The space to
think series
2024

Showstoppers

Elevating market research data quality into the limelight

INFOTOOLS
HARMONI

What's inside

Data quality is a perennial issue in the business world, one that market researchers have been more aware of than many other organizational functions.

Across the last two years, the team at Infotools has been talking with industry experts on a number of subjects that are important to the insights industry through our *Now that's Significant* podcast. Throughout these discussions, the matter of data quality has been constantly raised.

This publication is a distillation of these conversations, supported by additional data quality knowledge that we've acquired across our 33 years in the insights space. We hope you find this useful in your pursuit of data quality excellence.

An introduction

Storytelling is one of the most powerful tools in the world. And a number of industries have made considerable fortunes from it. Take film and television for example: these mediums turn carefully crafted assortments of data (i.e., words, visuals, and music) into engaging, enthralling, and entertaining spectacles. Randomly throw a combination of those three together and you'll get an incoherent mess that won't even register on Rotten Tomatoes or IMDB scales. Putting this into perspective, the average person crafting a story with those three points together won't get much better. In fact, even professionals don't always get it right.

Market research has several parallels to the film and TV industry. Whereas film and TV series have words, visuals, and music for their data, market research has quantitative, qualitative, behavioral, neuro, transactional, and more methodologies that feed into the data it provides. Experts in either field are able to distill those raw elements and then showcase them in such a way that takes others on a captivating journey towards a specific destination or action. But they're only able to do so according to the quality of their data.

While data quality has always been a hot topic within the research community, the rising action surrounding this challenge has yet to reach its denouement. What might surprise you, however, is that if you search for "data quality" in ESOMAR's vast library dating back more than 75 years, you'll find papers tackling this topic from as early as the 1970s, although the challenge has certainly been going on even longer than five decades. In 1972, the paper "Sources of error in the personal interview" made a plea for reducing errors in survey interviews and focusing attention on raising fieldwork standards. Sound familiar? Looking back even further, Professor Richard Millar Devens coined the phrase "business intelligence" in 1875 to mean gathering information to gain competitive advantage. Is this where our data quality woes began? The conversation has shifted and evolved over time, but the theme certainly remains.

Having been a part of the industry for 33 years and working alongside agencies and corporate insights teams, Infotools has seen data quality take center stage at varying degrees over that time, but it has never completely dropped off the radar. Being a conduit between agencies, data collectors, and corporate teams, Infotools has witnessed the wide-ranging contribution various parts of the industry have made to these quality issues. Although many fingers have been pointed, specific practices blamed, and band-aid solutions implemented, no one has been able to find the data quality magic wand.

Over this past 12 months, for the first time ever, a concerted global effort is underway that holds much promise to make headway on this critical challenge: the Global Data Quality Project. A cross-functional team made up of organizations from all around the world have joined forces to "combat ongoing and emerging risks to data quality in the market and social research, consumer insights, and analytics."

Currently, the Project is made up of the Canadian Research Insights Council (CRIC), ESOMAR, Insights Association, The Market Research Society (MRS), The Research Society (TRS), SampleCon, the Qualitative Research Consultants Association (QRCA) and The Association of Market Research Austria (VMÖ). They have joined forces with a shared mission: to combat ongoing and emerging risks to data quality in market and social research, consumer insights, and analytics. Open to any group from around the world that wishes to be part of the solution to the data quality challenge, the number of involved organizations will likely expand over time.

The Project is addressing a number of key tasks surrounding data quality, with each organization heading up specific data streams. To organize the data quality story in this paper, we have created categories loosely based on some of the themes that the Project is tackling.

- Prologue: Examining the language of quality
- Act one: Finding best practices for fraud detection and mitigation
- Act two: Achieving greater representativity
- Act three: Creating better respondent experiences

Under each of these broad and complex categories, we've tapped into the expert insights shared on our "Now that's Significant" market research podcast to help flesh out the discussion, as well as rounding up some of the commentary from the Project and the insights sector today. We acknowledge the contributions of these experts do not necessarily reflect the views of the organizations for whom they work or represent.

Whether you want to binge this publication all at once or draw it out over a few days or weeks, we'll leave it to you. Either way, we hope it will be useful for improving the outcomes your team helps to achieve.

Ant Franklin
CEO, Infotools

LinkedIn   antfranklin
X   @antfranklin99

How we talk about data quality matters

"So difficult it is to show the various meanings and imperfections of words when we have nothing else but words to do it with."
John Locke

The language the industry uses to discuss data quality is the foundation of the other work being done in the arena. By establishing data quality vocabulary that is used universally, while focusing on accuracy and transparency, further disconnects can be minimized. This is nothing new. Various groups have been working to outline this for years, creating teams to define data quality terms relevant to their specific domain or industry.

