HANDLE WITH CARE:
10 Steps to Good Data Stewardship

IAB Data Best Practices

FEBRUARY 2014

© 2014 Interactive Advertising Bureau

THE IAB DATA COUNCIL HAS DEVELOPED THIS DOCUMENT.

The following IAB member companies contributed to this document:

Acquireweb
Appsavvy
Cars.com
eXelate
Facebook
Google, Inc.
Krux
Mirror Image Internet
Pandora
PointRoll
SAS
Yahoo

Acknowledgements:

The IAB would like to thank the members of the Data Council for their time and effort on this, our organization's first Data Stewardship Best Practices. Special thanks go to Dan Blasko, Cars.com; Brian Buizer, Pandora; Michael Burke, about.me; Paul Cimino, Deep Data Solutions; Dan Johns, PointRoll.com; Khurrum Malik, Facebook; Ramsey McGrory; Ana Milicevic, SAS Institute; Alfred Gadbut, AcquireWeb Inc.; Tom Walsh, Mirror Image Internet. Thanks also to Nathan Abse (Consultant to IAB-NY) and Matthew Schumpert (Datameer).

The IAB lead on this initiative was Patrick Dolan. Contact Patrick@iab.net to comment on this document.

IAB Contact Information:
Patrick Dolan
EVP & Chief Operating Officer, IAB
212-380-4727
Patrick@iab.net

About the IAB's Data Council:
http://www.iab.net/member_center/committees/councils/data_council.
This Document Can Be Found on the IAB Website:
http://www.iab.net/datastewardship

Contents

Executive Summary

Data Stewardship—the collection, management, use, storage and disposal of data—is becoming an increasingly important part of any organization's operations. Maintaining ethical and efficient control over an ever-increasing data flow is also becoming a more demanding and necessary part of successful companies.

As the digital industry has grown and matured, digital publishers and advertisers have become increasingly dependent on data to drive their businesses. Interactive advertising's very foundation is data created and exchanged by online consumers about their interests and intentions. That data is essential to guiding the creation and serving of relevant messages. Increasingly sophisticated techniques—using more and more data—are improving the power of publishers and advertisers to reach specific audiences through their computers, tablets and mobile devices—offering ever more pertinent words and images to assist consumers and businesses in their purchasing decisions. Throughout the process of making messages more relevant, good data stewardship is increasingly indispensable—and increasingly challenging. These converging trends are why we created this document.

In online advertising, good data stewardship is crucial to managing an especially unwieldy supply chain, composed, or at least communicated, entirely in the form of data. "The supply chain by which digital advertising is created, delivered, measured, and optimized is so porous and perilous that it jeopardizes consumer trust and business growth," IAB president and CEO Randall Rothenberg wrote in a column published as we release this document. "The risk is so severe that the underlying innovativeness of the Internet itself is in danger of grinding to a halt—unless the interactive advertising industry agrees to police its own precincts and root out the malefactors." (For more please see: http://www.businessinsider.com/iab-randall-rothenberg-supply-chain-2014-2#ixzz2sYPv8HQj.) In fact, the greatest part of policing our own precincts is to engage in good data stewardship.

Our first goal in publishing this document is to help our colleagues in the industry by providing a guide to current data stewardship practices, organized into 10 steps that walk the reader through to best practices in the field—and warn against some questionable practices.

The 10 Steps to good data stewardship the IAB Data Council identified are:

1. Make your data policies transparent

2. Ensure and regularly verify data quality

3. Ensure data security: Use appropriate tools, training, policies and procedures

4. Ensure data protection: Use appropriate backup, replication, and other tools and measures

5. Ensure contracts are clear and detailed, and made with reputable partners

6. Define a reasonable time period for the retention of data

7. Employ appropriate site-tagging, related maintenance and other procedures

8. Clearly state the value proposition for online behavioral advertising

9. Market your data and its power to help reach specific segments, but do not exaggerate

10. Place useful, anonymous and non-proprietary data in the public domain

Our second goal for this document is to lay the foundation for productive and ongoing conversations regarding best practices for our industry. We hope the document also helps the industry pursue the following objectives:

- Identify additional proper and suboptimal practices in data stewardship in interactive advertising.
- Raise issues, on an ongoing basis, that will help clarify good data stewardship in the industry.
- Raise awareness of data stewardship in the advertising ecosystem, among businesses and consumers.

