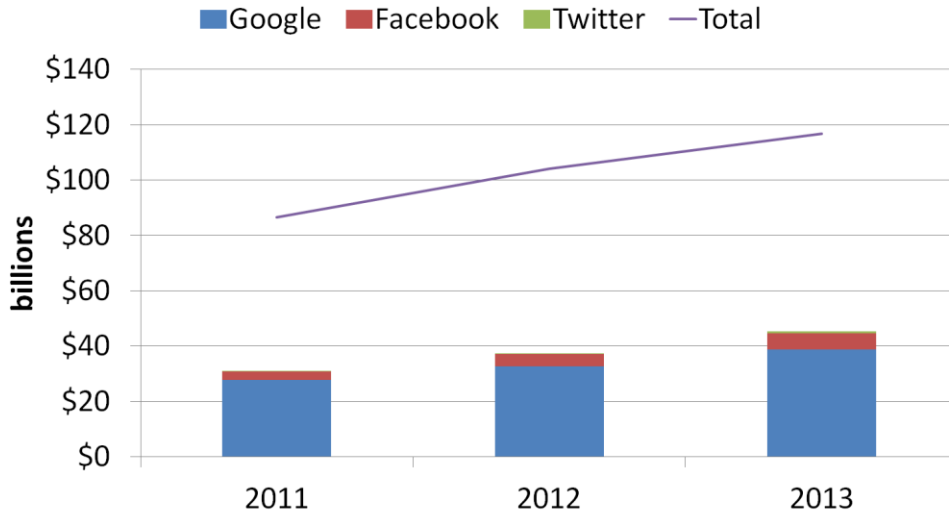


ビッグデータビジネスと程よいWebプライバシー

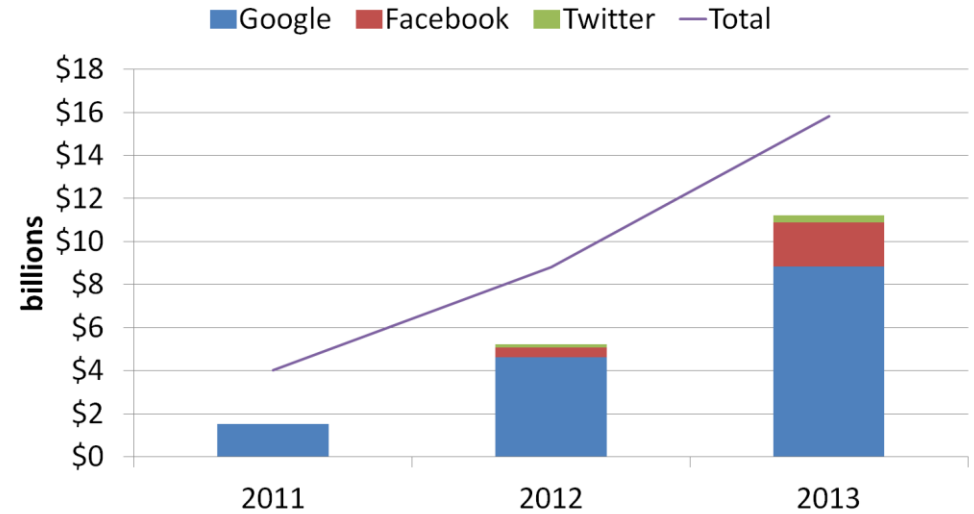
ソフトバンクテレコム株式会社
国際ネットワーク&サービス開発部
吉井英樹

ビッグデータ活用企業

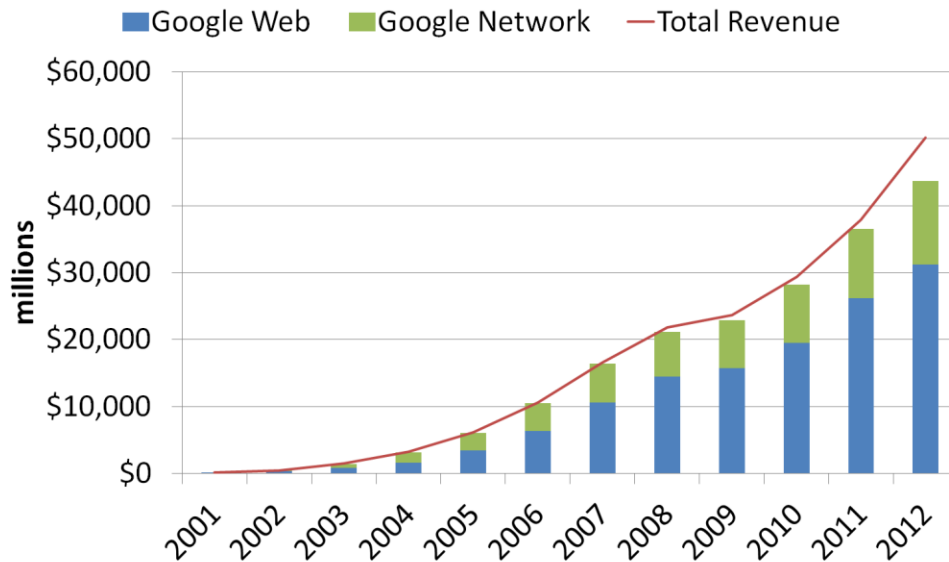
Net Digital Ad Revenues Worldwide, by Company, 2011-2013



