My Digital Footprint: a two-sided digital business model where your privacy is someone else's business! By Tony Fish

September 13th 2009 - FINAL

Issue date September 2009

Copyright and Trademarks

Published by: Futuretext, 36 St George Street. Mayfair, London, W1S 2FW.
http://www.futuretext.com/

Although great care has been taken to ensure the accuracy and completeness of the information contained in this book, neither Futuretext Limited nor its author, contributors, employees or advisors are able to accept any legal liability for any consequential loss or damage, however caused, arising as a result of any actions taken on the basis of the information contained in this book.

Certain statements in this book are forward-looking. Although AMF Ventures, the author and Futuretext believes that the expectations reflected in these forward-looking statements are reasonable, it can give no assurance that these expectations will prove to be correct. AMF Ventures, the author and Futuretext undertakes no obligation or liability due to any action arising from these statements.

For Nicky, Ellie and Millie,

my girls who always bring me

mugs of tea and plenty of sunshine

Table of Contents

FORWARD

This book illustrates Tony's comprehension of how media is consumed, and data is used, to in turn be able to provide a valuable service. The book raises a major debate around the digital footprint and the privacy and protection issues related to that. His experience – as an observer, an investor and a writer – puts the consumer at the forefront of all developments in media, by compelling companies and advertisers to use the power of data to benefit consumers.

To understand the future one must understand the past. Tony instinctively asks the right questions about the convergence of mobile and digital, demonstrating the emerging importance of mobile as a media platform and the rich data which it can provide for brands to leverage consumer behaviour.

As Tony himself says, 'I believe there is some value in this book but there is more value in the comments and community'. Throughout this work, certain questions deliberately remain unanswered which presents an opportunity; Tony has started an intriguing conversation. Not only are we ahead of the game in media by reading this book, but his inimitable style means we all have the opportunity to contribute to the debate.

Pekka Ala-Pietilä

Pekka is co-founder and CEO of Blyk. President of Nokia Corporation (1999-2005) and President of Nokia Mobile Phones (1992-1998). Pekka serves on the boards of SAP AG and Pöyry Oyj.

OPENING REMARKS – THE AUTHOR'S MOTIVATION AND BIAS

Marmite. Rather an odd first word for a book on digital footprints; however, some love it and some don't. Is your glass half full or half empty? Maybe not the best way to categorise human behaviour and our attitude to life, but one we understand. Cancer, faith, wealth, fame and world records: some people get it and some don't.

Identity, privacy, digital you, the social data revolution and this book on **MY DIGITAL FOOTPRINT** drive head first into the same upfront contention, some do and some don't. Some are advocates who want to embrace, contribute, collaborate and participate; others hate the idea of 'big brother': control, snooping and invasion.

As 'becoming digital' continues its pervasive invasion into every aspect of life, there will be a higher, wider and deeper realisation that **MY DIGITAL FOOTPRINT** is important for board and executive functions in marketing, operations, sales, finance, product development and legal; no-one is outside of the consequences.

The book has been written to provide the reader a framework to look at opportunities, strategy, governance and investments in this exciting area and to help provide a rationale as to where to focus time, energy and resource. It is not intended to convert or change professional beliefs, overcome resistance to change or answer all questions. Some of the issues raised are controversial, not for the mundane reason just to be controversial, but there are some underlying important strategic issues that need to be unpacked and discussed in a framework where ideologies can be put to one side and the value and concepts presented. Whilst not bold enough to say this goal will be achieved, I do hope it contributes, as exploitation of digital data is one of the most exciting areas of growth for the next five years.

The focus within the text is on Business to Consumer (B2C) mainly because it is easy to touch and see, but there is a distinct lack of the mention of Customer Relationship Management (CRM) and data mining. There is a passing mention of Business to Business (B2B) within a context that the same framework developed for B2C equally applies, and acknowledgment that there will be many businesses within the B2C value chain that are B2B and sit between the collection and value-adding outputs. The implications are that businesses have to decide what business they are in, no one business can do it all and how those businesses involved in creating value for an end user act together for each other's benefit (co-operation). The business models and frameworks are designed to make you think and are not prescriptive for any specific industry or economic model but embraces cross-subsidiary; two sides and freemium model are examples.

Whilst what I believe in is unimportant, I do subscribe to these points of view.

"The digital divide will not always be about access, but those who engage and participate and those who do not." Tony Fish, June 2009

"You may be disappointed if you fail, but you are doomed if you don't try." Beverly Sills

"Life's bad enough as it is without wanting to invent any more of it." Marvin the paranoid android. *The Hitchhiker's Guide to the Galaxy*

To those who have helped refine the story

In 1996, when I read *Being Digital* by Nicholas Negroponte, I kind of understood the idea of the slug trail. Since then **MY DIGITAL FOOTPRINT**, shadow, virtual identity, digital identity, digital exhaust, social revolution has been in and out of discussions with many of my social crowd. Many people have helped define and refine **MY DIGITAL FOOTPRINT** and continue to do so at the website. I am, however, indebted to: David Birch (Consult Hyperion); Nicky Hickman (Inglis Jane); Daniel Solomons (BLYK); Lord Erroll; Tom Ilube (Garlik); Jerry Fishenden (Microsoft); Gary Gale (Yahoo!); Jouko Ahvenainen (extract); Alan Patrick (Broadsight); Peter Cochrane (CA Global); Vic Keegan (Guardian), Professor James Woudhuysen, Peter Miles (subTV), Martin Leake (footprint graphic design) and Ajit Jaokar.

THE BIG PICTURE

This is a book about footprints – digital footprints to be precise. For the most part, unlike footprints in the sand, digital footprints are largely invisible and are about where we have been, for how long, how often and the inter-relationships. Digital footprints are a capture of memories and moments and are built from your interactions with mobile, web and TV. Digital footprints are not about your identity, your passport or bank account.

While the idea of digital footprints has existed for some time, this book extends the current body of knowledge on this subject in two specific ways, by considering the digital footprint as part of a real-time feedback loop and the impact of mobile devices on digital footprints. In an open loop system the digital footprint is part of behavioural or targeted marketing based on collecting and analysing data, but the closed loop model, as described within, is an entirely different animal and raises a provocative question: Is your digital footprint yours or someone else's business?

In conventional terms, digital footprints are the digital 'cookie crumbs' that we all leave when we use some form of digital service, application, appliance, object or device, or in some cases as we pass through or by, this happens regardless of whether we are actually cognisant of this. We intuitively accept, when forced to think about it, that these traces exist and we somehow expect that, over time, the waves will wash over the digital footprints to erase them like the ones on the beach – but they are not. Like everything on the web, digital data cannot be washed away, it remains forever, but we could actually benefit from taking control of our own digital footprint.

This chapter provides a context to identity, privacy and other widely debated issues that get bundled together as part of the digital footprint. There is both good and some very poor work available on identity and privacy; this summary is presented as an overview to provide the context of how identity (physical and digital) and privacy are related and connected to MY DIGITAL FOOTPRINT. In the simplicity of this big picture, I hope to highlight some of the key themes that lead to the black holes of debate and strong opinion without glibly or underemphasising some, rightly, contested issues.

Identity is not simple, but, at a very high level, it provides the person on the street with a name, driving licence, bank account, credit history, social security and certain certification from both government and non-government organisations. Whilst identity for the average person is seen as regulated and institution-controlled; privacy is emotional, a preference, determined by status but constrained by regulation and law. Digital versions of identity and privacy are in the most terms not even considered by Mr Average; leaving a digital footprint as a term that is not even contemplated. Defining MY DIGITAL FOOTPRINT is therefore complex, but this book suggests: collection, store, analysis and value created from digital data from mobile, web and TV. Additional collect points for data such as Near Field Connection cards (NFC) (e.g. Oyster in London) and non-web financial transactions will be added at a later stage.

The following chapter outlines the links between digital and physical identity connections and dependencies, as set out in Figure 1 below. This section then leads on to the connections between identity and MY DIGITAL FOOTPRINT.

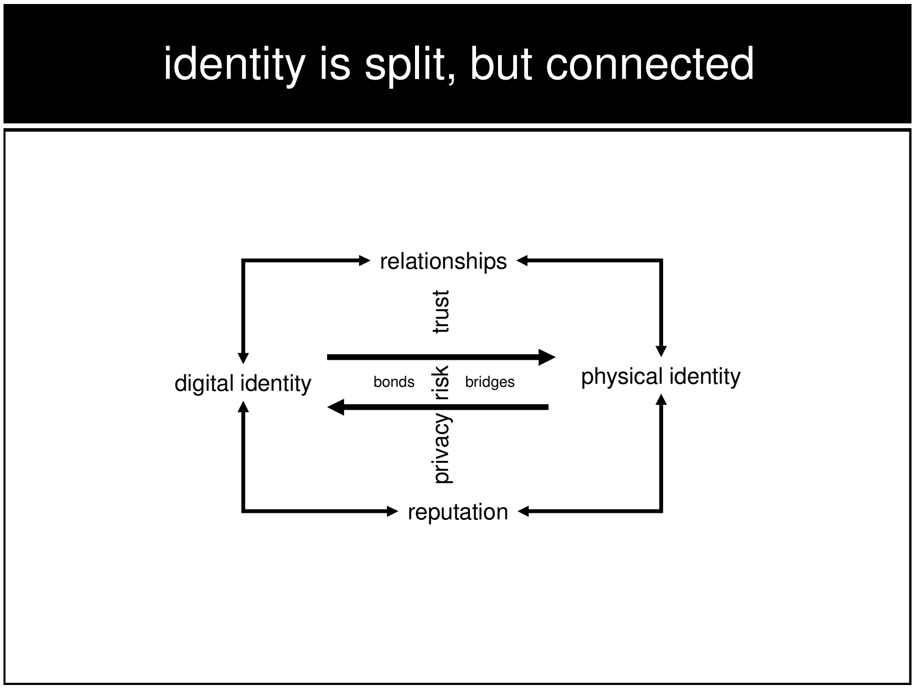


Figure 1 Digital and physical identity connections

Identity is split between ‘digital’ and ‘physical’. This split is important as there is a need to explore their different and similar characteristics. As explained later, **MY DIGITAL FOOTPRINT** (in this book) only relates to the digital side, but digital identity does have some strong dependencies on physical identity.

Digital and physical identities are related through certain bridges and bonds. These are couplings that allow the two forms of identity to have value. Bridges are direct connections that allow someone to use either form of identity for a single purpose, such as using your bank card for digital and physical transactions. Bonds are not the direct relationship between the digital and physical, but are bonds which allow the two to be related, as certain key information is held in both that allow the connection to be made. Both bridges and bonds have certain dependencies. These dependencies are privacy, risk and trust.

Digital and physical identities are also related to each other by either relationships or reputation. Facebook, MySpace, Friends reunited, Plaxo, Linkedin deliver relationships between the physical you (what you did and with whom) and the digital you. Indeed, you could see this as ‘bond and bridge’. As you trust someone from the past, you connect with them, this restores your relationship and improves your reputation (you are who you say you are).

The purpose of this book is not to explore these connections but to observe that they exist, as this is needed to help define what **MY DIGITAL FOOTPRINT** is, in relation to certain traits and characteristics of identity.

Figures 2 and 3 separate physical and then digital identity into certain traits and provide an outline of each.

physical identity has certain traits

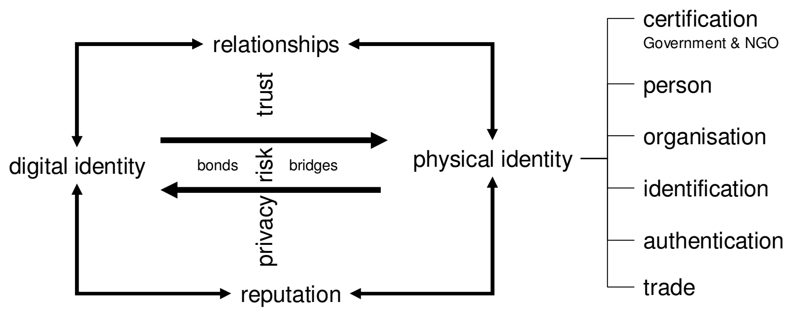


Figure 2 Traits of physical identity

Physical identity can be broken down into many traits; this list is only for example and highlights certain contentious issues. Again, this is not exhaustive coverage as there is an enormous body of work available.

Certification

The certificates produced by government (birth, marriage, driving licence, passport, social security). These documents are a stable proof of who you are. Fraud is possible and allows someone to take over your identity. Depends on original certification (provides evidence of history) and in some instances a match in a central database.

Person

The physical characteristics of you: eye colour, skin tone, iris, fingerprint, blood group, height, DNA, speech and style. Technology is continuously improving which allows you to use your body to identify and confirm that you are who you say you are, it does not relate you to others or history (aka certification).

Organisation

This trait of physical identity relates to organisation-based identification, such as your password and log-in ID, library card, entrance pass, phone number and utility bills. These organisations have satisfied themselves that you are who you say you are and provide an extension of your certificated proof. This has obvious and deep links with identification and authentication below.

Identification

A major part of identity is value. Value is achieved as you can provide identification to enable you to extend and expand the services you want. The eternal question is related to if you are providing evidence that is fraudulent and to what degree you can step back to warrant truth. Identification in this sense for true identity experts is validation.

Authentication

Authentication sits across several other physical traits as it is about the evidence that you are who you say you are. At a simple level this would include your card and PIN for accessing your bank account.

Trade

One of the purposes of identity is to enable you to start and continue to trade; either to buy goods and services for yourself or your organisation. As a trait it is important as it brings relevancy to the level of identity required by providing a hurdle to overcome and quantifies risk.

Our focus now turns to digital identity, which has different traits to physical identity. It is worth remembering that whilst digital is different, it is bridged and bonded to physical. Whilst exploring digital, part of the mind should remain cognitive of the fact that digital doesn't, in many aspects, exit without the physical.

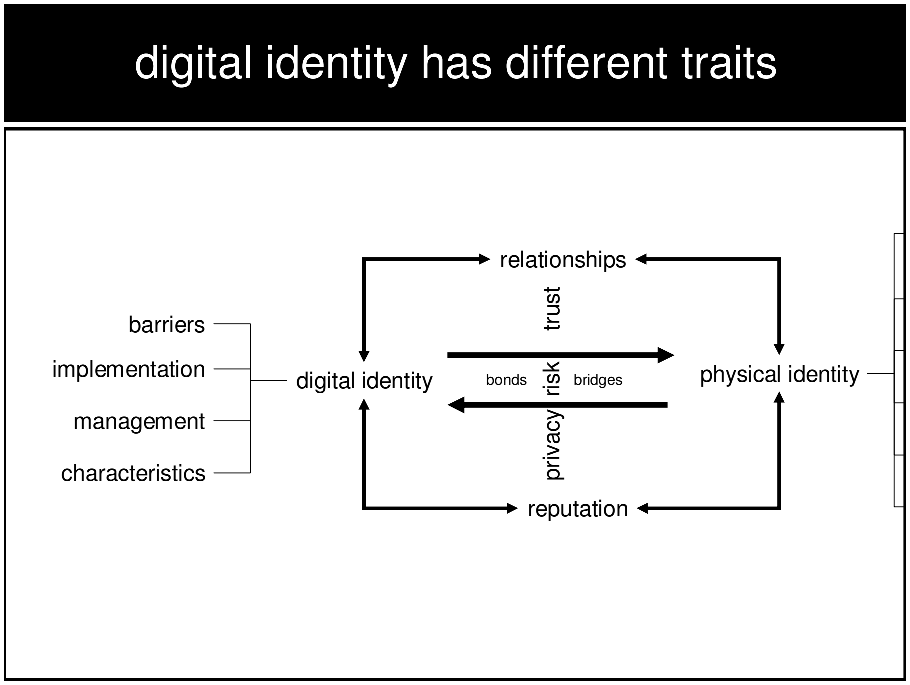


Figure 3 Traits of digital identity

Digital identity, like physical, can be broken down into several inter-related traits; this thirty thousand-foot list highlights certain contentious issues. Again, this is not exhaustive coverage as there is a wide range of detailed insights available. It is worth noting that digital identity allows users with the desire to build their own identity, a different identity or augmenting the one given to you by the church, society or government.

Barriers

These are some of the many objections to digital identity and include:

The Data Protection Act and similar law and regulation. This is a body of work designed to protect the user. In many cases it introduces a high overhead and often does not have the tools for enforcement.

Surveillance. Usually first up. Being watched and identified, this data then stored and used against you. Some benefits but fear, uncertainty and doubt reign.

Re-identification. A difficult topic, where data has been anonyised but can actually be reconnected to the original user. You don't want to give your home address, but your GPS-enabled device embeds information that allows re-identification. Usually takes some work, but in reality quite easy.

Tracking (things and people) like surveillance – tracking allows you and your objects to be found, identities and details of their routes and routines. 'Big Brother' state and companies trying to control your life.

Storage. If data is centralised it cannot be secure and has to be backed-up (copied) for business continuity processes. If de-centralised, who owns it and how is it managed? As a barrier, security is about complex balance between user experience and protection.

One ID. It is impossible to change passwords every week. Having one ID is what the user wants for user experience, but is the worst nightmare for security. A balance that is not in balance.

Implementation

This trait focuses on the technology and implementation. Some characteristics inter-relate with barriers.

Sign-in and passwords: Provide an essential part of digital identity. Management of sign-in and passwords is very difficult (open ID, one ID) etc, and have approaches to offer a better user experience, but this is at the risk of security

Overhead. Anyone who works for a corporate who has secure ID log-in will immediately identify with overhead. This is all about the layers of improving security at the cost of user experience.

Complexity. Every part of identity is complex and inter-related: there are no simple answers. Just imagine trying to ensure that your company has a database of all staff and that they have a physical reference related to each member that is checked, allocated, aligned and protected. This data is then guarded, protected and tracked. You now want to introduce a new access method to your buildings and network, how do you secure the old and the new, and guard against errors in the old data, clearing holes in the new?

Management. This captures the management requirement for the user to store and hide log-in, passwords and ID. [Technology management of solutions comes under implementation management.] Running digital identity for individuals only becomes more complex and more open to abuse. Management also covers executive management functions where current directors' fiduciary duties mean that their company brand and professional reputations are on the line as well as criminal proceedings for any breach.

Bypass. Without users this would disappear. The weakest part of any digital ID is admin bypass. Users lose log-in, ID and passwords, somehow recovery and access has to be administrated. The admin of bypass adds even higher levels of complexity.

Collection. Even at the most basic level, collection of data to build a map of behaviour is well established. Our finance market provides a warning about past performance and is not a good indication of future performance. Commercial and non-commercial organisations can gather data today without cost (Web 2.0) and this collection (what is collected, from who, the types of data and from what devices – mobile, web and TV) opens up new possibilities. Collection is not difficult!

Storage. As an implementation trait this is easy to understand (also see storage within the barriers section previously). Storage and the last topic, security, make storage a complex implementation issue. Federation and peer-to-peer look good at the outset, but the public has no idea what the industry is talking about or if the balance is in their favour.

Security. This is how to secure, and keep secure, digital identity.

Management

This trait is about what needs to be managed.

Person/persona. There is a difference for a user between something that is linked to the person, this being physical (money) and persona (second life, world of warcraft). Whilst both should be viewed to have equal value to the user and should be protected to the same degree, there is a difference in the linkage to what would happen if stolen. One is unbounded (access to a bank account gives many opportunities) and bounded (the world of warcraft is limited to selling ID or taking power).

Degree and level. To what degree or level is there a necessity to protect and manage digital identity? As with person/persona above, there is a variation, which links to implementation. However, if you could trade between the person and persona, and they had fundamentally different degrees of management, integrity, security, risk, trust, or privacy – should the trade be allowed or will it bring everything to the lowest common dominator?

Voluntary/forced. Forced is the requirements your bank and your corporate give you, for good governance reasons. Voluntary is open, and provides simplicity for ID access to your Plaxo, Linkedin, MySpace, Live and other 2.0 accounts. Management [forced] is easy to implement and deliver; voluntary, as per degree and level, may bring down secure systems if the two have to cross.

Transparent/forced. A topical issue with UK Government presently (May 2009). Expenses that have been hidden are easy to defraud. Transparent and open allows for inspection and the collective approval. Should your digital ID be approved not by you or an organisation (forced) but by your social network?

Open/closed. As per voluntary and transparent, should the system by which digital ID is verified and delivers integrity be open or closed? Open allows the possibility for bypass and creation of false identities as the system is understood. Closed, you would never know. Is open/closed the question, or should it be about development methods, requirements or agile iteration?

Control. As a management trait this is hidden in here as it is controversial. Control is required by the state and organisation as there is a belief that this provides for enforcement and delivers security. Control is interrupted by the user as controlling. There has to be a balance between integrity (below) and control.

De/centralised. As a management trait this is an important topic as centralised is easy to understand and as easy to break, and has links to storage and bypass. Decentralised offers an interesting alternative; however, introduces complexity and knowing your data may be stored on someone else's machine (even encrypted and only in part) may not engender participation and trust. Again a balance.

Integrity. Managing integrity (data, personal, corporate, governance), as per all the previous points, is a challenge especially if an inspection or regulation body is empowered to maintain standards (and what standards?).

Characteristics

This trait of digital identity brings out the aspects of value that can be created, within the context of the barriers, implementation and management traits mentioned over the

> last few pages.
>
> Benefits. Whilst understanding the conflicting issues mentioned in brief above, the rational for spending time on this topic and having a digital identity is to have benefits for the user. Benefits for the user include being able to trade and barter for goods and services in an easier and more user-friendly environment, reduce and prevent crime, improve medical diagnosis and recovery, remove virus and spam, and deliver context, personalisation and other benefits in a web, mobile and TV world (the screens of life).
>
> History. There is no evidence that the author attended Little Green Junior School, nor certificates, and in the past 30 years all records will have been destroyed. Personally there are a few text books left that could help. My Facebook profile shows a picture of me at age 10 with the class of 1977 on our French exchange trip. Suddenly I have history, developed and delivered and validated by my community.
>
> Reference. As per history, my Linkedin CV provides references about my work from the community, not made up by me.
>
> Anonymity. Digital identity does deliver personas and anonymity.

I hope that this highlights that whilst physical identities are centralised in their systems for design, development and control, they're mainly about an organisation or government keeping people out or getting people into a system; the emerging softer identities of the digital footprint are de-centralised 'starfish'-types of identities with peers and peer-groups providing reputational validation and authorisation, rather than authorities, (your boss, the government, the bank).

WHAT ARE THE LINKS BETWEEN IDENTITY AND MY DIGITAL FOOTPRINT*?*

Figure 4 provides a visual representation of the links between digital identity and MY DIGITAL FOOTPRINT. The purpose of this representation is to separate out the key topics and themes that will be explored in the remainder of this book and to highlight that some of the more controversial aspects of identity are dropped, as are the technical implementation arguments. This separation, it is hoped, will allow the reader to focus on the underlying important strategic issues that need to be unpacked and discussed in a framework where ideologies can be put to one side and the value and concepts presented.

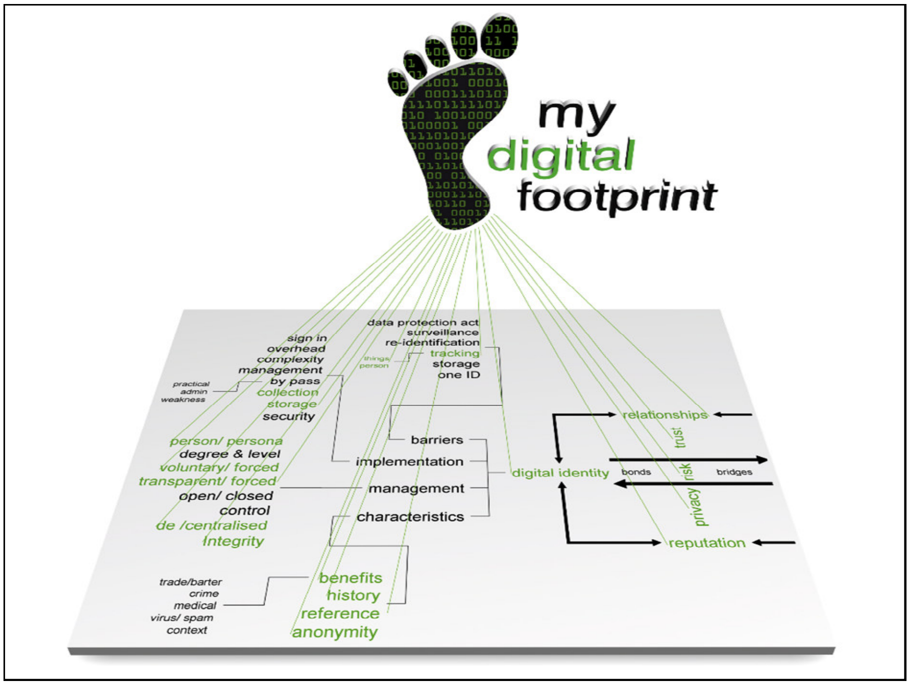


***Figure 4** Linkages between identity and* MY DIGITAL FOOTPRINT

These linkages form the basis for the framework presented in this book as per Figure 5. The core themes relate to data, dependencies, value and business models. Within the context of data, this book explores the collection, store and analysis of user data to create value. Dependences are threaded into the fabric of many aspects of this book as it brings out the bonds and bridges between **MY DIGITAL FOOTPRINT** and relationships, security, risk, privacy, trust, law, regulations and identity. Value has to be created from **MY DIGITAL FOOTPRINT** or it would not be a topic; therefore intent, reputation, discovery, recommendation, protection, personalisation, trade or barter and context are discussed in detail to explain value-creating concepts. From value comes a business model and eight business models are presented, but the most critical question is 'Who owns the data?' and this is explored.

Throughout the book there is a focus on two specific aspects: the digital footprint as part of a feedback loop and the impact of mobile devices on digital footprints.

My Tesco clubcard doesn't use my real name, Tesco own the data and it works perfectly. In other words, I can have a digital footprint that has value to brands as a commercial nexus but is not my only digital footprint. I believe that everyone will have a few different digital footprints, just like we have a few different credit cards, which will lead to an interesting model. Advertising, as one model, can be personalised without destroying privacy.

