



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Hausarbeit

Hans-Peter Sieg

„Lügen Daten“

Eine Odyssee durch den Datenschwungel

Hans-Peter Sieg

„Lügen Daten?“

Eine Odyssee durch den Datendschungel

Hausarbeit
im Studiengang Next Media
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer: Prof. Dr. Kai von Luck

Abgegeben am 29.2.2016

Hans-Peter Sieg

Thema

Lügen Daten?

Eine Odyssee durch den Datenschungel

Stichworte

Smart city, Luftqualität, Datenqualität, urban storytelling

Kurzzusammenfassung

Diese Hausarbeit beschäftigt sich mit den Zusammenhängen, Abgrenzungen und Problemen der Themenfelder Beurteilung von Informations- und Datenqualität, Datenbeschaffung, Datenerhebung und Datenkonsistenz am Beispiel einer konkreten Datenrecherche zum Thema Luftqualität in Hamburg im Rahmen des Projekts „urban storytelling - case:air“.

Im Weiteren wird beispielhaft ein Smart City Projekt zur alternativen Luftqualitätsmessung vorgestellt. Im Abschluss gehe ich kurz auf die Smart City Initiative der Stadt Hamburg ein.

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	2
2.1	Daten- und Informationsqualität.....	2
3	Datenrecherche	4
3.1.	Datenquelle.....	4
3.2.	Datenprobleme.....	5
3.2.1	Luftqualitätsindex (LQI).....	5
3.2.2	Datenerhebung.....	6
4	Alternativen.....	7
5	Resümee.....	9
6	Ausblick	10
7	Literatur	11

1 Einleitung

„Die chinesischen Behörden warnen vor Luftverschmutzung in ganz Nordchina. Am schlimmsten ist die Hauptstadt betroffen. Schulen und manche Betriebe bleiben geschlossen.“¹

Im asiatischen Raum nimmt offensichtlich die Luftqualität inzwischen lebensbedrohliche Ausmaße an. Aber wie sieht es in Deutschland und speziell in Hamburg aus?

„Schlechte Luft: BUND droht Hamburg mit Zwangsgeld - Naturschützer werfen Rot-Grün vor, das Gerichtsurteil von 2014 nicht umzusetzen. Stadt betont, die Belastung sei bei jeder zweiten Messstation rückläufig.“² Und auch in Europa ist das Thema Luftqualität von hoher gesellschaftlicher und politischer Brisanz.

Die folgende Arbeit beschreibt eine konkrete Datenrecherche zum Thema Luftqualität in Hamburg. Im Weiteren wird ein alternatives Konzept der Luftschadstoffmessung vorgestellt.

Als Ausblick wird die Idee angerissen, ob eine Verbindung von Luftqualitätsmessung, Bürgerpartizipation und der Smart City Initiative des Hamburger Senats möglich ist.

¹ Die Zeit, 18. Dezember 2015

² Hamburger Abendblatt, 27. Januar 2016

2 Grundlagen

Da „Qualität“ kein inhärenter Bestandteil von Daten ist, bedarf es eines Kriterienkataloges nach dem Daten und deren Qualität bewertet werden können. Denn Daten an sich sind nur „Darstellungen / Angaben über Sachverhalte und Vorgänge“.⁶

2.1 Daten- und Informationsqualität

Holger Nohr definierte 2001 den Begriff Daten- und Informationsqualität als „Gesamtheit der Anforderungen an eine Information bzw. ein Informationsprodukt, die sich auf deren Eignung zur Erfüllung gegebener Informationsbedarfe beziehen“.⁷