Net Mobile Internet Ad Revenues Worldwide, by Company, 2011-2013



Google Web & Network



•2013年、ネット広告市場は11.6兆円（1ドル＝100円）、モバイル広告市場は1.6兆円と予想されている。

•モバイル広告市場で、Google社はグローバルに50%以上のシェアを有する。

•2012年、Google社は4.6兆円の広告収入を獲得した。

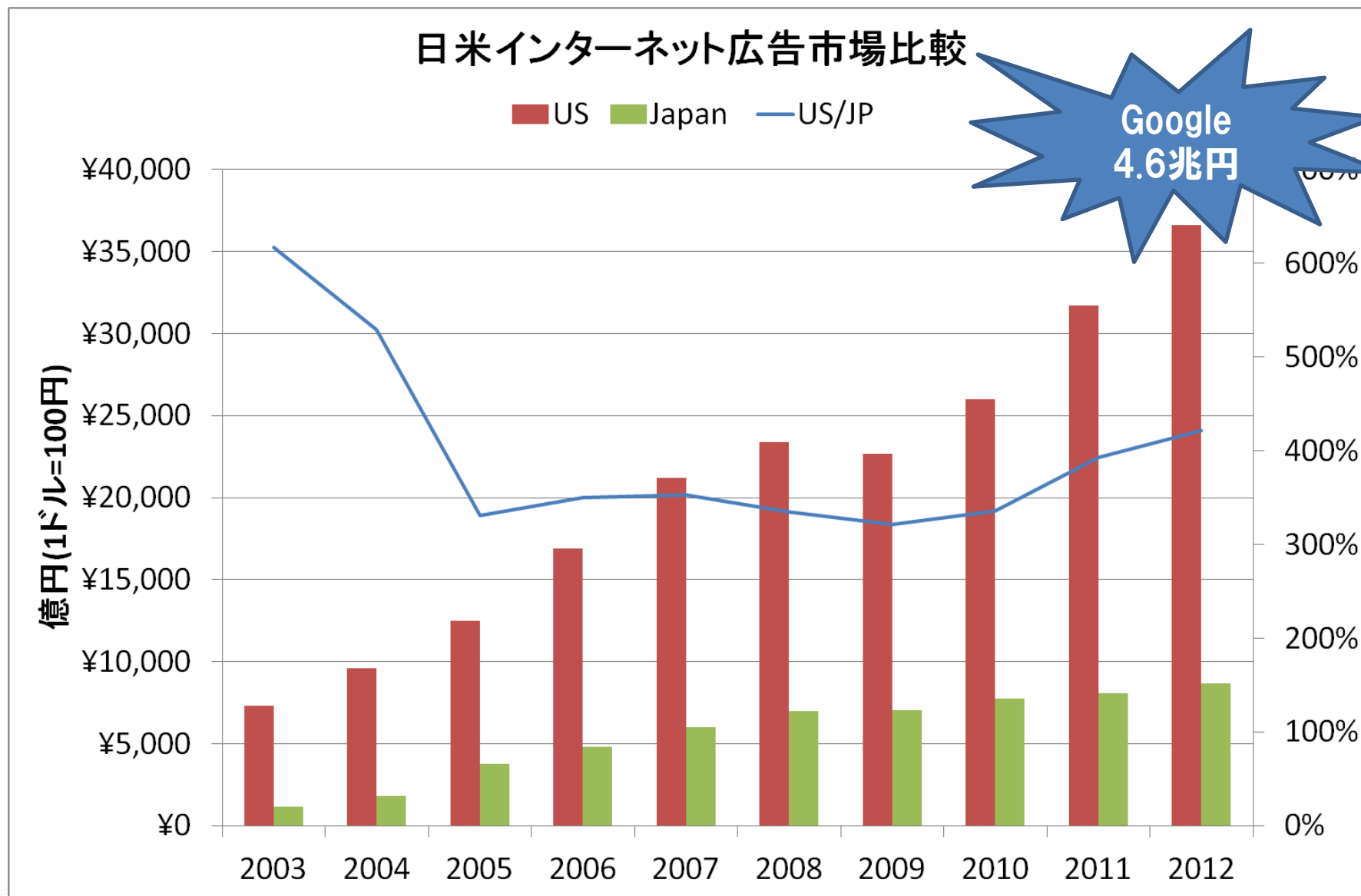
*1Net Ad Revenue is the revenue after companies pay traffic acquisition costs (TAC)

*2<http://www.emarketer.com/Article/Google-Takes-Home-Half-of-Worldwide-Mobile-Internet-Ad-Revenues/1009966>

