweckbestimmung von Daten, Panelpflege, Datenschutz und mit der Idee, auch komplexe Prozesse der Auswertungslogik, Algorithmen und Deep-Learning-Modelle so transparent wie möglich zu machen. Ein positives Wechselverhältnis von Datengebenden und -nehmenden widerspiegelt sich in einem emanzipierten und daher lohnenswerten Verständnis von Datenqualität für alle.

Literaturverzeichnis nach der Englischen Fassung

Möchten Sie mehr erfahren?

Erhalten Sie anschauliche Beispiele, Übungen und Best-Practice-Hinweise.

Nehmen Sie an unserem Webinar teil.

21. September 2022
09.00 - 10.15 Uhr, via Zoom

[-> Zur Anmeldung](#)

Die Autoren

Raffael Meier, MA, MSc – Mitgründer und CTO von onlineumfragen.com und Pionier der Onlinebefragungstechnologie. Er befasst sich mit gesellschaftlichen und methodologischen Aspekten von Daten und berät Kundinnen und Kunden mit dem Ziel «Empowerment».

Nina Gwerder, MA – ist Consultant bei onlineumfragen.com und spezialisiert auf die Beratung namhafter nationaler und internationaler Unternehmen rund um das Thema Online-Befragungen und deren effektiver Auswertung.

Kontakt

Raffael Meier, +41 44 500 5137
raffael.meier@onlineumfragen.com

Nina Gwerder, +41 44 500 5140
nina.gwerder@onlineumfragen.com

Dirty Data in Online Surveys. How to improve data quality before and after the field phase



Raffael Meier
Co-founder and CTO
onlineumfragen.com



Nina Gwerder
Consultant
onlineumfragen.com

More and more data are being processed faster and faster. Customer needs, time pressure and ethical standards are driving up relevance of data quality. How can we optimize data quality even before the field phase. And what if data have already been collected?

NASA's Challenger space shuttle, which exploded shortly after its launch in 1986 with seven astronauts on board^[1] and Lindsey Diaz's house^[2], which was accidentally demolished due to an error in Google Maps, are both consequences of poor data quality. As a multi-factorial, vague term, data quality is not a «measurement value» rather it is developed multi-modally: in online surveys specifically before the field phase (by means of elaborate sampling, questionnaire design, pretesting) and afterwards (by analysis and adjustment of the data obtained). And meta-aspects of data quality, such as the suitability of the data to the research question, process transparency, fair data acquisition or proactive data protection, simultaneously result in empowerment for customers.

Data quality – what is it?

With 57 million search results on google, the relevance of the term «data quality» becomes clear. In everyday language, the term «data quality» is often equated with the correctness (precision) of data. To increase this, more accurate measurement tools, rigorous data collection procedures and complex statistical methods for aggregating data are being developed^[3]. The focus of the measures presented below to increase data quality is also on improving the classic quality criteria of validity and reliability of survey results. Furthermore, concepts such as that of the University of Greifswald further differentiate data quality, here in the

dimensions of integrity, completeness and correctness (consistency and accuracy)^[4] – and there are numerous overlapping models in the literature^[5].

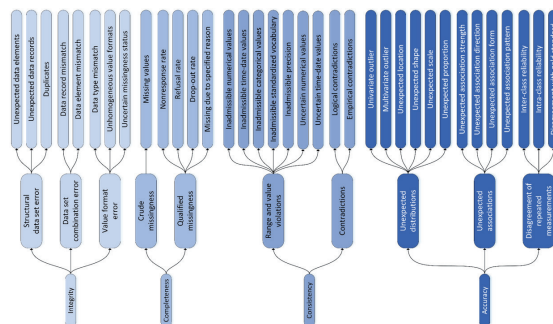


Figure 1: Data quality model of the University of Greifswald according to Schmidt et al. (2021)

Especially in the field of market and opinion research, at the end of data processing, clients or researchers are consumers of the data. Hence, it is particularly important to understand that data quality from the consumer's point of view often goes beyond the precision of values.