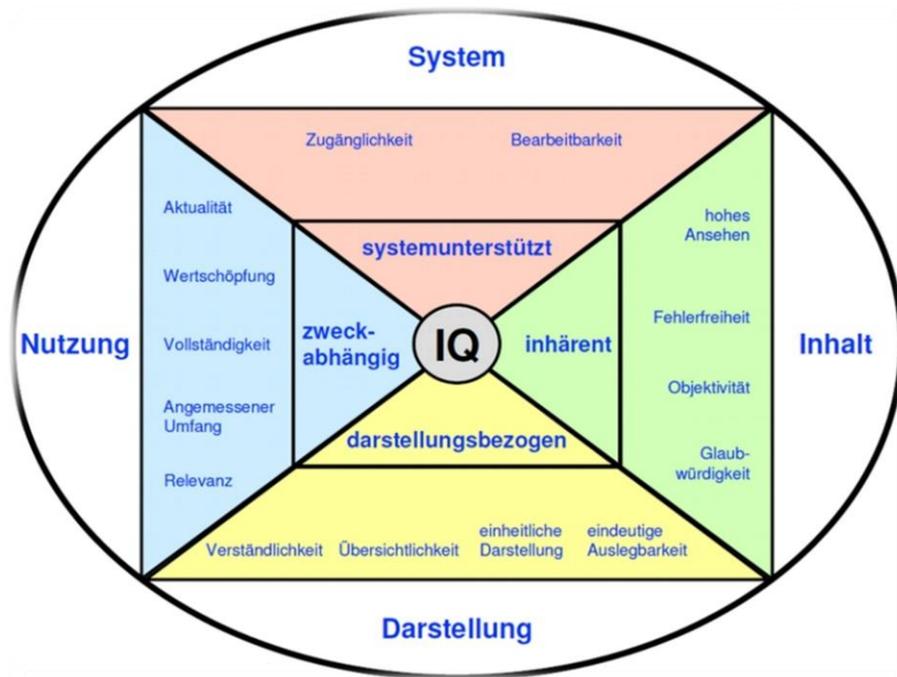
Kriterien zur Qualitätsbeurteilung sind bereits Mitte der 1990er Jahre von Richard Y. Wang und Diane M. Strong entwickelt worden und von Ihnen ist ein Kriterienkatalog aufgebaut worden, der bis heute als Standard zur Beurteilung von Daten- und Informationsqualität benutzt wird.⁸

Die Deutsche Gesellschaft für Informations- und Datenqualität (DGIQ) hat 2007 eine anwenderbezogene Visualisierung auf Basis dieser Forschungsergebnisse veröffentlicht.

⁶ Herbert Ernst Wiegand: Wörterbuchforschung: Wörterbuchforschung, Tl.1, Swalter de Gruyte, Reprint 2012, S. 163

⁷ Holger Nohr, Management der Informationsqualität. In: Michelson, M. & Riekert, W.-F. (Hg.): Informationswirtschaft: Innovation für die neue Ökonomie. Wiesbaden: Gabler, 2001. S. 6

⁸ Richard Y. Wang and Diane M. Strong. 1996. Beyond accuracy: what data quality means to data consumers. *J. Manage. Inf. Syst.* 12, 4 (March 1996), 5-33.
DOI=<http://dx.doi.org/10.1080/07421222.1996.11518099>



9

Die Kriterien sind in 15 Dimensionen, die wiederum in 4 Kategorien qualifiziert, auf deren Grundlagen Daten- und Informations-Qualität beurteilt werden kann und somit zu einem Informationsgewinn kommt.

Die 4 Kategorien sind:

- Systemunterstützte, also Zugänglichkeit und Bearbeitbarkeit
- Zweckabhängige, also Aktualität, Wertschöpfung, Vollständigkeit, Relevanz und angemessener Umfang
- Darstellungsbezogene, also Verständlichkeit, Übersichtlichkeit, Einheitlichkeit, Eindeutigkeit
- und inhärente, also hohes Ansehen, Fehlerfreiheit, Objektivität und Glaubwürdigkeit.