*3<http://investor.google.com/financial/tables.html>

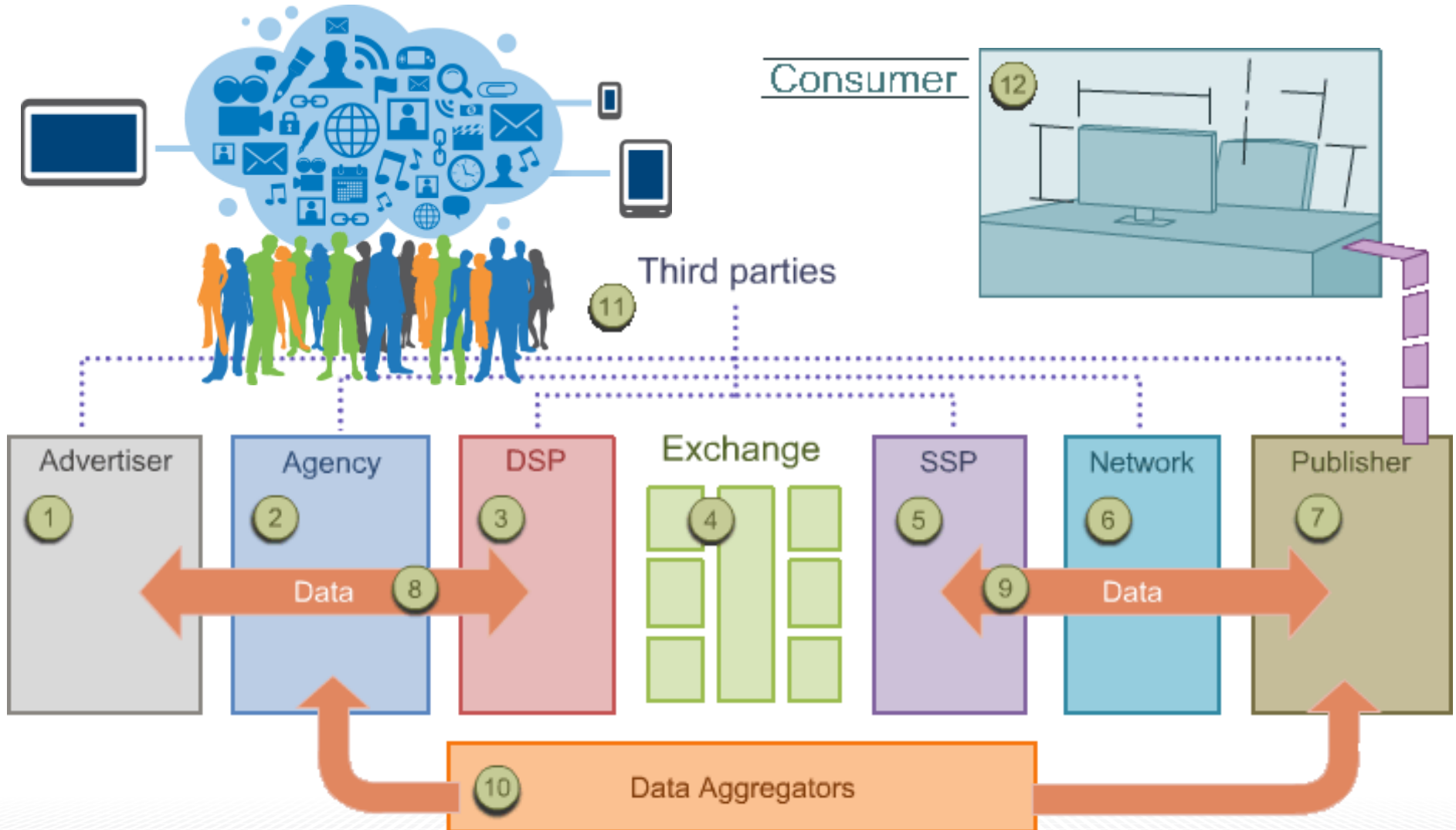
日米広告市場比較

SoftBank



Advertising Ecosystem

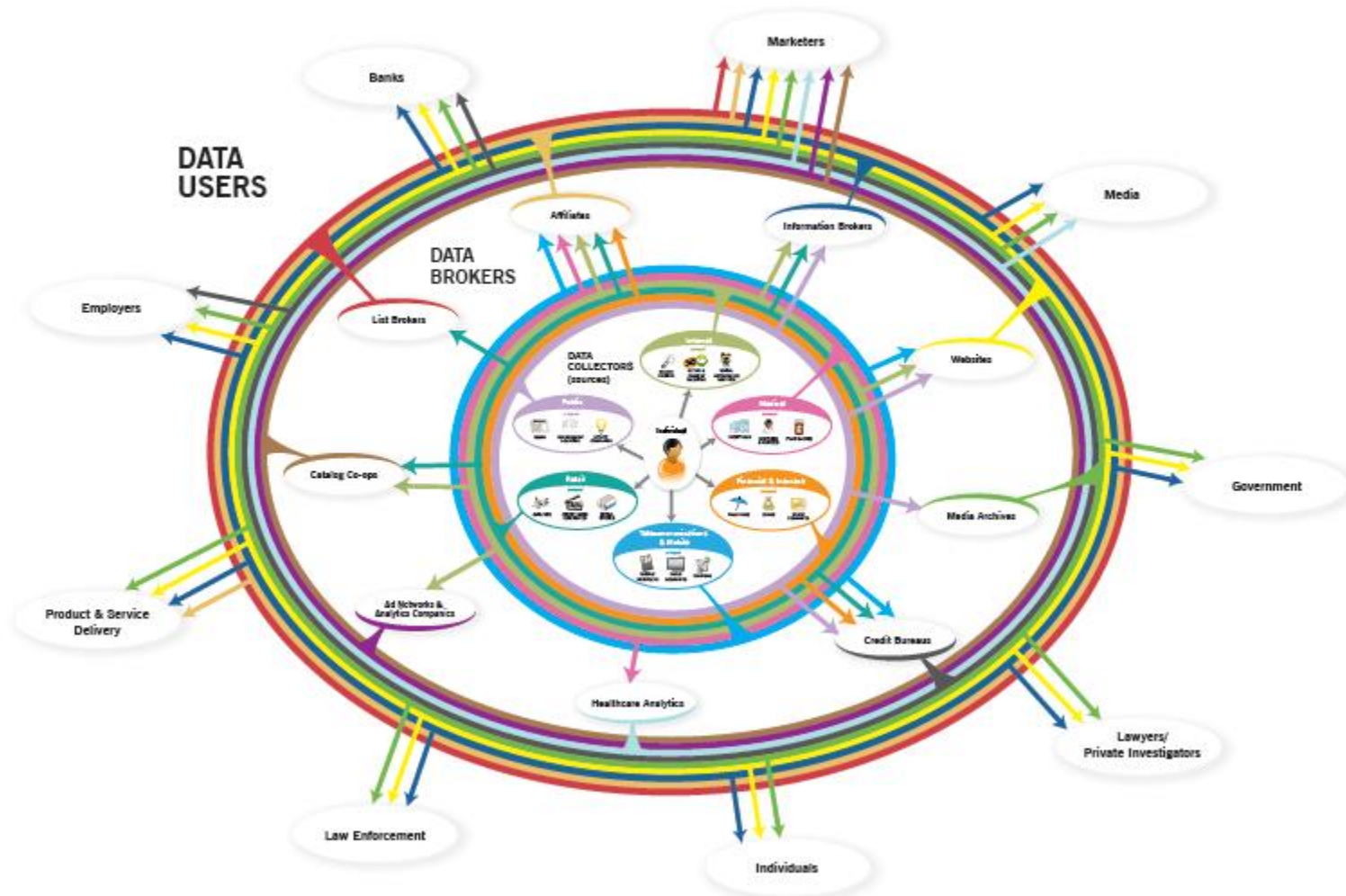
“The right offer, to the right person, at the right time”



Personal Data Ecosystem

個人とデータ収集者、データブローカー、データユーザによるエコシステムを想定

Personal Data Ecosystem

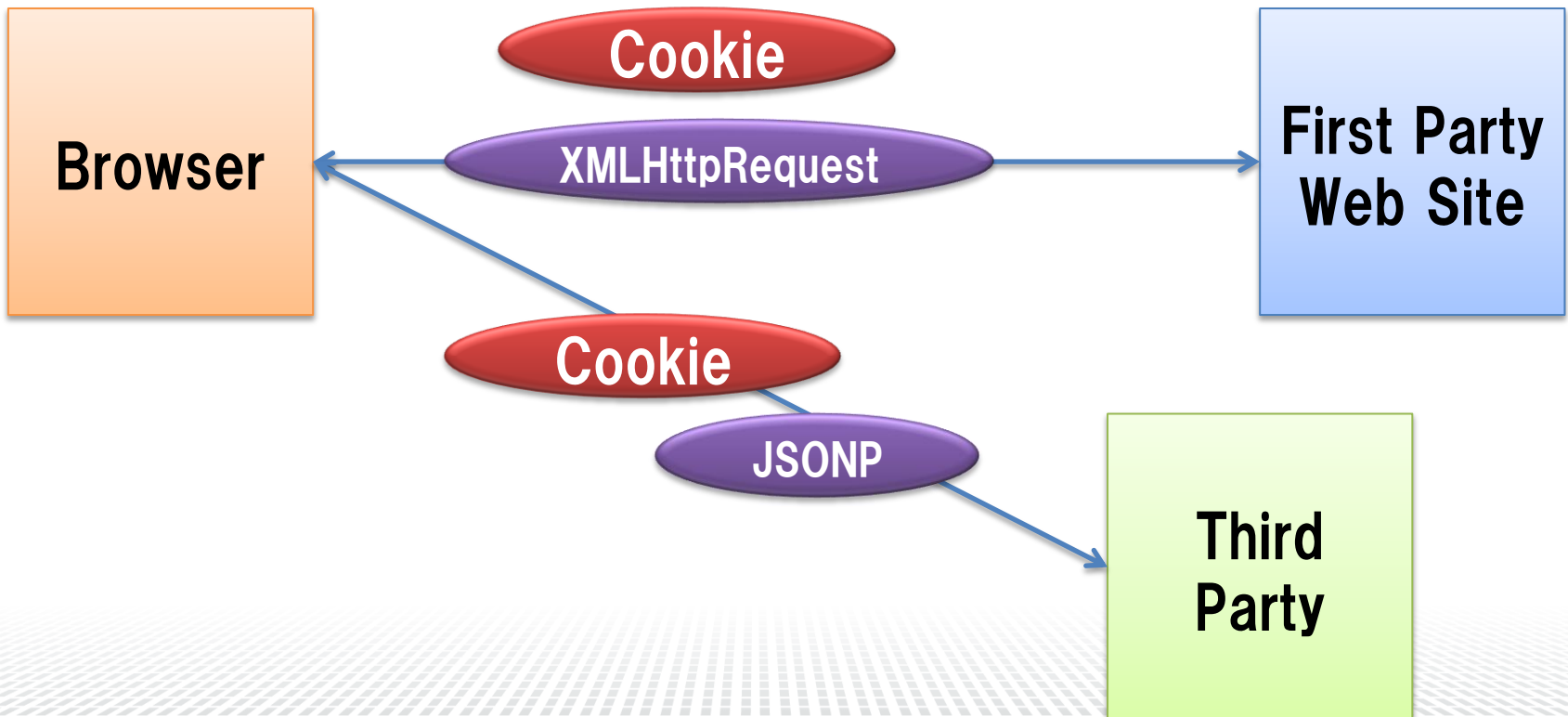


* FTC report, Protecting Consumer Privacy in an Era of Rapid Change, <http://ftc.gov/os/2012/03/120326privacyreport.pdf> 4