The consumer as a target value

Richard Y. Wang and Diane M. Strong examined which aspects of data quality are essential for consumers in a two-stage empirical study. Data quality was defined using the widespread 'fitness for use' concept^[6], which determines data quality through the suitability of the data for the purpose intended by the consumers. The result of the study was a four-dimensional model of data quality that operationalizes the existing concept empirically^[6].

The four dimensions are: intrinsic data quality, contextual data quality, representational data quality and accessibility data quality. They point out that the concept of data quality is complex for consumers and should not only be understood in terms of correctness. Aspects such as credibility, relevance, completeness, interpretability, consistency or accessibility of data are also decisive for consumers^[4].

Intrinsic Data Quality	Believability, Accuracy, Objectivity, Reputation
Contextual Data Quality	Value-added, Relevancy, Timeliness, Completeness, Appropriate amount of data
Representational Data Quality	Interpretability, Ease of understanding, Representational consistency, Concise representation
Accessibility Data Quality	Accessibility, Access security

Table 1: Dimensions of data quality with characteristics according to Wang and Strong (1996).

Before the survey

Despite the absence of an interviewer and the lack of primary interviewer effects (therefore often referred to as «distortion-free»), every online survey is always a situation in which, in extreme cases, a para-social interaction with an imaginary interviewer takes place, but at least triggers some kind of cognitive stimulus-response-process among respondents. This is divided into four simplified steps as Cognitive Aspects of Survey Methodology (CASM).^[7-10]:

1. Reading and understanding the question
2. Retrieving relevant information from memory
3. Assessment of the retrieved information regarding completeness and relevance, cognitive editing of the preferred answer to the desired answer format and appropriateness
4. Providing answers

Cognitive processing follows the dual-process models^[11-12] and takes place elaborately (intensively, centrally) or peripherally (superficially, quickly, parenthetically, heuristically). Thereby, among other things, the good comprehensibility of the question, the cognitive ability as well as the motivation of the respondents, sufficient time and few distractions

favor elaborate processing. Peripheral processing opens the door for distortion effects even further.

10 commandments for good questionnaires

This processing involves a lot of latent interference (Response Bias^[13]) that is difficult to control. Nevertheless, precautions can be taken in the design of questionnaires with the 10 commandments^[9], as far as possible and economically meaningful (whereby these serve less as rigid rules and more as a basis for reflection for the fine-tuning of questionnaires):

1. simple, unambiguous terms (understood very similarly by all respondents)
2. define unclear terms
3. no long / complex questions
4. no hypothetical questions
5. no double stimuli / double negation
6. no insinuations / leading questions
7. no questions for information that many respondents probably do not know
8. Questions with a clear time reference
9. Response categories are exhaustive and disjointed (non-overlapping)
10. The context of a question does not affect its answer

For example, a question like «Have any LP21 projects been implemented in your school?» may be asked to teachers, but not to parents (commandments 2 and 7). And «Experts think that CO2 measures are being implemented too slowly – do you think this view is right or wrong?» violates the sixth commandment. Even in scientifically respected studies, questions are asked that to a considerable extent cannot be remembered (mobile phone use in the last 6 months): «detailed questions were asked about the initial pattern of use, including network operator and average number and duration of calls, and any subsequent changes in usage patterns. Questions were also asked about the proportion of time the phones were used in urban, suburban or rural settings»^[14] (commandments 3, 7, 8).

According to Grice's maxims of conversation [15], the following also applies to formulations in questionnaires:

- Quality (true and no insinuation)
- Quantity (as detailed as necessary, as short as possible)
- Relevance (target group oriented, only ask and address what belongs to the topic)
- Modality (clear, unambiguous, compact and ordered)

Furthermore, transmission effects should be minimized: is the question «Which party will you vote for on election Sunday?» preceded by a question on environmental disasters or one on economic growth? Is meat 25% fat or 75% lean?^[16] Transmission and framing effects should be reduced with formulations and contexts that are as objective as possible.