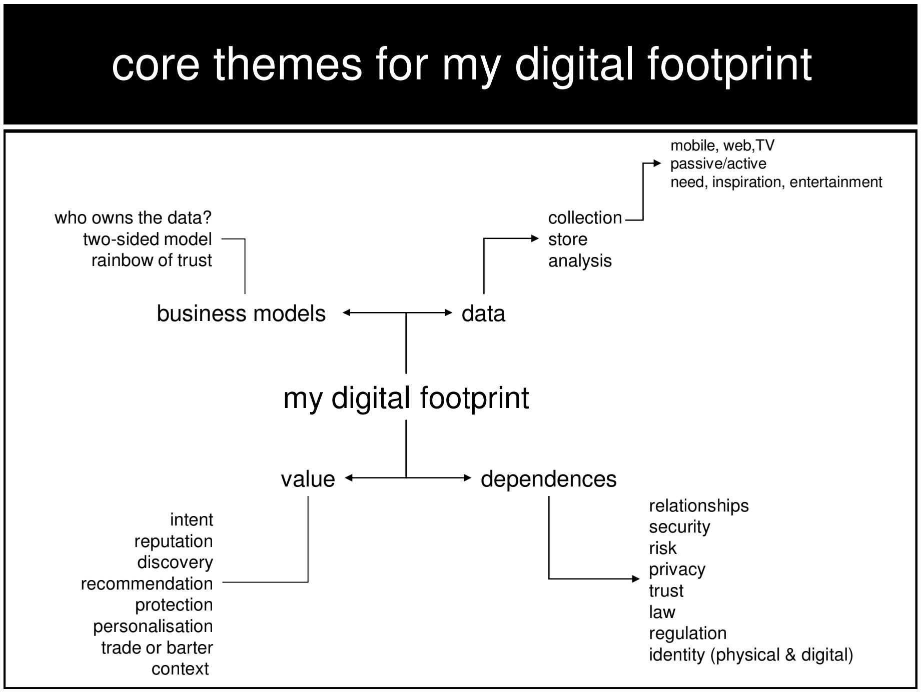


Figure 5 ***Core themes of*** **MY DIGITAL FOOTPRINT**

This book introduces the idea of a two-sided business model. Currently, we are used to the 'strong identity' model (the kind of identity mechanisms required by banks and immigration authorities). In contrast, a different form of web-based identity is emerging which could be described as complementary to the strong identity. This is your digital footprint. Both identity and digital footprints exist together. You cannot open a bank account with your digital footprint (yet!). So, the strong identity will exist for many cases and the impact of the digital footprint on new services and on trust is explored. Throughout this discussion, the emphasis is on the 'uniqueness of mobile' and increasingly mobile will contribute to a larger share of your digital footprint – which makes the mobile platform very significant.

This book cannot provide you all the answers, which is why the website http://www.mydigitalfootprint.com/ provides for comments and links, but it will make you think and I hope that you will join in the debate since it concerns every one of us and our future generations.

DIGITAL FOOTPRINTS

Like Neil Armstrong, whilst walking on the moon, and Nelson Mandela walking free from Robben Island, we all leave footprints. Footprints are more than identity. Footprints are about where we have been, for how long, how often and the inter-relationships, they are memories and moments. Therefore, digital footprints are not about your identity, your passport, bank account or social security number. Digital footprints come from your mobile, web and TV interactions and comprise the digital data and also the Metadata[1] (data about data) of who we are, the true value and why the ownership of this data class is the battleground to be won and lost.

However, the original web-based digital footprint and its digital data belonged to the individual at some point. But the individual is currently not empowered to hold or manage this digital footprint. Mobile adds a unique dimension to the digital footprint since mobile provides new content, Metadata and the social context for the digital footprint. In contrast, both TV and the web can provide some data – but the mobile device is unique in terms of its contribution to our digital footprint. This idea is illustrated in Figure 6. The taps and the volume of water that flows is a visual representation of the amount of value that can be created from user data. The water fills a pool which is the representation of the digital data that is stored about you from your interactions.

Data from Broadcast/Listen include: viewing times and schedules, preferences for channels and content, timing of programmes and presence (are you actually there, this could be determined through motion, channel hopping, fast forward or a secondary device, PC or mobile listening to your TV preference). Data from the web includes: attention, how long you read a page for, browsing history, search words and spelling, patterns and clicks, content created or viewed and purchases (consume). Data from a mobile would include: location, attention, browsing, search, time, who you are with (Bluetooth), proximity, clicks, creation of data and media, consumer, play lists and presence. The actual raw data could be location co-ordinate, a click, two-way interactions or a picture; the size of the tap and subsequent flow represents the volume of data that can be added to the digital footprint pool.

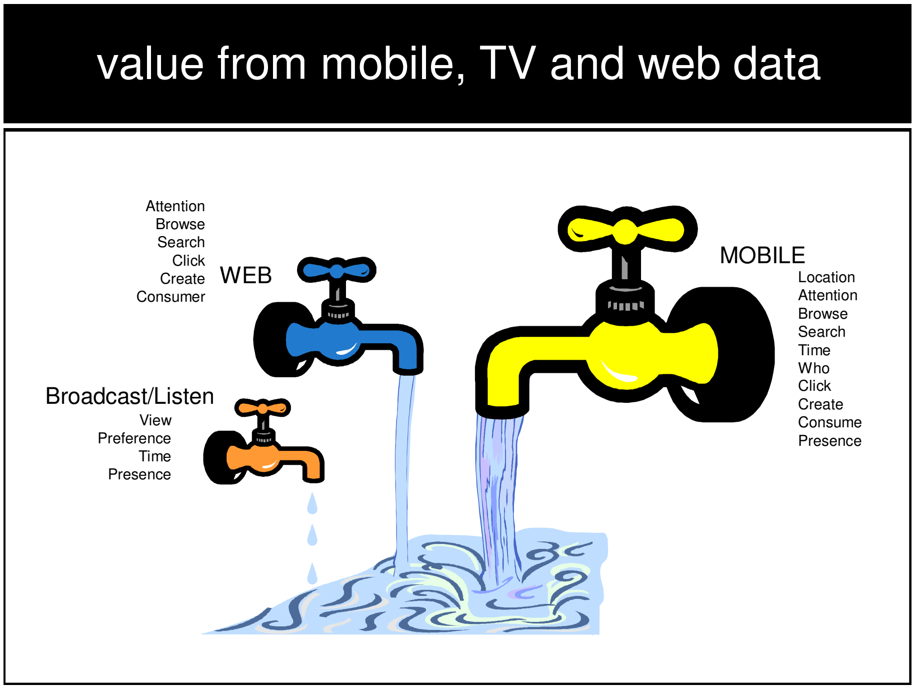


Figure 6 Value from mobile, TV and web

From digital footprints to MY DIGITAL FOOTPRINT

The idea of digital footprints has been discussed from a privacy or data protection standpoint. However, key commentators agree that we are increasingly leaving larger digital footprints over time, especially given the rise of popular social networks and mobile devices.

A digital footprint is the persistence of data trails online by a user’s activity in a digital environment – which Nicholas Negroponte called the ‘slug trail’ in *Being Digital* and John Battelle calls the ‘Clickstream exhaust’[2].

According to the Pew Internet report[3], there are two main classifications for digital footprints: passive digital footprints and active digital footprints. A passive digital footprint is created when data is collected about an action without any client activation (implicit) and include data from sensors; whereas active digital footprints are created when a user deliberately releases personal data for the purpose of sharing information about himself (explicit).

On the web, many interactions, such as creating a social networking profile or commenting on a picture on Flickr, leaves a digital footprint. In a mobile context, CDRs (Call Data Records) are the transactional data that constitute the user's digital footprint. But the mere availability of transactional data alone is not enough since privacy and data protection rules will apply to the usage of data, and rightly so. It is the ability to store, analyse and create value from the digital footprint that differentiates the study of digital footprints. In other words, if we all left digital footprints – and nothing happened to those footprints – then there are no concerns and no benefits.

Hence, in this book, we present a definition of digital footprints to include capture, store, analysis and value. The 'capture' phase includes not only your own activity – but also the activity of others related to that information element – for instance, the impact of your social graph and third parties on the digital footprint. This idea is depicted in Figure 7.

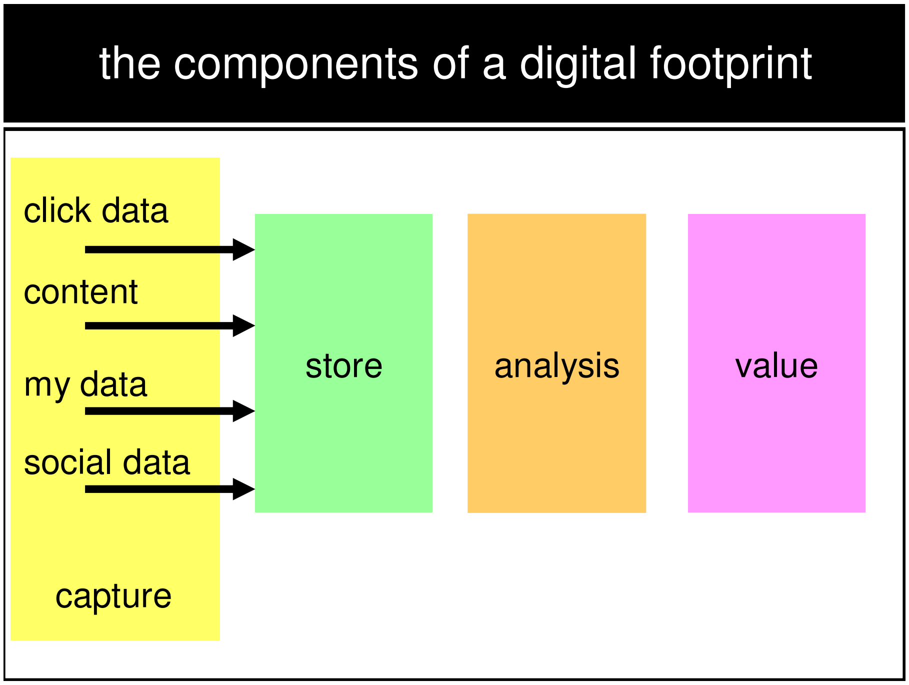


Figure 7 Defining the components that create a digital footprint in simple terms

The capture of data itself arises from multiple sources and importantly not only from the data created by the person but by sensors in or connected to devices. The storage of the digital footprint relates to where raw data is stored physically, it's the ownership and portability [the analysis is also stored but cannot be reversed to create raw data again and identify the user, the location, the service, the purchase or anything unique]. Analysis is the key differentiator in terms of where wealth could be created and in turn leads to the potential value available to both the user and the service providers.

Thus the data captured consists of the click data (the digital trace), and the content (the actual data created – e.g. the pictures uploaded from a mobile device). My data, as presented in Figure 8, is the automated data collection, which can be an embedded location in a picture or enabled by the user in some application. sensory.net is part of my data and is where the device acts as a sensor and collects information. The last input being the social component. The social component is not your data about you as per the other inputs. Social data is what your social group provide about you, as the individual; your references in Linkedin, your rating on eBay, your image tag on Flickr and Facebook.

There is a subtle but important distinction between 'digital footprints' – that is, between the raw data that is captured and stored for analysis and my digital footprint. The idea of my digital footprint extends the idea of raw data to the wider concept of capture, store, analysis and value created from data generated through digital engagement. The way in which my digital footprint is used is to focus on the system and value created.

This process of building **MY DIGITAL FOOTPRINT** is based on a structured approach incorporating inputs (collection) and outputs (value) and a feedback loop that governs the whole process. Who does the building is an important topic and is addressed later on the in the book under the business models chapter. This feedback loop progressively enriches and refines the outputs (value). The analysis phase is able to take raw data from various stored sources (which we refer to as the digital footprint) and from the analysis of this raw feed generates value (wealth and services), which takes the form of components such as personalisation, reputation or discovery. The output of the analysis stage we call 'behavioural DNA', as it provides a detailed description *of* you, but has not delivered any value *to* you.

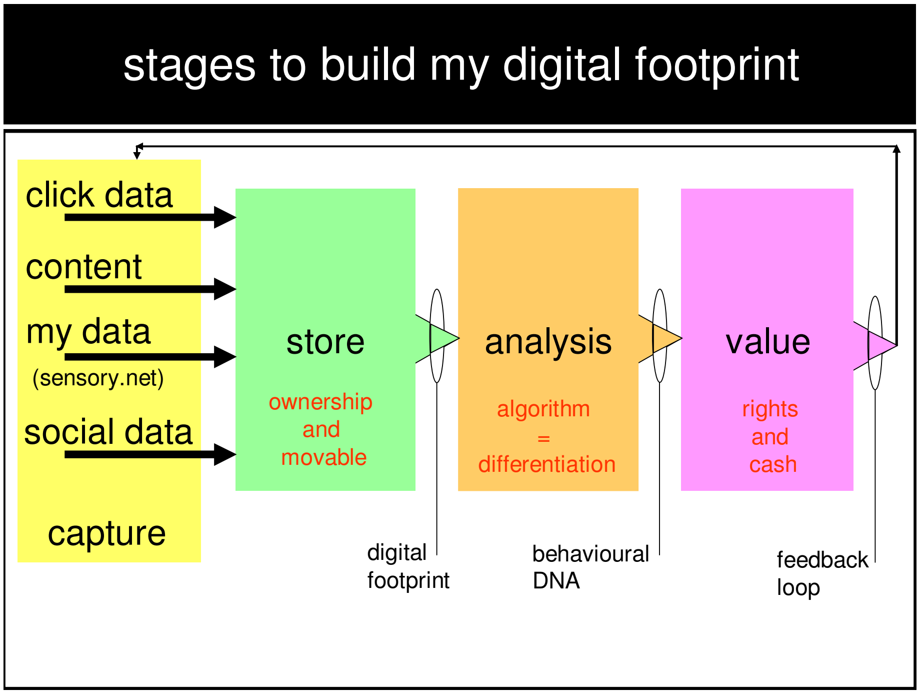


Figure 8 ***Defining the stages in the generation of*** **MY DIGITAL FOOTPRINT**

In the chapter on the two-sided business model there is a detailed description of the inputs and outputs and how the feedback loop works; however, here is a short description of the inputs and outputs. The inputs to **MY DIGITAL FOOTPRINT** are the data elements and the outputs are the value derived from the process which is in turn enhanced by the feedback loop.

Inputs into MY DIGITAL FOOTPRINT

Attention Data that indicates what you are doing, it is the provision of data that details what applications and services you are engaged with. This could be a widget on your desktop, mobile or set-top providing insight into which applications are open, how long you edited a document, which pictures you viewed, what music you listened to and how often. The attention data stream is the record of what you spend your time doing in a digital world on TV, web and mobile.

Location The data record of where you are. The live feed is collection, where you were (route taken) if stored.

Time

The time data record is both the time of day and also the period of time.

Search

The data string of search requests, currently the text words (and voice-based search on Google mobile), entered into a search engine, but progressing to automated search based on requests from 3D barcodes and local available intelligence.

Content (create)

The data record of type, context and information about the content you have created for text, voice, presentation, music, audio, images, video, blogs, tags and recommendation.

Activity

This is the dataset that defines what you are doing, whilst attention says you are looking at a web page, activity defines it that you are at a football ground. Location gives you the co-ordinates.

Outputs – value created from digital footprints

Intent

Intent is an output which provides predication about what you will do next based on what you have done, what your social graph does but also on what you have told/inferred/implied that you are about to do, such as your calendar, email or IM trail. Whereas context is about now; intent is about next.

Reputation

Reputation (digital) has many components. Reputation is both about a rating (good and bad) and about your propensity to do something, such as leave a comment. Reputation (digital) is therefore partly about your value to the community as a participant.

This output data produces a record which is your digital reputation.

Discovery

This output provides concepts, ideas, insight to enable the user to discover. Discovery is about risk and comes in the form of improvement to an existing service or discovery of a new service/application.

Recommendation

This is where, based on your digital footprint (you and your social graph), a service/application is able to make a recommendation about an existing or new product or service with a degree of confidence that it will be relevant. Where discovery is risk, recommendation is about trust.

Protection

This is where your data can be used to protect you and your data, in the same way fraud on credit cards works. Your data is a good predictor if you are the individual who is providing the data. This does depend on humans and certain social groups being creatures of habit.

Personalisation

This is where the application or service is personalised to a user for the particular instance or time. It is the modification of a generic service automatically, but based on what is known about you.

Trade or barter

This second order output function enables the user to trade or barter for goods or services. The trade or barter will not be for cash (this is payment) but for data or for insights, research, etc. This trade or barter is based on input data, analysis, intent and reputation.

Contextual adaptation

This is where the service or application will adapt to deliver a service that is unique to the individual’s requirements based on the existing environment.

The paradox of privacy

There is a paradox with privacy. On the one hand, everyone fears losing it. Scott McNealy of Sun Microsystems famously said that: “Consumer privacy is a red herring: we have zero privacy – and we should all get over it”[4]. This view has gathered credence after 9/11. Esther Dyson argues that we need more granular control over our data. She believes that the notion of privacy doesn’t fully capture the challenges of the current environment online. ”We need to stop talking about privacy and start talking about control over data,” she says. She argues that, “In the future, users are going to want more granular control over their data, making detailed decisions about what gets shared with whom. Users may be overwhelmed when first setting up an account, but when they get more comfortable with an application, they will exert more control.”

The idea of granular control over our personal information based on the work of Kim Cameron[5] and Stefan Brands[6] is worth reading if you have a particular need for detail on privacy. The kind of revised technical model enshrined in the laws of identity combined with the smart cryptography of minimal disclosure tokens provides (at least at the technical level) an important breakthrough in the way we think about engineering the design of digital products and services, and empowering the user to control their data in a highly granular and empowering way.

On the other hand, we all have an incentive to contribute data about ourselves, while reflecting on the manner in which we want to be seen, so as to be more visible within a digital context. For instance, even if we don't want to leave a digital footprint (traces, exhausts, trails, shadows), we do want to be searchable on the web and prefer the results to be seen to be favourable to ourselves.

So, either we are passively creating digital footprints or we are actively contributing information about ourselves. In either case, we are contributing to someone's view about us, but in doing so are giving up data and our digital footprint, which could be harnessed.

Who is harnessing your collective intelligence?

Web 2.0 taught us the concept of 'harnessing collective intelligence', which will be discussed in greater detail in this section. Harnessing collective intelligence is not a problem in itself. The dark side, however, arises if a business entices its audience (customers, clients, delegates, patients, friends) to give up their digital data, collect their digital footprint without their agreement, charge people to view their own data, or sell OUR data off with the sole expectation of making money though the one-sided route of exploitation. My [Tony Fish] mobile number is widely available on the web, I never get unwanted calls; my home number is ex-directory and only listed on private applications, once a week I receive an unwanted sales call, who sold my data?

On May 1 2009, Spock[7] was acquired by Intelius (a background check company). Spock is based on a robot which automatically creates tags for any person it finds. It trawls the web for sites such as Wikipedia, LinkedIn and others. It also allows users to enrich their data by letting them add tags of their own and add other data such as relationships between people. Thus, following the ideas of Web 2.0, the site gets better as more people use it. However, from the user standpoint, it gets tricky because:

- Spock trawls the web looking for our data;
- it creates a profile about us in their site without direct approval;

- it encourages us to enrich that information;
- it charges us to access our own information; and
- ultimately, it sells that same information to a background check company.

This leaves open the question, can I delete my own information in Spock? In many ways it now becomes more interesting. Spock says on deletion of information[8]:

‘If you'd like to remove yourself from Spock, please read the following information and click the link below. Before requesting removal, please make sure the original source of the information Spock found for you has been removed or made private (MySpace, blog, Friendster, etc). This will prevent you from being re-indexed on the site. Please note that you can only request removal for your Spock search result. When filling out your information please make sure to include your name, e-mail, a link to your Spock Search Result (http://www.spock.com/Tiger-Woods), and the reason why you'd like to be removed. The Spock Support Team will review your claim and get back to you within 24–48 hours.’

The implication is that, as a user, I have to ensure that the original sources of information that they (Spock) sourced (via a spider) the profiles from should also be made private (my blog, my Facebook profile, etc) or they will 'harness' me again. This is the Web 2.0 equivalent of harvesting email addresses and selling them on.

In future, legislation may extend to cover such practices and the idea of empowering the customer will gain more acceptance. As mobile devices become more common, the issue becomes more significant with Mobile Web 2.0. (By Mobile Web 2.0 – I mean the concept of extending the idea of harnessing collective intelligence to mobile devices, which are more attuned to capturing data along with the accompanying Metadata.) But on a more optimistic note, the use of user-generated data to create a better service is a good thing provided there is transparency and the control rests with the user.

Thus, whilst the principles of Web 2.0 were sound, some implementation of the business model may not be, especially is they are FREE. Responsible companies will come to use the principles and create better services (and will use the web and the mobile holistically). These may not be free, hence not 'Web 2.0' in the traditional sense. But they will be more honest in their relationship to their customers and more transparent in their usage of data from their customers. Users may also be wiser and more empowered. We will learn from the mistakes of Web 2.0 and create better engagement and trust based on converged/mobile-driven services.

Privacy – is your privacy someone else's business?

The above discussion raises the question: 'Is our privacy becoming someone else's business?' What I mean by this is; are companies exploiting my private data for their own gains, and the privacy I thought I had, either never existed or has been eroded.

These practices (exploiting data) have attracted the attention of regulatory bodies, such as the Federal Trade Commission in the USA[9]. It is widely accepted that the populace take the role of media and brands for granted. Marketers will tell you that 'brands need us and we need them'.

Much technology and innovation adoption is now driven by brands, think iPod, and the interests of the brands are not necessarily aligned with the interests of the customer. Hence, the notion that 'we need brands and brands need us' has to be tempered with the basic reality that the primary purpose of brands is to sell. And let's not forget that. As media becomes rich and complex, brands seek to engage with us and to measure that engagement for maximising their revenue. Hence, their interest in 'harnessing the digital footprint' and its consequent impact on privacy. Digital footprints and privacy, trust and risk are two sides of the same coin; they are bonded and bridged as discussed earlier. There is an extended chapter on these relationships towards the end of the book as the focus moves to the business model. Advertisers need a lot of data to make their advertising more personalised (and, by extension, to claim more money from the companies who use their advertising), but acquiring the data needs the customer to give up certain privacy rights, trust someone and take some measured risk, all in the interests of the advertiser.

Consider the industry's emphasis on 'convergence'. Convergence could take a much darker meaning in light of harnessing a digital footprint. If the media and advertisers were to indeed 'join the dots' between the various information elements left by us (cookie crumbs of information in different media), then advertising becomes powerful and personalised. This benefits the advertiser especially in a converged media scenario (where the same provider owns the TV, landline, mobile subscriptions, etc); however, this could lead to some questionable behaviour which could be currently legal but may soon be regulated. It could also lead to consumer backlash, as explained in the scenario below.

Suppose you are a fan of a rock group and you blog about it. Consider this scenario. Many TV companies are exploring ways to 'personalise' TV advertising to the home. For instance, they seek to gain viewing preferences from set-top boxes and other avenues and then (in an extreme scenario), to tailor the advertising to each home. The question becomes: Which data elements can be used to tailor this advertising?

Considering the ‘rock concert’ data element, which can be obtained from an RSS feed from my blog. It is easy to combine three sets of data: my home address and my name, which the cable company has in addition to the RSS feed from my blog (which ties to my name or, via packet inspection, my persona) and the phone book/voter registration (as a confirmation of my address). Knowing these elements, the cable TV company can join the dots (analysis in this book’s terminology) then 'personalise' the advertisement to tailor specifically to me (show me an advertisement of the next rock concert by the performer in the commercial break on the TV/cable). Presumably, this makes the advertiser 'happy' since they are 'personalising' the advertisement to me and I could even 'engage' with it by pressing the 'Buy it Now' button (advertising utopia!). On the other hand, it could be seen as a gross invasion of privacy and questionable 'Big Brother' tactics, which is why the first word in the book was ‘Marmite’ – some like this idea, others don’t.

This personalisation and engagement could be made progressively worse in the future based on the abundant availability of different datasets open to advertisers, all of which could be co-related from different datasets to gain new insights about us to 'sell' to us. This ability to re-identify or even identify does raise important privacy issues. You could call this 'micro-persuasion' and indeed it raises some questions about the ethics of advertisements and engagement (although none of this behaviour would be seen to be illegal). It also raises a genuine spectre of consumer backlash (e.g. if I were to see many such rock concert ads, I would know that my TV is watching me and I would take action, such as change channel or service provider).

Governments will also follow suit. Governments need to be involved in two ways:

- by creating regulation that benefits consumers in addition to the advertisers, especially in relation to new areas where regulation is sparse and consumers can be potentially exploited; and
- ensuring that the privacy rights of individuals are protected in the light of ever increasing encroachments from brands and advertisers.

At one level, we have laws such as the Data Protection Act[10] in the UK. However, government can also be part of the problem, for instance, the proposed law on 'data sharing' in the UK. Under the guise of 'mass exchange of data can offer some benefits' (to advertisers and governments), the UK Government is proposing legislation (source: *Telegraph* website[11]) by which data held by the police, the NHS, schools, the HMRC (tax), local councils and the DVLA could all end up in private hands, according to Privacy International. At the same time, information gathered by companies including hotel registrations, bank details and telecommunications data could be transferred to the Government as part of the provisions of the Coroner's and Justice Bill, it is claimed. The campaign group admits the ‘mass exchange of personal information has the potential to deliver some benefit’.

Yet another grey area is targeting minors and ethnic groups. Legally, there is no law that prevents the targeting of specific ethnic groups by advertisers. In fact, it can be profitable to do so, as per the benefits from the ad network JumpTap, which predicted that Hispanic centric campaigns would quadruple this year, with revenue increasing at least 20% in the segment[12]. There may be indeed nothing wrong in selling Hispanic-orientated content, music, etc, targeting a specific demographic, but change the model slightly and you get some serious privacy concerns. For instance, the South Asian population is genetically susceptible to Diabetes[13]. Does this mean that Diabetes medication advertisements should be targeted to South Asians in the UK? Again, this is not too difficult to do using current technology and increasing convergence and data availability. Where do we draw the line?