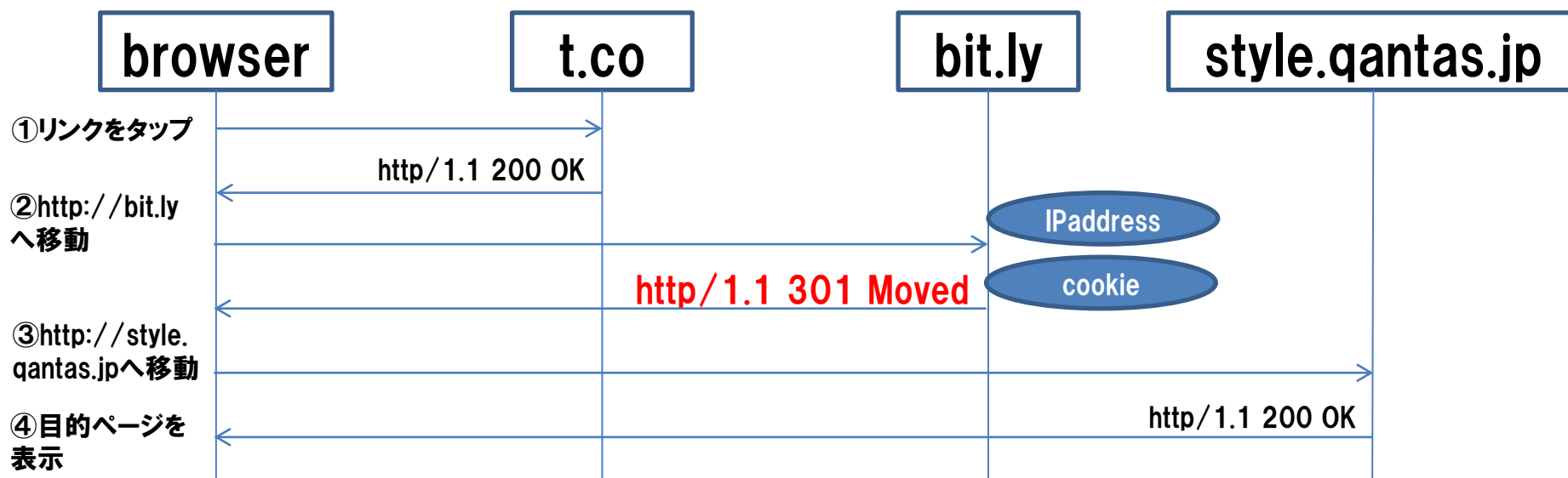
Webを使ったデータ収集

Webの建前と本音

- Same Origin Policy VS JSONP
- First Party Cookie VS `document.cookie [= value]`
- HTTP301 VS 電気通信事業者



Webを使ったData収集① 短縮URLと通信の当事者



・ドメインbit.lyは、Web上での行動履歴の一部を匿名化された個人に係る情報として収集することが可能である。

・bit.ly (HTTPのリダイレクト指示者) は、通信当事者なのか、通信を媒介するものなのかによって情報の取り扱いルールが変わるはずである。

Webを使ったData収集② ソーシャルプラグインとiframe

 @Asahi_Shakaiさんをフォロー

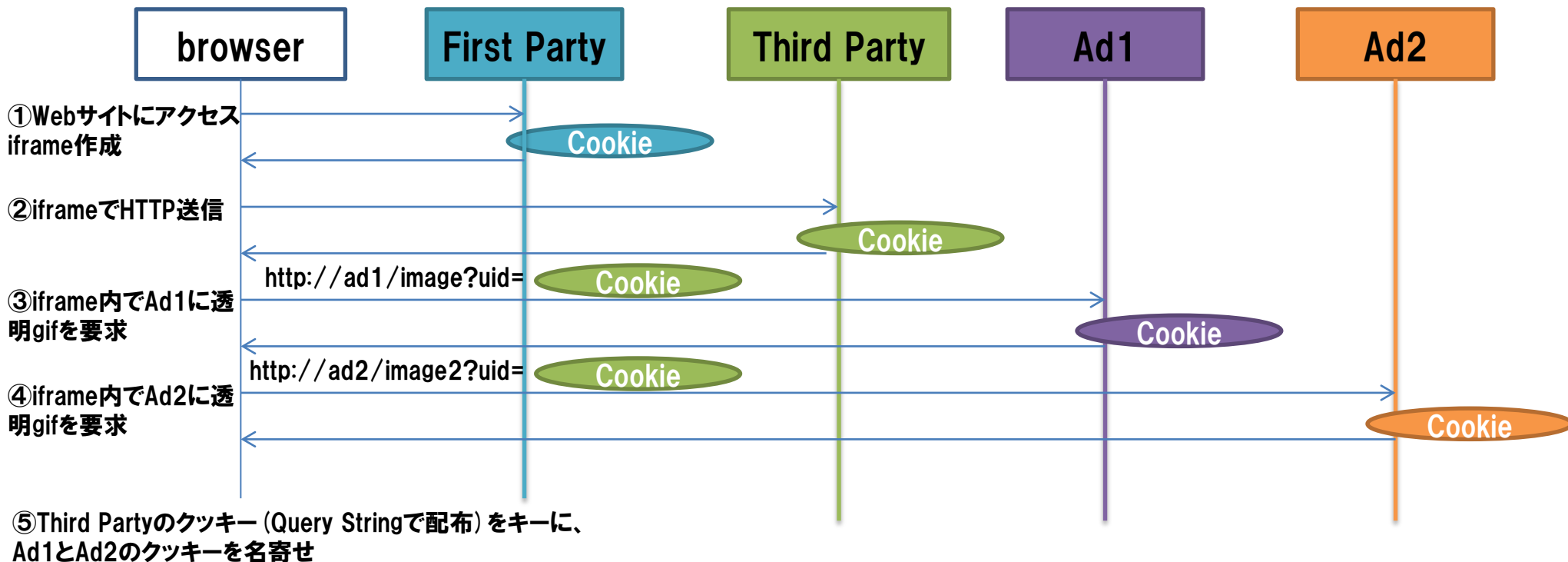
```
<iframe src="http://platform.twitter.com/widgets/follow_button.1371247185.html"></iframe>
```



- ・ソーシャルプラグインを採用しているWebサイトにアクセスした場合、認証トークンが失効していなければ、ソーシャルサイトにアカウント名を関連付けて閲覧履歴が収集される。
- ・ソーシャルプラグインを用いた場合、twitter上での行動履歴及び個人属性を用いたアクセス解析が可能である。