Privacy concerns relating to the use of consumer data also are an important facet of data stewardship in the interactive advertising industry. This document, though it touches on some of these concerns, covers a broader spectrum of issues surrounding data stewardship. For the IAB's latest exploration of consumer privacy issues, specifically, please see IAB's work in Public Policy at http://www.iab.net/public_policy/home. Here you will find up-to-date information on privacy issues, as well as more information about the industry's first self-regulatory program in this space, the Digital Advertising Alliance's Self-Regulatory Principles for Online Behavioral Advertising.

Introduction: Data Stewardship and the Future of Interactive Advertising

The following is an overview of why data stewardship matters—and some current and future trends highlighting key challenges posed by the need to manage—collect, use, store and dispose of—increasingly vast amounts of data used in interactive advertising.

Why Data Stewardship Matters

In December 2013, at the height of the holiday season, a data breach at one of the world's largest international retailers—Target—led to the infringement of as many as 110 million consumer credit card data records. Although this instance of credit card information abuse was discovered and set on a course of being contained, at root it was an unacceptable security violation—a large-scale successful deployment of malware and the retailer's failure to catch it.

Participants at all levels of interactive advertising—advertisers, agencies, ad networks and data aggregators, as well as publishers and consumers—were put on notice by the event. It reminded us all that we rely on good, accurate, appropriate—and secure—data to make interactive advertising what it is and can be, more and more relevant and helpful to making consumer decisions. Online and offline data must be collected and maintained properly to create and improve interest-based advertising, personalization, reporting, and ad and website optimization.

Security is just one part of data stewardship. Many more considerations are part of data stewardship: How much time different kinds of data should be maintained prior to deletion, how and when to effect deletion, how to properly tag and track advertising, how to secure data that needs to be preserved intact, among others. With so many considerations in play, good data stewardship requires all levels of an organization to engage the issue.

HANDLE WITH CARE: 10 STEPS TO GOOD DATA STEWARDSHIP

We have developed 10 steps that cover the most important dimensions in data stewardship. Bad or questionable practices to guard against also are covered, where appropriate, under the best practice area they most directly violate. Questionable—and sometimes unethical or illegal—data stewardship practices range from those that can hurt businesses directly by rendering less useful or inaccurate data, to those that hurt consumers by intruding on their privacy or failing to explain clearly how their personal data will be used.

Questionable practices can take a heavy toll on your organization. For instance, poor disclosure on the use of data, or unclear instructions or Web copy regarding data, can steer customers and consumers away from your products or services. In the worst cases, inadequate data security or unclear terms of use can result in compromises to data that can create negative public relations, legal problems, consumer backlash and even government intervention. When this happens to any organization in the online space, it can indirectly affect the entire interactive advertising industry.

The 10 Steps to Good Data Stewardship:

1. Make your data policies transparent. Policies governing the use of internal and external data should be clearly and transparently defined and codified, and those policies should be easy to understand to persons inside and outside of your organization—and, once adopted, these policies must be adhered to.

