October **2018**

Knowledge Management - A primer

Marc Maxmeister,*
Caroline Fiennes,
Natalia Kiryttopoulou

**Corresponding author: marc@keystoneaccountability.org*

keystone
accountability for social change

Giving Evidence
Enabling **giving** based on sound **evidence**

Knowledge Management (KM) efficiently handles an organization's information and resources, in pursuit of answering its most important[1] questions, while having mission-related impact. All institutions create a lot more knowledge than they use, and most do a poor job of organizing information[2] so that it can be effectively put to work by people during strategy redesign and programme implementation. This primer will cover only the "need to know[3]" aspects of KM and define what a KM system is, what it is not, and provide guidelines for designing good systems.

We envision creating / improving KM as a five-step process.

1. **Understand** what those key questions are. This requires a clear theory of change.
2. **Define:** determine what information you need.
3. **Collect**: plan how your team and systems will collect this information.
4. **Interpret**: make sense of the information, using computers to facilitate and accelerate interpretation.
5. **Act** on what the organization has learned - to manage adaptively - changing programmes and approaches as new information comes to light.

Many other people have proposed various five-step processes for KM, shown in Figure 1[i]. Our guidelines focus on how to improve Steps 2, 3, and 4 (Define, Collect, and Interpret). The other two steps focus on an organization's overall strategy and management, separate from the knowledge.

Figure 1: The Knowledge Management Cycle

Step	Keystone & Giving Evidence	Zack	Bukowitz & Williams	WIIG	McElroy	Colantino & Harmon[ii]
1	Identify key questions	Acquire	Get	Create	Formulate problem claim	Come up with a strategy
2	Define evidence required	Refine	Use	Source	Learn and Validate	Incentivize contributions
3	Devise a plan for collecting	Store	Learn	Compile	Acquire	Create a culture
4	Interpret and Summarize	Distribute	Contribute	Transform	Integrate	Start simple with document collection
5	Act on the evidence	Present	Assess	Apply	Complete	Standardize and streamline your forms

1 By important, we mean an institution's ability to have the impact it aims to have on some global issue, while also sustaining itself financially.
2 The concept of information assumed here is that of Claude Shannon's "Shannon Information Theory". His seminal and prescient 1948 paper (http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf) defines information in a way that makes it quantifiable in any context - from noise on phone lines (he worked in Bell Labs), to encryption, to our understanding that information density of video, audio, and narrative exceeds that of spreadsheets with M&E indicators by orders of magnitude. Some thoughts on how it applies to aid agencies here and here.
3 Any "neat to know" ideas will be relegated to footnotes.

In Step One, KM designs must fit with a clear framework or taxonomy surrounding an institution's *raison d'être*. Only then will the initial line of questions become clear. The framework should highlight what an institution needs to know and learn.

Steps Two and Three are about **information flow**: the team delves into the specific information it needs and how to collect it. What relevant information is available? How do people provide that information? What new information must you collect? Figure 2 offers some relevant data questions[iii] that can help in designing a plan. This can help you choose between paper-based and web-based collection approaches, and be prepared to take advantage of future technology improvements in machine learning, automation, and interoperability.

Figure 2: Important questions to ask in Step 3 - Devise a plan for data collection

Access to data		
Accessible: You have the data.	**Available:** You need to call someone to get it.	**Aspirational:** You want the data, but can't get it (yet).
Prioritize: Decide which data are critical, and ensure they are managed first / more carefully.		
Data Logistics		
Timing of data collection: At point of contact? Quarterly? Yearly? Something else?	**Who** collects the data? Which partners need to be engaged to access the data?	**How** is the data formatted? Does it need to be converted or standardized before interpretation?
How tidy or messy is the data?	**How rich** is the data (for further, deeper analysis)?	**Challenges** with data: Describe any other access issues.

It is important to know how data is structured. ***Information becomes knowledge*** when it is structured to address the business demands of an organization[4]. Because restructuring incoming data is the greatest barrier to data interpretation in Step Four, one must plan for how data will be (re)organized and how they will be aligned with the organization's needs *before* collecting it in Step Three. If you know that your incoming data will be unstructured, how messy will it be? Will reports include a clear reference to named organizations? Or will someone be "scraping" websites and social media mentions and guessing at the organizations that deserve attribution? Will the people in a data set have names and unique identifiers, such as a phone number? Or will they be anonymous comments from communities meetings where nothing can be connected back to organizations? Knowing the data's limits allows an organization to invest in improving data, and these investments are justified if the data answers key strategic questions. The messiness within a data set and the lack of consistency from one collection period to the next is quantifiable, it turns out. We recommend Hadley Wickham's **Tidy Data**[iv] as a useful treatise on how to detect data tidiness and how to refine / interpret messy data.