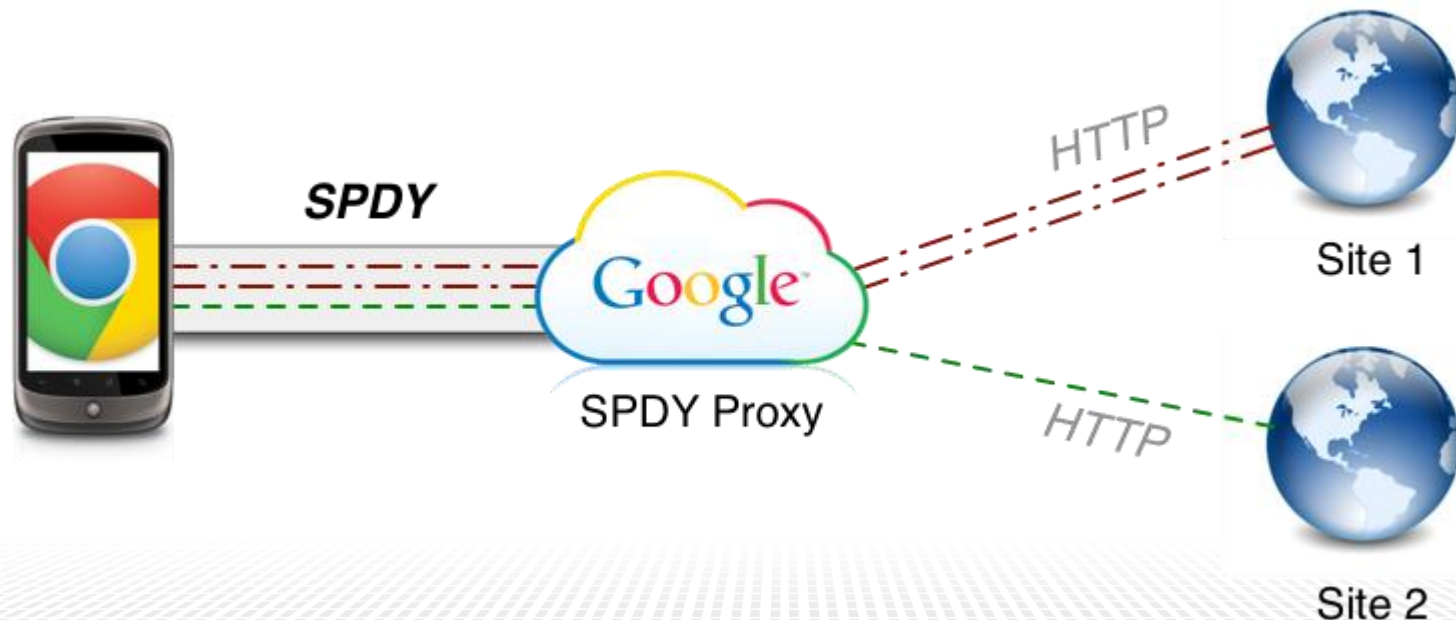
Webを使ったData収集③ Cookieの名寄せ

- Browser側に保存されているCookieを、Query Stringに記載して、第三者に送信することが広く行われている（代表例：Google Analytics）。
- このテクニックを応用すれば、異なるドメインのクッキーを名寄せすることが可能である。
- 一般的には、Invisibleなiframeと画像を使って実現されるが、このようなiframeを提供するのは通信を媒介するものになるのか？ FTCが述べているデータブローカーと呼べる？



Webを使ったData収集④ Google SPDY Proxy

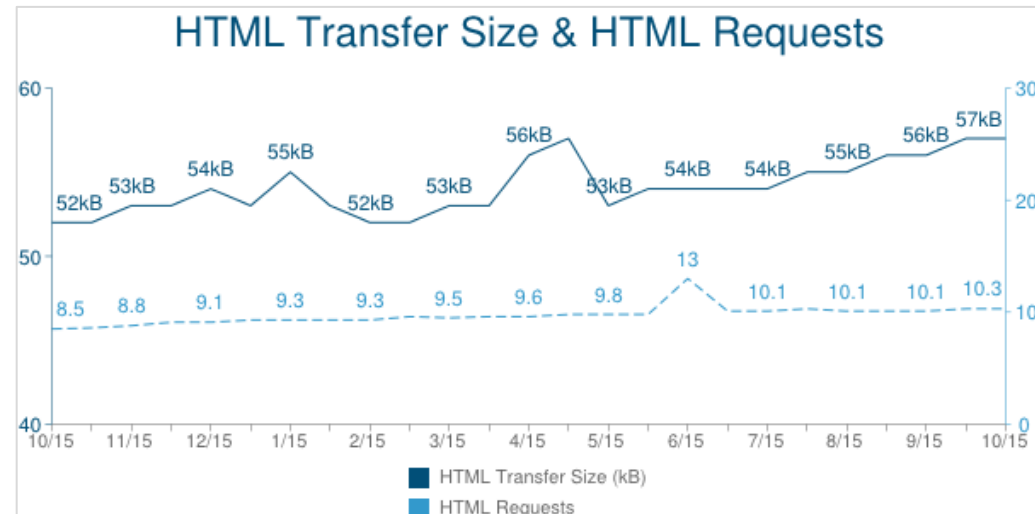
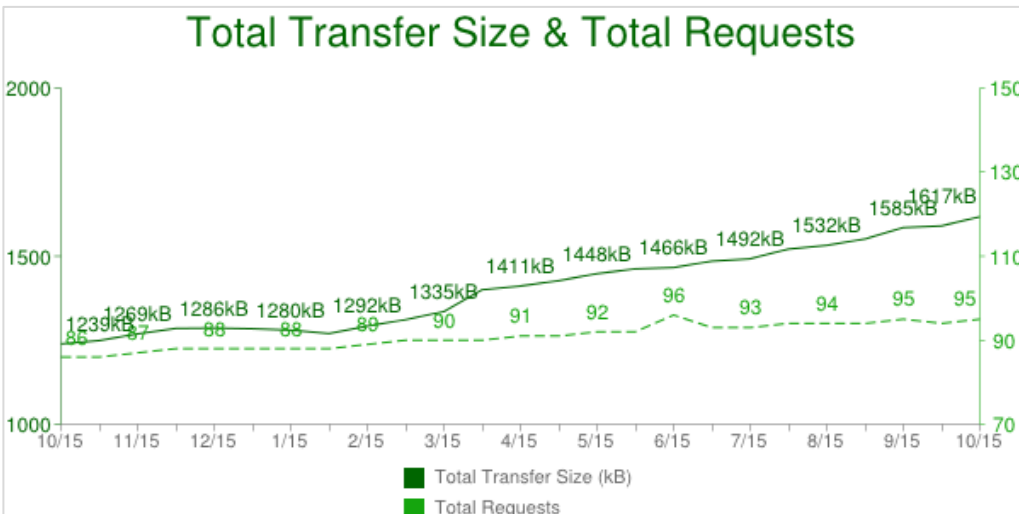
- SPDYとは、Google社がWebアクセスの遅延時間を短くすることを目的に開発中の通信プロトコルであり、現在、IETFでHTTP2.0として標準化が進んでいる。
- このSPDYを用いたProxyサービスは、ユーザにWeb高速化というインセンティブを提供し、グローバルにモバイルキャリアをダムパイプ化する可能性を秘めている。
- 例えば、端末とProxy間はSSLで暗号化されるため、セキュリティレベルが高くなる。一方、現在、キャリア側が提供している青少年向けフィルタリングやブロックングは、Proxy側でしか対応できないことになる。
- 仮に、このProxyサービスが普及した場合、Google社はより多くのデータを収集することが技術的には可能となる。