Scales are corsets that have to fit

Rating scales as such can also support bias, as participants responses are «made to fit». «Trash answers» can be reduced by offering an alternative category «no answer». Unfortunately, some valid answers maybe are lost as a result, but the hidden data contamination when simply «anything» is ticked due to the lack of an alternative category^[17] decreases. A medium expression such as «partially» is also recommended^[18]. Rating scales with 5 to 7 fully verbalized answer categories are mostly considered advisable (too few categories differentiate insufficiently, too many are difficult to distinguish in terms of content, missing verbalization is more difficult to interpret)^[18].

Some other bias effects that should be taken into account in online surveys are acquiescence tendency (general tendency to agree in scales), non-attitude (arbitrary answers in case of disinterest, lack of time), social desirability (conformity to expectations), sponsorship effects (answers «for the sake of the client»), situation effects (context and place of completion, presence of acquaintances, helpers), tendencies towards

the middle (avoidance of extremes), location effects (primacy/recency effect: first and last answering options in multiple choice lists are better remembered and preferred because they are more salient) or anchor effects (transference, influence by previous questions, selective answer universe, priming). Schwarz^[19] already established in 1999 that given answers set implicit anchors. Using the scale on the left (see Figure 2), only 16% of those questioned answered «more than 2 ½ hours». With the scale on the right, on the other hand, a total of 38%, although the same question was asked. The «extreme position» of the response option «more than 2 ½ hours» in the left scale causes the participants to underreport, on the right hand side the opposite effect is the case.

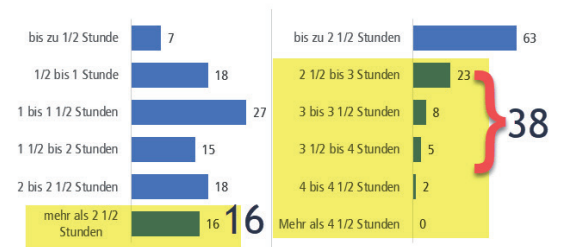


Figure 2: Anchor effect of the response specifications in Schwarz (1999)

The *question formats* (open text fields versus scales versus semi-open formats, validations directly in the questionnaire, director's notes, visual representation of scales, mobile representation, technical influences such as sliders that are difficult to operate) and challenges of certain target groups (children, digitally inexperienced users) also have a significant influence on the data quality. Pre-testing as well as director's notes regarding the expected data entries in terms of scope, type and format of the entries are almost always an advantage for the questionnaire. Some problems of data quality can therefore be addressed by intensive reflection, involvement of questionnaire consultants, extensive pretesting (e.g. also with exploration or split-half setups of different question formats) and revision of already existing questionnaires.

After the survey

Measures to improve data quality can also be taken afterwards. These mainly relate to distorting effects that participants create through their behavior in the questionnaire.

While some of these effects, such as *outliers*¹, *speeders*² or *incomplete questionnaires* occur independently of the survey mode (e.g. also in CATI, Face-to-Face), so-called «response sets» (tendencies of participants to answer a series of questions in a certain pattern) are becoming more frequent, especially in written and online surveys^[20].

Classic response patterns that can be observed in written surveys with rating scales are also response ranges (use of only a certain area of the scale, regardless of the content of the question or the orientation of the answer options), extreme checking style (also called extreme response style^[21] ERS: alternately selecting the left and right extreme points of a scale), pattern tickers (ticking patterns such as diagonals or arrows in table questions) and straightliners (zero-variance response behavior, i.e. selecting a specific scale point, regardless of scale width, orientation and question wording). The latter in particular are popular with unmotivated participants. In an open online survey on young people's media use, Jandura identified zero-variance response behavior by a quarter of respondents in at least one out of nine answered questions^[20]. However, it must be qualified here that «straightlining» is valid under certain circumstances, for example if an item battery has a high internal consistency and all items are formulated in the same direction^[22]. In order to prevent valid straightlining, reversing the polarity of one item per question block could be considered before the field phase.