For many of us who travel to the United States, we see drug companies advertising medication on TV. This is illegal in many countries – especially in Europe. The message from the advertisements seems to be 'Call your General Practitioner (GP) and ask him to recommend our drugs'. Broadcasting drug company advertisements raises the ethical issue of the advertising company influencing the doctor's judgement for commercial reasons (selling their products). In many countries, this is an ethical question and under regulatory scrutiny.

Yet another area is the protection of minors especially in an era dominated by mobile and social networking. Social networks, mobile and other emerging mediums offer the possibility of pushing the boundaries of advertising to target kids. Again, this practice is not illegal (yet!), but it is certainly morally questionable. As per the *Guardian* newspaper’s[14] review of the book *Consumer*

Kids by Ed Mayo, which summarises the case study of seven-year-old Sarah, who has been recruited through Dubit.com to act as a brand ambassador for Mattel and promote her Barbie MP3 player to school friends. In exchange for keeping the sought-after shiny pink gadget, her job description includes creating a fan site where she blogs about the product, taking pictures of her sales missions and posting them back to Dubit, where she is rewarded.

As more and more mobile devices are able to purchase goods and services, extending the above discussion, we enter into the realm of the ethics of impulse purchasing. Impulse purchasing is not un-ethical in itself. Supermarkets, for instance, regularly encourage impulse purchases though product placements. However, with a mobile device, new problems could arise. Consider the example of the phone 'reminding' you to buy a related product. This would be based on 'opt-in' so it's not spam. So far, so good. At worst a minor irritation, at best a useful recommendation.

Now extend this further. Knowing the person/object they are looking at (based on location, e.g. they are standing in front of a car showroom) and their credit history (available on the web), can we offer a 'One Click' loan to 'engage' with the person and 'encourage' them to buy the car? Legally and technologically it is not banned. However, morally and ethically it may be considered dubious. Note that all this precise engagement and personalisation can be enabled by co-relating different datasets.

If we consider advertising: to what extent does advertising dictate content? It is an intriguing question and most media channels will deny that their content is influenced by advertising. However, there are indicators that this may be the case based on the limited and advertising-led range of content. For instance, advertisers would favour entertainment-led content since it places the viewer in a more receptive mood to buy, in contrast to the more serious documentary-based content (which does not). The question of profiles is also interesting and raises some questions.

For instance, consider the abstract of the following patent filed by Google (source: search engine journal)[15]. 'Personalized advertisements are provided to a user using a search engine to obtain documents relevant to a search query. The advertisements are personalized in response to a search profile that is derived from personalized search results. The search results are personalized based on a user profile of the user providing the query. The user profile describes interests of the user, and can be derived from a variety of sources, including prior search queries, prior search results, expressed interests, demographic, geographic, psychographic, and activity information.’

Such a profile would appear to be recording all our activities in cyberspace and tying them individually to us (to be used for the purposes of advertising). This practice does raise privacy, trust and risk concerns. On the other side are anonymised profiles which seek to anonymise personal data and then create 'templates' of user behaviour, which may be used to predict future behaviour based on past behaviour. For instance, it may be used to identify in advance who will churn (move off) from a social network. In this case, rather than getting an individual profile, we get audience segments. Audience segments are not tied to individuals (of course in a very small segment, for example, a segment of one, it would be a direct link).

When it comes to the mobile platform, the mobile operators generally have a good reputation for managing data and preventing misuse from advertisers. Misleading promotions, such as the Crazy Frog ringtone[16] in Europe were not created by telecom operators but rather by mobile marketing companies. Certainly, most operators take privacy seriously. Over time, mobile operators and the industry will face new challenges and will work with new forms of advertising as indicated in the discussion above. Whatever the direction we choose, 'mobile', due to its unique, personalised nature, will have to go beyond 'opt-in' and may need higher standards beyond statutory regulation based on moral and ethical integrity with a view to protect consumer interests. The future of privacy will lie in customer empowerment. Some of the mechanisms for privacy that we discuss later include:

- anonymisation;
- revocation;
- vendor relationship management; and
- full disclosure.

For marketers, the temptation to treat social media and the digital footprint as a 'channel' is strong, along with the desire to retrofit the new world of communication to the familiar world of brands, traffic, audiences, growth and value. However, this is not (always) in the consumer's interest. The pendulum of legislation will shift from an emphasis on brands to empowering the consumer, and the debate is just beginning.

Perceptions of privacy

It is interesting to see the public's reaction to the privacy issue. Pew International[17] gives some insights about customer perception. According to their analysis: 'Online adults can be divided into four categories based on their level of concern about their online information and whether or not they take steps to limit their online footprint:

- Confident Creatives are the smallest of the four groups, comprising 17% of online adults. They say they do not worry about the availability of their online data, and actively upload content, but still take steps to limit their personal information.
- The Concerned and Careful fret about the personal information available about them online and take steps to proactively limit their own online data. One in five online adults (21%) fall into this category.
- Despite being anxious about how much information is available about them, members of the Worried by the Wayside group do not actively limit their online information. This group contains 18% of online adults.
- The Unfazed and Inactive group is the largest of the four groups – 43% of online adults fall into this category. They neither worry about their personal information nor limit the amount of information that can be found out about them online.'

Thus, I see a range of perceptions of privacy and digital footprints. As we shall see going forward, the issues, benefits and perceptions are all going to change significantly in the near future; especially in relation to mobile.

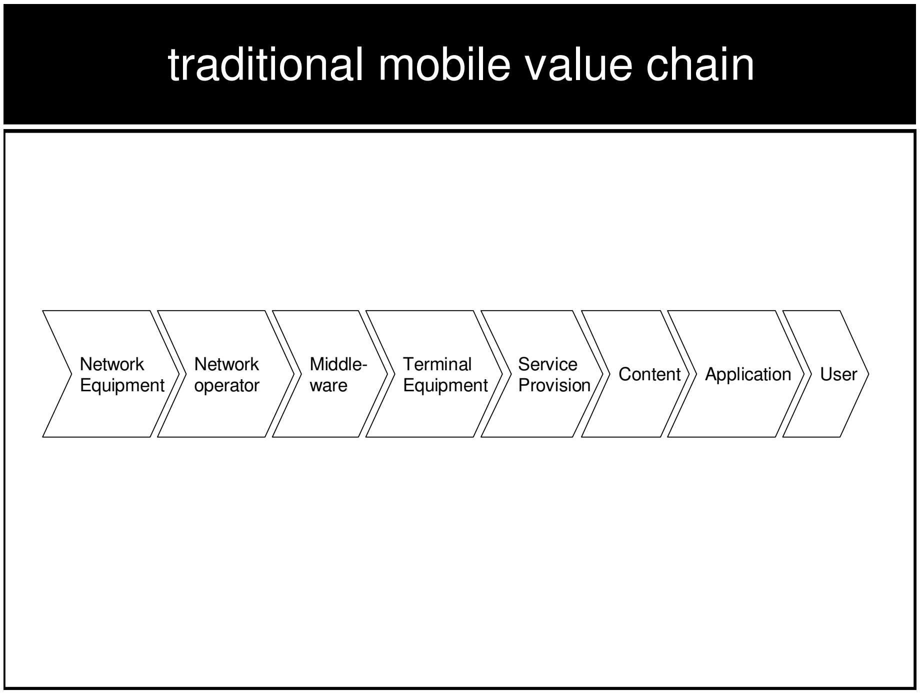
WEB 2.0 AND MOBILE WEB 2.0

We mentioned in the previous chapter that the mobile device will play an important part in the debate about the digital footprint. This chapter focuses on the implications of Web 2.0 on mobile but starts by reviewing the mobile value chain and then reflects on how this will change as value is migrating.

The traditional mobile value chain of the access industry is rooted in voice and the consumption of content. Thus, the main players are:

- network operators: who run the telecoms network;
- device manufacturers: who provide the device or set-top, which is often subsidised by the service provider/operator;
- service providers: who create a service;
- content creators: who sell the content often through the operator portal (e.g. ringtones) and directly for TV; and
- consumers: who consume content and services and primarily do not create professional content.

In this ecosystem, voice and text (SMS) are the predominant applications with content taking mainly the form of packaged content. The traditional value chain is depicted in Figure 9.



traditional mobile value chain

Network Equipment

Network operator

Middle-ware

Terminal Equipment

Service Provision

Content

Application

User

Figure 9 Traditional value chain

Since the rise of Web 2.0 and Mobile Web 2.0 (which is explained in detail below), we have seen a fundamental shift in the mindset of the customer, which affects the value chain of the industry. The main difference is: customers are now creators of content (point of inspiration) and not merely consumers of content (point of entertainment). Customers (and the mobile devices through which they interact with the network) are at the centre of the ecosystem and the value chain is no longer dominated by consumption but rather by creation, as shown in Figure 10. Access happens and users don’t care how it works as long as it does, it is the services that are available that drives choice.

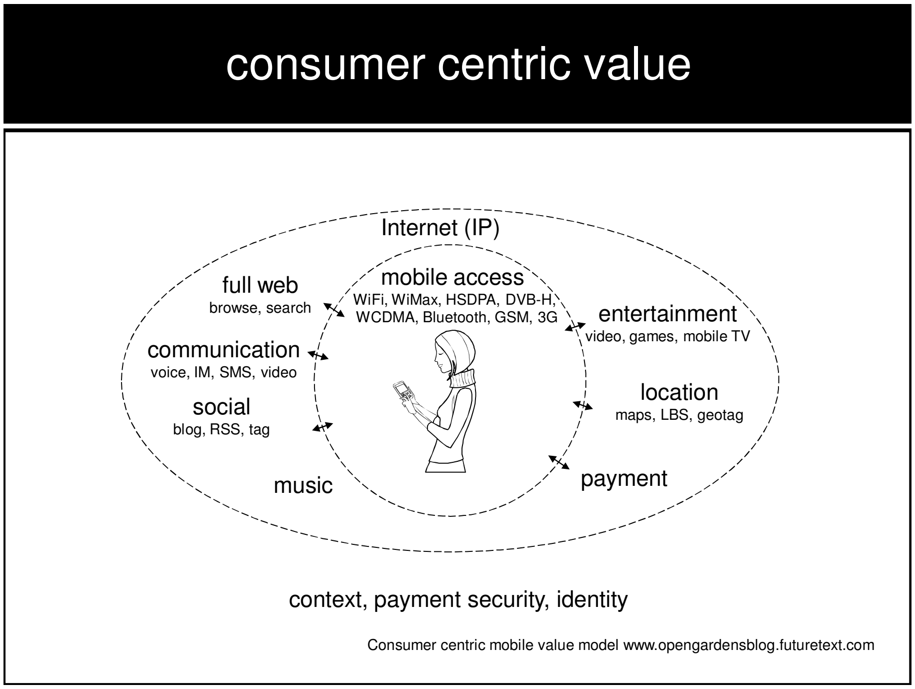
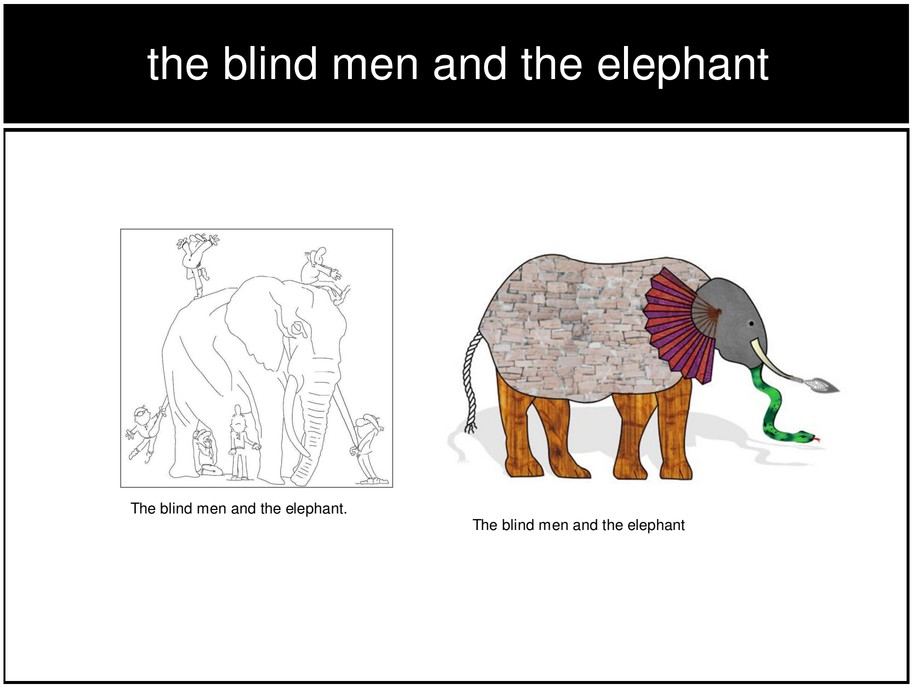


Figure 10 Consumer centric value chain

Web 2.0 – the creation web

In usage terms, the web took off globally in the mid- to late-1990s. Until the dot-com bust (2001), the web was primarily treated as a consumption medium. From 2002, we see the rise of Web 2.0, which is based on the ideas of social media and the creation web.

Web 2.0 was outlined by Tim O'Reilly in his seminal document in September 2005[18]. After its identification and naming it as a trend, Web 2.0 created considerable controversy, although now it is accepted globally. Many people take a one-dimensional view when it comes to Web 2.0 as with the story of the proverbial blind men and the elephant[19]; they look at only one facet of Web 2.0 and insist that it's the whole (see Figure 11). If the blind men looked at the elephant in isolation, they would not perceive it correctly (e.g. the elephant's trunk could be thought of as a snake and so on). It is only by considering the totality of all the aspects that they can reach the correct conclusion.



the blind men and the elephant

The blind men and the elephant.

The blind men and the elephant

Figure 11 The blind men and the elephant[20]

The seven principles of Web 2.0

What is Web 2.0? The long answer is: a service that follows all (or as many as possible) of the seven principles of Web 2.0. We will discuss a simpler definition later. These seven principles are outlined in Tim O'Reilly's original document[21] as per Figure 12.

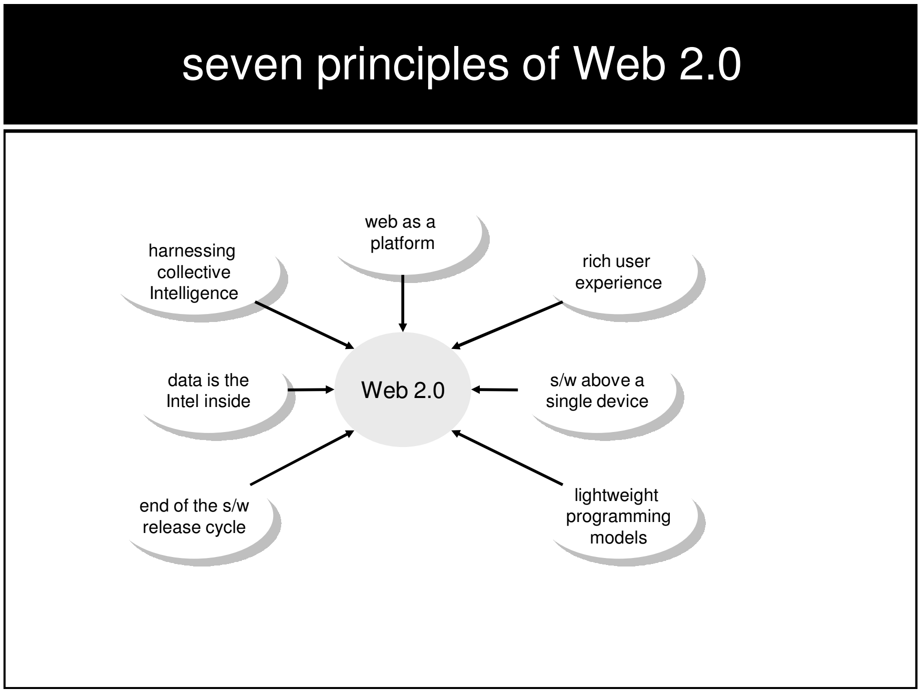

seven principles of Web 2.0
harnessing collective Intelligence
web as a platform
rich user experience
data is the Intel inside
Web 2.0
s/w above a single device
end of the s/w release cycle
lightweight programming models

Figure 12 Seven principles of Web 2.0

Principle one: web as a platform

Web 2.0 services make the fullest possible use of the web. The principle of ‘web as a platform’ encompasses the concept of Software As A Service (SAAS).

A Web 2.0 service is a combination of software and data. The term ‘web as a platform’ is not new. Netscape was an early example of using the web as a platform. However, Netscape used the web in context of the existing ecosystem: thus becoming the ‘web top’ instead of the prevailing ‘desktop’.

While Netscape was still ‘software’, in contrast, Google is ‘software plus a database’, right from the start. (In this context, we use the term ‘database’ generically to mean that ‘Google is managing some data’.) Individually, the software and the database are of limited value, but together they create a new type of service. In this context, the value of the software lies in being able to manage the (vast amounts of) data. The better it can do it, the more valuable the software becomes.

The term ‘long tail’ refers to the vast number of small sites that make up the web as opposed to the few ‘important’ sites. Harnessing the ‘long tail’ is illustrated by the ‘DoubleClick vs. AdSense/Overture’ example. The DoubleClick[22] business model was not based on harnessing the vast number of small sites. Instead, it relied on serving the needs of a few large sites (generally dictated by the media/advertising industry). In fact, their business model actively discouraged small sites through mechanisms like formal sales contracts. In contrast, anyone can set up an AdSense/Overture account easily. This makes it easier for the vast number of sites (long tail) to use the AdSense/Overture service.

In general, Web 2.0 services are geared to harnessing the power of a 'large number of casual users who often contribute data implicitly' as opposed to 'a small number of users who contribute explicitly'.

Tags are an example of implicit contribution. Thus, the Web 2.0 service must be geared to capturing 'many implicit/Metadata contributions from a large number of users' and not a small number of contributions from a few 'expert' users.

Principle two: harnessing collective intelligence

In this context, 'collective intelligence' can mean many things:

- Yahoo! as an aggregation of links;
- Google PageRank;
- blogging;
- tagging[23] and collective categorisation (e.g. Flickr[24] and del.icio.us[25]);
- eBay buyers and sellers;
- Amazon reviews; or
- Wikipedia.

And so on.

All of the above are examples of content created by users that collectively adds value to the service (which, as we have seen before, is a combination of the software and the data). In addition, harnessing collective intelligence involves understanding some other aspects like peer production, the wisdom of crowds and the network effect. Value is created from understanding the data and Metadata, and the connections.

Peer production is defined by Professor Yochai Benkler's [26] paper, *Peer Production*[27].

A concise definition from Wikipedia is:

‘A new model of economic production, different from both markets and firms, in which the creative energy of large numbers of people is coordinated (usually with the aid of the Internet) into large, meaningful projects, largely without traditional hierarchical organization or financial compensation.’

We discuss the motivations underlying peer production in the section on Maslow’s hierarchy of needs. The most prominent example of peer production is ‘Linux’. However, the same rationale also drives many individuals to contribute in a small way; for example, reviews on Amazon. Collectively, these small contributions lay the foundation for the ‘intelligence’ of Web 2.0, also called the ‘wisdom of crowds’.

The wisdom of crowds is fully discussed in *Wisdom of Crowds* by James Surowiecki[28]. The central idea of the wisdom of crowds is: large groups of people are smarter than an elite few, no matter how brilliant the elite few may be. The wisdom of crowds is better at solving problems, fostering innovation, coming to wise decisions, and even predicting the future. We discuss the principle of the wisdom of crowds in greater detail in the section on the unified definition of Web 2.0.

And finally, the network effect generated from user contribution. This is where a user can add value (knowledge) easily and subsequently for their contributions to flow seamlessly across the whole community, enriching or refining the whole body of knowledge. A collective brain/intelligence of the ‘web’ if you will, made possible by mechanisms such as RSS[29].

Principle three: data is the next ‘Intel inside’

Data is the key differentiator between a Web 2.0 service and a non-Web 2.0 service. A Web 2.0 service always combines function (software) and data (which is managed by the software). Thus, Web 2.0 services inevitably have a body of data (Amazon reviews, eBay products and sellers, Google links, etc). This is very different to a word processor for example, which comprises of only software (and no data).

While data is valuable, the company need not necessarily own the data. Although in most cases the company serving the data (e.g. Google) also ‘owns’ the data (e.g. information about links), that may not always be the case. In the case of Google Maps[30], Google does not own the data. Mapping data is often owned by companies such as NavTeq[31] and satellite imagery data is owned by companies such as Digital Globe[32]. Google Maps simply combines data from these two sources. Such combination of data from two or more sources is called a 'mashup'. A mashup is defined as a website or web application that seamlessly combines content from more than one source into an integrated experience[33]. A mashup could be seen as a 'web API' (Application Programming Interface).

Taking the 'chain of data' further, sites such as Housing Maps[34] are a mashup between Google Maps and Craigslist[35]. The more difficult it is to create the data, the more valuable the data is (e.g. satellite images are obviously valuable). At the time of writing, there are still grey areas in the ownership and creation of mashups. For instance, if a company makes its data available for 'mashing up', does it control the functionality of ultimate mashup itself? Conflicts could arise if the mashup creator does not use the service in a way approved by the body releasing the API. Nevertheless, there is a significant industry momentum behind mashups.

For clarity, data only becomes information and then knowledge through analysis and processing, and crucially, when it is in the context of a business process which has an economic value.

Principle four: end of the software release cycle

Web 2.0 services do not have a software release cycle. While Google re-indexes its link indices every day, Microsoft releases a major software release every few years. That's because there is no 'data' in Windows 95, Windows XP, etc. It's pure software. Not so with Google. Google is data plus software. It has to re-index its 'data' every day, otherwise it loses its value. Thus, operations are critical to a Web 2.0 company and there is no 'software release' as such. The flip side of this coin is, there are widespread beta releases and users are treated as co-developers.

Principle five: lightweight programming models

Web 2.0 services could be seen as a lighter form of SOA (Service Oriented Architecture). Functionally, a Web 2.0 service acts as a distributed application. Distributed applications have always been complex to design. However, they are central to the web. Web services (SOA) were deemed to be the ideal mechanism to create distributed applications easily. But, web services, in their full incarnation using the SOAP[36] (Simple Object Access Protocol) stack, are relatively complex. RSS is a simpler (and quicker) way to achieve much of the functionality of web services.

Simpler technologies, such as RSS and Ajax (explained in detail later), are the driving force behind Web 2.0 services. These lightweight technologies are designed to ‘syndicate’ rather than ‘orchestrate’ (orchestration is one of the goals of web services). Because lightweight programming models are oriented towards syndicating data, they are contrary to the traditional corporate mindset of controlling access to data. They are also designed for reuse. In this context, ‘reuse’ indicates ‘reusing the service’ and not ‘reusing the data’ (they make it easier to remix the service through mashups). As a result of this architecture, grassroots innovation is given a boost because a new service can be created using existing services through mashups.

Principle six: software above the level of a single device

In essence, Mobile Web 2.0 is about the sixth principle. Hence, we do not discuss it in detail here.

Principle seven: a rich user experience

For the web to be truly useful, we need a mechanism to improve user experience. In comparison to platforms such as Windows, the web offered a relatively limited user experience. Technologies like ActiveX and Java applets attempted to improve the user experience, but these were proprietary. The main technological driver for an enhanced user experience on the web is Ajax. Ajax uses web technologies (non-proprietary technologies). Ajax was outlined by Jesse James Garrett in an article published online[37]. Ajax is being used in services like Gmail, Google Maps and Flickr, and it already provides the technology to create a seamless user experience combining many discrete services.

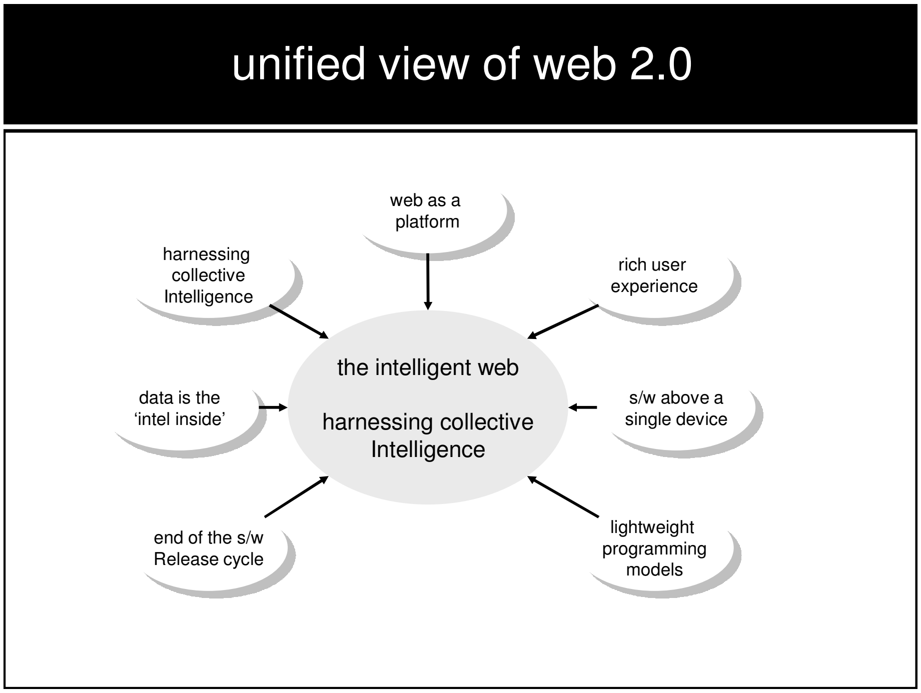
Web 2.0 – a unified view based on harnessing collective intelligence

Having discussed the seven principles of Web 2.0, let us now look at a simpler definition of Web 2.0. We call this definition the ‘unified view’ of Web 2.0.

If we reconsider the seven principles, we observe that the second principle (harnessing collective intelligence) encompasses the other six.

Thus, we can view Web 2.0 as ‘harnessing collective intelligence’ or ‘the intelligent web’. What kind of intelligence can be attributed to Web 2.0? How is it different from Web 1.0? This is shown in Figure 13.