SPDY Proxyへの追い風 = Webサイトの高度化(複雑化) — SoftBank

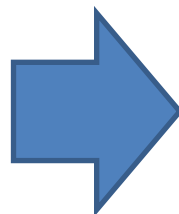
■1つのWebアクセスには、平均で約100のHTTPリクエストを伴う。

■そのうち約10リクエストはHTMLを取得するものである。



asahi.comのあるページのHTML文書構成

情報種別	個数
HTTP Requests	182
Images	102
Scripts	49
Documents (iframe)	13
Cookies	51



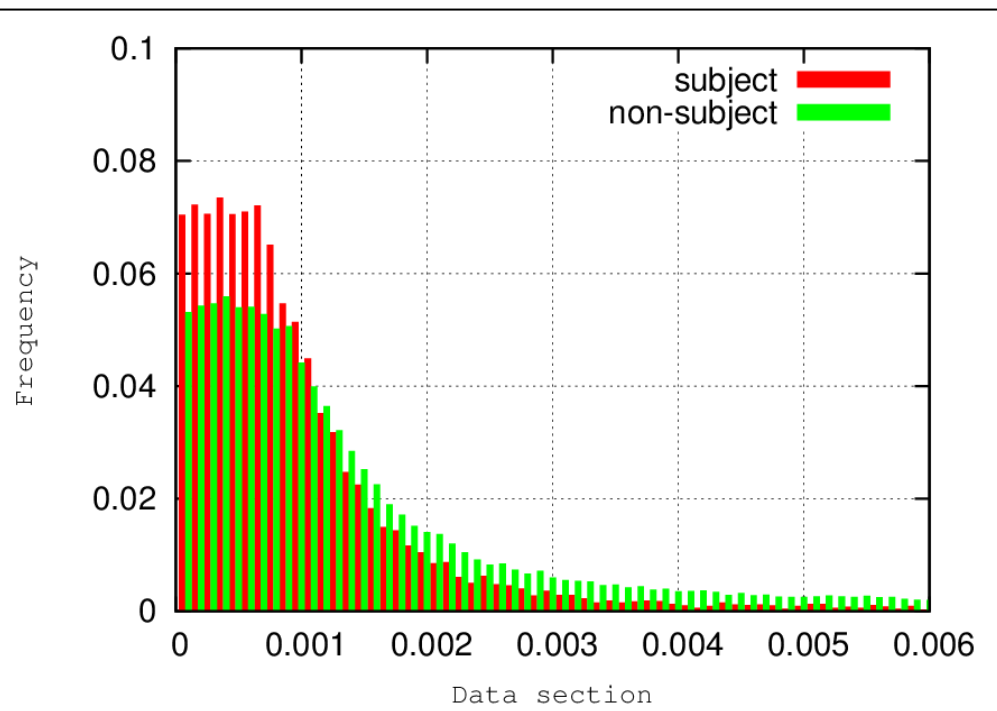
•Webサイトの複雑化に比例して、SPDY Proxyの価値は向上する。
 •Proxy上で情報を制御することを許せば、Privacy上の懸念まで制御できる。

*グラフは、<http://httparchive.org/trends.php>より、2012年10月15日から、2013年10月15日までのトータルリクエストとHTMLリクエストのグラフをコピー 11

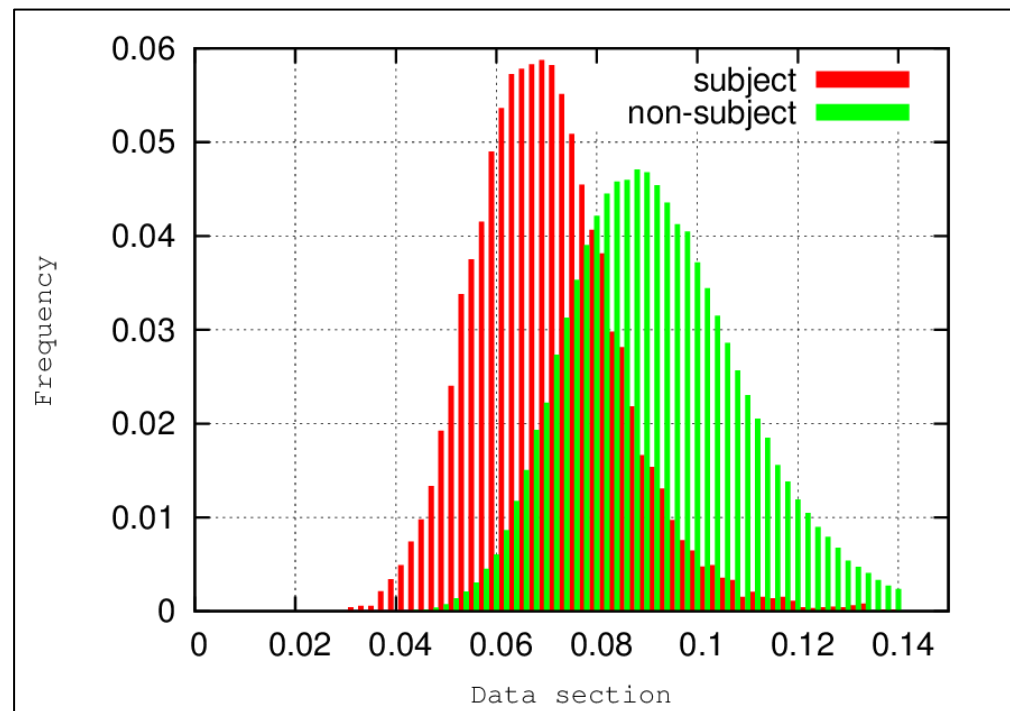
實質的個人識別性

ライフログを用いた個人識別性に係る課題

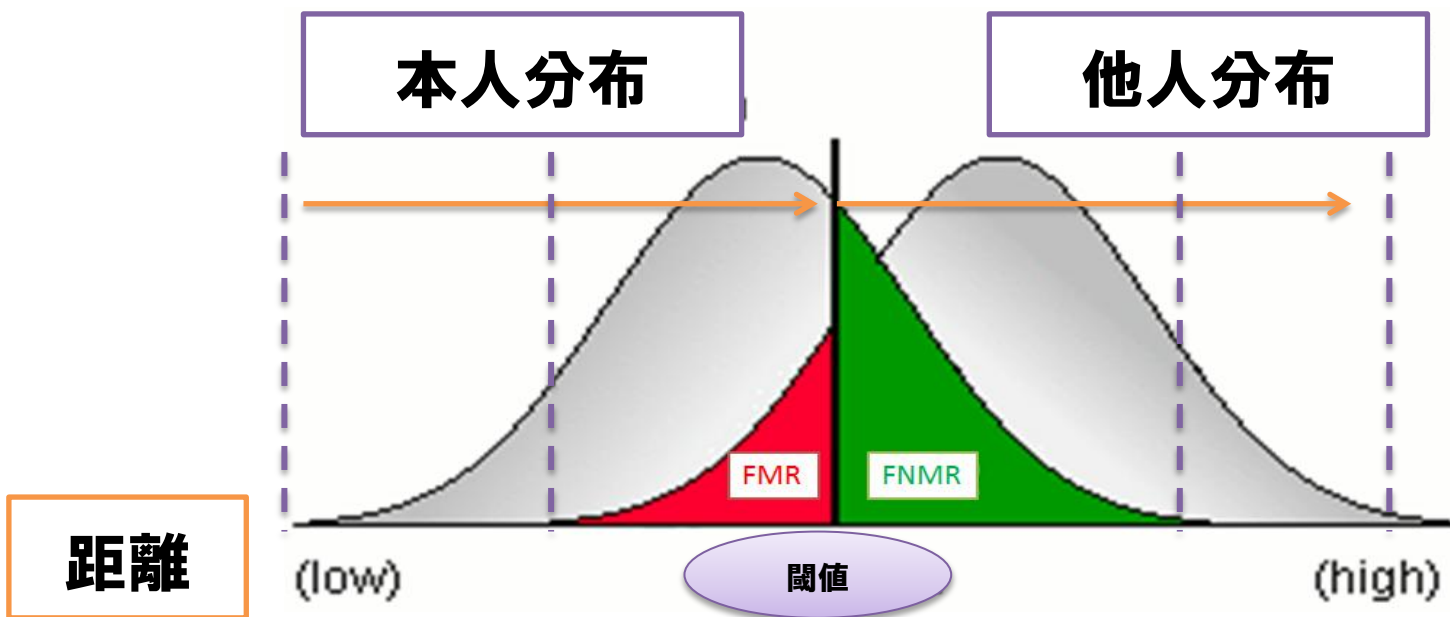
- センサデータ（位置、加速度）や顔画像等から獲得される個人識別性について、FMR (False Match Rate) やFNMR (False Non-Match Rate) の概念が制度面の議論でもなされるべき。
- 他人と間違われる可能性をどう扱うかは、制度、技術両面の課題である。