In order to increase the quality of the results of a survey, cases with the described effects should be examined and, if necessary, excluded from the analysis. While this is relatively easy

for some effects (e.g. speeders can be quickly identified based on the processing time and incomplete questionnaires based on the missing answers), a closer look is required for other effects (see Table 2). The process can be supported by the computer with multifactorial, statistical procedures. However, the decision whether a specific case should be excluded as a quality fail or not should be made by trained staff after reviewing the corresponding questionnaire and not on the basis of certain criteria alone.

Speeders	Impossibly low processing time
Incomplete Questionnaires	Missing answers, early termination of the survey
Outliers	Scatterplot of all data points, significant deviation from sample mean.
Response-Range	Low standard deviation across all items.
Extreme Checking Style	Quotient (number of times an answer choice was selected divided by number of items in the corresponding table question) is approximately 0.5 (exactly 0.5 not possible for tables with odd number of items)
Pattern Tickers	Visual inspection of the answers (e.g. of participants who have lingered comparatively briefly on the corresponding table questions)
Straightliners	Quotient (number of times an answer choice was selected divided by number of items, in the corresponding table question) = 1 Important: Straightlining may be valid under certain circumstances.

Table 2: Recognition of distorting effects in the data set, based on Jandura (2018) and Reuning and Plutzer (2020).

Conclusion

In a short article, the topic of data quality can only be touched on with selected aspects. The essential points can be summed-up as follows:

- Aligning data quality with customer groups is worthwhile.
- The course for data quality is set before the field phase.
- Data quality analysis and cleaning up responses after the field phase are technically possible and useful.

These are important findings for more valid analysis, for customer communication, for the project structure and the industrie image.

Data quality also goes hand in hand with topics such as collection and use ethics, purpose of data, panel maintenance, data protection, and with the idea of making even complex processes of evaluation logic algorithms and deep learning models as transparent as possible.

A positive interrelationship of data givers and takers is reflected in an emancipated and therefore rewarding understanding of data quality for all.

¹ Data point that is significantly different from the rest of the data points.

² Participants who complete the questionnaire in a time that is not plausible for a serious reading through of the questions and answer options.

The authors

Raffael Meier, MA, MSc – co-founder and CTO of onlineumfragen.com and pioneer of online survey technology. He deals with social and methodological aspects of data and advises clients with the goal of «empowerment».

Nina Gwerder, MA – is a consultant at onlineumfragen.com and specializes in advising well-known national and international companies on the subject of online surveys and their effective evaluation.

Contact

Raffael Meier, +41 44 500 5137
raffael.meier@onlineumfragen.com

Nina Gwerder, +41 44 500 5140
nina.gwerder@onlineumfragen.com

Would you like to find out more?

Receive clear examples, exercises and best practice tips.