Web 1.0 was hijacked by the marketers, advertisers and the people who wanted to stuff canned content down our throats. The dot.com bubble was the end for many who took the approach of broadcast content. What’s left is the web as it was originally meant to be – a global means of communication.



unified view of web 2.0

web as a platform

harnessing collective Intelligence

rich user experience

the intelligent web

harnessing collective Intelligence

data is the ‘intel inside’

s/w above a single device

end of the s/w Release cycle

lightweight programming models

Figure 13 Unified view of Web 2.0

The intelligence attributed to the web (Web 2.0) arises from us (the collective/people) as we begin to communicate. Thus, when we talk of the 'intelligent web' or 'harnessing collective intelligence' we are talking of the familiar principle of the ‘wisdom of crowds’. In order to harness collective intelligence:

- information must flow freely;
- it must be harnessed/processed in some way – otherwise it remains a collection of opinions and not knowledge; and
- from a commercial standpoint, there must be a way to monetise the 'long tail'.

Our essential argument is: if we consider Web 2.0 as 'intelligent web' or 'harnessing collective intelligence' (principle two), and then look at the other six principles feeding into it, Web 2.0 is a lot clearer. Since the wisdom of crowds is so important, let's consider that in more detail. From the Wikipedia entry for the wisdom of crowds[38] which is a wonderful irony, are all crowds wise? No. They are not. The four elements required to form a 'wise' crowd are:

- diversity of opinion;
- independence: people's opinions aren't determined by the opinions of those around them;
- decentralisation: people are able to specialise and draw on local knowledge; and
- aggregation: some mechanism exists for turning private judgements into a collective decision.

Conversely, the wisdom of crowds fails when:

- decision-making is too centralised: the Columbia shuttle disaster occurred because the hierarchical management at NASA was closed to the wisdom of low-level engineers;
- decision-making is too divided: the US intelligence community failed to prevent the September 11, 2001 attacks partly because information held by one subdivision was not accessible by another; and
- decision-making is imitative – choices are visible and there are a few strong decision-makers who, in effect, influence the crowd.

Now, based on this background, let's look at the seven principles again.

The web as a platform

The web is the only true link that unites us (those with access) all together, whoever or wherever we are in the world. To harness collective intelligence and to create the intelligent web, we need to include as many people as we can. The only way we can do this is to treat the web as a platform and use open standards. You can't harness collective intelligence using the IBM ESA/39039, no matter how powerful it is.

Harnessing collective intelligence

This now becomes the 'main' principle or the first principle.

Data is the next ‘Intel inside’

By definition, to harness collective intelligence, we must have the capacity to process massive amounts of data. Hence, data is the 'intelligence' (Intel).

End of the software release cycle

This pertains to SAAS (Software As A Service). Software as a 'product' can never keep up-to-date with all the changing information. A Web 2.0 service includes code as well as data. Thus, SAAS keeps the data relevant (and the harnessed decision accurate) by accessing as many sources as possible. In some examples the data may change and the APIs that expose that data will have a release cycle. Enterprise customers continue to demand release processes.

Lightweight programming models

The heavyweight programming models catered for the few. In contrast, using lightweight programming models, we can reach many more people. Hence, we are working with many more sources of information, leading to an intelligent web, for example, from the seven principles[40].

Amazon.com's web services are provided in two forms: one adhering to the formalisms of the SOAP web services stack, the other simply providing XML data over HTTP; in a lightweight approach sometimes referred to as REST (Representational State Transfer). While high value B2B connections (like those between Amazon and retail partners such as ToysRUs) use the SOAP stack, Amazon reports that 95% of the usage is of the lightweight REST service.

Software above the level of a single device

More devices to capture information and better flow of information between these devices leads to a higher degree of collective intelligence.

Rich user experiences

A rich user experience is necessary to enable better web applications leading to more web usage and better information flow on the web: leading to a more 'intelligent' web.

Web 2.0 – summary

- What is Web 2.0? It's the intelligent web (the second principle: harnessing collective intelligence).
- What makes it intelligent? We do.
- How does it happen? By harnessing collective intelligence.
- What do you need to harness collective intelligence? The other six principles.

Mobile Web 2.0

Based on our understanding of Web 2.0, let us consider the implications of extending the definitions of Web 2.0 to Mobile Web 2.0. As we have seen previously: Mobile Web 2.0 is focused on the user as the creator and consumer of content 'at the point of inspiration' and the mobile device as the means to harness collective intelligence. Mobile Web 2.0 is happening on a set of 'restricted devices' that have more limited functionality than a PC and we need to think specifically about which 'restricted devices' are important to Mobile Web 2.0 as the seemingly simple idea of extending Web to Mobile has many facets, for instance:

- What is a restricted device?
- What are the implications of extending the web to restricted devices?
- As devices become creators and not mere consumers of information – what categories of intelligence can be captured/harnessed from restricted devices?
- What is the impact for services as devices start using the web as a massive information repository and the PC as a local cache where services can be configured?

Mobile Web 2.0 is happening on a set of 'restricted devices' that have more limited functionality than a PC and we need to think specifically about which 'restricted devices' are important to Mobile Web 2.0.

A broad definition of a 'restricted device' is not easy. The only thing they all have in common is – 'they are battery-driven'. But then, watches tend to have batteries. A better definition of restricted devices can be formulated by incorporating Barbara Ballard's carry principle[41], which introduces the concept of an information device. Using the carry principle, a restricted device could now be deemed as:

- carried by the user;
- battery-driven;
- small footprint (by definition);
- probably multifunctional but with a primary focus;
- a device with limited input mechanisms (small keyboard);
- personal and personalised; but
- not wearable (that rules out the watch). But, there is a caveat, a mobile device in the future could be wearable and its capacities may well be beyond what we imagine today. The input mechanisms in the future may not be keyboards. So, this is an evolving definition.

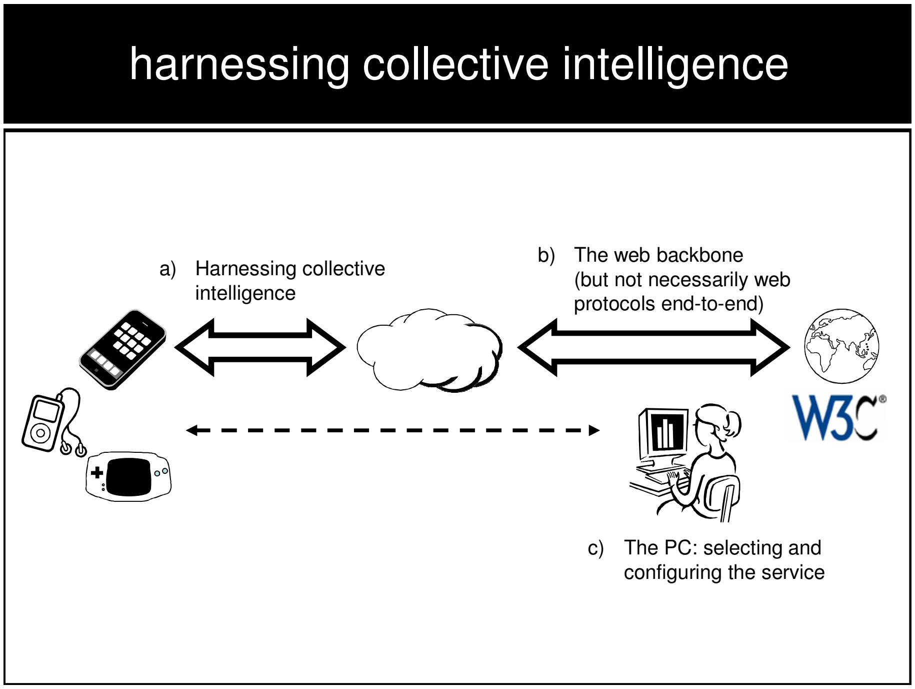
Finally, there are differences between a small screen, ‘carry/portable’ device such as mobile phones and a device which is ‘carried’, such as in-flight entertainment, which we call a ‘carried/ported’ device. For example, in a car, a GPS navigator is a ‘small screen-based mobile device’ and in a plane, the in-flight entertainment screen is also ‘mobile’. However, both these devices are not ‘carried by a person’ and do not have the same screen/power restrictions as devices that people carry. However, whichever way you look at it, it's clear that the mobile phone is an example of a restricted device. From now on, we use the definition of mobile devices interchangeably with ‘restricted devices’ and the meaning will be clearer in the context.

Extending the web to restricted devices

It may seem obvious, but Web 2.0 is all about the ‘web’ because it could not have been possible without it. Thus, in a ‘pure’ definition, Web 2.0 is about ‘harnessing collective intelligence via the web’. When we extend this definition to ‘Mobile Web 2.0’ – there are two implications:

- the web does not necessarily extend to mobile devices; and
- even though the web does not extend to mobile devices, intelligence can still be captured from mobile devices.

The seven principles of Web 2.0 speak of this accurately when they discuss the example of the iPod/iTunes. The iPod uses the web as a back-end and the PC as a local cache. In this sense, the service is ‘driven by the Web and configured at the PC’ but it is not strictly a ‘web’ application because it is not driven by web protocols end-to-end (iPod/iTunes internals are proprietary to Apple) as shown in Figure 14.



harnessing collective intelligence

a) Harnessing collective intelligence

b) The web backbone (but not necessarily web protocols end-to-end)

c) The PC: selecting and configuring the service

Figure 14 Harnessing collective intelligence from mobile devices

Thus, the characteristics (distinguishing principles) of Mobile Web 2.0 are:

- harnessing collective intelligence through restricted devices (a two-way flow where people carrying devices become reporters rather than mere consumers);
- driven by the web backbone, but not necessarily based on web protocols end-to-end; and
- use of the PC as a local cache/configuration mechanism where the service will be selected and configured.

MY DIGITAL FOOTPRINT is not Web 3.0[42], it is firmly in the domain of Web 2.0 as the value is still centred on harnessing collective intelligence. The next phase of the web will be the move to intelligence in the network (Semantic Web); **MY DIGITAL FOOTPRINT** is only improving how it works today.

Mobile Web 2.0 – value lies in getting data out from a device

We have discussed Mobile Web 2.0 and Web 2.0 extensively above. We have seen that Web 2.0 could be viewed as harnessing collective intelligence and, by extension, Mobile Web 2.0 could be viewed as harnessing collective intelligence from mobile devices. This is depicted in Figure 15.

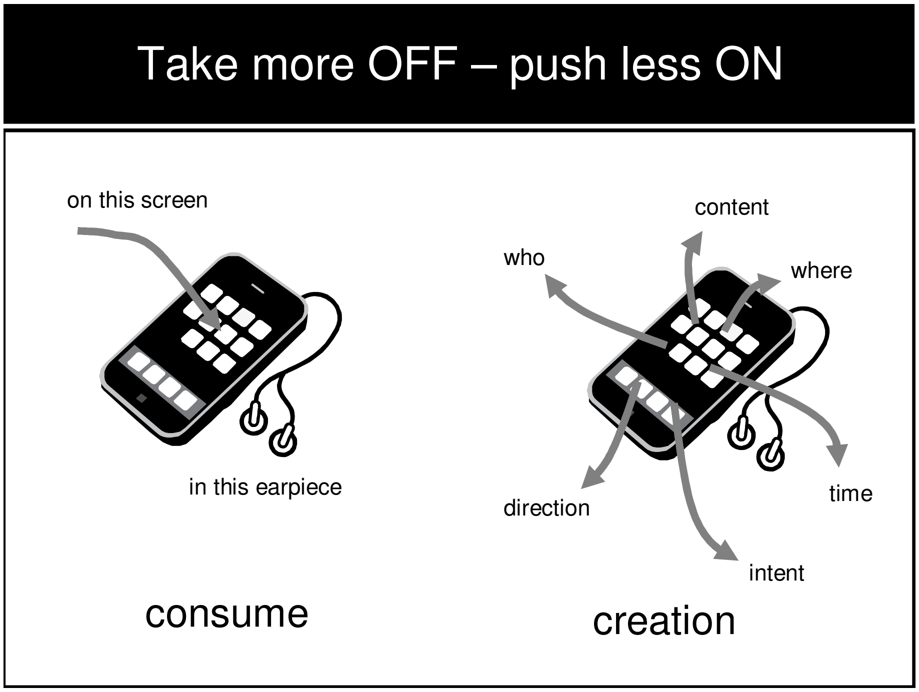

Take more OFF – push less ON

Figure 15 Moving focus to getting more data off a device than on to it

The ability to get data out of or off a mobile device lends itself to the unique advantage a mobile device has.

We explore this idea in greater detail in subsequent chapters.

Considering this concept, that there is more value in getting data off a mobile, let's consider that sensors (acceleration, temperature, noise level) can easily be placed in or attached to mobiles. Further, a user can send information from their device, by voice, IM (Instant Message) or text, to a centralised service point. Both sensors and people can provide vital data during a disaster-relief operation or outbreak of a disease, for example, could not be gathered. It is known that aid agencies are building systems that use handsets to sense, monitor and even predict population movements, environmental hazards and public-health threats. InSTEDD (Innovative Support to Emergencies, Diseases and Disasters), a non-profit group based in California USA, focuses on the use of mobile-gathered data to improve developing countries' ability to respond in emergencies. Funded with seed money from Google's philanthropic arm, it has released a suite of open-source software to share, aggregate and analyse data from mobile phones. It is being used in anger in Cambodia, where health-workers can send an SMS, of observations and diagnoses, to a central number.

FootPath, a system developed by Path Intelligence (UK), aggregates and analyses signals picked up from mobiles as people move through a particular area. The data can be used to optimise the flow of pedestrians through high density areas, such as railway stations and airports. Data can determine, for example, whether customers visit a specific shop. This will be linked to marketing at some point to close the loop.

Dr Alex Pentland, at MIT, describes 'X-raying entire organisations, cities and countries' by collecting data in the two ways described: passive (no user intervention) and active (user interaction). Dr Pentland's algorithms can already cluster information from thousands of mobiles and divide people into 'tribes' of like-minded folk. He calls this 'reality mining', something I explore later as the 'rainbow of trust'. Dr Pentland's company, Sense Networks, is working with Vodafone and other collaborators to build an early-warning system for modelling and predicting the spread of tuberculosis in South Africa.

There is more value in getting data off the mobile!

MY DIGITAL FOOTPRINT

To reiterate the previous argument, the idea of **MY DIGITAL FOOTPRINT** extends the idea of raw data to the wider concept of capture, store, analysis and value created from data generated through digital engagement. This process is based on a structured approach incorporating inputs and outputs, and a feedback loop that governs the whole process. This feedback loop progressively enriches and refines the outputs (value) over time. The analysis phase is able to take raw data from various sources (which I refer to as the digital footprint) and generate value in the form of services, such as personalisation, reputation or discovery – the output from the analysis process I call 'behavioural DNA'. The value derived from the process is **MY DIGITAL FOOTPRINT**.

There are two central ideas which underpin this book (and this section): the feedback loop which enriches the digital footprint and the role of the mobile device in enriching the value from **MY DIGITAL FOOTPRINT**.

Identity is not a digital footprint, but a digital footprint has to be related to a digital identity in order to be interpreted and given value of some kind. It must also be related (at some point) to an identity in order to unlock its potential for control and contribution from the user.

The Nobel Prize winner with a poor reputation

Let us first understand the difference between identity and reputation from an imaginary anecdote.

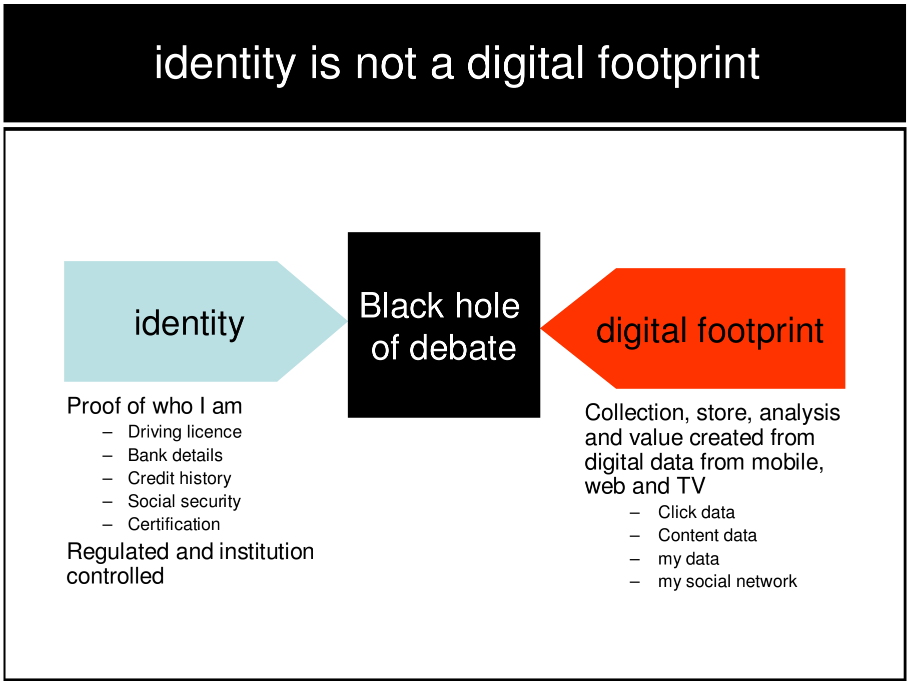
Many social network sites depend on explicit reputation mechanisms. These can be implemented in many ways, such as the members rating each other 'good' or providing testimonials. However, such explicit recommendations have limitations and can be ‘gamed’. We illustrate the drawbacks of explicit reputation by considering the problem of the ‘Nobel Prize winner with a poor reputation’.

If a person’s ‘reputation’ depends on updating someone’s site (or depends on my friends updating the site) then it could mean: if I win the Nobel Prize (which presumably means I am the best in my field), but I forget to update that site, then my reputation is low, irrespective of my achievements.

This is a simplistic example but it illustrates the concept. Reputation should be implicit and should be reflected by your actions (which in turn are automatically captured in transactions/data), in other words, reputation should be a by-product of activities on a network.

Identity and digital footprints

The above example illustrates the difference between identity and reputation – in that identity (a passport, an ID card or even the Nobel Prize) is explicitly conferred by an external body and is recognised within its jurisdiction. Reputation (or digital footprint) is a much softer concept with a different use case as presented in the earlier work. Identity is based on documents, such as a driving licence, bank details, social security, certificates, etc. The digital footprint on the other hand is captured in the forms of data streams such as click data, content data, my data and social data – generated both by the individual but also by the social context of that individual as per Figure 16.



identity is not a digital footprint

identity

Black hole of debate

digital footprint

Proof of who I am

- Driving licence
- Bank details
- Credit history
- Social security
- Certification

Regulated and institution controlled

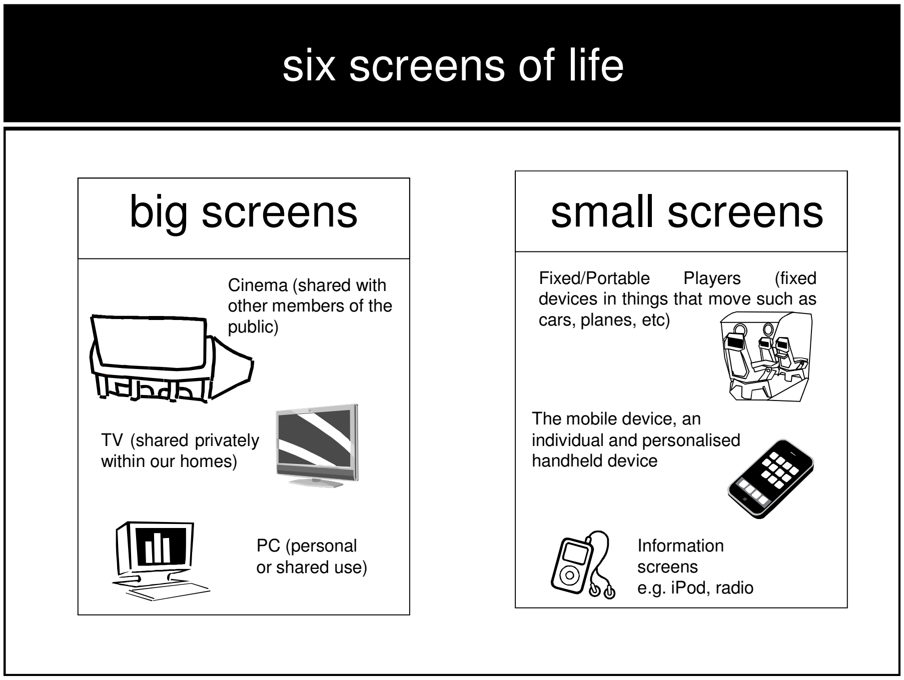
Collection, store, analysis and value created from digital data from mobile, web and TV

- Click data
- Content data
- my data
- my social network

Figure 16 Identity is not a digital footprint

The six screens of life

For the most part, we are consumers of content. In our daily lives we consume professionally created, produced and edited content from traditional and new media providers on our ‘six screens of life’. These screens are divided into two broad categories, big screens and small screens, each with three subgroups as show in Figure 17.



six screens of life

big screens

Cinema (shared with other members of the public)

TV (shared privately within our homes)

PC (personal or shared use)

small screens

Fixed/Portable Players (fixed devices in things that move such as cars, planes, etc)

The mobile device, an individual and personalised handheld device

Information screens e.g. iPod, radio

Figure 17 The six screens of life (for digital engagement)

Both for big and small screens, the user has traditionally been a passive receiver of content (content has been broadcast to the user) or the user has been seen as a member of a carefully controlled and managed audience (e.g. voting) – but not as a primary creator of content. For instance:

- both TV and cinema need users to consume (view); and
- a website needs users to consume/interact in most cases.

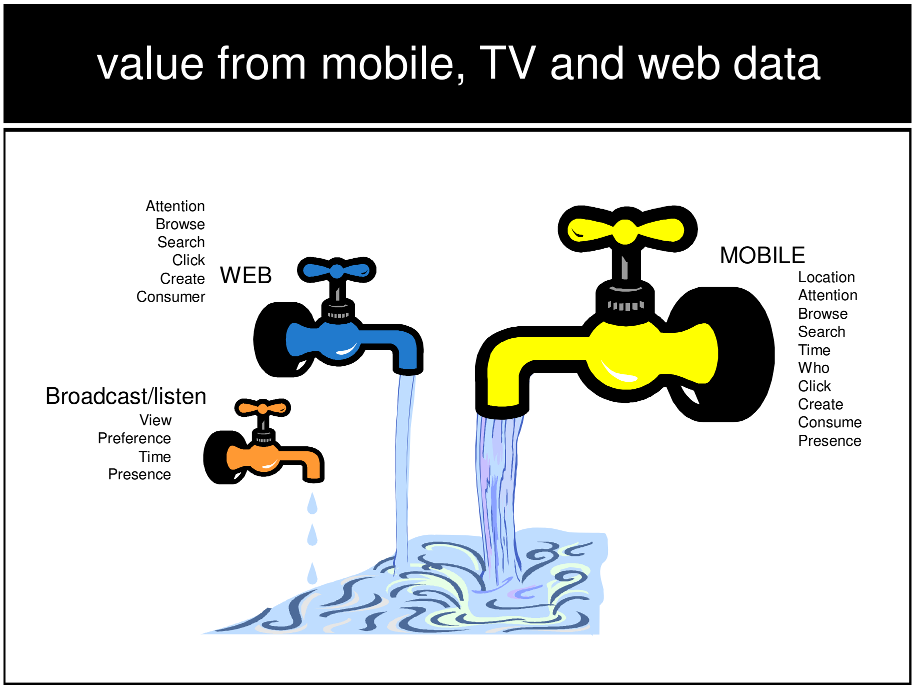
However, according to Forrester, “Monolithic blocks of eyeballs are gone. In their place is a perpetually shifting mosaic of audience micro-segments that forces marketers to play an endless game of audience hide and seek.” As advertisers lose the ability to invade the home, they will have to wait for invitations, and this means they have to learn how to adopt and understand the user, a good reason for understanding the impact of digital footprint.

Reflecting the above trend, most of the content on the mobile device to date has also been the ‘re-presentation’ and ‘reproduction’ of existing material delivered to the mobile screen. The mobile device, however, is changing from being a primary consumer to a major creator of content. It is worth noting at this point that the separate screens of life don’t have to be managed or offered by separate service providers, devices will no longer control what you can do and where, instead the screen will be able to come under the control of the user, with services relevant to the size of the screen. However, least we forget that the interactive age of communicates and social media delivers diversity and innovation, broadcast amplifies.

The click streams of life

Whilst the broadcast medium dealt with consumers in isolation, solitude and separation, Web 2.0 has brought relationship, engagement and conversation and now have a greater influence of the medium they want to create and consume.

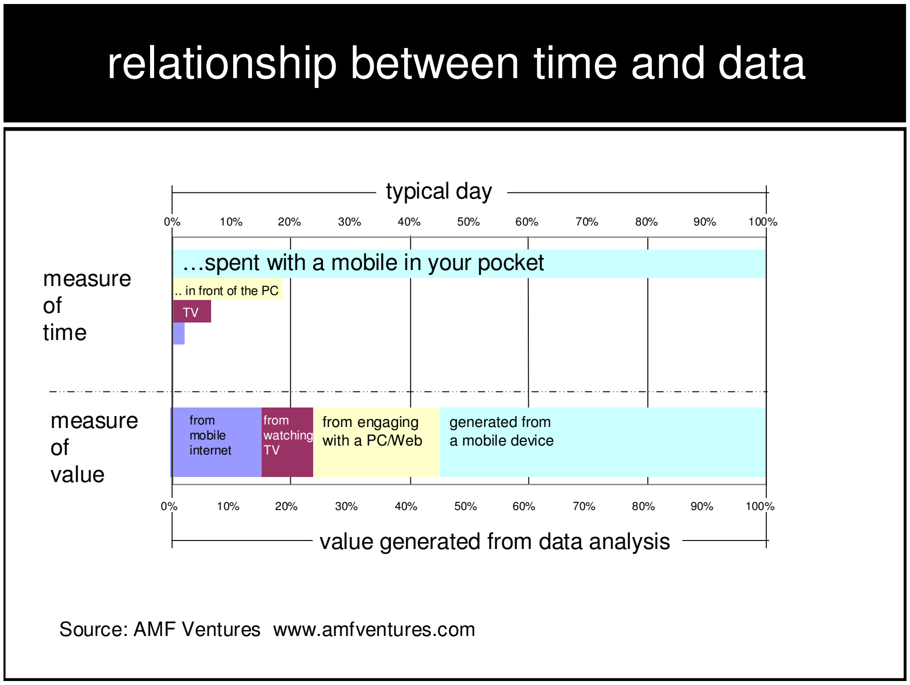
Related to the idea of the 'screens of life' is the concept of 'click streams of life'. By this, we mean the ability of each platform to capture increasingly greater information (as opposed to merely consume it). For instance, web captures attention, browsing patterns, search patterns, click information, content creation, content consumption, as per Figure 18. TV captures only viewing and time preferences. Mobile captures a lot more, for instance: location, attention, browsing patterns, search patterns, time, who we consumed that information with, click stream information, content creation and consumption patterns, etc.



value from mobile, TV and web data

Figure 18 Data types from mobile, web and TV

In addition, there is a proportional relationship between the time we spend with the mobile device and the data captured by the device, as shown in Figure 19.