特徴量Aを用いた本人・他人分布



特徴量Bを用いた本人・他人分布



■ 閾値を最小から最大までスライドさせる

- FMR (False Match Rate) : 他人受入率 (%)
- FNMR (False Non-Match Rate) : 本人拒否率 (%)

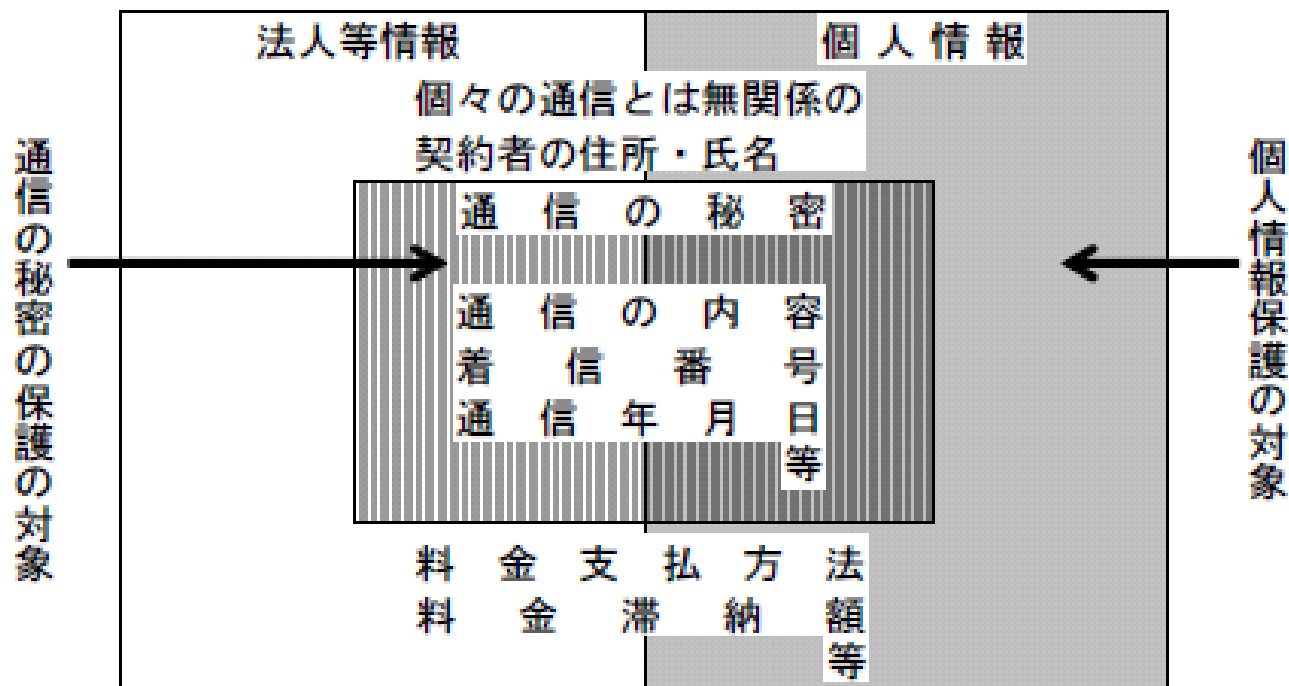
■ FMRとFNMRの散布図を描く⇒ROC曲線

■ 閾値を決めるのは、サービス事業者（とユーザ、すなわち私人）

電気通信事業法と個人情報保護法

- 一方、通信の秘密に係る電気通信事業法第4条その他の関連規定については、通信の秘密に属する事項については、個人の情報であるか、法人その他の団体の情報であるかの区別なく保護されるものであることから、法人その他の団体に関するものも保護の対象となる(下記図参照)など、その対象及び規律の内容について、本ガイドラインの範囲を超える場合がある。

●個人情報と通信の秘密との関係



		電気通信事業者		登録・届出を要しない電気通信事業を営む者*
		登録電気通信事業者	届出電気通信事業者	
電気通信事業法 第四条（秘密の保護）	第1項	電気通信事業者の取扱中に係る通信の秘密は、侵してはならない。		
	第2項	電気通信事業に従事する者は、在職中電気通信事業者の取扱中に係る通信に関して知り得た他人の秘密を守らなければならない。その職を退いた後においても、同様とする。		