Take part in our webinar

21st september 2022
09.00 - 10.15h, via Zoom

-> Registration

Literatur / Literature

- [1] Fisher, C. W., & Kingma, B. R. (2001). «Criticality of data quality as exemplified in two disasters», *Information & Management*, 39(2), 109–116. [https://doi.org/10.1016/S0378-7206\(01\)00083-0](https://doi.org/10.1016/S0378-7206(01)00083-0)
- [2] *Wrong house gets torn down based on a Google Maps error*. Engadget. (n.d.). Abgerufen am 17. August 2022, von <https://www.engadget.com/2016-03-24-texas-wrong-house-torn-down-google-maps.html>
- [3] Keller, S., Korkmaz, G., Orr, M., Schroeder, A. und Shipp, S. (2017). «The Evolution of Data Quality: Understanding the Transdisciplinary Origins of Data Quality Concepts and Approaches», *Annual Review of Statistics and Its Application*, 4(1), S.85-108.
- [4] Schmidt, C. O., Struckmann, S., Enzenbach, C., Reineke, A., Stausberg, J., Damerow, S., Huebner, M., Schmidt, B., Sauerbrei, W., & Richter, A. (2021). «Facilitating harmonized data quality assessments. A data quality framework for observational health research data collections with software implementations in R», *BMC Medical Research Methodology*, 21(1). <https://doi.org/10.1186/S12874-021-01252-7>
- [5] Haug, A. (2021). «Understanding the differences across data quality classifications: a literature review and guidelines for future research», *Industrial Management and Data Systems*, 121(12), 2651–2671. <https://doi.org/10.1108/IMDS-12-2020-0756>
- [6] Wang, R.Y. und Strong, D.M. (1996). «Beyond Accuracy: What Data Quality Means to Data Consumers», *Journal of Management Information Systems*, 12(4), S.5-33
- [7] Cannell, C. F., Miller, P. v., & Oksenberg, L. (1981). «Research on interviewing techniques», in S. Leinhardt (Ed.), *Social Methodology*. Jossey-Bass Publishers.
- [8] Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The Psychology of Survey Response*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511819322>
- [9] Porst, R. (2014). *Fragebogen*. Wiesbaden: Springer Fachmedien. <https://doi.org/10.1007/978-3-658-02118-4>
- [10] Tourangeau, R. (2018). «The survey response process from a cognitive viewpoint», *Quality Assurance in Education*, 26(2), 169–181. <https://doi.org/10.1108/QAE-06-2017-0034>
- [11] Petty, R. E., & Cacioppo, J. T. (1986). «The elaboration likelihood model of persuasion», *Advances in Experimental Social Psychology*, 19(C), 123–205. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- [12] *Dual-process theories in social psychology*. PsycNET. (n.d.). Abgerufen am 17. August 2022, von <https://psycnet.apa.org/record/1999-02377-000>
- [13] Bogner, Kathrin und Landrock, Uta (2015). *Antworttendenzen in standardisierten Umfragen*. Mannheim, GESIS Leibniz Institut für Sozialwissenschaften (GESIS Survey Guidelines). https://doi.org/10.15465/gesis-sg_016
- [14] Cardis, E., Richardson, L., Deltour, I. et al. (2007). «The INTERPHONE study: design, epidemiological methods, and description of the study population», *Eur J Epidemiol*, 22, 647–664. <https://doi.org/10.1007/s10654-007-9152-z>
- [15] Grice, H. P. (1975). «Logic and Conversation», in P. Cole, & J. L. Morgan. (Eds.), *Syntax and Semantics*, Vol. 3, Speech Acts (pp. 41-58). New York: Academic Press.
- [16] Levin, I. P., & Gaeth, G. J. (1988). «How Consumers are Affected by the Framing of Attribute Information Before and After Consuming the Product», *Journal of Consumer Research*, 15(3), 374. <https://doi.org/10.1086/209174>
- [17] Berekoven, L., Eckert, W., & Ellenrieder, P. (2004). *Marktforschung*. Wiesbaden: Gabler Verlag. <https://doi.org/10.1007/978-3-663-05734-5>
- [18] Menold, N., & Bogner, K. (2015). «Gestaltung von Ratingskalen in Fragebögen (Version 1.1)», *GESIS Survey Guidelines*, 13. https://doi.org/10.15465/GESIS-SG_015
- [19] Schwarz, N. (1999). «Self-reports: How the questions shape the answers», *American Psychologist*, 54(2), 93–105. <https://doi.org/10.1037/0003-066X.54.2.93>
- [20] Jandura, O. (2018). «Fake Data? Zur Trennung von sauberen und verschmutzten Daten bei selbst-administrierten Befragungsmodi», in Rössler P. & Rossman, C. (Hrsg.), *Kumulierte Evidenzen*. Wiesbaden: Springer VS, S. 207-223.
- [21] Greenleaf, E. A. (1992). «Measuring Extreme Response Style», *The Public Opinion Quarterly*, 56(3), 328–351. <http://www.jstor.org/stable/2749156>
- [22] Reunig, K. und Plutzer E. (2020). «Valid vs. Invalid Straightlining: The Complex Relationship Between Straightlining and Data Quality», *Survey Research Methods*, 14(5), S.439-459.