Source: AMF Ventures www.amfventures.com

Figure 19 Relationship between time and data

Figure 19 shows two results: one for the amount of time a user spends with a certain screen of life (TV, mobile and PC), and one as a measure of value generated from the same devices. The top graph shows that we tend to have the mobile on and with us for the majority of our waking hours. However, there is evidence that young people are rejecting this model and choosing to leave the mobile at ‘home’ to avoid parental control; which is ironic, as in some original research by Norman Lewis for Orange in 2004, young people adopted the mobile as it was a place where there was no paternal control. The lower graph presents a measure of value generated from the different devices and our interaction, shown as a normalised percentage. This assumes that all the value generated from mobile, web and TV creates 100% of value. About 15–17% will be generated from data from our interaction with the mobile Internet, even though this will only account for 2 or 3% of time on our chosen devices. A further 5–7% will be generated from the analysis of data that we will provide from our interaction with our TV (including voting). Watching TV accounts for an average of 17% of the total available time. Some 20% of total value will come from the analysis of data from the web, leaving a balance of about 52–55% of all value to be created from data that our mobile devices pick up. This is data that our mobile generates without the user doing anything. The battleground is who can collect this data, who can store it, who can analyse it and who can create value from it?

Converged click streams, Wikipedia button on my TV remote

Related to the two ideas of greater amounts of information being captured by mobile devices and the increasing amounts of time that we spend with mobile devices is the concept of how information captured from mobile devices can be used? Using the web analogy of mashups, information captured on one platform can be combined with data from another service to create a new application or service and a new dataset.

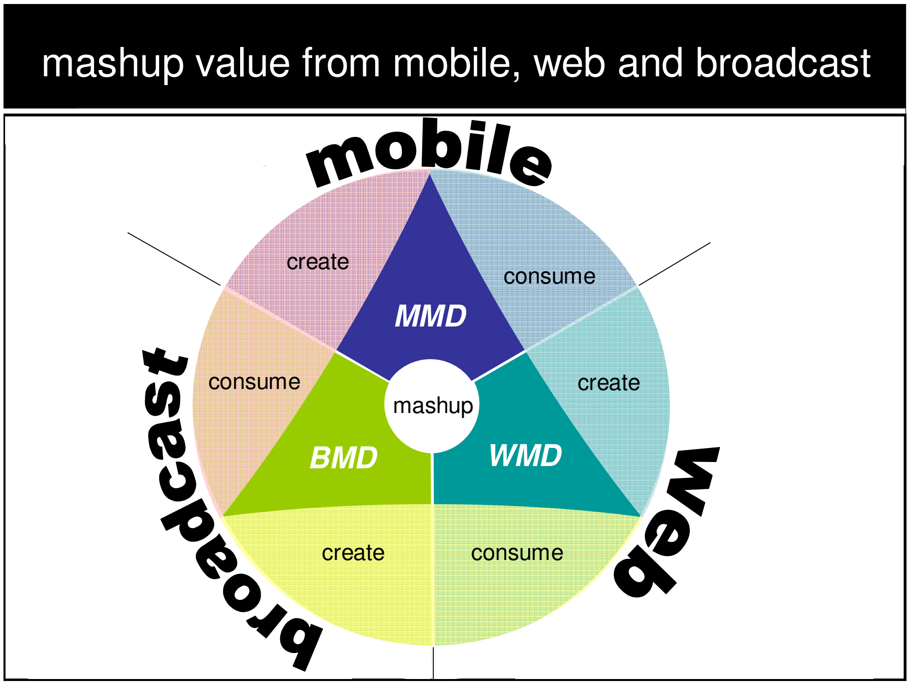


Figure 20 Value from the mashup of data from mobile, web and TV

In Figure 20, it is assumed that each of the platforms (web, mobile and broadcast) have a 'create' and 'consume' component. MMD is Mobile Metadata, WMD is Web Metadata and BMD is Broadcast Metadata. In each of the cases the platform Metadata comprises patterns about consummation and creation. The cross platform concept of a mashup is the analysis of these datasets together. The output or value can be that the service that started on one platform is improved on that platform, not only by that platform's own Metadata, but as a result of Metadata from another platform, and it also includes the possibility that an experience that starts on one platform improves a service on another. Additional user data such as NFC cards (e.g. Oyster in London) and non-web financial transactions can be added either by the mobile device being a component of the actual transaction, or being used as a sensor, or the user may download the data and do a manual add. However, given that the mobile will have location, NFC data for travel may not actually add any new information.

Mashups are not new and the idea of cross-platform mashups is a natural extension of the mashup concept. Initial synergies between platforms are based on relatively simple structured mechanisms, such as SMS voting on TV shows, (voice-based) fixed to mobile convergence, etc. However, a deeper integration means that data gathered from one platform can enhance services on another platform, in our case mobile data enriching a web platform. This deep integration of data is happening slowly and causes a conflict in business models as some are based on access or subscription whilst others on advertising and product sales.

Most traditional mediums, such as TV, are trying to incorporate some form of 'controlled interactivity'; for instance, SMS voting. Interactivity is interesting but it is old stuff, especially if it is 'managed' or ignored in the interest of editorial prowess. A truly converged medium would be ‘non-linear’ and boundaries between platforms would blur. For instance, the web has caused most people to absorb knowledge in a non-linear manner. By extension, a truly useful feature would be a Wikipedia button on our premium subscription TV remote. Traditional media will not allow it since they like (need) linearity for the advertisement business model. They fear that I may 'go away' – which I well might. But that's how the new mind may work; the user will return to the provider who allows freedom. From a digital footprint point of view, as customers begin asking for such converged services, it creates an opportunity to create truly integrated services based on understanding the digital footprint. We explore these ideas in greater detail below.

After you have trodden this path come the analysts and the anthropologists

As we have discussed so far, we all create digital footprints as we engage with digital platforms. Platforms like mobile devices will create a larger share of that footprint. Digital footprints will be cross-platform and will be ‘mashed up’ across platforms (for the lack of a better word). We have extended the idea of 'digital footprints' to '**MY DIGITAL FOOTPRINTS**'. The concept of **MY DIGITAL FOOTPRINT** is therefore complex, but this book suggests it is a system and process for the 'collection, store, analysis and value created from digital data from mobile, web and TV'.

Storage and analysis of digital footprints raises some important questions. Who analyses the digital footprint? Who stores it? What value is derived from the process and for whom?

Humans have always left traces of their activity. The oldest human footprints found date back to about 3.6 million years ago at Laetoli, Tanzania[43]. The ancient human beings who left those footprints would not have known that 3.6 million years later, we modern humans, would analyse them, photograph them, categorise them and draw new insights from those footprints about the people who created them. The difference now is that there is no need to wait 3.6 million years. As soon as those digital footprints are created, there is a host of online companies analysing the cookie crumbs and immediately creating new insights (to be used for commercial reasons).

As we discussed previously, harnessing collective intelligence is the key idea behind Web 2.0. This is not a problem in itself, but the dark side arises if a business entices its audience (customers, clients, delegates, patients, friends) to give up their digital data, collect their digital footprint without their agreement, charge people to view their own data, or sell our data off with the sole expectation of making cash though the one-sided route of exploitation.

On the other hand, the value of digital footprints lies in using the analysis of data and to complement services. However, this use (exploitation) is likened to the most fundamental components of digital identity, that of risk, privacy and trust. These three inter-related components bond and bridge all the characteristics of a digital footprint and identity, as we will see later in business models.

Figure 21 pictorially shows the aspects of the model that will be explored in the next chapter. It shows that on one side there is the collection of data these are the inputs (click data, content data, my data and social data) from web, TV and mobile. These inputs create the digital footprints (raw data) which are stored. This stored digital footprint is analysed to create behavioural DNA. This is what your data says about you. This behavioural analysis is then used to create output, such as service discovery, service improvement and trade or barter.

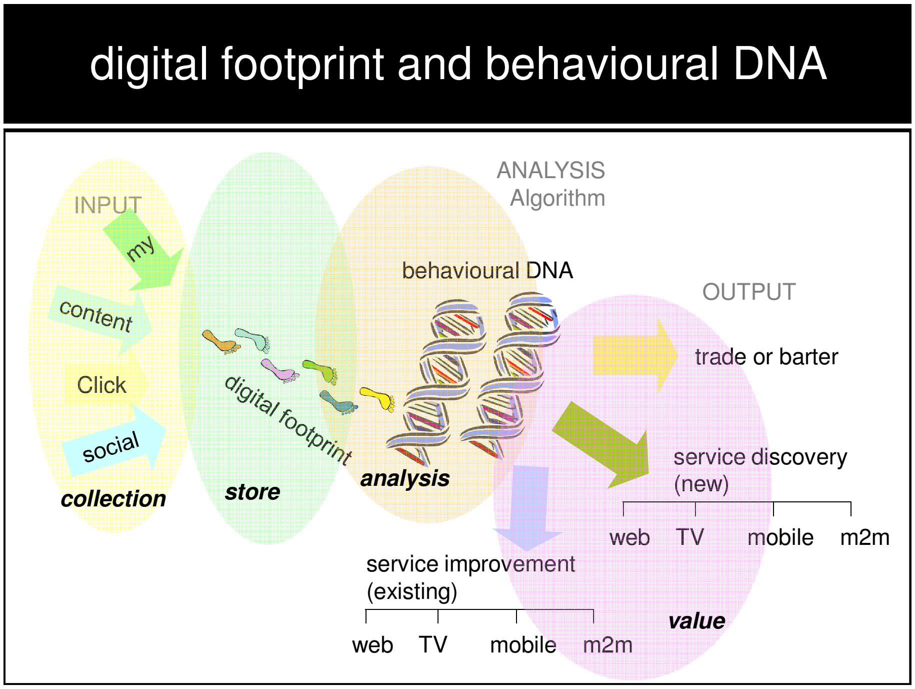

digital footprint and behavioural DNA

***Figure 21 Components of* MY DIGITAL FOOTPRINT**

The connection between the inputs and the outputs is the algorithm.

As previously mentioned, the algorithm is the component that creates the value; the outputs are how that value is realised. The algorithm that computes, combines, compares and analyses the digital footprint is the differentiator for a service provider. A good algorithm can produce success, a poor one can bring a company down. Whilst a company can implement the same algorithm, the way it is presented to the community will also lead to success or failure. This provides the bridges and bonds to risk, trust and privacy and how governance and the culture of the company, led by the CEO, will bring some brands down and others to new heights. Considering the algorithm is important. It is a very complex component and the part of the process that will bring differentiation.

As shown in Figure 22, there is a comparison between the credit card algorithm and what a Web 2.0 **MY DIGITAL FOOTPRINT** algorithm would look like. On the left is a traditional credit card algorithm. Original data collected from transactions was used to build an algorithm for the purpose of predicting if a transaction is fraudulent. When you sign-up and start to use a new credit card, normalised data provides a traditional pattern of spending for your income group in your location and your profession (the behavioural model). Over time this is complemented with your own data, which sets up triggers and thresholds. If one of these triggers or thresholds is broken, there is a simple action of alert and a person steps in to determine the next action. This is a complex algorithm based on a lot of historical data, which works very well but has a singular function – fraud detection. Yes, the same datasets are used on another system applying another algorithm and determine if you pay on time, if you should get increased credit and if you may purchase new financial products, but the core of the value is to reduce fraud. The norms do provide a very good prediction, it is socially acceptable to do this and we enjoy the value that comes from giving up this data and the rights to it.

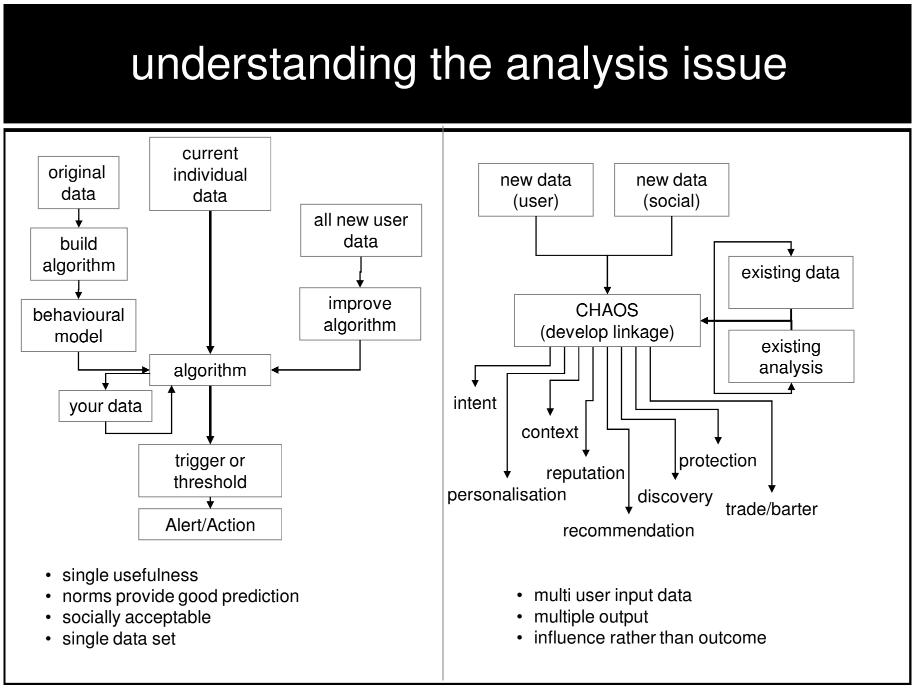


Figure 22 Understanding the analysis issue

On the right is an entirely new and more complex algorithm. It is not data mining principally as the output has moved from a trigger or threshold to prediction of intent, providing: personalisation (filters), delivering context at a specific moment in real-time, determining reputation, providing recommendations, pushing discovery, adding protection, and adding an ability to trade and barter. In this case, your data and others from social groups is combined, compared, contrasted and, through a 'chaos' algorithm, develops linkages, bridges and bonds to deliver value. It also takes the immediacy of the output reaction to the value back into the system to refine it; a continuum of feedback, honing and improving services and applications. As we will see in the business models chapter, my feedback produces focus and depth, and my social groups data produces colour and breath; components in real life that are at odds with each other. This is why the algorithm is complex, but is the part that will deliver differentiation.

Of course not all digital footprints are created equal, both in terms of the person to whom it is related and the actual value of different data types that could be collected. Whilst it is obvious that some people are worth more to certain brands (lifestyle and segmentation works), it is less obvious which types of data have value and why it varies. Some intrinsic value in **MY DIGITAL FOOTPRINT**, the output of the analysis stage, lies in how difficult it is to change or swop provider of a service. If a provider is applying good analysis tools and delivering value through the output mechanisms, it is hoped that the propensity to change/swop/migrate/move/try will be reduced. This means lower customer churn, higher retention, lower cost of customer acquisition and gaining access to more of the lifetime Net Present Value (NPV) of the customer. In this context, the lifetime NPV value of a customer translates into the longer a customer is with you, the higher the value you as a business have achieved from that customer, which will affect your share price. Therefore, what data types are there and which data could affect the ability of a brand to keep a customer loyal? In another direction the question could be, what data is hard or takes a long time to capture and would benefit the outcomes from analysis in a positive way, so as to ensure customers are more loyal?

As an example, let's review the types of data a mobile operator could have of its customers and see if any of the data could be used to make you more loyal. Figure 23 provides an overview.

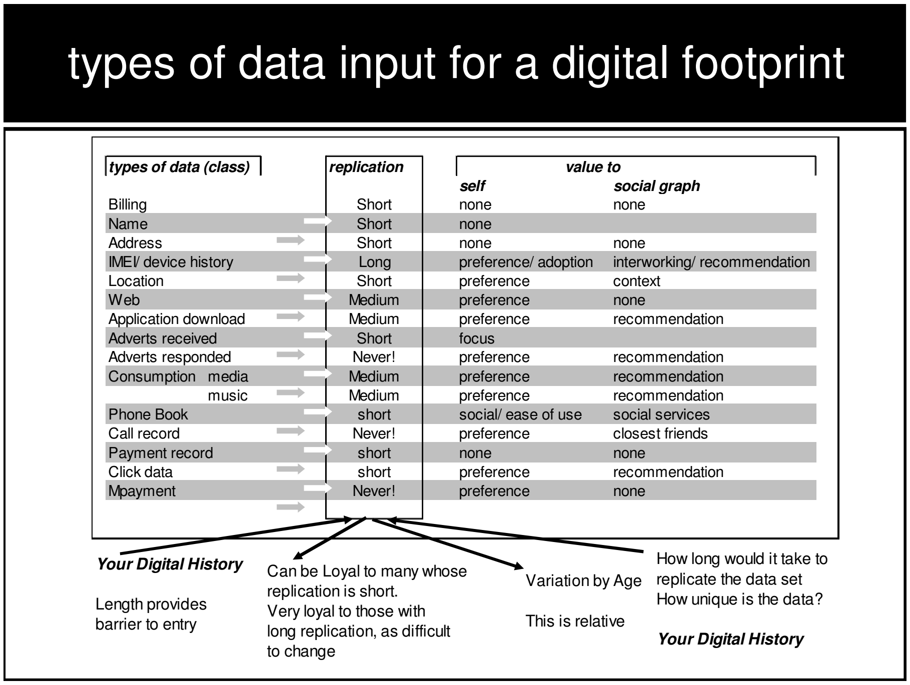
The left-hand column shows the types of data that a mobile operator could collect, starting with the obvious ones on billing (monthly invoicing or pre-paid spend pattern), user's name, and billing (home) address. In reality these data types are very fast to replicate and, for an operator, it provides little competitive advantage. Another operator would be able to replicate this very quickly (in a short period). There is some value from the analysis of this data (classic segmentation) but the value is biased to the operator.

Looking down the list, the majority of data that a mobile operator can collect is quick to replicate and would take a short period for someone else (another interested service provider) to collect the same data if the customer churned. Some data would take a little longer to replicate, such as your mobile web browsing history, your application download history (assuming an operator-controlled portal) and what media you consume. These latter data types, if analysed with a ‘good’ algorithm (one that creates value for the user), will provide some additional loyalty or stickiness.

The operator challenge, however, is that they don’t have access (under existing terms and conditions) to most of these data classes, other than snooping, as users go off portal and depend on the operator as an access (IP) provider. Some data types in the list would take a long time to replicate, an example would be the IEMI/device history. However, whilst it would take a long time to replicate, there is very limited value in knowing this data.

Not all data, even though it can be collected, has value. Indeed, the operator does have access to some data that will not be/cannot be replicated, such as adverts responded to, past call record patterns and payments used by SMS or m-payment (assuming the operator was partly to the transaction). However, whilst such non-replication data types could produce loyalty and value, these classes of data are both held by other third parties and, on their own, have limited value. Knowing you used the m-payment service for a ticket once is not probably as valuable as the data your bank has on you, as the user’s bank (transaction provider) can see all transactions.

Certain types of data are very hard to replicate as it takes a long time (and hence could be very valuable), but this should not be assumed. Indeed, from this view, the mobile operator will find it difficult to retain customers and improve loyalty as they own and have access to low value data types.



types of data input for a digital footprint

types of data (class)	*replication*	*value to*	
		self	***social graph***
Billing	Short	none	none
Name	Short	none	
Address	Short	none	none
IMEI/ device history	Long	preference/ adoption	interworking/ recommendation
Location	Short	preference	context
Web	Medium	preference	none
Application download	Medium	preference	recommendation
Adverts received	Short	focus	
Adverts responded	Never!	preference	recommendation
Consumption media	Medium	preference	recommendation
music	Medium	preference	recommendation
Phone Book	short	social/ ease of use	social services
Call record	Never!	preference	closest friends
Payment record	short	none	none
Click data	short	preference	recommendation
Mpayment	Never!	preference	none

Your Digital History

Length provides barrier to entry

Can be Loyal to many whose replication is short.
Very loyal to those with long replication, as difficult to change

Variation by Age

This is relative

How long would it take to replicate the data set
How unique is the data?

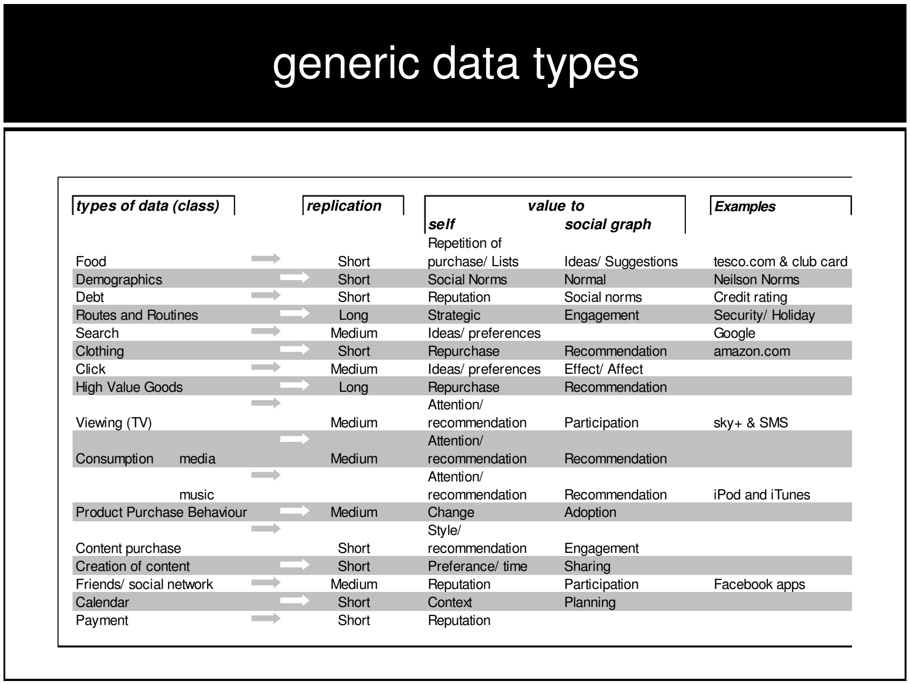
Your Digital History

Figure 23 Types of data inputs that make up a digital footprint

Creation of services therefore requires the user/customer to share data with the service provider. In sharing the data with the service provider, the services should become better, more accurate and more efficient, as this is the exchange for giving up some value, the customer should receive value. However, not all services are equal. Specifically services which have a higher cost of data acquisition or take longer to acquire the data will be seen to be more valuable as they present the opportunity to create barriers to entry.

It is worth a brief look at some other data types of a more generic nature to determine if there are any obvious data types that could create loyalty. Figure 24 provides such an overview. Again, the left-hand column shows different classes of data types, such as food, clothes and content purchased. It is possible to see that most to the generic data types are quick to replicate. By this it is meant it would not take long for two major food retailers to have the same profile of a user based on till/online spend, unless the user is very discerning and only buys specialist or selective foods from each store. The stores, however, know this as they can compare individual spend to the norm of other buyers in their store. Some data types, such as web searches, browsing, TV viewing and media consumption, take longer to pick up and hence are slower to replicate. It would take a generic medium length of time to replicate the data between two providers of the service. This starts to open up the opportunity for a provider to add value and hold the user. Therefore, a good predictive algorithm could help one provider make a user more loyal, and gain that additional slower-to-replicate data, as long as the user perceives value has been added. Very slow-to-replicate data, which takes a long time to build, can have significant value. However, some legislative jurisdictions are seeking companies to only hold data for a short period, which would block out the value from this data class. Ignoring these possible hurdles, high value goods are only purchased every several years, it is useful to know when re-purchase (new or preventative) is required or to offer additional insurance.

Another class of data that takes time to build is routes and routines. This class of data is about where you have been and what you do. To generate this data class from the source of location takes time. In a chapter later in the book we look at how this data class (routes and routines) could be used to offer a secure mobile service, based not on passwords but on your behaviour.



generic data types

types of data (class)		*replication*	*value to*		*Examples*
			self	*social graph*	
Food	→	Short	Repetition of purchase/ Lists	Ideas/ Suggestions	tesco.com & club card
Demographics	→	Short	Social Norms	Normal	Neilson Norms
Debt	→	Short	Reputation	Social norms	Credit rating
Routes and Routines	→	Long	Strategic	Engagement	Security/ Holiday
Search	→	Medium	Ideas/ preferences		Google
Clothing	→	Short	Repurchase	Recommendation	amazon.com
Click	→	Medium	Ideas/ preferences	Effect/ Affect	
High Value Goods	→	Long	Repurchase	Recommendation	
Viewing (TV)	→	Medium	Attention/ recommendation	Participation	sky+ & SMS
Consumption media	→	Medium	Attention/ recommendation	Recommendation	
music	→		Attention/ recommendation	Recommendation	iPod and iTunes
Product Purchase Behaviour	→	Medium	Change	Adoption	
Content purchase	→	Short	Style/ recommendation	Engagement	
Creation of content	→	Short	Preferance/ time	Sharing	
Friends/ social network	→	Medium	Reputation	Participation	Facebook apps
Calendar	→	Short	Context	Planning	
Payment	→	Short	Reputation		

Figure 24 Generic data types and comments on value

This chapter has focused on **MY DIGITAL FOOTPRINT**'s reputation and identity, and how our interactions with the six screens of life will both help generate data for our digital footprints (raw data) and also how value will be made through cross-platform integration, co-operation and convergence of services. Finally, we have seen that not all digital footprint data has equal value and just collecting data does not create any competitive advantage, but will introduce serious governance issues.

Fish tails, a model to categorise data

Having reviewed which types of data it is possible to collect and if value can be reached, this section provides a model to look at these data types in a framework and suggests two forms of classification, in addition to the idea of replication time. The first observation is how the data arrives if collected, the second is a classification. These tools I use to help companies undertake a simple data survey and determine where value will be generated.