⁹ DGIQ IQ_Definition 2007, <http://www.itwissen.info/definition/lexikon/Datenqualitaet-data-quality.html>, Abruf 10.1.2016

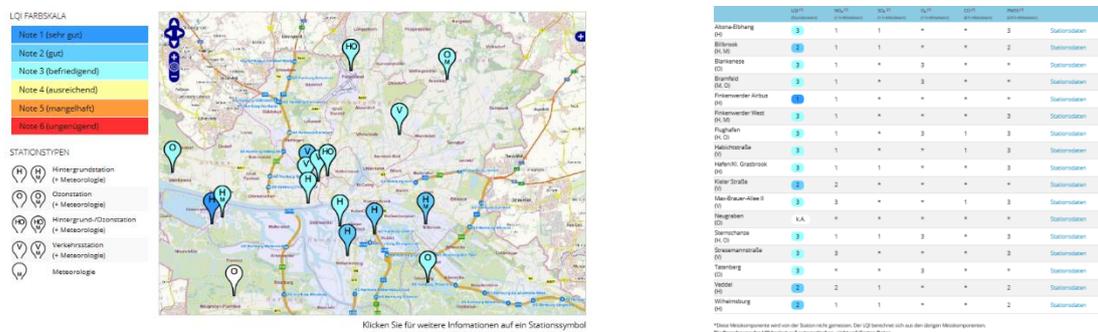
3 Datenrecherche

Ziel der Datenrecherche ist aussagekräftige Informationen über die Luftqualität Hamburgs zu erhalten.

Daten über Luftqualität müssen laut Umweltinformationsgesetz (UIG) Abschnitt 2, § 3.Abs. 1 veröffentlicht werden. „Jede Person hat nach Maßgabe dieses Gesetzes Anspruch auf freien Zugang zu Umweltinformationen.“¹⁰

3.1 Datenquelle

Die Seite des Hamburger Luftmessnetzes (HaLm)¹¹ bietet sowohl aktuelle als auch archivierte Luftschadstoffmessdaten. Es wird dem Bürger eine übersichtliche interaktive Karte angeboten, genauere Informationen zu den einzelnen Luftmessstationen sind verfügbar und es wird für jede dieser Stationen ein Luftqualitätsindex (LQI) angezeigt. Daten können für einzelne Stationen im CSV-Format für einen Zeitraum von 3 Monaten rückwirkend heruntergeladen werden.



¹⁰ http://www.gesetze-im-internet.de/uig_2005/, Abruf 10.1.2016

¹¹ <http://luft.hamburg.de/>; die Abbildungen sind Screenshots dieser Seite, Abruf 10.1.2016

3.2. Datenprobleme

Bei intensiverer Analyse der vom Luftmessnetz Hamburg angebotenen Daten ergaben sich einige Auffälligkeiten. Beispielfähig führe ich im Folgenden zwei Problemstellungen auf.

3.2.1 Luftqualitätsindex (LQI)

Ein erstes Problem taucht bei dem angebotenen Luftqualitätsindex (LQI) auf. Der LQI wird auf der Grundlage der gemessenen Luftschadstoffe: Schwefeldioxid (SO₂), Kohlenstoffmonoxid (CO), Stickstoffdioxid (NO₂), Ozon (O₃), Feinstaub (PM10) ermittelt. Diese Schadstoffe werden im LQI in Indexklassen eingeteilt, die entsprechend einem Schulnotensystem mit den Klassen 1 („sehr gut“) bis 6 („sehr schlecht“) eingruppiert werden. Der Übergang der Klasse 4 („ausreichend“) in die Klasse 5 („schlecht“) entspricht dem Überschreiten des EU-weit gültigen Grenz- oder Informationswertes für diesen Schadstoff. Für eine Messstation wird immer die höchste Belastungsklasse der an der Station gemessenen Schadstoffe als Indikator für die Schadstoffbelastung an dem Standort angegeben.¹²

NO ₂ 1-h-MW (µg/m ³)	SO ₂ 1-h-MW (µg/m ³)	CO 8-h-MW (mg/m ³)	Ozon 1-h-MW (µg/m ³)	PM10 24-h-MW (µg/m ³)	LQI Note	Bewertung
200	350	10	180	50		Grenzwerte
0–25	0–25	0–1	0–33	0–10	1	sehr gut
>25–50	>25–50	>1–2	>33–65	>10–20	2	gut
>50–100	>50–120	>2–4	>65–120	>20–35	3	befriedigend
>100–200	>120–350	>4–10	>120–180	>35–50	4	ausreichend
>200–500	>350–1000	>10–30	>180–240	>50–100	5	schlecht
>500	>1000	>30	>240	>100	6	sehr schlecht

13

Bei genauer Betrachtung einzelner Stationen ist festzustellen, dass das HaLm auch für die Stationen einen LQI ermittelt, bei denen nicht alle für die Ermittlung notwendigen Werte vorhanden sind.