4 It's worth mentioning here that in statistics there are ***three*** types of error (most college classes leave out the zeroth case): **Type 0 error** is mistake of answering the wrong question. Type 1 is a false positive result, and Type 2 is false negative result.

As data is collected (Step Three), a good KM approach needs to restructure and align it with the needs of the organization. The data pipeline should also archive data in its original form and link these archives with its representation in the system. As an example of this, read about how AirBnB designed their KM system[v].

Step Four: Summarizing and interpreting data are increasingly becoming about restructuring data to fill dashboards, trigger alerts, enable machine learning (automated pattern detection), and similar forms of automation. A KM system that merely archives the unprocessed "raw" data will fail to provide enough value to the organization to justify the investment. There are levels of sophistication in how an organization stores and commoditizes data:

1. Data is in **paper** form, spread out in many desks and filing cabinets, with no way of locating it.
2. Everything is saved on personal desktop **computers**.
3. Everything is in **cloud**[5] storage, such as google drive or dropbox.
4. Everything is stored in a single place and there's a global index or logical folder structure one can navigate to locate most data - i.e. in the cloud and **indexed**.
5. Cloud storage, indexed, and full text contents of documents are **searchable**[6].
6. Cloud + index + search, and incoming data is being pulled regularly by ***humans*** to produce dashboards that display key performance indicators (KPIs) that address key strategic issues[7].
7. Cloud + index + search, and ***computers*** regularly produce dashboards that display key performance indicators (KPIs) that address key strategic issues.
8. Cloud + index + search + computerized dashboards + KPIs, and humans regularly run **machine learning**[8] algorithms on the raw data underlying these KPIs to glean deeper insights and learning about patterns that lie outside current assumptions and the theory of change - i.e., implicit inquiry.
9. Future iteration on level 8 is likely to produce deeper learning that drives action - currently in research mode with such examples as Watson, Google Deep Dream, Tensor Flow, and Genetic Algorithms. (If you don't know what any of these are, don't worry. You may never have to, if you achieve the other levels.)

An organization operating at levels 1-4 isn't likely to get value from its KM system; its staff are feeding data into a system that they cannot easily search. Yet in 2018, most organizations still struggle to achieve "level 5" - where a KM system begins to provide value.

For interpretation purposes (Step 4), having tidy data in a central place is necessary, but not sufficient to render it useful and insightful[vi]. Separate from data tidiness is the question of the richness of the data. Data-richness relates to how many insights one can draw from it, and it is closely related to what Shannon information theory calls information density - a measure of how much information a stream of data can possibly contain. As KM capacity improves, so too does the organization's ability to take full advantage of richer data sources.

5 Let's not split hairs about "the cloud:" If most data is on an office server, and desktops and laptops use a shared drive, and everyone can access a general larger data repository in some way, you've met the criteria for level 3.
6 Indexed versus searchable: Indexes list file **names** and you can only search them by name and date. Windows explorer used to only use indexes. "Searchable" means that the inner contents of every file are also indexed, and there is more metadata available, such as the author on the team and the project it relates to.
7 Alignment between KPIs and strategic questions is what's critical at level 6. This takes years of iteration.
8 Nick Hamlin adds another step in here: KM systems should be designed from level 6 onward to generate labeled training data sets deliberately, as a consequence of their built-in workflows. This is the single greatest hurdle to learning from data at scale. People don't realize how hard it is to create good training sets until after they've been sitting on years of raw collected data.

Figure 3. Data richness:

international development — tech world & the web

rich data

Spreadsheets < SMS < story/text < audio < video

Some spreadsheets lack the data richness needed to answer a question. Some sources of data are too reductive[9] to provide enough information to answer questions about the root causes of complex social problems, or even describe events in sufficient detail. As shown in Figure 3, a large messy data set (of mostly text) can be cleaned and transformed into a far more reliable signal, and used to answer a key question. The same is impossible with just a spreadsheet of indicators - it lacks sufficient information density. The role of statistics, data reduction[10], computer algorithms, and machine learning[11] is to evaluate rich data and structure it to provide tidy, reduced data sets that best answer key strategic questions. To understand the world and measure change, you need *both* indicators and rich data that can be mined[12]. These are combined to produce knowledge. The system that supports this KM process is the knowledge management system (KMS).