As already mentioned, not all data is created equal in terms of its value, further not all data is created in the same way, but it is even more complex than this as the value of data changes depending on the context. An online bill for payment in 30 days is less important (has less value) than that same online bill when the deadline is just 1 day. However, location data is a continuous feed, food is short bursts of information, either daily for lunch or once a week for a household. High value goods and services, holiday, cars and a tax accountant present a small amount of data very infrequently. Content and music both provide two types of data, one-off dataset related to purchase data as previously mentioned and the other dataset is usage or consumption, which varies by time. Some of these data types are presented in Figure 25.

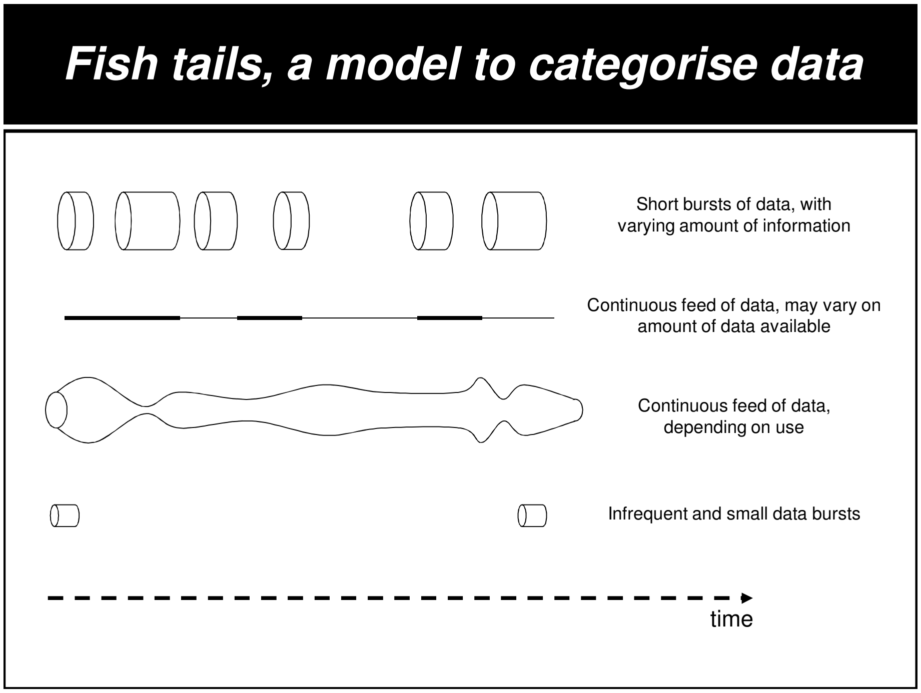
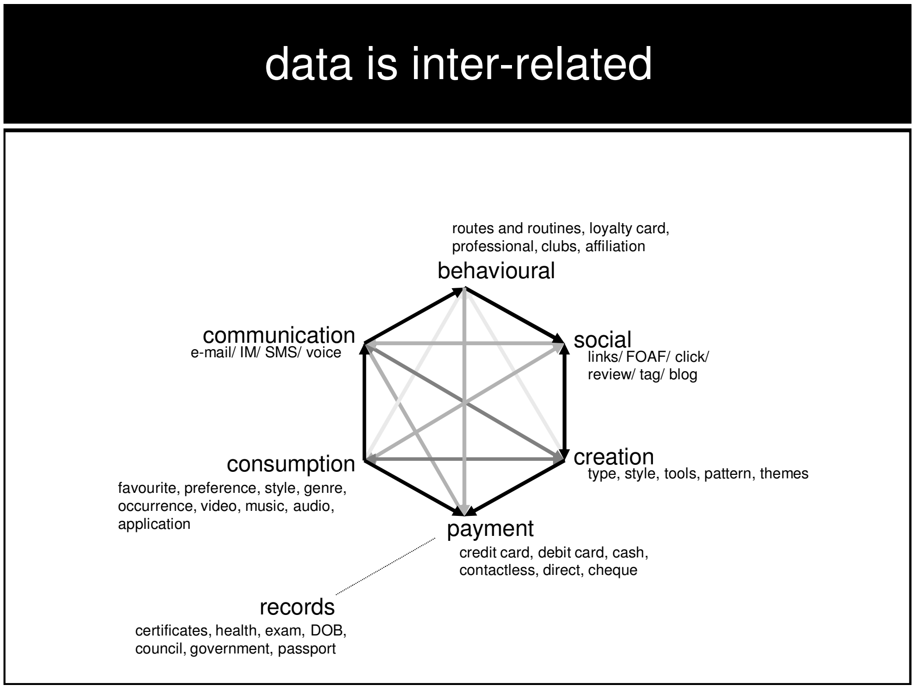

Fish tails, a model to categorise data
Short bursts of data, with varying amount of information
Continuous feed of data, may vary on amount of data available
Continuous feed of data, depending on use
Infrequent and small data bursts
time

Figure 25 Fish tails, a model to categorise data

Taking this obvious view that data is created in different ways, when this creation process is mapped onto the knowledge that data types are all inter-related it becomes clear, or very confusing, that the inter-relationships, bonds and dependencies mean that many companies with different data collection capabilities can reach the same value at outputs. By this I mean, two companies can have different datasets from the same user, but reach the same conclusion via analysis as to where value lies. This presents an interesting dilemma for application companies – should I ask the user for their home address or should I find it from records, depending on a relationship to a social group and behavioural data. However, if I show that I know your home address, will that spook the user? This is why trust, risk and privacy are so important. As an applications company, who should I buy data from and should I reveal my data sources? Figure 26 shows some of these inter-relationships between data types and, as can be seen, collecting some types of data will revel other material facts; however, gaining access to all of that class of data could be difficult, especially payment. Increasingly with a move to mobile payment, micropayment and near field pre-pay card payment, no one provider has all the data, indeed it could be important that the user is able to collect this data class and offer it through an open API to other providers as it will reveal all of the payment behaviours.



data is inter-related

routes and routines, loyalty card, professional, clubs, affiliation

behavioural

communication

e-mail/ IM/ SMS/ voice

social

links/ FOAF/ click/ review/ tag/ blog

consumption

favourite, preference, style, genre, occurrence, video, music, audio, application

creation

type, style, tools, pattern, themes

payment

credit card, debit card, cash, contactless, direct, cheque

records

certificates, health, exam, DOB, council, government, passport

Figure 26 ***Data is inter-related***

I do like the work in progress by Leafar[44] and Fred Cavazza's[45] 2006 framework for mapping our digital identity and, to quote some translated work, ‘increasingly, footprints appear on the digital sands over which we don't exercise any control: people blog about you, take and publish pictures, and if someone searches for your name in Google or GoogleBlogsearch or Clusty or IceRocket or Technorati, up comes whatever comes, and that's also defining your digital identity – each one of those search results composes one of your digital personas, giving information and hints about what you do, your character, your opinions, your network, in short, who you are. And there are more: just think of Second Life avatars.’ This work is a good framework for items we create about ourselves, but not for data created by others on us, via mashups or automated data from sensors.

Therefore, data collection from the screens of life (mobile, web and TV) is in some ways easy, from both explicit and active (the user providing) and implicit and passive (automatic via sensors) sources. However, not all data has equal value and having some data can create barriers to entry and loyalty. In the long term, data collection will be a commodity business and data will be traded. The value will lie not in the position to collect but in the analysis and the algorithm that drives the outputs and value. Today’s simple CRM and data mining is not up to it; however, the skills that underpin them will bring the advantage.

MY DIGITAL FOOTPRINT *AND THE TWO-SIDED BUSINESS MODEL*

Having introduced ideas in previous chapters, it's time to work through the conversion of concept to business model. The basis of the model is that it is bit-based (electronic ones and zeros) not an atom-based model (physical things you can touch). The model will be developed using the user as both the provider of the data and as the consumer of the analysis in the form of value. Surrounding this provision and consumption is a feedback loop that allows for the refinement and improvement of services. This creates a two-sided business model with the user on both sides of the business as shown in Figure 27. The two-sided model described here should not be confused with the one which has two revenue streams – one from the consumer and one from the corporate, where the intermediaries sell consumer data to a corporate, generating a new revenue stream.

In a traditional commerce business model there is no immediate direct or indirect feedback loop. Some improvement is possible from customer feedback and market research, but this is not in real-time. In the model we explore in this book the feedback is immediate, explicit and provider specific. This basic loop model is rapidly developing based solely on personal feedback (direct) but is extending to encompass social group feedback (indirect). In both cases of feedback, services are enhanced as the business knows more about us and our communities' reaction through the feedback. Direct feedback is somewhat easier to understand as it is our action or reaction to certain services that are directly offered. Indirect feedback is the reaction to that service due to our influence on our social groups and our social group influencing us. Indeed, our influence on our community can be compared to the norms of others' reactions to determine how we are influenced by, or influence, our community.

two-sided digital business

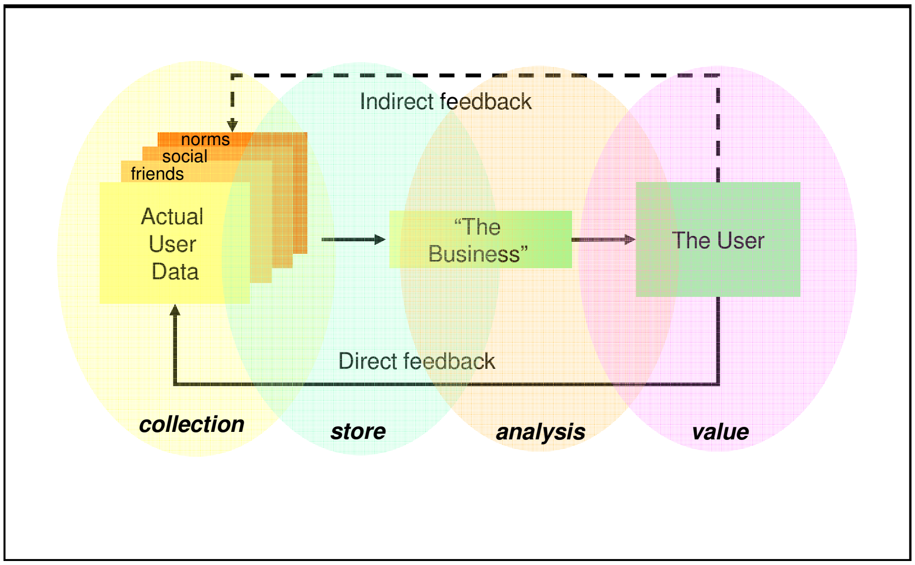


Figure 27 The two-sided business model

It is possible now to understand that data collected from a user builds a profile, which is used by a business. As long as that business can now collect data on how that user uses and interacts with the service, there is now a feedback loop that closes the system. This closed loop brings stability to the service improvement as there is direct and immediate feedback, making the perpetual Beta model of continuous improvement even more attractive. This closed direct feedback loop is enhanced if further data can be collected from the user's social cloud, friends and norms. This model is now contrasted to a view of the traditional open system business model and how the two worlds will collide.

Figure 28 provides two models, a traditional business model on the left and the closed loop business model on the right. Whilst I address these models as unique or stand alone, in reality they co-exist and they will increasingly become co-dependent, but the motivations and drivers will remain. The left-hand side of the diagram, the traditional approach, is founded and protected by law and policy. Traditional business will not change as law, regulation and governance determine that this is currently how business will operate. The motivation for policy enforcement is to control and limit the user's action within the context of the provider's requirements, and to ensure that businesses know who they are dealing with, both for B2C and B2B. As this model is controlled by law, the inputs for this process are based on data collected to provide proof that you are who you say you are, for example, log-in (approved), authentication (prove again), proof (original documentation) and trust (that you as a business can control the input against fraud). On the basis of this information, the provider is able to validate an identity for the customer (i.e. the provider 'knows you', you are who you are, or you want to repeat using a service). The identity validates against the company's own internal systems. Having now established you are who you say you are, you can now access the value (services you want), these are the outputs. Examples of value or outputs are access to services, banking, content purchase, being able to allow a provider to bill you, etc. The beneficiary of this entire process is the 'corporate'. The reason being that this is set up to ensure the directors fulfil their obligations, fiduciary duties and meet good governance rules. Without this policy, enforcement, proof, identity and value model the corporate world would grind to a halt. We accept it as it is the way we do business. This first model on the left-hand side of the equation is the classic Identity Management system.

Before we consider the right-hand side of Figure 28, there is a summary on how the identity model works.

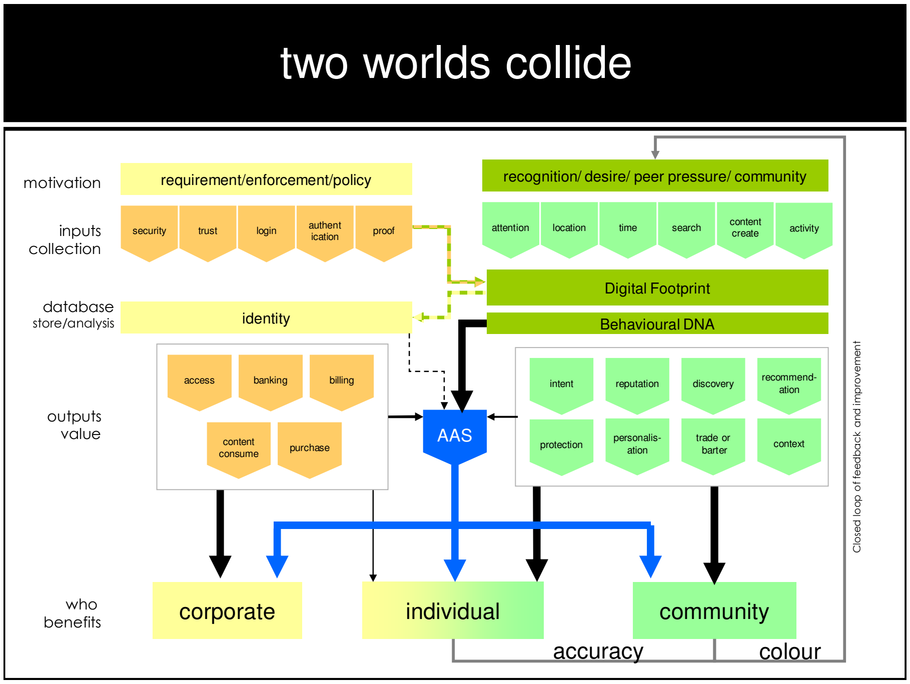


Figure 28 Two worlds competing for the same two-sided model

Identity – a telecoms perspective

The following section is based on *Identity and Security* by Rakesh Radhakrishnan[46]. The discussion below is outlined from the perspective of a telecoms provider. I expect that some readers may skip this section as it is more technical but it adds an important component to the overall story, which is a key difference between OpenID and IMS-based identity is that the first is user-centric and relatively weak, the second is device–centric and relatively strong. IMS can provide an identity layer in the network and this could be critically important in how networks are designed in the future.

Traditionally, networks were not integrated and were built and operated in silos, as depicted in Figure 29. In that scenario, limited insights can be gathered from data and also each system needs its own identity services.

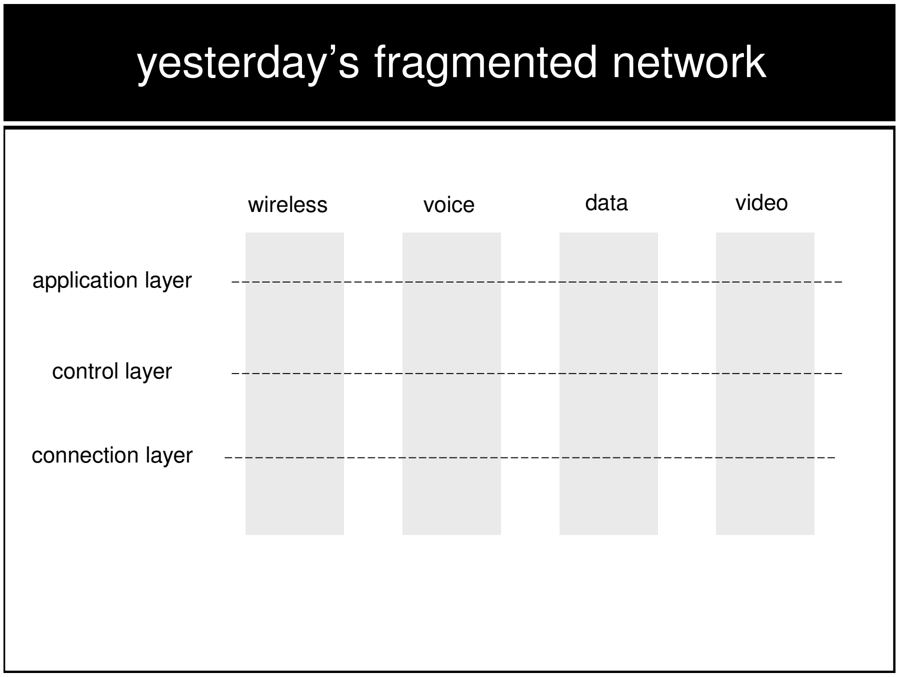


Figure 29 Identity-enabled IMS services

In contrast, today's networks are based on IP (Internet Protocol) as per Figure 30. More importantly, they have a single access and connection layer, which should allow for a user to use all services by one log-in (user ID and password or SIM).

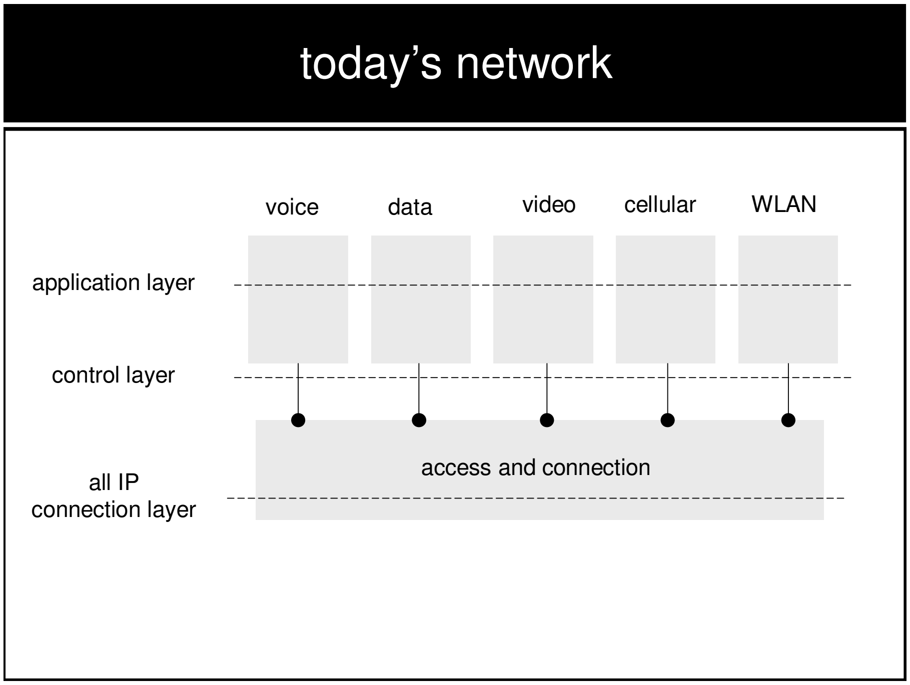


Figure 30 Today's network

In an all-IP world, the telecoms network has the following subcomponents which comprise the identity network as shown in Figure 31:

- An user identity /login management ecosystem based on schemes like OpenId[47], Liberty Alliance[48] etc;
- IMS – IP Multimedia Subsystem[49], which is concerned with deploying the IP protocol in the telecommunications system;
- SIP – Session Initiation Protocol[50] to manage the session;; and
- HSS – Home Subscriber Server to manage the profile information.

In this scenario, identity becomes a telecoms service with SIP to manage sessions, HSS to hold profile information and all tied to the phone, which in turn ties back to a federated identity system.

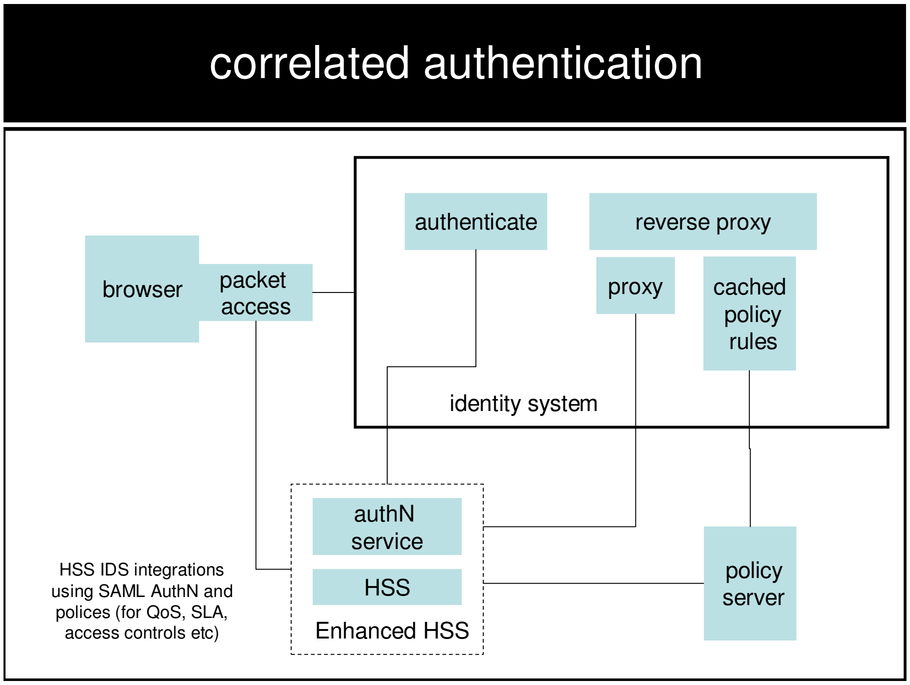


Figure 31 Identity and security from a user perspective

From a user perspective, the basic technical validity of the client (requestor) device involves checks for patches, security fixes, updates, worms, viruses, thefts and attacks (e.g. spyware and phishing) on the client device. The user is then authenticated by the provider including bio-metrics, certificates, etc.

Once the client and the individual are validated, a secure session is established. Note that the authentication and level of authentication are only relevant and valid for a given session and, sometimes restrictions are applied within the session as well.

Within the context of a session, we also have:

- Roles: What rights does the user have?
- Rules: Who can do what?
- Resources: The target of Roles and Rules (the content or the service being accessed);
- Federation: Sharing beyond the regulatory entity;
- Regulation: Conformance to government regulation; and
- Logging, feedback and intervention.

From a telco perspective, this mechanism is based on IMS. IMS is a standardised Next Generation Networking (NGN) architecture for telecom operators that want to provide mobile and fixed multimedia services. It uses a Voice-over-IP (VoIP) implementation based on a 3GPP and a standardised implementation of SIP, and runs over the standard Internet Protocol (IP). Existing phone systems (both packet-switched and circuit-switched) are supported. The aim of IMS is not only to provide new services but all the services, current and future, that the Internet provides. Users will be able to execute all their services when roaming as well as from their home networks. A multimedia session between two IMS users, between an IMS user and a user on the Internet, and between two users on the Internet is possible.

IMS merges the Internet with the cellular (mobile) world; it uses cellular technologies to provide ubiquitous access and Internet technologies, to provide appealing services. IMS helps the Network evolve. User identity takes on a much greater role in an IMS environment and the HSS (Home Subscriber Service) is at the centre. It’s helpful to think of the HSS as the nexus between the IT and the core network worlds. HSS acts as the central repository for user-related information such as security information (who am I?), location information (where am I calling from?) and user profile information (what services am I subscribed to?) to name a few. The session is a way to provide call control and manage multimedia sessions over IP networks. Today, when you make a phone call, a connection is established and cut off the instant you hang-up or attempt to do something else, like send an SMS or take a picture. There’s no ability to combine these services together. The objective, going forward, is we need a single IP-based structure that combines the services based on IP.

The HSS, or User Profile Server Function (UPSF), is a master user database that supports the IMS network entities that actually handle calls. It contains the subscription-related information (user profiles), performs authentication and authorisation of the user, and can provide information about the user's physical location.

From the user perspective, a mechanism like OpenId will be used as a decentralised single sign-on system. Using OpenID-enabled sites, web users do not need to remember traditional authentication tokens, such as username and password. Instead, they only need to be previously registered on a website with an OpenID ‘identity provider’, sometimes called an i-broker. Since OpenID is decentralised, any website can employ OpenID software as a way for users to sign in; OpenID solves the problem without relying on any centralised website to confirm digital identity.

In addition, identity could be federated. For example, a traveller could be a flight passenger as well as a hotel guest. If the airline and the hotel use a federated IMS, this means that they have a contracted mutual trust in each other's authentication of the user. The traveller could identify him/herself once as a customer for booking the flight and this identity can be carried over to be used for the reservation of a hotel room.

Federated identity, or the 'federation' of identity, describes the technologies, standards and use-cases which serve to enable the portability of identity information across otherwise autonomous security domains. The ultimate goal of identity federation is to enable users of one domain to securely access data or systems of another domain seamlessly, and without the need for completely redundant user administration[51].

The above sections describe a 'big picture view' of a more traditional Identity management system. We use this to contrast with the digital footprint mechanism which is a more 'lightweight' form of managing value from data.

Digital footprint: inputs and outputs

Having studied the left-hand side of Figure 28, let's now consider the digital footprint (right-hand side) of that diagram. We study digital footprints from two perspectives: the feedback loop and mobile. We do this by studying the inputs and outputs of **MY DIGITAL FOOTPRINT**. We have introduced the concept of inputs and outputs for **MY DIGITAL FOOTPRINT** and we recap them here: the inputs to **MY DIGITAL FOOTPRINT** are the data elements and the outputs are the value derived from the process which is in turn enhanced by the feedback loop.

It is worth noting that attention (input) and reputation (output) are often called the 'currencies of the web'. Economic value is created from these two from the ability to trade and barter (output). Given that time is scarce, attention is scarce.

Inputs into MY DIGITAL FOOTPRINT

Attention

Data that indicates what you are doing, it is the provision of data that details what applications and services you are engaged with. This could be a widget on your desktop, mobile or set-top providing insight into which applications are open, how long you edited a document, which pictures you viewed, what music you listened to and how often. The attention data stream is the record of what you spend your time doing in a digital world on TV, web and mobile.