Eklatantestes Beispiel ist die Station Kieler Straße, eine Verkehrsmessstation, die an einer Straße mit sehr hohem Verkehrsaufkommen steht. Hier wird der für die Beurteilung der Luft durch

¹² <http://luft.hamburg.de/wir-ueber-uns/4242598/hamburger-luftqualitaetsindex/>, Abruf 10.1.2016

¹³ <http://luft.hamburg.de/wir-ueber-uns/4242598/hamburger-luftqualitaetsindex/>, Abruf 10.1.2016

Verkehrsbelastung entscheidende PM10-Wert nicht gemessen. Dennoch wird ein LQI errechnet und angegeben.

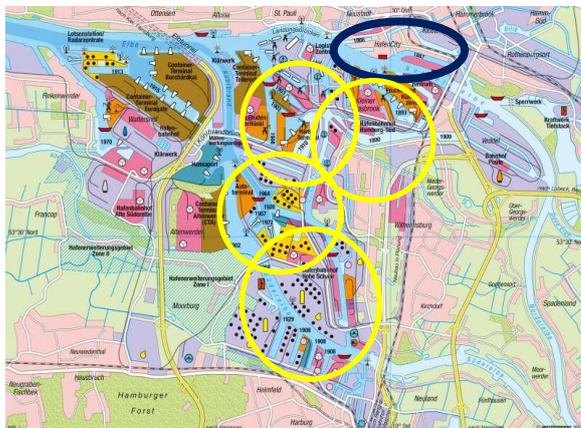
Der LQI wird vom HaLm beschrieben als „...ein stundenaktueller Indikator für die allgemeine Luftqualität. Es wird hiermit versucht, die Kurzeiteinwirkung von Luftschadstoffen auf die menschliche Gesundheit zusammenzufassen und übersichtlich darzustellen.“¹⁴

Die Kriterien Glaubwürdigkeit, Objektivität, Vollständigkeit und Relevanz (s. Kriterien in Kapitel 2) sind hier nicht erfüllt. Eine Errechnung UND visuelle Darstellung des Luftqualitätsindex sehe ich als zumindest fragwürdig an.

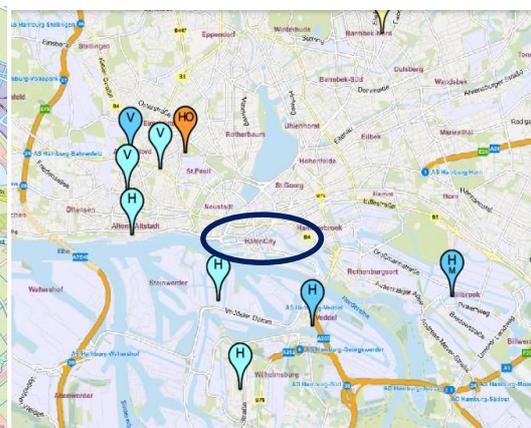
3.2.2 Datenerhebung

In Hamburg wird seit 2001 mit der Hafencity ein komplett neuer Stadtteil gebaut. Es werden dort 6.000 bis 7.000 Wohnungen für ca. 14.000 Bewohner und für bis zu 45.000 Arbeitsplätze bis 2020 entstehen. Dieser Stadtteil grenzt an großen Hafen- und Industrieflächen, die sich süd- und Südwestlich befinden. Weiterhin befinden sich in südlichen Teil der Hafencity zwei Kreuzfahrtterminals, die im Jahr 2015 von 84 Kreuzfahrtschiffen angelaufen wurden.^{15 16}