For example, data governance frameworks include <u>DAMA International Data Management Body of Knowledge (DMBOK)</u> and the <u>Data Governance Institute's Data Governance Framework</u>. Heavily regulated industries have established clear requirements to guide data quality, such as the healthcare sector's use of frameworks like HL7 and SNOMED CT to define terms related to medical data. Organizations in sectors like this must adhere to accepted definitions and incorporate them into glossaries. In addition, many for-profit businesses provide publicly available glossaries and term definitions surrounding data quality and its associated practices, especially if delivery of reliable data is part of their value proposition.

In the market research space, much of the work being done now on the data quality terminology front is building on existing initiatives from both the public and private sector. Concerted efforts like that of the Global Data Quality Project in the insights space are, in part, a result of increased challenges due to AI and an overall decline in data trust and confidence. They know that if everyone uses the same words when discussing things like fraud, duplicates, and survey cleaning, it can help ensure everyone involved is on the same page, and faster progress can be made.

The Insights Association has already published a glossary of terms related to everything from the behavior of participants in research (are they bad actors or problematic? Or are they valid?) all the way to the words used to talk about bots, bias, types of sample, and fraud detection solutions. This kind of terminology can start to build a common ground for talking about the subject of data quality in insights.

Because the data quality language framework is so critical, part of the Insights Association's work on the Data Quality Project is to refresh and expand global definitions in partnership with the Market Research Society. So far, the joint taskforce has identified upward of 50 terms related to emerging technologies, such as those surrounding bots and AI, to establish a shared vocabulary that truly articulates the challenges being faced. The definitions will provide context, giving practical and precise meanings to these terms, promoting a common understanding among practitioners within the research sector. The two organizations are working together to provide, in context, widespread guidance to the industry.

As the industry continues to rapidly evolve, there is recognition among all groups working on data quality language that this is an area that requires continuous improvement. Data quality glossaries should not be static, but rather evolve as understanding of data quality matures and as new data sources and technologies are introduced.

Insights Association Definition of Research Respondents

Bad Actors: Individuals who misrepresent themselves to qualify for surveys.

Fraudulent Participants: Groups of people attempting to collect survey rewards, often through multiple accounts.

Problematic Participants: Those using automation tools.

Inattentive Participants: Respondents who engage without adequate consideration.

Mischievous Participants: Individuals providing intentionally false or misleading information.

Professional Participants: Participants who complete surveys without regard for eligibility.

Good Completing Participants: Respondents who complete surveys at a high rate with good quality.

Valid Participants: Attentive respondents meeting requirements.

Detecting and mitigating fraudulent activities

"Rather fail with honor than succeed by fraud."
Sophocles

Fraud is complex and dynamic, and the tactics leveraged by cyber fraudsters continue to morph and change creating a very reactionary position for the insights industry. Any tools we have at our disposal to fight against their behaviors are matched on the fraudsters' end with increasingly sophisticated ways to "trick the system." Whether it is humans using new approaches to conduct fraudulent behavior, such as accessing the same survey multiple times to gain more rewards, or using "bots" to answer surveys and then taking home the incentive themselves - all of these behaviors add up to fraud. Just to complicate things, bots are better at emulating real human behavior than they used to be and AI is even answering open-end questions with plausible responses.

According to Accenture, 68% of businesses today feel that their cyber security is at risk on multiple fronts and this is expected to rise in future years as fraudsters become increasingly sophisticated. This is driving the fraud detection and prevention sector to grow with the pace of technology. According to a recent report by Technavio, this sector will grow by more than 47 million USD in the next five years. Another report by Fortune, predicts the growth will be even more accelerated. Both reports name major players in the fraud prevention space such as Experian, IBM and ACI Worldwide that are making investments in the space.

While most market research suppliers employ a number of solutions to combat fraud, some companies focus specifically on fraud mitigation within the insights ecosystem. We spoke with Vignesh Krishnan, the Founder and CEO of Research Defender, on our podcast last year. Research Defender is employed by a number of companies in the industry to help reduce fraud in market research sample through a variety of methodologies. Vignesh says, "Fraud keeps challenging us, and we keep challenging fraud."

He reiterates the importance of employing consistent methods to combat fraud that include all the basics, such as digital fingerprinting, tracking blacklists and similar techniques, as well as more advanced practices. He says companies need to make an investment in technology and human expertise and that a comprehensive data quality program is essential for long-term success.