- **Create clear disclosures and contracts regarding data use and stewardship.** Among the most important elements of data transparency is the disclosure that a site—whether Web or mobile Web—offers visitors and users. Using plain English—taking care to employ language that is readily understandable to laypersons and technical visitors alike—is especially useful to maintaining and building trust online. In any case, companies and organizations should make opting out of tracking or other data relationships an explicit—and uncomplicated—option for users, to earn that trust.
- **Adhere to disclosures and contracts on data use and stewardship.** An organization can offer the clearest disclosure and the most explicit contract describing just how it intends to respect the user's data. But if it does not actually put in place procedures and adhere to these agreements—no matter how well-conceived—good data stewardship will not be maintained. Although this point appears obvious, it bears detailing. Compliance with the IAB Code of Conduct—which closely reflects the DAA Self-Regulatory Principles for Online Behavioral Advertising—is required of IAB members and recommended for all advertisers and publishers engaged online. The Code prescribes ethical and consumer-friendly advertising practices, in disclosures and throughout the online space. Equally important here is the legal dimension: any disclosures made to consumers on privacy policies or terms of use must be strictly adhered to—whether they concern data retention, third-party data sharing, or any other activity involving the data. Under the Federal Trade Commission's Section 5 authority, if companies deviate from their written privacy policies or terms of use, they can become subject to the FTC's jurisdiction.

Bad practices:

Automated manipulation of cookie and other privacy settings in browsers and mobile browsers. *Industry experts report that some companies and organizations use automated means to alter user cookie and other privacy settings in their browsers, particularly in Mobile settings. These entities do this in order to gain access, and to cookie or otherwise track the user and serve messages, as the user engages online. The extent of this bad practice is unknown. In some cases, this bad practice has resulted in legal action against parties that engage in it, leading to expensive consent orders and sanctions.*

Use of Mobile or other log data without clear disclosure or user consent in order to re-target or otherwise serve ads. *Some parties use mobile log data to track users and serve messages (re-targeting), without first gaining user consent. Some—often smaller, less established—companies reportedly do not use the data themselves, but leverage their own trusted, first-party relationship with their users to gain valuable raw mobile log data and sell it to other companies. Other parties reportedly use—again, without consent—other, non-mobile log data, and resell or reuse it for commercial purposes. All such data grabs are bad practices. Remember: all parties should be transparent in their data policies—and using user data without consent is a bad practice and can be against the law.*

Failure to develop or follow standardized data stewardship and use guidelines. *The widespread absence of standard guidelines at many companies about data often results in those companies adhering to only parts of data privacy best practices—perhaps abiding a user's cookie preferences selected on their PC, but meanwhile making use of other data without the user's fully informed consent, such as "raw log" data on mobile devices or IP addresses on PCs and home networks, to target ads to users.*

2. Ensure and regularly verify data quality. Data goes through numerous phases as it is obtained and used by your organization: planning, acquisition, processing, use, storage and preservation, and disposal. Persons who produce, define and manipulate data for an organization should also be considered or defined as data stewards. Although in many organizations, data stewardship falls to a limited number of specialized persons, it is necessary that all who touch the data—often a wider group—act as good data stewards. One primary task in data stewardship is that data stewards set requirements and standards for data used in any part of that organization's work or mission. Along with setting the standards, which vary as needed, data stewards maintain the data, to ensure it is good enough for its intended use. Damaged or inferior quality data must be repaired or augmented, or excluded from use. All organizations that deal in data should draft and circulate clear, written policies regarding responsibility for data stewardship---and ensure that data stewards do their jobs, and the data is maintained to adequate quality.

Bad practice:

Failure to maintain adequate routine data stewardship training and procedures. *Good data stewardship means maintaining data quality—but that requires that data stewards are trained to carry out complex, often-changing procedures. Some organizations fail to provide ongoing training—which is necessary to keep data stewards current, and data high quality.*

3. Ensure data security: Use appropriate tools, training, policies and procedures. Data security, guarding against unapproved use or corruption, is ensured by numerous procedures, including data-securing measures-and the use of appropriate tools. These include anti-malware software, firewalls, encryption, audit logs, and passwords. Tools also can include appropriate commercial data management software. Staying current and flexible in choosing and maintaining the appropriate tools to suit a given organization's data purposes is also a crucial part of good data stewardship throughout the interactive advertising industry. All data—even the least sensitive kind—should be secured against unauthorized modification or destruction. Sensitive data—such as personally identifiable information (PII)—must be especially well-secured. Certain kinds of sensitive data—for example, financial data, health data, or data affecting advertising to children—fall under specific laws that require special handling. For more on this, please see "IAB's Digital Advertising Regulation 101" (http://www.iab.net/public_policy/digitalad101).