Step Five is about taking action based on what the information and data is telling you. Information has no value if no one sees it and learns from it, and KM doesn't matter unless an organization acts on what it learns. For what is learning other than knowledge being put to work? Ultimately, most organizations fail at Step Five, according to Dennis Whittle - founder of Feedback Labs - because they don't build measurement and KM systems that necessitate timely follow up, or that "close the loop" with the people, communities, and programmes they measure.

What follows are some of our observations on how KM systems can help facilitate better, more timely, insights at lower cost; increasing the chance of leaders taking action.

9 For example, "reductionist" data would be a statistic showing the percent of students who attended class. A full data set would show which person was present on what day, and include names, so that this data could be reinterpreted and augmented by others. If given a list of names alone, it is now possible to predict the age, sex, and ethnicity of the group - even if these were not originally recorded. But nothing further can be done with the statistic alone. The principle for making data useful, from Jake Porway (founder, DataKind consulting) is to never assume people will use the data they way we think they will - https://www.oreilly.com/ideas/five-principles-for-applying-data-science-for-social-good.
10 Statisticians most often apply a technique called "principal component analysis" to reduce the complexity of data; data scientists use machine learning to train models that classify data to achieve similar ends.
11 A statistician and a data scientist answer different questions. The data scientist builds models to discover and explain patterns in the world about "what happens." A statistician applies tools to explain "why things happen," and decides whether these patterns are consistent enough to form the basis of policy (based on statistical tests of significance). Both must clean their data, though a data scientist typically requires a larger sample of data to detect useful patterns.
12 In a prior review of Peace Corps, we noted that despite the agency's sophisticated system for indicators tracking all forms of outputs, these mostly served to provide outside stakeholders with metrics and fulfill contractual reporting requirements. Most country-level managers preferred to read all the narratives provided in reports from volunteers in making decisions about strategy and project redesign, and look at the indicators to confirm their conclusions.

KM design guidelines

Automate data transfers and transforms

Using a wider variety of data sources and automating one's interpretation is now feasible. In 2018, new algorithms and data products from Amazon, Google, IBM, and Microsoft allow organizations to batch convert images of words (even handwriting)[vii] into text, translate documents, transcribe audio and video into text, and extract faces and emotions from images. All of these are examples of extract-transform-load (ETL) pipelines - a necessary part of creating usable data out of vast amounts of available raw or "wild" data. Other types of algorithms can now detect and interpret patterns, make predictions, and summarize. As an example of these, "deep learning" algorithms have been writing the sports column in most newspapers[viii,13] for several years now. Nearly all commercial websites use algorithms like these to suggest products to consumers and detect inappropriate content. Even Keystone's own FeedbackCommons.org uses artificial intelligence (AI) to suggest key findings from relationship surveys, feeding the charts and narratives into its "robo-reports."

Both humans and AI will be "consumers" of data

KM design should optimize for information being read by both people and computers. Their needs differ: people prefer narrative and summary charts of trends. In contrast, AI requires full raw data sets in standard formats[14]. In this context, Excel documents have the greatest discrepancy between what a human sees and what a computer "sees"; JSON[15] objects appear to be essentially identical to both humans and computers - making them the better choice, and the most popular data format in use today. For an extensive review of interoperability best practice, refer to Keystone's USAID white paper on the subject[ix].

Humans are really good at consuming information and forming conclusions and summarizing it as key insights - but they introduce many biases into these interpretations. For example, **recency bias** means that a person is more likely to regard instances they seen recently as unduly prevalent. There is also **courtesy bias** in both numbers and text - where people give inflated scores or give tactful, artful feedback, respectively. People are also biased by their personal experience, called **confirmation bias:** people are likely to disproportionately weight information that is consistent with their preexisting view, and may ignore (or overlook) information that contradicts it. The most problematic bias of all for foundations and philanthropy is **survivor bias**[x]: nobody ever publishes examples of failed ideas and bad investments. An effective KMS must capture both examples of successes and failures in an obvious way[16]. Despite these biases, many institutions rely heavily on people to extract meaning from large amounts of data, and tend to disregard rich data sources, because they require the most time to explore, and lead to biased impressions when only small samples are examined.