Lisa Wilding-Brown, Chief Executive Officer at InnovateMR agrees, recognizing that there will always be fraudsters in the market research space and that the way they present themselves is constantly evolving. She reiterates that there is no "magic bullet" or 100% solution to combat this reality, and that insights companies need to be constantly vigilant. Lisa says "My big takeaway is that the DNA of fraud is changing…there are new permutations of fraud for us to be mindful of."

While the entire industry, buyers and suppliers alike, need to prioritize fighting fraud and focus on data quality, much of the immediate responsibility lies in the laps of the sample providers. Karine Pepin, Sr. Vice President at 2CV, said on our podcast that sample has become a commodity, in part due to automation and programmatic technologies. But this "race to maximize traffic and minimize cost cannot go on forever." She maintains we must start to treat sample as an asset, not a commodity.

Jason Buchanan of Verve says that researchers would be surprised if they knew who was in their sample makeup, and maintains that it is important to weed out nefarious players before a survey - rather than trying to identify fraudulent behavior after the fact. Unless a field supplier is focused on making sure that real and verified people are participating, it's a problem and it will continue to be a problem.

It is becoming painfully clear that this critical part of the research process deserves more attention and respect. Multimillion-dollar decisions hinge on participants' self-reported identities, behaviors, and preferences; in other words, they depend on the quality of sample - and the absence of fraudsters.

There is no easy fix to the problem of fraud. Jim Longo of Discuss Research recently gave some practical tips for reducing fraud in an article for Quirk's, including integrating fraud detection software; increasing use of video responses in surveys (hard to fake for bots); using behavioral vetting questions asked in several different ways; and incorporating guidelines from data integrity initiatives - something that researchers can look to the new Global Data Quality Project to provide as the group continues to work toward its objectives.

Overcoming bias to ensure greater representativity in research

"I think unconscious bias is one of the hardest things to get at."   Ruth Bader Ginsburg

The concept of bias has recently taken center stage as AI rises to the forefront of the global conversation. Many are concerned that bias, an unreasoned and unfair distortion of judgment in favor of or against a person or thing, might become baked into AI systems if the "training data" that they use is intrinsically biased from the start. In a recent Forbes article, Monika Mueller of Softensity writes that "machine learning algorithms have the power to amplify these biases, and unless we actively check for and address them, we risk perpetuating societal prejudices unintentionally."

Bias can also sneak into market research. David Paull of Dialsmith talks about this on our podcast, where he covers cognitive biases, such as framing effect, self-reference effect, confirmation bias and things like that. Cint also outlines some common types of bias in quantitative research on its blog, including everything from social desirability bias, or respondents seeking to answer questions in a way that makes them look good, all the way to cultural bias, or making assumptions on data when viewed through a cultural lens. This piece also mentions "ethnocentrism, which is using your culture's standards to judge another."

We also spoke about this topic with Nancy Hernon of G3 Translate, including the importance of culturally relevant survey design. She gave us some international survey design pillars including:

- Ensuring your language is culturally relevant;
- Your demographic information is correct for the target audiences;
- Your questions are relevant to the target audience;
- And your design feels fluid and natural to the target audience without introducing bias.

Taking a one-size-fits-all approach to surveys can introduce bias, and she says "If you really insist that each market receives the same exact questionnaire design for all markets involved, you run the risk of bad data."

Improving representation in research

The Market Research Society Representation in Research Group recommends "expanding profiling points for respondents to go beyond things like gender and age, and include other points like identity, region, social grade, ethnicity, sexual orientation, and physical disability and/or mental health."

In addition, being culturally sensitive equates respect as well as delivering better data quality, says Anne Brown of Gazelle Global, "It's downright rude to expect someone to do a study in English just because you don't want to pay for a translation. A person is going to give their best answer, their best feeling, their most true response, in their native language."

Appropriate survey design is just one small factor in reducing bias. One of the key ways to avoid bias in market research is to ensure that the sample itself is representative of the population, including traditionally underrepresented groups. Sabrina Trinquetel of Measure Protocol has long been a champion for representativity in research, maintaining that data quality is intrinsically linked to this concept. She says that researchers who are conducting the research need to consist of a diverse group of people to gain multiple perspectives and ensure inclusivity, and the respondents we are collecting insights from need to be truly representative of the right population.

The Market Research Society's Representation in Research Group is addressing this, looking to improve representation of groups that are often underrepresented. They recommend expanding profiling points for respondents to go beyond things like gender and age, and include other points like identity, region, social grade, ethnicity, sexual orientation, and physical disability and/or mental health.