Bad practices:

Neglecting to keep anti-virus, anti-malware and other data security necessities up-to-date. *Although most organizations maintain strong data security, others—large and small—fail to keep current with necessary data security procedures, resulting in compromises to data integrity.*

Failure to maintain adequate routine data security training and procedures. *Data security relies not merely on software and other technology, but proper, ongoing training and related resources to remain effective months or years after new programs are implemented.*

4. Ensure data protection: Use appropriate backup, replication, and other tools and measures. Protection from physical loss or unintended deletion of the data is ensured by the use of appropriate backup systems, replication of data and other measures that provide redundancy in information systems used in interactive adverting processes. "Data protection" is a term that can mean different things to different people—it could mean one thing to a consumer, but carry a different definition to an organization in the industry—say, a publisher or data technology vendor. We mean by this term specifically the production and maintenance of backup copies of data to ensure against information loss.

5. Ensure contracts are clear and detailed, and made with reputable partners. Ensure that all relationships that involve other parties are with reputable organizations, and are contractually documented. Ensure that all applicable laws and industry standards are properly followed in your contracts. Vendor contracts and the provisions in those contracts and other agreements should specify all relevant details about the data exchanged, providing parties recourse if needed. Enforce contracts—try to negotiate remedies to problems, but terminate relationships that violate laws or standards.

6. Define a reasonable time period for the retention of data. The appropriate duration of retaining data for later use or potential use—with different types of data held for varying periods prior to deleting or anonymizing—should be determined and disclosed, if possible, in advance. Industry experts note that many organizations tend to retain more data, over more time, than reasonable use justifies—especially considering that most online data documenting consumer behaviors and intentions at a given point in time soon becomes stale and inaccurate. However, the parameters governing the length of time that data should be retained will vary depending on the kind of data at hand (for example, in some instances, old data can be useful to running long-term analytics.) Data retention policies—especially those concerning consumer data—must vary not only according to the organization's needs, but also according to laws and regulations on privacy.

7. Employ appropriate site-tagging, related maintenance and other procedures. Organizations must hold themselves to appropriate tagging practices, because tracking and targeting—kept within the bounds promised by disclosures and contracts, and as required by law and best practices—are cornerstones of properly conducted interactive advertising. Good tagging practices also are integral to good data stewardship, since proper tags are key of ensuring security to the data. If data is not tagged with the correct level of sensitivity or for access by the correct parties, data leakage can occur. For the IAB's latest on site tagging, please see our Site Tagging Best Practices (http://www.iab.net/media/file/SiteTaggingBP_final.pdf).

Bad practices:

Inappropriate tagging. *Attaching certain types of inappropriate tagging can, in less damaging cases, impair performance or, in more damaging cases, lead to undesired data capture or transfer to or by a third party.*

Inappropriate collection and use of mobile phone data—such as mobile phone tags, raw phone log data, and geo-data. *Some companies and other parties have engaged in the inappropriate use of tags of all kinds to track and serve ads to users, and make inappropriate uses of raw mobile phone log data—for instance, the unconsented reuse and sale of phone geo-data to outside parties to help them to re-target users. For more on this subject, please see the Digital Advertising Alliance's "Application of Self-Regulatory Principles to the Mobile Environment" (http://www.aboutads.info/DAA_Mobile_Guidance.pdf).*

Inappropriate use of ad exchanges, such as "cookie-licking." *Some parties take advantage of the security limitations of tagging procedures and ad exchanges by engaging in inappropriately siphoning off valuable data. In "cookie-licking," for example, a party abuses an exchange by skimming off data about users' interests without compensating the publisher or any other party for the data they accessed.*