17



18



¹⁴ <http://luft.hamburg.de/wir-ueber-uns/4242598/hamburger-luftqualitaetsindex/>, Abruf 10.1.2016

¹⁵ http://www.schiffe-in-hamburg.de/de/_calls_2015.htm, Abruf 10.2.2016

¹⁶ Für die kommenden fünf Jahre erwartet die Hamburg-Port Authority eine Zunahme der Schiffläufe von ca. 30%, <http://www.hamburg-port-authority.de/de/presse/studien-und-berichte/Documents/Kreuzfahrtprognose%20Hamburg%202030.pdf>, Abruf 10.2.2016

¹⁷ Hafen-Gewerbekarte: <http://www.diercke.de/content/hamburg-hafen-978-3-14-100800-5-34-2-1>, Abruf 10.1.2016

¹⁸ <http://luft.hamburg.de/>, Abruf 10.1.2016

Die Hauptwindrichtung in Hamburg liegt zwischen 190 und 210 Grad, das heißt: zwischen Süd und Süd-Südwest.¹⁹ Ein Blick auf die Lokation der Messstationen zeigt, dass sich in der Hafencity keine Luftmessstation befindet. Aber auch in den Stadtteilen, die in der Hauptwindrichtung liegen, sind keine Messstationen eingerichtet.

Die Kriterien Glaubwürdigkeit und Objektivität, (s. Kriterien in Kapitel 2) sind hier nicht erfüllt. Eine Landesbehörde gefährdet dadurch nach meiner Beurteilung ebenfalls sein „Hohes Ansehen“ (s. Kriterien in Kapitel 2).

4 Alternativen

Forscher des Institute for Advanced Architecture of Catalonia in Barcelona haben zusammen mit dem in Kalifornien ansässigen Unternehmen Acrobotics Industries ein „SMART CITIZEN KIT“ entwickelt. Kernstück ist der mit Sensoren ausgestattete Open-Source-Mini-PC „Arduino“, der vom Benutzer an einen geeigneten Standort – zum Beispiel den eigenen Balkon – gesetzt wird, um lokale Umweltdaten aufzuzeichnen. Es werden Temperatur, Luftfeuchtigkeit, Luftschadstoffe (CO und SO₂) aufgezeichnet.²⁰ Die ermittelten Daten werden an eine Plattform gesendet, dort validiert und sind für jeden abrufbar.

¹⁹ Winddaten Hamburg: <http://ecowetter.de/ort/klima/hamburg/>, Abruf 10.1.2016

²⁰ <https://smarcitizen.me/>, Abruf 10.12.2015

In Barcelona,²¹ Manchester,²² Amsterdam²³ und seit Ende 2015 auch in Wien²⁴ werden im Rahmen der jeweiligen Smart-City-Projekte diese SMART CITIZEN KITS getestet.



25

Ziel des SMART CITIZEN KIT-Projektes ist, das Thema Luftqualität und Luftmessungen zu demokratisieren und dem Bürger so die nötigen Werkzeuge an die Hand zu geben, um ihn zu befähigen selber Daten zu sammeln und Luftqualitätsdaten nicht, wie bisher punktuell, sondern aus der Fläche zu erheben.

Es gibt im Rahmen des Amsterdamer SMART-CITIZEN-KIT-Projekts Arbeiten, um auch die SO₂-, O₃- und Feinstaubwerte messen und einbinden zu können.²⁶

²¹ <http://us2.campaign-archive1.com/?u=d67ba8deb34a23a222ec4eb8a&id=439df431b6&e=06fb8b7939>, Abruf 10.12.2015

²² <http://futureeverything.org/projects/smart-citizen/>, Abruf 10.12.2015

²³ <http://amsterdamsmartcity.com/projects/detail/id/69/slug/smart-citizen-kit#>, Abruf 10.12.2015

²⁴ <https://schoolofdata.wiki.zoho.com/public/ViennaSensor.html>, Abruf 10.12.2015