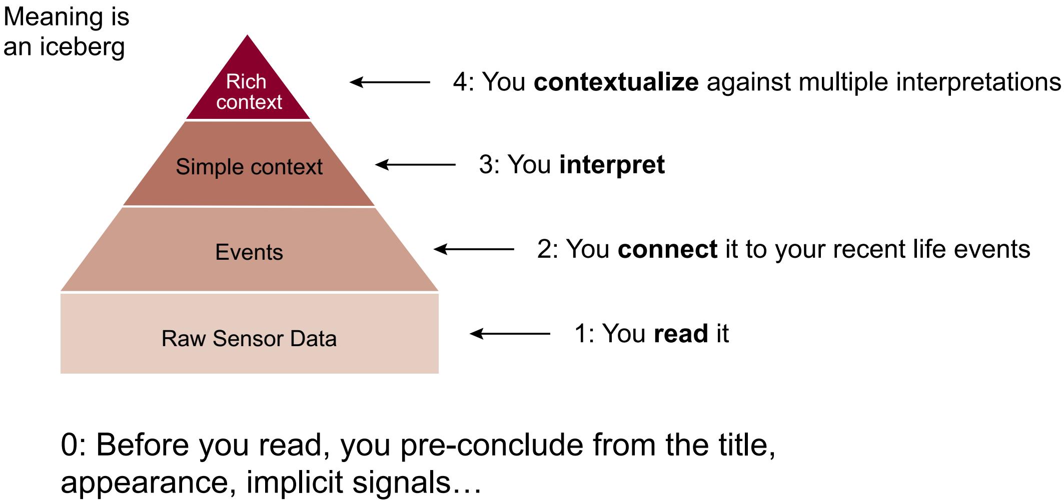
13 Heliograf wrote 850 articles for Washington Post in 2017, tracking sports and local elections around the country. This allowed the paper to cover a broader swath of American politics at no additional cost. https://digiday.com/media/washington-posts-robot-reporter-published-500-articles-last-year/

14 JSON (javascript object notation) is the best and cleanest format, but XML and CSV are acceptable fallback options. CSVs are prone to breakage and cannot handle nested and hierarchical data. XMLs are harder to parse and have no standard way of representing data structures like JSON does. Excel documents are the worst, because computers read them differently (unicode breakage) and different programmes read the same file differently; large parts of the "meaning" in excel documents is not machine readable - such as different font sizes, box colors, and borders.

15 JSON stands for JavaScript Object Notation. This standard "lightweight data-exchange" format is used in the tech world and is more versatile than other forms. It supports nested data (imagine a spreadsheet where cells contain smaller spreadsheets inside them), unicode (can represent all symbols in existence), retains the original structure of data, and can be easily stored in databases, or interpreted as instructions by code.

16 The US clinical (drug) trial registry - https://clinicaltrials.gov/ct2/about-site/background - is an example of a system to counter survivorship bias in medicine. Since 2000, it has logged over 280,000 clinical trials that were initiated. Once the FDA required drug companies to register any drug trial for its results to be considered as part of FDA approval, the science world was able to see how often these trials are undocumented failures. Few trials ever receive mention in science journals, but at least the name of every treatment tested is now recorded for posterity. Companies can avoid repeating the same mistakes.