Pre-testing questionnaires before they launch can be another tool in the "bias reduction" toolkit, says Lynn Pellicano of Simon Kucher, "There's inherent bias with just taking surveys in general. So if someone refuses to answer a question, that's bias in itself - there could be questionnaire design flaws, maybe you're using inappropriate scales or poorly worded questions. I always encourage teams to conduct a soft launch before they release their survey to a wider population because it really gives you a chance to evaluate, not only the quality of the data, but that you're getting the right respondents in and that your questions make sense to them."

Lauren Isaacson of Curio Research also brings attention to the potential for bias in qualitative research. She champions making research more accessible for those with disabilities, a group which represents approximately 20% of the population. "By ignoring people with disabilities, you're ignoring a significant portion of the population" not to mention that many countries have regulations surrounding inclusivity, so research companies would do well to ensure that they are compliant. She outlines several ways to be more inclusive, and reduce bias, including the way online research is conducted.

Tackling bias and ensuring representativity is also a key focus of Data Quality Project, with a core initiative covering the "identification and mitigation of bias from sample frame and representativeness". Reducing bias runs throughout all the workstreams, with an interesting angle led by The Research Society in Australia regarding the potential for bias when using incentives, alongside how incentives can lead to more extreme cases of fraud.

Avoiding bias in market research is a crucial piece in the data quality conversation, as insights professionals seek to ensure that research results accurately represent the target population and provide actionable insights. Reducing bias requires a combination of approaches to enhance the accuracy and reliability of research findings. Some practices can include using sampling and data collection methods that reach diverse audiences, testing for potential bias before projects launch, and most importantly seeking continuous improvement and a commitment to diversity, equity, and inclusion.

Finding & mitigating bias

"Avoiding bias in market research is a crucial piece in the data quality conversation, as insights professionals seek to ensure that research results accurately represent the target population and provide actionable insights. Reducing bias requires a combination of approaches to enhance the accuracy and reliability of research findings."

Improving respondent experience and engagement

"Experience is the teacher of all things."
Julius Caesar

As the story of the survey has moved from in-person to paper to phone to online to mobile, the narrative tension remains the same: respondents need to be engaged and have good experiences, or it negatively impacts data quality. Fraud aside, when real respondents are frustrated with the survey experience, they tend to display behaviors that deliver bad data such as straightlining (answering every question the same), speeding through their answers, answering with bias (see above, such as providing socially desirable responses), and simply not completing a survey because it is so tedious.

And when respondents aren't engaged, it can affect data quality, says Anne Brown on our podcast. "I think persistent issues in data quality are related to the way that we treat respondents. We don't treat them as humans and with respect." The Data Quality Project gives a nod to this concept, calling for better survey experiences to avoid poor behavior by respondents who might otherwise give insightful answers. Karin Pepin of 2CV says researchers need to be extra careful in designing a shorter, better survey experience to minimize participant dropout, among other fresh approaches.

Practices that have been extensively covered by experts in the field that can help to create positive and meaningful interactions between researchers and survey participants include things outlined on the following pages.

1 Survey design

Attention spans are dropping. A study led by Microsoft Canada found that since the year 2000, the average attention dropped from 12 seconds to 8 seconds. This doesn't lend itself well to lengthy, complicated surveys, which can quickly cause respondent fatigue and lead to lower response rates. Beyond concerns about length, creating user-friendly and engaging designs is also important. This can include using clear and simple language and optimizing question formatting, including using varied types of questions to create a more interesting experience, among other things.

2 Mobile-first approaches

This should be a no-brainer. In a survey by Google, 94% of respondents reported using their smartphones to take surveys. If surveys are not created to work seamlessly and responsively in a mobile environment, response rates will drop as frustration with clunky experiences rises. Dynata provides some practical tips on its blog about how to create better mobile surveys, including designing them for landscape and portrait views, reducing the number of questions, avoiding Flash and rolling designs, and more.

3 Proper rewards

According to research-on-research, the majority of survey respondents do it for the money - nearly 65% according to GRBN. While there is some debate on whether or not incentives can increase or decrease quality, there is no doubt that they do boost response rates. Payouts for survey takers have been dismally low in the past; rewarding individuals properly for their time is an important piece of the survey experience. Some say incentives can negatively affect data quality, although this is generally disproved by research-on-research, including this recent peer-reviewed academic paper published by the Public Library of Science.

Respondent respect

4

Treating survey participants as humans is one of the first steps toward better data quality. Communicate up front about survey purpose and length; use language that is grateful and considerate throughout the survey; use culturally appropriate approaches; and focus on building positive relationships. And stop the endless routing. Anne Brown says, "We route them mercilessly in online samples from survey to survey. We're asking them the same questions over and over again. We have endless attribute lists." These simple acts of respect can foster trust and encourage participation. She continues, "We'll get better responses if the respondent feels like they're in an ecosystem where they can trust that they're being treated properly."