Location

The data record of where you are. The live feed is collection, where you were (route taken) if stored.

Time

The time data record is both the time of day and also the period of time.

Search

The data string of search requests, currently the text words (and voice-based search on Google mobile), put into a search engine but progressing to automated search based on requests from 3D barcodes and local available intelligence.

Content (create)

The data record of type, context and information about the content you have created for text, voice, presentation, music, audio, images, video, blogs, tags and recommendation.

Activity

This is the dataset that defines what you are doing whereas attention says you are looking at a web page, activity defines that you are at a football ground. Location gives you the co-ordinates.

Outputs – value created from digital footprints

Intent

Intent is an output which provides predication about what you will do next, based on what you have done, what your social graph does and also on what you have told/inferred/implied that you are about to do, such as your calendar, email or IM trail. Whereas context is about now; intent is about next.

Reputation

Reputation (digital) has many components. Reputation is both about a rating (good and bad) and about your propensity to do something such as leave a comment. Reputation (digital) is therefore partly about your value to the community as a participant.

This output data produces a record which is your digital reputation.

Discovery

This output provides concepts, ideas, insight to enable the user to discover. Discovery is about risk and comes in the form of improvement to an exiting service or discovery of a new service/application.

Recommendation

This is where, based on your digital footprint (you and your social graph), a service/application is able to make a recommendation about an existing or new product or service with a degree of confidence that it will be relevant. Where discovery is risk, recommendation is about trust.

Protection

This is where your data can be used to protect you and your data, in the same way fraud on credit cards works. Your data is a good predictor if you are the individual who is providing the data. This does depend on humans and certain social groups being creatures of habit.

Personalisation

This is where the application or service is personalised to a user for the particular instance or time. It is the modification of a generic service automatically (without user intervention, such as time zone updates) but based on what is known about you. Often the spectrum of personalisation is Vanilla, Tailored, Personalised and Customised.

Trade or barter

This second order output function enables the user to trade or barter for goods or services. The trade or barter will not be for cash (this is payment) but for data or for insights, research, etc. This trade of barter is based on input data, analysis, intent and reputation.

Contextual adaptation

This is where the service or application will adapt to deliver a service that is unique to the individual's requirements based on the existing environment.

It is worth re-stating that the mobile device enriches the digital footprint because the mobile can be viewed as a platform that contributes additional data to the digital footprint, and not just as a determinant of how content is consumed on the device. There is more value created by getting data off a mobile device than on to it.

If we view Figure 28 from the perspective of the customer, the story depicted on the left-hand side is how a typical business operates today and has done so for a long time and will continue to do so. The right-hand side shows this new world of 2.0 thinking: participation, network effects, real-time, collaboration, participation, collection, creation and treating the web (and mobile) as a platform.

Knowing that the mobile is in your pocket from the moment you wake to the time you return to sleep, that same device can know what you are doing, where you are doing it, the applications you use and who you are with. Thus, the mobile device can become the method for collecting data that could be helpful for other services that I would like to consume. The automation of this collected data can become my CV, it could be my reward calculator and indeed could complement a whole host of other documents and certificates that we use today. **MY DIGITAL FOOTPRINT** will never 'replace' actual documents or certificates, but rather complement them by providing a method of determining proof. For instance, **MY DIGITAL FOOTPRINT** may not be my Passport, but could contribute to my Passport. More importantly, in the near future, it could act as a means to personalise my services – provided I can influence it.

The enrichment feedback loop

The importance of the feedback loop is critical to the success of a digital footprint as it has the capability to enrich the outputs. We discuss this idea in greater detail below by looking at some of the outputs and how it becomes enhanced by the feedback loop.

As a practical measure of enrichment, could the feedback help with control of my kids' pester power? My daughter comes home and tells me that all her mates have one, it is the item to have and she wants one too. Would it be beyond the world of privacy and value to say, "well let's have a look!" I then would go to a browser application that could have a look to see who from my social group has purchased one or who has the intention to purchase or who has one and says it is a waste of money.

Contextual adaptation

Meaning

Context (especially mobile context) is the intersection between location, time, presence, handset capabilities and social context according to C Enrique Ortiz[52]. A big picture of the elements of context can be depicted as below and includes:

- positioning: spatial/location information, and related (surroundings);
- point in time;
- presence and related status (online, offline, available, busy, etc);
- handset status and capabilities (capabilities of my handset vs. other handset capabilities);
- personal context (user preferences, calm behaviour);
- information genre, descriptor tags, allows for categorisation and context-based;
- processing capabilities; and
- social context, which is represented by sets of intersections and relationships amongst you and the people you interact with. The social context consists of the person's social circle or context, and related attributes and actions, including: the friends and family augmented or live address book; the relationship distance or degrees of separation; social information such as events (calendar, location, other); inbound/outbound social media channels; social actions – find friends/family, meet, invite/introduce, share content, and this is represented by a social graph.

Impact of the feedback loop

The analysis of the mobile context leverages the digital footprint by providing a feedback loop to services such as:

- informative services
- timely services
- accurate information (accuracy)
- useful information (relevant)
- connected (to friends and family, and other)
- dynamic (always changing)
- adaptive (to current circumstances)
- transformational (promotes behavioural change).

Reputation

Meaning

Reputation and trust are important since they are the foundations of so many other elements within social networks. Reputation is also a currency of the web, something you can trade on. For instance, product recommendations follow a trusted path, Alpha users have a high reputation and are trusted and so on. Trust and reputation are significant for three reasons:

- content flows along trusted paths (recommendations);
- a trusted node (profile) is more valuable than a non-trusted node from both a social and a commercial perspective; and
- trust and reputation can now be quantified through online social networks. Hence, a computation of trust and reputation is now possible – which is different from the pre-social network era.

However, trust and reputation are not easy to define, let alone quantify. A good read is *The Future of Reputation: Gossip, Rumour, and Privacy on the Internet* by DJ Solove. Trust and reputation can be defined as[53]: Trust, a peer's belief in another peer's capabilities, honesty and reliability based on its own direct experiences; reputation, a peer's belief in another peer's capabilities; honesty and reliability based on recommendations received from other peers.

Reputation can either be centralised (like your eBay reputation or a credit score system), or it can be decentralised; computed locally and then aggregated. Trust and reputation are used to evaluate the trustworthiness of a node. They have the following characteristics:

- They are specific to a context: for example, we may trust a doctor for medical reasons but not to fix a car.
- Trust and reputation are multifaceted: They are based on a number of characteristics of the person being evaluated (we may trust a specific characteristic of a person and at the same time distrust other characteristics of the same person).
- Dynamic: trust and reputation are dynamic; they may change over time.
- Trust is asymmetric (if person A trusts person B that trust may not be reciprocated).

- Trust is not necessarily transferable. If person A trusts person B, and if person B trusts person C, then it does not follow that person A trusts person C. But person A is more likely to trust person C than if the relationship between A & B hadn't existed. Probably if person A also knows person D who trusts C, then the likelihood is that person A will trust person C increases, and so on.

The evaluation of trustworthiness itself can be multifaceted; we can identify five facets of trust[54]:

- the feedback a peer obtains from other peers: or direct transactional feedback;
- the feedback scope, such as the total number of transactions that a peer has with other peers. The proportion of positive feedback is more important, the average rating over a range of transactions with a reflection of the total number of transactions;
- the credibility factor for the feedback source: how credible is the source giving the feedback;
- the transaction context factor: for instance, a few mission-critical transactions are more important than a number of smaller, less important transactions; and
- the community context factor for addressing the community: for instance, only a few nodes in a community may actually give feedback, thus skewing the reputation mechanism in the whole community.

Trust can be calculated using a range of models[55] (note that while these are peer-to-peer models, the same principles can be applied in other scenarios):

- Central Control Model through 'leader peers': this technique is used in certificate management and the leader peers are assigned digital certificates by the Certifying Authority (CA). Public Scheme architecture (PKI) is an example of such a scheme[56].
- Local Based Recommendation Model: in this model, one node asks a group of other local nodes to obtain trust information about the target node (the node whose reputation is being evaluated). The dependence on local trust is a limitation of this model but the advantage is that the method operates on a partial trust graph.
- Digital Signature Model: this model relates to the integrity of a file (content) and does not care about assigning trust to the sender of the file. When a file (content) is used and is found to be good, the user would sign the file with a positive rating based on a private key. The more positive feedback the file has, the greater its positive reputation.
- Global Based Recommendation Model: in this model, a peer's trust is defined through the evaluation of all the peers which transact with it.

Impact of the feedback loop

Trust and reputation are foundations of a number of services. The ability to quantify trusted relationships through a range of online mechanisms – especially online social networks – is the key difference in the current ecosystem. Thus, the various elements of the digital footprint (e.g. the contextual elements) can contribute towards the reputation and can lead to a trusted relationship. The benefits to services include the fact that content flows along trusted paths (recommendations) and a trusted node is more valuable than a non-trusted node.

PageRank[57] from Google is one of the best examples of a reputation system based on a feedback mechanism used to create a new service (search). Google describes PageRank as: 'PageRank relies on the uniquely democratic nature of the web by using its vast link structure as an indicator of an individual page's value. In essence, Google interprets a link from page A to page B as a vote, by page A, for page B. But Google looks at more than the sheer volume of votes, or links a page receives; it also analyzes the page that casts the vote. Votes cast by pages that are themselves "important" weigh more heavily and help to make other pages "important".'

PageRank is thus a vote/ballot from users about the importance of a page (hence, a reputation system). A link to a page is a vote of support. This is a recursive mechanism and gets better with feedback (with more usage). Through mechanisms like Android, the same process is repeated on mobile devices. The more data passes through Google's search engines, the greater the feedback/enrichment and the better the process (search).

Discovery

Meaning

Content discovery can be viewed as 'searchless' search and reverts the concept of search to an 'agent' which fetches information based on a set of parameters. In addition to notification, discovery could have the following features: additional delivery mechanisms (Twitter alerts); filter on content source (selecting content sources); personalisation and prioritisation of results; summarisation of results; negotiation (trade) – e.g. eBay; follow-up action on behalf of the creator based on rules; voice activation; dynamic rules engine (powered by recommendations including past interactions and social recommendations, etc).

Impact of the feedback loop

Many elements contribute to discovery. An agent needs to be automatically aware of the preferences of its user and contextually change as the user's context changes. In addition, content needs to be 'semantic' in order to match the intent with the content most accurately. Thus, the digital footprint contributes to this process.

Recommendation

Meaning

Recommendation engines are closely related to search and are evolving significantly. They will be an important part of search going forward. The Wikipedia entry [58] defines 'recommender systems' as 'a specific type of information filtering (IF) technique that attempts to present information items (movies, music, books, news, images, web pages, etc.) that are likely of interest to the user'. Recommendations arise from past interactions (e.g. past purchases) and social recommendations (others in a peer group).

There could be three different approaches to recommendation engines (source: Hyveup blog:[59]):

- Deep structural analysis of an item for its recommendations (e.g. Pandora). 'Together our team of fifty musician-analysts has been listening to music, one song at a time, studying and collecting literally hundreds of musical details on every song. It takes 20–30 minutes per song to capture all of the little details that give each recording its magical sound – melody, harmony, instrumentation, rhythm, vocals, lyrics ... and more – close to 400 attributes' (Pandora, about deep page) [60]
- Intensive social behaviour analysis around an item (e.g. Strands): 'Strands develops technologies to better understand people's taste and help them discover things they like and didn't know about. Strands has created a social recommender engine that is able to provide real-time recommendations of products and services through computers, mobile phones and other Internet-connected devices.'(Strands)[61]
- Structural analysis of an item, paired with behavioural analysis around the item (e.g. Aggregate Knowledge)[62]. For instance, by visits (traffic source, landing page, semantics, visitors' demographics...) and the behaviour of the visitors (page views, clicks, time spent...).

An alternative approach to view a recommendation engine is outlined by Alex Iskold in the four main approaches to recommendations[63]:

- personalised recommendation: recommend things based on the individual's past behaviour;
- social recommendation: recommend things based on the past behaviour of similar users;
- item recommendation: recommend things based on the item itself; and
- a combination of the three approaches above.

Impact of the feedback loop

Both Amazon and Google are extensive users of recommendation engines, for instance PageRank is based on social recommendations (who links to a webpage). Recommender systems are especially relevant to mobile devices since they deliver exactly what the user wants based on preferences (either their own or which their provider has identified). In the mobile space, Xiam (now owned by Qualcomm) provides recommender technology used by Orange[64].

Does feedback work. Quoting from *Nudge,* Thaler and Sunstein. "There is not the slightest doubt. In all eight worlds, individuals were far more likely to download songs that had been previously downloaded in significant numbers, and far less likely to download songs that had not been popular. Most strikingly, the success of songs was quite unpredictable, and the songs that did well or poorly in the control group, where people did not see other people's judgements, could perform very differently in the 'social influence worlds'. In those worlds, most songs could become popular or unpopular, with much depending on the choices of the first downloader's."

Protection

Meaning

This is where **MY DIGITAL FOOTPRINT** can be used to protect you and your data, in the same way fraud on credit cards works. The base assumption is that your routes and routines are a good predictor if you are the individual who is providing the data. This does depend on humans and certain social groups being creatures of habit.

The concept of **MY DIGITAL FOOTPRINT** being used to offer some protection is outlined in Figure 32. Raw data is collected from your device and usage, which is stored and analysed. The analysis provides your behavioural DNA about your locations, routes, routines and timing. The proposition is that this analysis can be used to protect users since the device/system can learn and 'know' about normal behaviour. Thus, if the device can learn and detect normal behaviour, it can offer the output value of protection. In this case let's assume that overnight the device becomes more locked due to no user intervention. Assuming the device wakes up at its 'home' location (can be more than one as long as it is regular) the device will gracefully open and provide access to services. As the device determines that everything is normal (based on similar patterns) it will unlock until the user can use the most secure services without having to enter a pin or password. If the device detects a deviation from the normal routine – then it would progressively 'shut down' or lock-up. Thus, the digital footprint becomes a mechanism for device access based on behavioural DNA. Lock-up could be caused by a change of route, but some services could remain open if the device is travelling with another known and friendly device, or has access to a calendar and knows that it should be on that route.

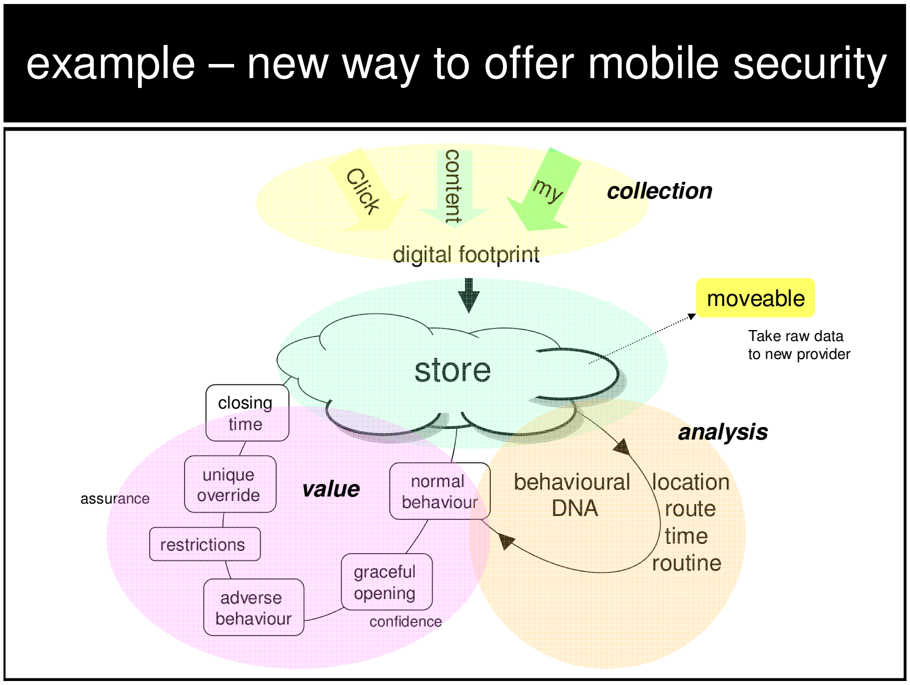

example – new way to offer mobile security

Figure 32 How protection could be offered from **MY DIGITAL FOOTPRINT**

Taking the same ideal of collecting data, user profiling based on size of your calling circle, mobility (how much you travel with your phone), preferred method/volume of communicating at different times of day (including PC IM) and monthly spend, all goes toward building up a profile of everyone on every network. This dataset can be combined with postcode-based datasets and targeted focus groups to build a really clear picture of your likely attitudes based on the way you use communication. At the moment all this (CRM) work goes into selling you more, although I'd rather it was used for health reasons, protection, recommendation, trade and barter, etc.

Impact of the feedback loop

The feedback loop in real-time allows this service to learn and improve. An open system, or non-real-time one will be slower to learn and react. The immediacy of real feedback is critical. The assumption is that the lock and open are based on more than location and time, but also use, such as which applications are regularly opened first and even in which order. I known mine is SMS, email, Twitter, Facebook and then a variety of test and beta versions that I have been asked to play with every morning, irrespective of location and time.

Trade or barter

Meaning

This second order output function enables the user to trade or barter for goods or services. This will not be for payment in terms of money but for data or for insights, research, etc. This trade of barter is based on input data, analysis, intent and reputation.

Like protection, this is probably best described by an example. Suppose a hotel chain wants to reduce it overheads. In reality, after physical running and financing costs, staff is the next largest cost. How can the staff costs be reduced? Youth Hostels in the UK have one staff member and the guests do the cleaning and cooking for themselves. The trade here would be offering my cleaning services for half a day, in exchange for a room overnight. On the assumption that hotels are usually less than 90% full, it may be possible to have 10 guests staying in exchange for services, maintaining a skeleton cleaning rota.

As a paying guest, I assume that the hotel will have vetted the staff and therefore they will not steal my possessions. But how would a hotel chain be able to trust a random person arriving for a free room in exchange for services? A closed loop real-time feedback loop could help, where the person’s reputation (web currency) and skills are available for review, not by a distorted CV but by a group who are interested in the ability to maintain.

Not every hotel chain will buy in, leading premier hotels that pride themselves on service may find this untenable, but a certain chain may find this an attractive model either where maintaining staff is difficult or there is a low occupancy rate.

Nexon is a good example of trade given that they currently have a top line of over $100m per year by selling upgrades to gamers' avatars. They provide the facility to enable the player to modify their persona, such as buying 'flames' for your car, as such it is virtual cosmetic surgery. What Nexon sells is staying cool, providing the online version of having Armani in the real world.

Another example of trade could be allowing a third party to pay you in food vouchers to look at your data on the understanding that you will accept advertising. These are the free models that Chris Anderson presents in his book on *Free*.

Impact of the feedback loop

The ability to trade and or barter **MY DIGITAL FOOTPRINT** for what I want is attractive. It is within my control to do a good job and improve my rating, with the hope that this enables me to do more valuable barter deals. The requirement for transparency, independence, openness, based on a variety of measures, make this attractive to some. However, this value (output) will only be open to those who participate.

Customer Relationship Management (CRM)

The rhetorical question is, "is this not just advanced CRM?" to which the answer is "probably". However, the business framework for **MY DIGITAL FOOTPRINT** is about the service. From the collection, through store, analysis, value creation and the feedback from the user and the social graph, it is about creating value from improvement, colour and refinement. Our CRM world today is about the collection, store and analysis, and then aims to find single points of value to add, usually with an output of selling or up-selling a product or service. So yes, **MY DIGITAL FOOTPRINT** is CRM on steroids; however, it is also not, as the benefit of CRM is the corporate – they own your data – law and regulation is about protecting data use and abuse. **MY DIGITAL FOOTPRINT** is about the value resting with the user, the user has command and control over their data. It is a CRM world, but will or can CRM tools and companies adapt?

Population profiling on the other hand is a technique that uses data to analyse our attitudes and behaviours and predict what we might do in the future. Often called "customer insight" and is dependent on analysis derived from personal data. Richard Webber[65], probably the father of profiling, discovered that there are some things that we share with people who live in similar neighbourhoods (such as susceptibility to disease or owning similar cars) and others we do not (such as criminal behaviour or aptitude for different sports). He likes to show people photos taken at random in different streets, which he has sorted into “herds” by the data. It is eerie how similar some are, from choice of houses to the style of patios. Webber’s model finds herds by linking people and postcodes. He created Mosaic when working for Experian, a system incorporating anonymised data from more than 400 sources which generate different postcode maps that predict people’s behaviour and the analysis allows insight into key questions such as will this “herd” buy designer clothes or save? Mosaic’s 61 herds allow supermarkets to target stock to local populations, and let Saga write tailored letters to every man and woman in the country as they turn 50. It revolutionised retail and direct marketing industries and this is now about to move forward again.

B2B and MY DIGITAL FOOTPRINT

This text has purposefully focused on B2C (Business 2 Consumer). In reality the business that sits between the collection and creation of value will be a value chain of many businesses. Specialist companies will focus on collection and store and these are likely to become low value commodity models, full of rules and regulations leaving high wealth creation opportunities from analysis and its subsequent application. The implication is that businesses will have to decide what business they are in, no one business can do it all, and how those businesses involved in creating value for an end user act together for each other’s benefit (co-operation). Further, pure B2B will benefit from the same process improvement (the feedback model in **MY DIGITAL FOOTPRINT**) in the supply chain, where real-time data is gathered based on usage of components and services, not just for stocking, logistics and just-in-time production where feedback is already deeply embedded, but for the improvement of the products and services based on the data from its ultimate usage. At one time this was market research, sampling and CRM in linear non-real-time.

Everything is a service: including advertising

In Figure 28 there was a pictorial view of the traditional business model as a linear single direction through benefiting the corporate and this was complemented by a view of **MY DIGITAL FOOTPRINT** with a feedback loop that enabled both focus and breadth. These two sides are joined by data that can be fed from the regulated world of identity and married with collected data in the form of a digital footprint and data from the analysis creating behavioural DNA. This joining point is AAS (As A Service) – charging a subscription for a service.

The discussion of a service in the previous section also includes advertising, since advertising could also be treated as a service. Currently, advertising is 'shotgun' (shoot at prospects, shout loudest, interrupt content) it is paid by the brand and sold as a product. But that trend has started to change with initiatives like Google Interest-based profiling[66], Feeva[67], The Now Factory[68] and Blyk[69] which could also be treated as ‘Advertising As A Service’. To explain this AAS model using advertising as an example, it is possible to see the model as:

- an ongoing relationship between the provider and the advertiser;
- based on preferences/profiles;
- based on an agreement (for the lack of a better word) between the provider and the advertiser;
- the agreement will evolve as the service progresses (a self-learning service); and
- the arrangement may span platforms (mobile, web and TV) with probably mobile and web as the leaders/early adopters.

The question now becomes: how do you use this to create a virtuous circle of services?

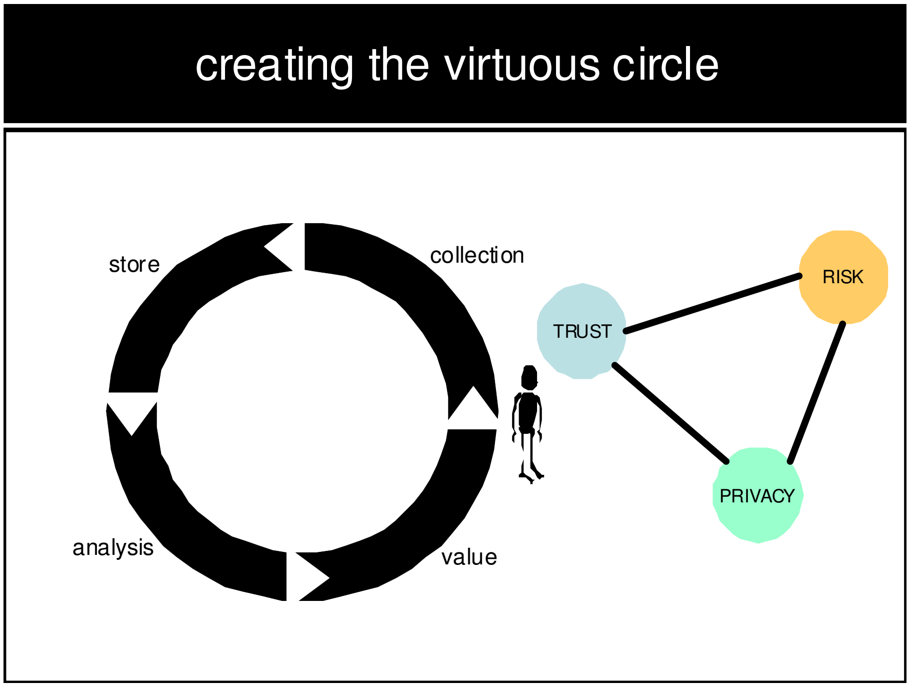
This does raise an interesting question for companies like Google, Yahoo, Microsoft and Fast, and where they go next. Do they sell their algorithm to companies who will match buyers and sellers in a market, or do they do it themselves? If they try to do both, surely there is a conflict and the company may not sell the best algorithm trying to keep it for themselves as it is more effective and creates more wealth or even try ideas on others first to discover, rather than damage, their brand. It could be seen that this developing market is wide open for someone to come in on one (advertising as a service or algorithm development) or both sides and compete.

Creating the virtuous circle

Be under no illusion that this (creating a virtuous circle) is either simple or easy. Mark Zuckerberg, the founder of Facebook, said when commenting on Beacon: “We’ve made a lot of mistakes building this feature, but we’ve made even more with how we’ve handled them. We simply did a bad job with this release, and I apologise for it.” This was in response to the 70,000 users on Facebook responding to Facebook’s new Ad and Beacon features in December 2007. The Facebook Ad followed a well-trodden path of purchase goods, PIN codes, getting free extras online. Fun and not a big fuss, Beacon, however, was different. Beacon would look at what you do and as such has deep roots in behavioural marketing based on targeting (open loop); however, it took your data and told your friends what you had done, really without any due care or thought. If you looked at someone's profile, it told your friends. Rather hard to hide the fact you were looking for 'fitties'. Buy a book or CD; well your friends may want it as well. In so many ways this is **MY DIGITAL FOOTPRINT**, but perhaps in the rush to monetise the knowledge, no rules were put in place. Ideas are cheap, implementation is hard.