²⁵ <https://smartcitizen.me/>, Abruf 10.12.2015

²⁶ <http://futurezone.at/digital-life/smart-citizens-daten-sollen-handlungen-ausloesen/124.314.354>, Abruf 10.12.2016

5 Resümee

Nach den von Richard Y. Wang und Diane M. Strong ermittelten Kriterienkatalog weisen die vom Hamburger Luftmessnetz veröffentlichten Messdaten erhebliche Mängel auf. Durch fehlende und nur punktuelle Datenerhebung (s. Kap 3.1.2) und die (zumindest fragwürdige) Datenauswertung (s. Kap 3.2.1) können zum jetzigen Zeitpunkt keine aussagenkräftige Informationen über die Luftqualität Hamburgs gewonnen werden. Dass dringender Handlungsbedarf besteht, ergibt sich auch aus der Tatsache, dass, obwohl nur lückenhafte Daten vorhanden sind, die Grenzwerte für Feinstaub (PM 10) im Jahr 2015 116-mal, im Jahr 2014 134-mal überschritten wurden (es sind laut EU maximal 35 Überschreitungen p.a. erlaubt²⁷).

Welche Relevanz das Thema hat, lässt sich auch aus dem Gerichtsurteil des Verwaltungsgerichts Hamburg ermesen. Die Stadt ist 2014 zur „Verpflichtung einer schnellstmöglichen Umsetzung von weiteren Maßnahmen zur Luftreinhaltung“²⁸ verurteilt worden. Der BUND, der zusammen mit Anwohnern der Max-Brauer-Allee diese Klage eingereicht hat, hat am 26. Februar 2016 einen Antrag auf Zwangsgeld gegen den Hamburger Senat gestellt, da die Fachbehörden der Stadt trotz der rechtskräftigen Verurteilung an Ihrem Zeitplan festhalten will und erst im September 2017 ein neuer Luftreinhalteplan vorlegen will.²⁹

²⁷ <http://eur-lex.europa.eu/legal-content/DE/TXT/PDF/?uri=CELEX%3A32008L0050&qid=1419250736800&from=DE>, Abruf 10.2.2016

²⁸ Urteil des Verwaltungsgerichtes Hamburg (VG Hamburg 9 K 1280/13) vom 06. November 2014

²⁹ http://bund-hamburg.bund.net/nc/presse/pressemitteilungen_hamburg/detail/artikel/luftschadstoffe-bund-stellt-antrag-auf-zwangsgeld-gegen-hamburger-senat/?tx_ttnews%5BbackPid%5D=4655&cHash=9c8eefac63f6f84547a15e1cc1dff87a, Abruf 27.2.2016

6 Ausblick

Hamburg ist eines von den 12 Bundesländern die ein Informationsfreiheitsgesetz auf Länderebene haben „um über die bestehenden Informationsmöglichkeiten hinaus die demokratische Meinungs- und Willensbildung zu fördern und eine Kontrolle des staatlichen Handelns zu ermöglichen.“³⁰

Der Senat hat im April 2014 eine „Smart City Initiative“ zusammen mit Cisco-Systems gestartet.³¹ In dieser Initiative sind eine Einbindung des Bürgers und damit eine „Demokratisierung“ des Smart-City-Konzeptes bisher nicht erkennbar.

Wenn der Wille zur Öffnung hin zu demokratischer Einbindung des Bürgers der „Smart-City-Initiative“ zusammengebunden würde mit Initiativen wie z.B. das SMART-CITIZEN-KIT-Projekt, wäre die Chance Luftdaten auf breiterer Basis erheben zu können und damit die Kriterien nach Daten- und Informationsqualität besser erfüllen zu können, möglicherweise höher.

Aber auch dem Gesetzauftrag nach Förderung „[der] demokratische Meinungs- und Willensbildung“ und damit Bürgerinnen und Bürgern eine stärkere Partizipation an gesellschaftlicher Teilhabe zu ermöglichen, würde womöglich besser entsprochen.