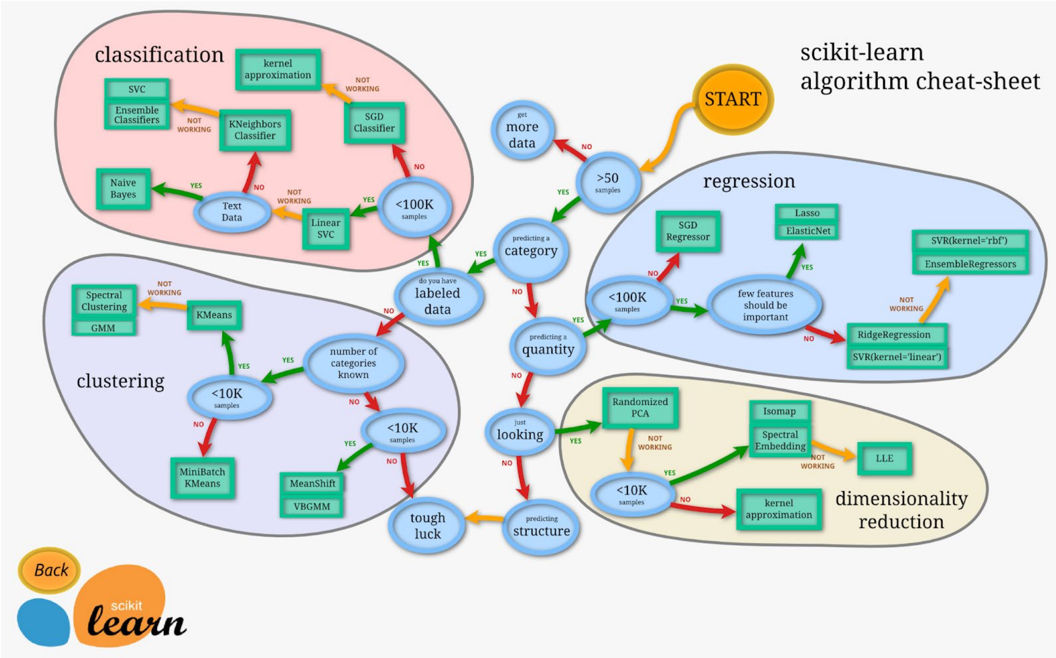
Figure 4: The "iceberg" of constructing meaning from data.



Computer programmes and AI also introduce their own biases, but they tend to be complementary to the biases that humans inject. An algorithm can read millions of documents in minutes and summarise them, but these summaries are prone to missing nuance in language that a human would catch. Language is not math, and is full of exceptions to rules, but recent "natural language processing" (NLP) tools have largely solved these problems in most of the languages in use on the Internet today[17]. Computers interpret text or numbers according to models, and models, like human staff, can be "trained" in a variety of ways[18,19]. The rule of thumb here is that "all models are wrong, but some are useful." Computers provide accurate and powerful answers when the model is appropriate for the question - further emphasizing why frameworks and assumptions sit atop the KM design hierarchy. The following is a visual of the scope and variety of training approaches that data scientists use to extract meaning from information:

17 Primer.ai recently assembled a massive data set with millions of science papers and related news articles and used it to train a computer to write and update Wikipedia entries for the world's scientists. It effectively doubled the useful knowledge overnight, adding 40,000 notable scientists. Source: https://blog.primer.ai/technology/2018/08/03/Quicksilver.html
18 For example, algorithms encode the racial and gender biases inherent in the training data. See: Facial Recognition Is Accurate, if You're a White Guy and How to Fix Silicon Valley's Sexist Algorithms, and Bias In Maternal AI Could Hurt Expectant Black Mothers.
19 Detection algorithms are only as unbiased as the data we use to train them, and human judgment often goes into creating the training data sets.

Figure 5: Common types of machine learning.



These models can answers such questions as: What are the most often expressed (and unmet) needs of people in a community?[xi] What is the most profitable item to feature at supermarket check-out, if a hurricane is about to hit? (answer: Pop-Tarts[xii]). What are the patterns in domestic abuse cases that could help society prevent it?[xiii] Or to stop sex-trafficking?[xiv] In summary, when designing a KM system, one should ensure that useful, relevant information flows into it and that it can be understood by both humans and computers in the formats they need it; avoid losing data by storing only an interpreted summary (reduced form) of it; make information useful by allowing humans to access it in the context of doing their work, and being able to see a summary of it instead of turning on a firehose of information. And finally, continuously compare the workflow, the information flow, and the key questions to ensure they are aligned.

Integrate KMS within workflows

Systems for organizing knowledge should be built ***into*** - not on top of - existing workflows and procedures. If someone has to regularly copy data out of whatever programmes or systems they use into some external KMS, the system will fail to capture information or provide knowledge. This, incidentally, is one of the major added costs into the US healthcare system; with dozens of competing proprietary record systems and no national, standardized records system - healthcare consumed 17.9 percent of the US economy (GDP) in 2016[xv].

Most hospitals employ more than three parallel electronic record management systems (EMS), functionally analogous to a KMS. It is usually possible to automate data transfer where necessary, ideally with simultaneous additional data **transformation** at this stage. If videos need to be archived, why not also produce and store transcripts in English and the original language at the same time? Thinking about workflows and pipelines as the entry point for a KMS enables project data to feed into analysis. Sometimes, workflow improvements can vastly improve productivity. Going "paperless" is one example - where avoiding the need to record information on paper accounted for some of the US GDP growth in the 1990s, according to former Chairman of the US Federal Reserve, Alan Greenspan[xvi] .