Swiss Insights

Swiss Insights ist der Verband und die Interessensvertretung aller Unternehmen, die Daten und prädiktive Modelle im Rahmen von Marketing, Innovationsprozessen, Kundenservice, Angebotsgestaltung, Kommunikation und Zielgruppendefinitionen erheben, analysieren, einsetzen und daraus Handlungsempfehlungen ableiten.

Swiss Insights pflegt einen aktiven Dialog mit politisch und gesellschaftlich wichtigen Akteuren und fördert den Austausch mit anderen nationalen und internationalen Fachorganisationen.

Eine der Hauptaufgaben des Verbands ist die Förderung der Markt-, Meinungs- und Sozialforschung im Allgemeinen und der Wissenschaftlichkeit im Besonderen. Er entwickelt, definiert und unterhält strenge Leitlinien zur Qualitätssicherung und grenzt sich im Bereich

der Markt- und Sozialforschung klar von Werbung und Direktmarketing ab. Hierzu führt der Verband das Qualitätslabel «Market & Social Research by Swiss Insights».

Darüber hinaus engagiert sich Swiss Insights dafür, dass die Nutzung von Daten und die Anwendung von datengetriebenen Modellen transparent, nachvollziehbar und in diesem Sinne fair gestaltet wird. Hierfür wurde das Label «Data Fairness by Swiss Insights» geschaffen.



Herausgeber und Kontakt

Swiss Insights, Swiss Data Insights Association, Gruebengasse 10, 6055 Alpnach, Switzerland
+41 44 3501960, info@swiss-insights.ch, www.swiss-insights.ch

Unsere Mitglieder

SWISS INSIGHTS Institute Member

Im Verband sind alle relevanten **Markt- und Sozialforschungsinstitute** unter einem Dach organisiert. Alle Mitgliedsinstitute unterliegen einem strengen Regelwerk von schweizerischen und internationalen Normen.

Mitgliedsinstitute dürfen das Label **Market & Social Research by SWISS INSIGHTS** und je nach Tätigkeitsgebiet das Label **Data Fairness by SWISS INSIGHTS** tragen.

Die Mitgliederliste finden Sie auf der nachfolgenden Seite.

SWISS INSIGHTS Corporate Member

Die Corporate Mitglieder sind **Unternehmen**, die sich für den fairen Umgang mit Auskunftspersonen und Auftraggeber sowie für den Schutz der Privatsphäre engagieren.

Corporate Member, die das Label **Data Fairness by SWISS INSIGHTS** tragen, stehen für den wissenschaftlichen, seriösen und respektvollen Umgang mit Daten ein.

Die Mitgliederliste finden Sie auf der nachfolgenden Seite.

SWISS INSIGHTS

Institute Member

amPuls Market Research

Hirschengraben 49, 6000 Luzern 7
+41 41 612 14 14 / info@ampuls.ch
www.ampuls.ch

amrein+heller MarktforschungsTreuhand AG

Südweid 7, 6274 Eschenbach
+41 748 63 70 / contact@ah-feedback.ch
www.ah-feedback.ch

Bilendi Schweiz AG

Reinhardstrasse 19, 8008 Zürich
+41 79 801 88 80 / contact.ch@bilendi.com
www.bilendi.ch

Boomerang Ideas AG

Sihlquai 131, 8005 Zürich
+41 44 500 88 60 / raphael@boomerangideas.com
www.boomerangideas.com

Constant Dialog

Alte Steinhäuserstrasse 33, 6330 Cham
+41 41 310 05 40 / info@constant-dialog.ch
www.constant-dialog.ch

DemoSCOPE Data + Research

Klusenstrasse 17, 6043 Adligenswil
+41 41 375 40 00 / demoscope@demoscope.ch
www.demoscope.ch

dr-ouwerkerk ag – just-medical!