³⁰ HMBTG, § 1, Abs. 1, <http://transparenz.hamburg.de/das-hmbtg/>

³¹ <http://www.hamburg.de/smart-city/4306386/cisco-smart-city/>

7 Literatur

- Devarakonda, Srinivas, Parveen Sevusu, Hongzhang Liu, Ruilin Liu, Liviu Iftode, and Badri Nath. 2013. Real-time air quality monitoring through mobile sensing in metropolitan areas. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (UrbComp '13). ACM, New York, NY, USA, , Article 15 , 8 pages. DOI=<http://dx.doi.org/10.1145/2505821.2505834>
- Doering, Michael. 2011. High-resolution large-scale air pollution monitoring: approaches and challenges. In Proceedings of the 3rd ACM international workshop on MobiArch (HotPlanet '11). ACM, New York, NY, USA, 5-10. DOI=<http://dx.doi.org/10.1145/2000172.2000177>
- Hu, Ke, Timothy Davison, Ashfaqur Rahman, and Vijay Sivaraman. 2014. Air Pollution Exposure Estimation and Finding Association with Human Activity using Wearable Sensor Network. In Rahman, Ashfaqur, Jeremiah Deng, and Jiuyong Li (Eds.). Proceedings of the MLSDA 2014
- Klein A. and W. Lehner. 2009. Representing Data Quality in Sensor Data Streaming Environments. J. Data and Information Quality 1, 2, Article 10 (September 2009), 28 pages. DOI=<http://dx.doi.org/10.1145/1577840.1577845>
- Leonardi, Chiara, Andrea Cappellotto, Michele Caraviello, Bruno Lepri, and Fabrizio Antonelli. 2014. SecondNose: an air quality mobile crowdsensing system. In Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational (NordiCHI '14). ACM, New York, NY, USA, 1051-1054. DOI=<http://dx.doi.org/10.1145/2639189.2670273>
- Li, Jason Jingshi, Boi Faltings, Olga Saukh, David Hasenfratz, Jan Beutel, Sensing the Air We Breathe — The OpenSense Zurich Dataset; <http://www.aaai.org/ocs/index.php/AAAI/AAAI12/paper/view/4896/5158>
- Lucke, Jörn von. 2014. Open Societal Innovation. In Proceedings of The International Symposium on Open Collaboration (OpenSym '14). ACM, New York, NY, USA, , Pages 14 , 3 pages. DOI=<http://dx.doi.org/10.1145/2641580.2641598>

Madnick, Stuart E., Richard Y. Wang, Yang W. Lee, and Hongwei Zhu. 2009. Overview and Framework for Data and Information Quality Research. *J. Data and Information Quality* 1, 1, Article 2 (June 2009), 22 pages. DOI=<http://dx.doi.org/10.1145/1515693.1516680>

Nohr, Holger, Management der Informationsqualität. In: Michelson, M. & Riekert, W.-F. (Hg.): *Informationswirtschaft: Innovation für die neue Ökonomie*. Wiesbaden: Gabler, 20012nd Workshop on Machine Learning for Sensory Data Analysis(MLSDA'14), ACM, New York, NY, USA, , Pages 48 , 8 pages. DOI=<http://dx.doi.org/10.1145/2689746.2689749>

Wang, Richard Y. and Diane M. Strong. 1996. Beyond accuracy: what data quality means to data consumers. *J. Manage. Inf. Syst.* 12, 4 (March 1996), 5-33.
DOI=<http://dx.doi.org/10.1080/07421222.1996.11518099>

Websites

<http://www.siemens.com/innovation/de/home/pictures-of-the-future/digitalisierung-und-software/von-big-data-zu-smart-data-dossier.html>

http://ec.europa.eu/environment/europeangreencapital/index_en.htm

<http://www.wired.com/2008/06/pb-theory/>

https://www.sit.fraunhofer.de/fileadmin/dokumente/studien_und_technical_reports/Big-Data-Studie2015_FraunhoferSIT.pdf

Versicherung über Selbstständigkeit

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Hilfsmittel benutzt habe.

Hamburg, den 29. Februar 2016

Hans-Peter Sieg