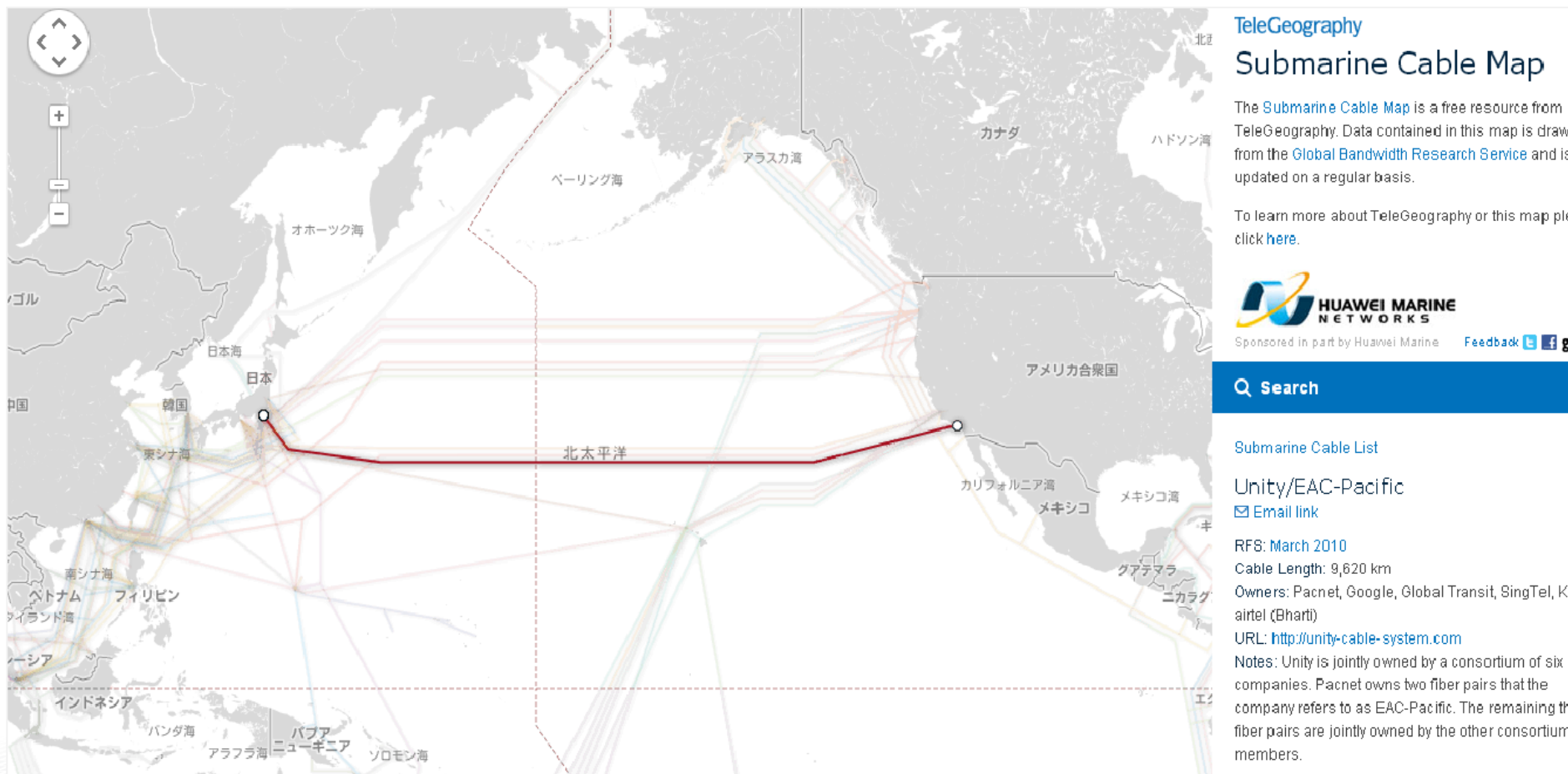
疑問：

グローバルビッグデータ事業者は実質的に、秘密の保護第2項に係る情報を活用したサービスを実施しているのではないか？

*「通信の秘密について」(<http://www.win-cls.sakura.ne.jp/pdf/35/02.pdf>)より、以下抜粋
 Web サイトのオンライン検索(狭義のポータルサイト)の提供や電子掲示板(BBS)の開設は、他人の需要に応ずるためのものであるから電気通信事業に当たる。もっとも、これらのサービスは「電気通信設備を用いて他人の通信を媒介する電気通信役務以外の電気通信役務を**電気通信回線設備**を設置することなく提供する電気通信事業」であるため、原則として同法の適用を受けず、従って登録・届出も不要ではあるが(164条1項3号)、検閲禁止及び通信の秘密保護の規定に限っては、「電気通信事業を営む者の取扱中に係る通信について」適用されることが定められている(同2項)。つまり**同法にいう「電気通信事業者」であるか否かにかかわらず、これらのサービスには通信の秘密の保護規定が適用される**ことに注意が必要である。

電気通信回線設備とは

- 送信の場所と受信の場所との間を接続する伝送路設備及びこれと一体として設置される交換設備並びにこれらの附属設備
- 日本～米国間光海底ケーブル「Unity」の当事者:KDDI (日本)、Bharti Airtel (インド)、Global Transit (マレーシア)、Google (米国)、Pacnet (香港)、SingTel (シンガポール)



*日本～米国間光海底ケーブル「Unity」の運用開始について、http://www.kddi.com/corporate/news_release/2010/0401a/

*図は、次のURLより筆者がコピー。<http://www.submarinecablemap.com/#/submarine-cable/unityeac-pacific> 18

グローバルビッグデータ事業者はグローバルAS番号を持つ SoftBank

例えば、Google(15169)、Facebook(32934)、Amazon(16509)、Twitter(13414)らは、Equinix Tokyo(品川)に、Ethernetのポートを用意している。

Company Information		Public Peering Exchange Points					
Company Name	Google Inc.	<u>Exchange Point Name</u>	<u>ASN</u>	<u>IP Address</u>	<u>Mbit/sec</u>		
Also Known As	Google, YouTube, DoubleClick	AMS-IX	15169	2001:7f8:1::a501:5169:1	100000		
Company Website	http://www.google.com/about.html	AMS-IX	15169	2001:7f8:1::a501:5169:2	100000		
Primary ASN	15169	AMS-IX	15169	195.69.144.247	100000		
IRR Record	AS-GOOGLE	AMS-IX	15169	195.69.145.100	100000		
Network Type	Content	BCIX	15169	193.178.185.100	10000		
Approx Prefixes	250	BCIX	15169	2001:7f8:19:1::3b41:1/64	10000		
Traffic Levels	Not Disclosed	BigApe	15169	198.32.238.22	10000		
Traffic Ratios	Mostly Outbound	BigApe	15169	2001:0458:0026:2::980	10000		
Geographic Scope	Global	BIX	15169	193.188.137.163	20000		
Looking Glass URL		BIX	15169	2001:7f8:35::1:5169:1	20000		
Route Server URL		BNIIIX	15169	206.130.61.18	1000		
	Please visit https://peering.google.com/about/peering_policy.html for Google's (generally open) peering policy and more information about our traffic delivery and management practices.	BNIIIX	15169	2001:504:20:81::18	1000		
	Related ASNs: Google also manages the following ASNs: AS36040, AS43515, AS36561	1 2 3 4 5 of 13 Next > Last >>					
Notes	Peering Requests: Contact peering@google.com	Private Peering Facilities					
	Peering Operational Issues: Contact noc@google.com 24x7	Facility Name	ASN	City	Country	SONET	Ethr
		Equinix Tokyo (TY2)	15169	Tokyo	JP	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		Equinix Toronto	15169	Toronto	CA	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		Equinix Zurich (ZH1)	15169	Zurich	CH	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		euNetworks (Global Voice)	15169	Dublin	IE	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		Global Crossing Sao Paulo	15169	Cotia	BR	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		Global Switch (London 2)	15169	London	UK	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		Global Switch Singapore	15169	Singapore	SG	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
		Global Switch Sydney	15169	Sydney	AU	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		Infomart	15169	Dallas	US	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		InterXion Frankfurt 1	15169	Frankfurt	DE	<input type="checkbox"/>	<input checked="" type="checkbox"/>
		InterXion Frankfurt 3	15169	Frankfurt	DE	<input type="checkbox"/>	<input checked="" type="checkbox"/>
Protocols Supported	Unicast IPv4 <input checked="" type="checkbox"/> Multicast <input type="checkbox"/> IPv6 <input checked="" type="checkbox"/>						
Date Last Updated	2013-10-21 20:28:27 UTC						
Peering Policy Information							
Peering Policy URL	https://peering.google.com/about/peering_policy.html						
General Policy	Open						

<https://peering.google.com/about/peering.html>

<http://as15169.peeringdb.com> 19

まとめ

■ネット広告産業は、「Highly sophisticated money-making privacy-concerned distributed computing system」であり、現在、グローバルに成長及び寡占が進んでいる分野だと考えている。

- その中で、日本企業は十分に能力を活用できていないと考えている。
- この原因の1つに程よい議論の不足があると考えている。
 - Webトラッキング層を体系化し、共有知識とすべきである。
 - プライバシーポリシーは、法務担当者とプログラマーの共著にすべきである。

■グローバルなイコールフットリング環境の必要性

- 結局、良いプログラムを書いて、グローバルに普及させたものがルールを作っている。
 - 例えばHTMLやHTTPの標準化

■皆さんと、「程よい」解決策を模索したい。