Put privacy first

5

As concerns around personal data protection and security rise, researchers must ensure that survey participant data is protected and handled in compliance with privacy regulations. Clear privacy statements and secure data handling practices should be prioritized and communicated with individuals. According to GRBN, only 1-in-10 people have a high level of trust in market research companies, so survey takers need to be well- informed about how their personal data is being collected, stored, and used to start building confidence and trust. Sabrina Trinquetel shared a study by Measure Protocol that indicated that 50% of individuals felt concerned about privacy of their data when taking surveys, and aren't sure what is being done with their data.

Positive experience respondent experiences not only lead to more accurate data but also contribute to maintaining a willing and engaged pool of respondents for future studies.

Prioritizing respondents

Practices such as improved survey design, mobile-first approaches, proper rewards, respect and communication and a focus on privacy can help lead to positive respondent experiences, not to mention more accurate, complete data.

In conclusion

"Find the journey's end in every step."
Ralph Waldo Emerson

The data quality conversation is far too large to fit into a single document, or this challenge would have been solved long ago based on the sheer number of academic papers, articles, presentations, training, education and thought leadership surrounding the subject.

Today, the Global Data Quality Project is one of the industry's first concerted efforts to tackle this in a practical and comprehensive way. Its formation was driven by a confluence of factors that have caused a drastic erosion of trust and confidence in market research insights. Shifting this perception is critical to the future success of the industry.

While it's a seemingly insurmountable task, it is possible when everyone plays their part in lifting the overall quality data of the industry. Hopefully, this publication has helped underline the need for collaboration. Because when people and insights teams do, it's far more likely that stakeholders will want to binge as many market research projects as they can. Surely that would be a win-win for both consumers and market research teams alike.

We'd like to say a big thanks to the cast who helped make this piece possible: Vignesh Krishnan, Lisa Wilding-Brown, Karine Pepin, Jason Buchanan, David Paull, Nancy Hernon, Sabrina Trinquetel, Lynn Pellicano, Lauren Isaacson, Anne Brown, Marie Melsheimer, and everyone else who is making a difference to lift the data quality bar in the market research and insights industry.

And this brings us to our final point.

It's now time to take [or keep taking] *3... 2... 1... action!*
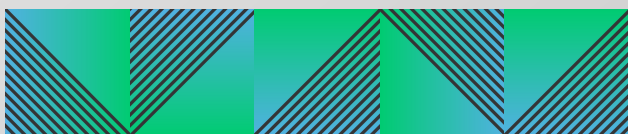
The space to think series

2024

Infotools was created by curious market researchers who wanted to uncover new ways to better understand the world. And we're still just as curious. We're acutely aware of how deep insights require time, and can't be rushed. That's why everything we do at Infotools is dedicated to giving market researchers more space to think. We trust this and other papers in this series will do just that. If you're interested in other publications in this series, feel free to check them out below.

The buyer's guide to market research analysis and reporting software **>**

Insights and innovation: The market research software & tech landscape **>**

The Inclusion, Diversity, Equity, and Accessibility Handbook **>**

Showstoppers: Elevating market research data quality into the limelight **>**

Harmoni at work in your world

Infotools Harmoni enables your team to:

Increase productivity. Spend less time processing research data, and more time analyzing it. Our platform uses automation to make your job easier.

Deliver real value. It's easy to provide valuable, actionable insights to stakeholders.

Rely on many levels of support. From our excellent Customer Services Team to our extensive Knowledge Base, a searchable site full of video tutorials and helpful "How to" tips.

Process your data. This is a data agnostic platform that understands market research nuances, supports most file types, and has APIs to Decipher, Voxco and others.

Report. Create beautiful charts, reports and dashboards with the click of a button. Highlight significant data points, suppress low sample sizes, and more.

Meet growing expectations. Market research budgets are shrinking, but expectations are ramping up. Slice and dice your data on the go, as the questions are being asked.

Analyze. With our platform processing complex tasks like calculating statistical significance, you have more time to dig and discover those magic insights.

Share. Share your insights with clients either in dashboards, story views or even exporting to PowerPoint or Excel. Plus, keep them informed of the latest updates with mobile data alerts.

Connect with us

Request a demo today

🌐 infotools.com
✉️ hello@infotools.com
in infotools
🐦 infotools
f infotools1

INFOTOOLS®
HARMONI