Failure to remove or deactivate tags that should no longer be active. *Most tags are placed by partner organizations. Once a partnership with another party has ended—whether because a contract expires, or a party engages in unapproved activities, or any other reason—the tags placed by that other party must cease to report data to that organization. Organizations must take care that tags placed by such former partners are removed.*

8. Clearly state the value proposition for online behavioral advertising. Online behavioral advertising—or interest-based advertising—offers advertisers and marketers advantages in lowering the costs and increasing the effectiveness of messaging. Messages are far more effectively and efficiently directed at audience segments with a demonstrated interest in the products and services offered. Consumers also are great beneficiaries of interest-based advertising—they receive messages about goods and services that interest them, and can take advantage of more competitive pricing. Make the facts of this value proposition clear and explicit to users when soliciting consent to use their data. For example, the IAB offers the Privacy Matters campaign website (http://www.iab.net/privacymatters/campaign.php) which offers an outline of how online and mobile user data are used in interest-based advertising—leading to more relevant ads and significant discounts.

Bad practice:

Failing to disclose your targeting criteria. *Non-disclosures about your targeting criteria make consumers uncomfortable. Wherever possible, transparency about how you collect and use data is always the best practice.*

9. Market your data and its power to reach specific audiences, but do not exaggerate. If your data can help accurately direct interest-based targeting to reach highly-focused segments, market your services accordingly. But if your capacities are more limited, your organization should be as transparent in its pitches to businesses as it should be in disclosures and contracts to users and consumers. Some organizations instead try to "pump up" their data with information that's far less powerful than claimed. Some detractors call this weak but overrated data "pink slime."

Bad practices:

Too much targeting. *On the other side of targeting that's too weak (too little targeting) is targeting that's too strong (too much targeting). Your organization has to find the right balance so that your consumers don't find your ads creepy or intrusive.*

Over-segmentation of the data. *Over-segmenting of the data often becomes highly inefficient to maintain and execute at a campaign level.*

Click-fraud. *Various means used to fraudulently boost statistics on an ad's success—or "click-fraud"—must be combated throughout the online advertising space.*

10. Place useful, anonymous and non-proprietary data in the public domain. Data that is non-proprietary and offers a public good, should be considered for publication and maintenance in the public domain, with the caveat that great care must be taken first to strip any such data of potentially identifying information. There are numerous government, academic, and other nonprofit organizations and websites that offer their services for less commercially viable data, since maintaining such data often does not serve the mission of a given company or organization.

Conclusions:

Interactive advertising has profoundly changed in recent years, and continues to evolve. All activities that go into interest-based advertising—from the business end of conceiving, planning, developing and implementing new and improved messaging, to consumer participation in life-enhancing aspects of this space—are dependent on the integrity of online data, first- and third-party.

Maintaining ethical and efficient control over data flow—good data stewardship—is becoming ever more demanding and critical in this industry. Good data stewardship is a cornerstone of ensuring higher quality data, because it provides needed protection against everything from low-quality or inappropriate data, to data leakage and data breaches. In protecting against these problems, good stewardship practices can also earn organizations greater consumer trust and protection from potentially wasteful and innovation-inhibiting government regulation.

The IAB hopes that this document clarifies the foundations of good data stewardship in interactive advertising. The best practices outlined herein go well beyond any one narrow issue. Notably, the data being collected and stewarded is growing so quickly, both in volume and complexity, that all stakeholders—commercial parties and consumers alike—must be aware and invest time and effort into effecting these best practices. We trust our 10 steps offer a clear start on learning what you need to know—whether data stewardship is your specialty, or more likely, only a piece of your data-driven world that we all now participate in.

The IAB believes adherence to these best practices will reduce confusion and increase trust in online users and between businesses and their partners. We hope this document serves as a guide and anchor to all.