Blegistrasse 5, 6340 Baar
+41 41 766 11 55 / info@just-medical.com
www.pharmaagentur.ch

Gallup AG

Reinhardstrasse 19, 8008 Zürich,
+41 78 891 31 15 / office@gallup.swiss
www.gallup.swiss

gff Swiss Research Services

Baarerstrasse 25, 6300 Zug
+41 41 560 01 60 / gut@gff.ag,
www.gff.ag

GfK Switzerland AG

Suurstoffi 18A, 6343 Rotkreuz
+41 41 632 91 11 / info.ch@gfk.com
www.gfk.ch / www.gfk.com

gfs.bern. Menschen. Meinungen. Märkte.

Effingerstrasse 14, Postfach, 3001 Bern
+41 31 311 08 06 / info@gfsbern.ch
www.gfsbern.ch

gfs-zürich, Markt- & Sozialforschung

Riedtlistrasse 9, 8006 Zürich
+41 44 360 40 20 / gfs@gfs-zh.ch
www.gfs-zh.ch

gfs-befragungsdienst

Schaffhauserstrasse 491, 8052 Zürich
+41 44 360 26 40 / info@gfs-bd.ch
www.gfs-bd.ch

GIM Suisse AG

General-Wille-Strasse 10, 8002 Zürich
+41 44 283 18 18 / info@g-i-m.ch
www.g-i-m.ch

Happy Thinking People AG

Staufacherstrasse 101, 8048 Zürich
+41 44 204 16 26 / contact-zurich@happythinkingpeople.com
www.happythinkingpeople.com

INNOFACT (Schweiz) AG Research & Consulting

Flurstrasse 50, 8048 Zürich
+41 43 931 77 82, Info@innofact.ch
www.innofact.ch

Instight Institute AG

Bergstrasse 138, 8032 Zürich
+41 44 387 90 90 / info@insightinstitute.ch
www.insightinstitute.ch

intervista

Optingenstrasse 5, 3013 Bern
+41 31 511 39 00 / anfragen@intervista.ch
www.intervista.ch

IPSOS Suisse SA

11, Chemin du Château-Bloch, 1219 Le Lignon
+41 22 591 06 00 / Contact_Switzerland@ipsos.com
www.ipsos.com/de-ch

Kantar Media Switzerland AG

Bahnhofstrasse 4, 3073 Gümligen
+41 31 537 79 00 / ch.panel@kantarmedia.com
www.kantarmedia.com

SWISS INSIGHTS

Institute Member

LINK

Baslerstrasse 60, 8048 Zürich
+41 41 367 73 73 / zurich@link.ch
www.link.ch

Marketagent.com Schweiz AG

Seefeldstrasse 19, 8008 Zürich
+41 43 555 06 50, schweiz@marketagent.com
www.marketagent.com

M.I.S. Trend SA

Pont Bessières 3, 1005 Lausanne
+41 21 320 95 03 / info@mistrend.ch
www.mistrend.ch

onlineumfragen.com GmbH

Kernserstrasse 15, 6056 Kägiswil
+41 44 500 50 54 / info@onlineumfragen.com
www.onlineumfragen.com

POLYQUEST AG

Flurstrasse 26, 3014 Bern
+41 31 335 64 00 / info@polyquest.ch
www.polyquest.ch

Publicom AG

Alte Landstrasse 55, 8802 Kilchberg
+41 44 716 55 11 / publicom@publicom.ch
www.publicom.ch

Qualitest AG, Institut für Marketing- und Sozialforschung

Rosenberghöhe 3, 6004 Luzern
+41 41 712 12 21 / qualitest@qualitestag.ch
www.qualitestag.ch

SensoPLUS

Industriestrasse 16, 6300 Zug
+41 41 710 71 61 / info@sensoplus.ch
www.sensoplus.ch

NielsenIQ (Switzerland) GmbH

Park 4, 6039 Root D4
+41 41 445 64 64 / nielsen-ch@nielsen.com
www.nielsen.com

TransferPlus AG Market Research

Haldenstrasse 11, 6006 Luzern
+41 41 618 33 11 / transfer@transferplus.ch
www.transferplus.ch

SWISS INSIGHTS

Corporate Member

BSI Business Systems Integration AG

Täferweg 1, 5405 Baden
+41 58 255 90 00, info@bsi-software.com
www.bsi-software.com