Here are some important questions to ask in redesigning a workflow to automatically feed knowledge generation:

1. [**do we need it**] When considering how each legacy document fits into a workflow, ask yourself: is this required by law / contract / policy compliance / reporting? Can we get this information another way, or from a different source? Is it redundant?
2. [**does it need to be paper**] If this source document is collected on paper, could it be turned into a webform? (there are survey options that support both paper and web with merging possible) Could it be a mobile / tablet app?
3. [**who reads it and why**] For the data you collect, why do you collect it? Who is the audience? And what are they doing with the information?
4. [**completeness**] Is all the data of this type stored in the same place? Can all of that data be accessed, if needed? Avoid spreading your data across many devices.
5. [**consistency**] If data comes from multiple sources, are all sources using the exact same data structure? If recording trip reports or evaluations in MS Word, start by creating standard data tables with the same headings in each column. Treat the sections of the report as immutable. Put them in the same order and keep the exact same wording in headings. Use the "h1,h2,h3" style headings in Word and Google documents instead of changing font sizes, as these are machine readable.
6. [**analysis**] What data is needed for each KPI, or to answer key top level questions?
7. [**frame of reference**] When formulating questions, aim to have at least a few questions that define change in terms of S.M.A.R.T.[20] goals with a frame of reference that defines "typical" amounts of success, so that you can show whether you are on track with expected results. Some questions are purely descriptive in nature and cannot be benchmarked, but most questions can be. Does your organization know about sector indicators and standards, even if you don't follow them (yet)?

20 SMART goals are Specific, Measurable, Achievable, Relevant, and Time-bound.

References

i Extended descriptions of these procedures can be found here https://www.tutorialspoint.com/knowledge_management/models_of_km_cycle.htm

ii https://www.law.com/legaltechnews/2018/04/24/5-steps-for-beginners-to-implement-a-knowledge-management-system

iii https://www.livingcities.org/resources/313-data-inventory

iv Hadley Wickham. Tidy Data. "A huge amount of effort is spent cleaning data to get it ready for analysis… The advantages of a consistent data structure and matching tools are demonstrated [herein]." https://vita.had.co.nz/papers/tidy-data.pdf

v Explained here https://medium.com/airbnb-engineering/scaling-knowledge-at-airbnb-875d73eff091. Their system allows many teams to interpret the same data. Teams can author and review each other's work and present competing graphical interpretations. The system stores each version of an analysis separately and enables others to adapt and extend these analyses/interpretations. Most of the analysis and visualization tools are cutting-edge yet reusable. Airbnb recognized this as the key step in solving their "knowledge scaling" problem. Specifically, poorly designed research environments manifest as a "knowledge cacophony" full of parallel analyses and data sets, where teams only read and trust their own research and don't bother validating it with others. This accurately describes the current global state of non-profit knowledge, and Airbnb's approach would be a useful solution.

vi For a longer write up of this issue, see https://chewychunks.wordpress.com/2013/09/03/evolution-international-development/. For a competing view on the limits of big data, see https://www.ted.com/talks/tricia_wang_the_human_insights_missing_from_big_data

vii A worthy demonstration is found at https://cloud.google.com/vision/

viii https://www.wired.com/2012/04/can-an-algorithm-write-a-better-news-story-than-a-human-reporter/

ix http://keystoneaccountability.org/interoperability-white-paper-released-oct-16-2018/

x https://en.wikipedia.org/wiki/Survivorship_bias

xi https://chewychunks.wordpress.com/2016/10/07/a-complete-ngo-ecosystem-map-for-kenya-2010-2013/ and https://chewychunks.wordpress.com/2016/10/10/phylogenetic-tree-ngo-aid/

xii https://www.nytimes.com/2013/06/11/books/big-data-by-viktor-mayer-schonberger-and-kenneth-cukier.html

xiii http://www.bigmountaindata.com/

xiv https://www.usnews.com/news/articles/2015/01/14/how-big-data-is-being-used-in-the-fight-against-human-trafficking

xv https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html

xvi https://www.brookings.edu/research/the-internet-and-the-new-economy/