At this point there is a need to bring back some themes that we skipped past in the early part of the book. These are the bonds and bridges between privacy, risk and trust. Figure 33 shows how these bonds and bridges relate to the **MY DIGITAL FOOTPRINT** feedback model of collection, store, analysis and value. The bonds and bridges of risk, privacy and trust are the connection fabric that ensures that the user continues to enjoy the experience or will cause the user to stop using the service. How do our experiences enforce the benefits that mean we participate more, or damage the benefits and head off into the dark side, causing lockdown and disengagement? Confidence, referral, recommendation, privacy, trust and risk are all key aspects to unlocking the virtuous circle.

Within this start phase of discovery there will no doubt be areas that will push boundaries. Whilst we understand legal and illegal as boundaries, there are often instances where within a grey area of illegal but acceptable, sometimes referred to as 'socially acceptable' [allowing your underage teenager to see a film with a different classification, providing your bank PIN to a partner when they jump out of the car in the rain to get some cash from an ATM, allowing your child to try alcohol at home], or 'in the public interest' in the case of journalism where the boundary is crossed, but no one would be too upset if the outcome is good, positive or beneficial. As the user is the provider and consumer of data, the damage will be caused when, as a trusted provider (brand), one crosses the line of protector to exploitation agent. A lot of work and debate is needed on unacceptable and inappropriate interpretation of data and Metadata. As a director, governance will be difficult, balancing these intangible issues along with wealth, competition and value, as this is where the opportunities lie; it is something I spend a lot of time worrying about.



creating the virtuous circle

Figure 33 Creating the virtuous circle

To help in the understanding of why trust, privacy and risk are bonds and bridges, I am going to use the concept of capital (trust capital, risk capital and privacy capital) but not in the strict economics definition sense. Trust capital is how much, based on previous experience, you will be prepared to trust a service provider you have or have not previously used. Risk capital is how much, based on previous experience, you are prepared to risk (chance) using a service provider you have or have not previously used. Privacy capital is how much, based on previous experience, you are prepared to open up your privacy or personal data to a service provider you have or have not previously used. Capital, in each case, can be built and destroyed. In the next section we are going to briefly look at how you can build and erode capital.

Privacy capital

The premise of privacy capital is that people will be born with no privacy capital and over one's life this capital will be built or eroded based on experience. Good experiences of seeing your privacy protected or, in general, government protecting your privacy through action, law and regulation, the nation's privacy standard will build privacy capital. Your social group, your family and their experiences will build privacy capital which is good. These good experiences provide positive feedback as shown in Figure 34. Positive experiences mean that you will have a higher propensity to engage and more hope by doing so that you will get better services and experiences. This lowers your fear, uncertainty and doubt (FUD). The more you see your privacy is protected, the more you may wish for someone to use your data as you do not fear them abusing it. You become more willing to share patterns, preferences, routes routines, shopping lists, click data, location and other inputs that will improve services. The converse being true; the more you see someone abuse your privacy, your friends, your social groups and generally the media invading privacy, the lower your FUD becomes, and your privacy capital is not eroded.

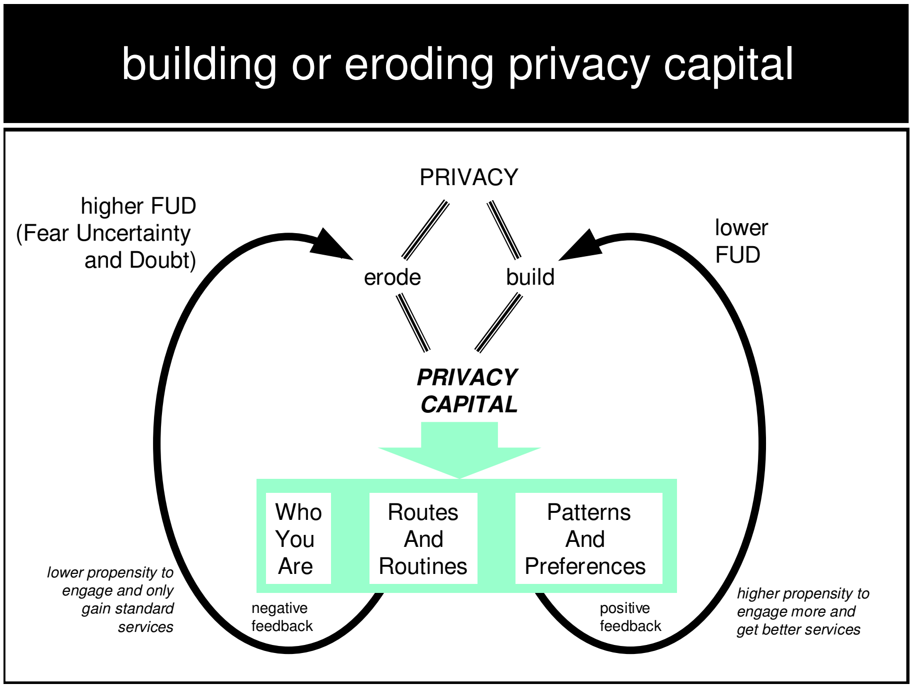


Figure 34 Building or eroding privacy capital

It is expected that your capital will even out at some point after you are 21 and will only be moved by major events.

Risk capital

The premise of risk capital is that people will be born with some differing levels of propensity to take risk and over one's life this capital will be built or eroded based on experience. When you take some risk and reap the reward for it, you will take more risk, hoping to receive greater reward or satisfaction. Risk itself will vary depending on the situation and the actual activity and likely reward. People having the propensity to take risks at home with their own financial future is different to what they may take with someone else's resources. As with privacy, risk that is rewarded, which increases capital, lowers FUD and allows us to engage more and demand better services as per Figure 35. Risk taken that fails or creates damage to our reputation or relationships will erode our capital and our willingness to take a chance on the next service.

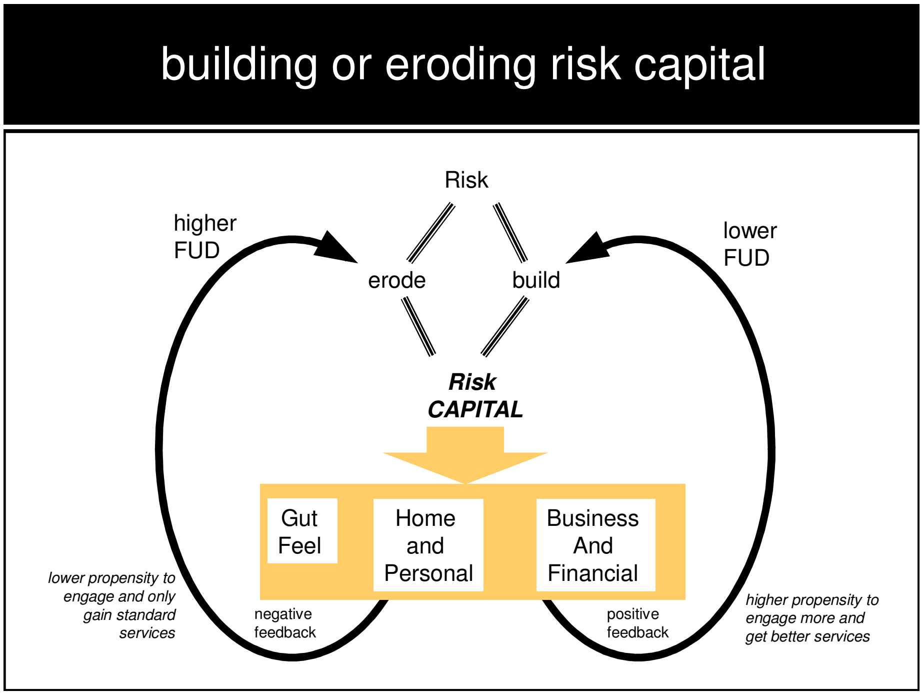

building or eroding risk capital
Risk
higher FUD
erode
build
lower FUD
Risk CAPITAL
Gut Feel
Home and Personal
Business And Financial
lower propensity to engage and only gain standard services
negative feedback
positive feedback
higher propensity to engage more and get better services

Figure 35 Building or eroding risk capital

It is expected that your capital will even out at some point after you are 21 and will only be moved by major events, but will vary depending on the situation and the reward.

Trust capital

The premise of trust capital is that people will be born with some differing levels of propensity to trust others and over one's life this capital will be built or eroded based on experience. I clearly remember the falling back game on management training and being called stupid as I trusted so implicitly that anyone would catch me, if I knew them or not. As with risk and privacy, trust capital can be built or eroded. Trust, risk and privacy are bonded and bridged and, as such, poor experience on one will affect the others. Trust in this case is about trusting a service provider with **MY DIGITAL FOOTPRINT**. A good experience of sharing your data builds trust capital (akin to that of a brand). The more you trust them with your data, the higher the risk, the more privacy you impart, the better the services you get. You share more of your routes, routines, patterns, preferences, reputation and recommendation in the hope that the services will continue to improve and this builds your trust capital, lowers your FUD, as per Figure 36. This will also move you to influence your social group, and your good or poor experiences will build or erode their capital reserves.

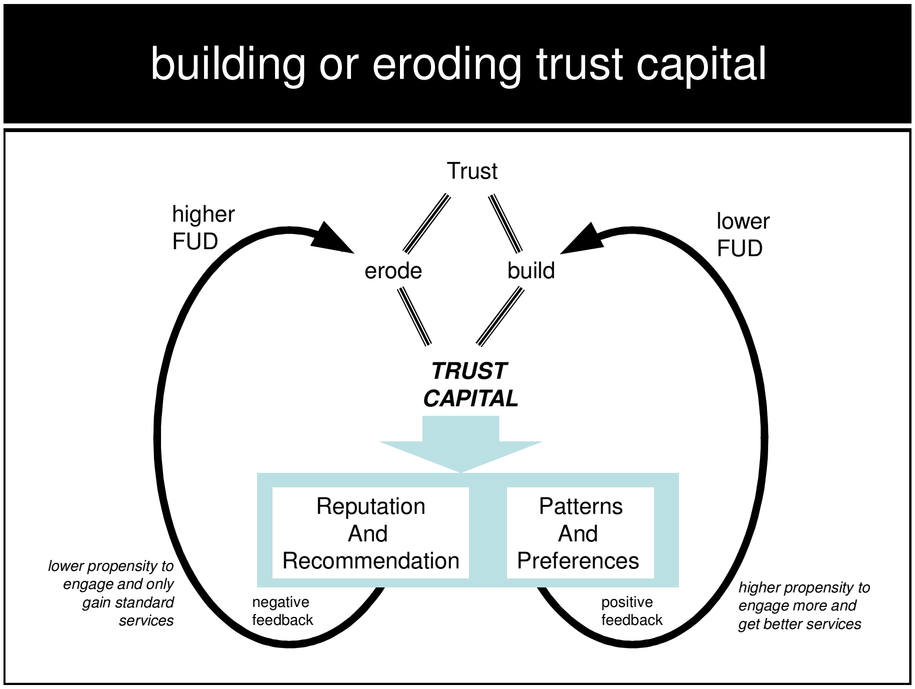

building or eroding trust capital
Trust
higher FUD
lower FUD
erode
build
TRUST CAPITAL
Reputation And Recommendation
Patterns And Preferences
lower propensity to engage and only gain standard services
negative feedback
positive feedback
higher propensity to engage more and get better services

Figure 36 Building or eroding trust capital

It is expected that your capital will even out at some point after you are 21 and will only be moved by major events, but will vary depending on the situation and the reward.

Bonded but not related

The theme being explored in this chapter is the business models, which has led us to look at the creation of the virtuous circle, where the user and their social group continue to add data, which produces improvement to services and benefits. These benefits mean that the user is more willing to participate and in doing so will trust more, be more open (less private) and take more risks. The last part looked at how trust, risk and privacy are bonded and bridged and how good experiences build capital, which leads to more engagement, participation, connection, conversations and relationships.

Figure 37 shows this concept, but looks at the bonds between privacy, security, identity, risk and trust. The proposition is that identity (digital and physical) is a lynch pin between all the others. Therefore, it is split so that above the line is trust and risk; below the line is privacy and security. In Maslow's Hierarchy of Needs[70] one would look at privacy and security as hygiene factors, meaning that trust and risk (capital) could not be built without them in place. However, the purpose of this diagram is to focus on why some commentators focus on the 'dark side' of identity, locking it down, protection, privacy invasion and everyone wanting to control you. Those who start their argument from privacy and security tend to argue that there is a need for increasing control and present more reasons why there are barriers and objections.

Those who start from the basis that privacy and security have to be in place and must work, but look at trust and risk (capital) as an opportunity to grow, focus on the benefits, and indeed, a more open system, transparency and less-centralised control.

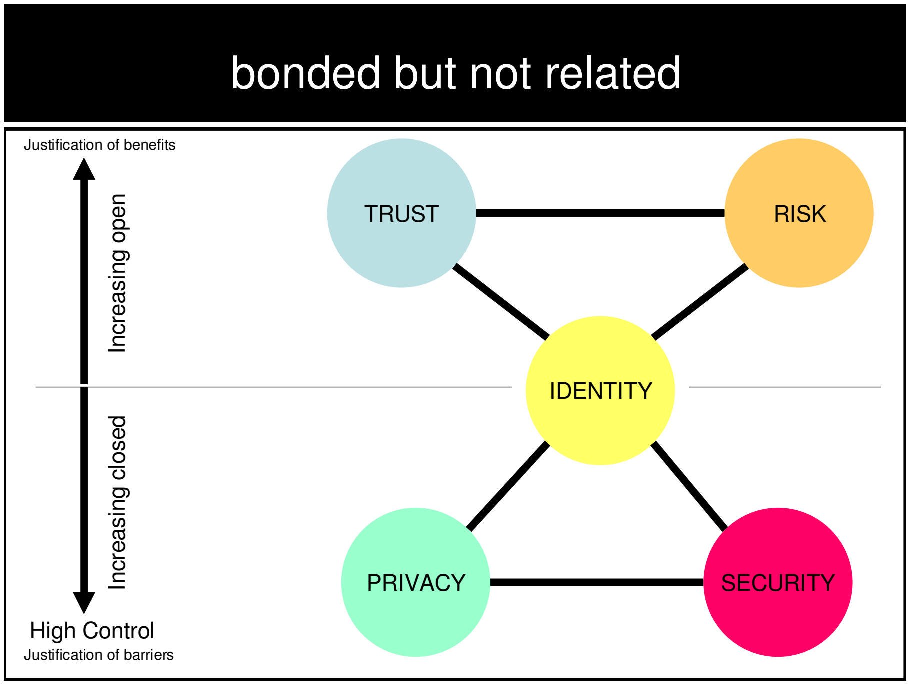


Figure 37 Bonded but not related

The reason for pointing this out, is not to make a case for one or other side being right, just that there are differing views about identity and how the bias of identity and its value is based on the direction from which it is viewed. This applies to **MY DIGITAL FOOTPRINT** as well, and depending on one's risk, privacy and trust capital will partly prescribe which side of the argument one would start from, and this basic stance will determine if someone buys into the idea of a virtuous circle or not, irrespective of whether it creates value or can be implemented. In the next chapter we take this one stage on to look at the possibility of converged services and trust-based segmentation to determine who has a higher propensity to buy into this model.

RULES, LAW AND REGULATION

Rules, law and regulation, including implemented, interruption and what is planned is vast and obviously varies by country and viewpoint. This section on the website http://www.mydigtialfootprint.com/ has suggestions and links to great resources, comments and insights. You will need to register to see these comments and links.

MY DIGITAL FOOTPRINT *AND CONVERGED SERVICES*

In the previous sections, we have introduced a number of concepts relating to the implications of digital footprints. Building on that foundation, we now discuss the practical implications of the concepts especially in relation to new services in a converged ecosystem.

Convergence and converged services

An introduction to convergence

What is digital convergence?

Digital convergence is a much-maligned concept. Mention digital convergence, and it conjures up images of the 'intelligent fridge'; a concept most people think they have no need for. But, digital convergence is an idea whose dawn is near. There is a lot of confusion about what exactly is meant by digital convergence. When people talk of digital convergence, they could actually mean different things, including:

- co-mingled bits: the original definition of digital convergence as outlined in Nicholas Negroponte's 1995 book *Being Digital*[71];
- device convergence: one device to rule them all. Think the iPhone (a combination of the iPod and the mobile phone), Nokia N-gage (a combination of a gaming device and a phone), etc;
- fixed to mobile convergence;
- service-level or application-level convergence; and
- devices being able to speak to each other and share intelligence, leading to a new service aka the 'intelligent fridge' or home networks.

Consider also: ‘If all you have is a hammer, everything looks like a nail.’ – as Mike Langberg so aptly put it in his article soon after CES[72]. By that, we mean: your tools (focus) determine your viewpoint of the world. The ‘nail’ in this case, is digital convergence. The ‘hammer’ is the viewpoint (strengths) from which each player is approaching digital convergence. For example, (as per the article referenced above or Microsoft), convergence is a software problem; to be solved using an upgrade of the Windows operating system (Microsoft’s strength). Intel sees convergence as a ‘microprocessor problem’, to be solved with a branding programme called ‘Viiv’ [73], Cisco sees convergence as a home networking problem, to be solved with, guess what, networking. Yahoo! and Google see convergence as an online services problem. To them, the solution lies through the web browser, a common element in all devices. Sony sees convergence as a consumer hardware problem, to be solved with consumer devices, new standards built around its own strengths, like the PlayStation. No wonder there is confusion.

It’s important to note that the only things in common between all these definitions are:

- digitisation; and
- communication.

In other words, information must be digitised and it must flow freely. This leads to new services, which are greater than its parts (greater than what the devices could provide on their own).

Digital convergence: definitions

Let’s first discuss the definitions above in a little more detail.

Co-mingled bits: the first definition was proposed by Nicholas Negroponte in his 1995 book *Being Digital*. Negroponte’s definition of digital convergence is: ‘Bits co-mingle effortlessly. They start to get mixed up and can be used and re-used separately or together. The mixing of audio, video and data is called multimedia. It sounds complicated, but it's nothing more than co-mingled bits.’

Another way to put it is: to a computer, there is no difference between a symphony, a voice call, a book, a song, a TV programme, a shopping list, etc, as long as they are all digitised. The factors driving digital convergence/co-mingled bits include the rapid digitisation of content, greater bandwidth, increased processing power and the Internet. Digital convergence brings four (previously) distinct industry sectors in collaboration/competition with each other. Thus, we have media/entertainment, PC/computing, consumer electronics and telecommunications industries all interacting more closely with each other than ever before. This version of digital convergence is happening all around us.

Terms like ‘triple play’ or ‘quadruple play’ are a part of this scenario. Triple play involves voice, broadband and TV, and quadruple play enhances TV with digital and adds mobile to that mix. These are different from the origins of the cable industry ‘dual play’ (one build, two income streams). It now refers to one infrastructure with many services. Whatever name you call it, here are co-mingled bits in action. If everything has become digital, then the boundaries between the providers fade away.

Device convergence: addresses the age-old question: Will we carry one general purpose device or will we carry many specialised devices? Boundaries between devices are fading fast and devices are now capable of performing more than one function. It is unclear if customers would really want a single device. Most people have a view on this, and so do the device manufacturers.

Fixed to mobile convergence: this is a relatively new area. It has emerged because fixed-line telecoms operators and mobile telecoms operators are each vying for customers in each other’s traditional domains. Telecoms access networks are converging due to the emergence of new technologies. Thus, mobile network providers can provide fixed network services and vice versa. Services could also be converged. Thus, a user could access the same service from either a fixed or a mobile network. Fixed to mobile convergence could be seen as a larger concept called ‘seamless mobility’ – the overall idea being that a customer should be able to ‘roam’ seamlessly between different network types (fixed, mobile, WiFi, etc). Bodies such as UMA (Unlicensed Mobile Access) are driving the standards for seamless mobility.

Service-level or application-level convergence: means bringing together online services and providing a management platform to support all services. This provides the biggest technical challenge in convergence, but also the most significant benefits to customers. Service-level convergence also enables operators to offer all services to all customers, irrespective of device, access, connection, login, services, applications and payment. This is the level where common identity or digital footprint is most useful and also where significant convergence value (cost saving) is derived.

Services

A ‘service’ constitutes an ongoing relationship between the customer and the provider. Unlike a product, which is bought and consumed, a service implies a trusted relationship between the provider and the customer. On the web there is a shift or move towards software being sold as a service and not as a product. Similarly, many traditional products (e.g. books) are increasingly being sold as a service through sites like Amazon. Similarly, mobile applications are moving to a service model. These services will be converged, that is, spanning platforms (web, mobile, TV). In doing so, the underlying data and the digital footprint will constitute the bedrock of new services.

The significance of mobile services

Until recently, devices were very simple and mostly comprised only voice and SMS. In this scenario, supply chain efficiencies and lowest common denominator/mass market devices prevailed. Differentiation was achieved through price plans, device aesthetics, etc. With the increasing proliferation of Smartphones, devices are becoming more complex and sophisticated. Ironically, at one level, they are being commoditised (the megapixels are no longer a differentiator since most devices are roughly similar). Nor is the network itself a differentiator since we expect good coverage at all places (within reason). Also, with the uptake of OpenSource and other technologies, there are many more device manufacturers who are giving the customer greater choice. So, there are many devices and many more new ones to come.

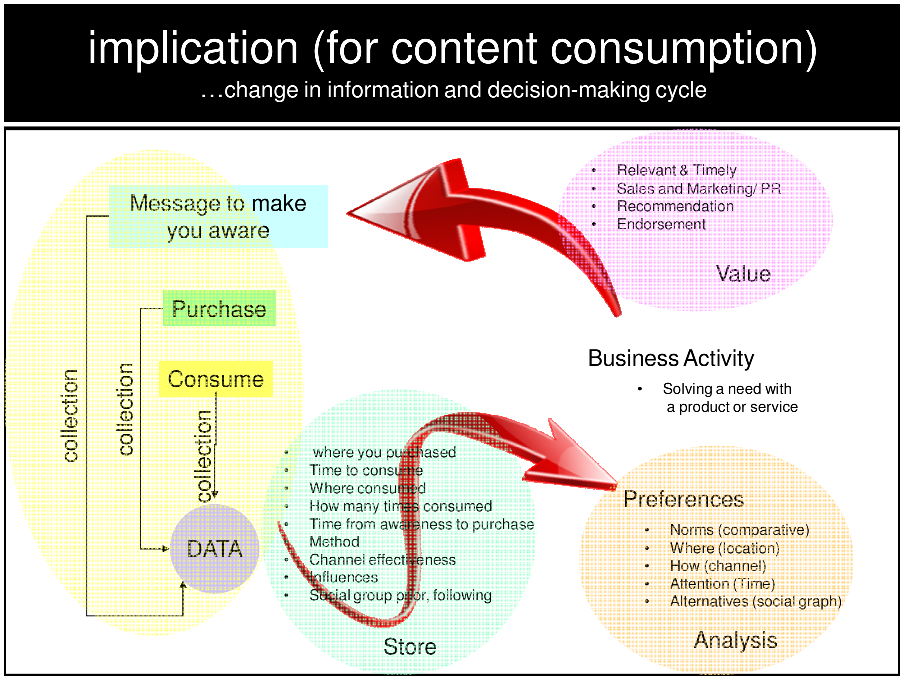
In addition, operators cannot charge for the value of the packet in IP traffic but many (secretly) wish that they could do. Add to this desire the impossible situation of guaranteed coverage, which the user believed they have paid for. Consider the scenario where the customer pays more for a ‘high value IP session’. He then walks under a bridge, loses connection and sues the operator. If we are worried about 'impact on helpdesk' then that's a MUCH greater worry if you think you can get away with differential IP charging on mobile devices. Thus, the result is: when it comes to upgrades and selection of device/operators today customers no longer want to upgrade for network connectivity or more megapixels, etc. They want a step change. They want differentiation that they can understand and are already familiar with. This means new services.

By ‘services’, we mean a mechanism that the user expects from an ongoing application, perhaps on a subscription basis. The business model for services is well known, for instance, the freemium model[74], where you provide a basic service for free but charge a premium for extra features. To use a common example, a dating site would charge no money to set up a profile, but you would have to pay to see who viewed your profile.

With mobile services, we have a more complex equation. The traditional way to classifying mobile services is a classic ‘telco only’ view (look at the capabilities of a network, add a service on top of it and pray that the services created from them will be useful to the customers, for example, person-to-person video calling which never took off). The web approach is much more sophisticated and relies on knowing the customer better with each transaction.

Content service example

Marketing (behavioural, targeted, segmentation, demographics) has been based on some quite blunt analysis tools after the action and there has traditionally been a significant lag between product idea, launch, purchase and research or evidence to support actions. Today there are some very capable real-time tools that measure specific actions, which, via analysis, are able to provide reasonable predictions. However, moving forward, the expectation is to be able to gather relevant and specific data that will enable accurate and real-time change, action, resolution, update, modification or improvement to products and services. As shown in Figure 38, the collection of data is moving from just knowing that a product or service has been purchased, to knowing what influenced the purchase, when it was purchased, where it was purchased, how it was purchased and how often it has been consumed (music/video). This data can be stored and analysed to provide preferences for both style and taste, but also to channel, time of day, location and even activity. The business activity will move from the application of blunt wasteful mass media campaigns to focused, insightful, relevant and meaningful media campaigns. The value for the user of giving up data is that services are relevant, timely and endorsed or recommended by social networks. The data on how you then react can hone messaging and focus, providing an ever-improving service and value. It becomes obvious at some point that it will be possible to measure how much you influence and are influenced by your social group. The more you can influence, the higher value you have to a brand.



implication (for content consumption)

…change in information and decision-making cycle

Message to make you aware

Purchase

Consume

collection

collection

collection

DATA

- Relevant & Timely
- Sales and Marketing/ PR
- Recommendation
- Endorsement

Value

Business Activity

- Solving a need with a product or service

- where you purchased
- Time to consume
- Where consumed
- How many times consumed
- Time from awareness to purchase
- Method
- Channel effectiveness
- Influences
- Social group prior, following

Store

Preferences

- Norms (comparative)
- Where (location)
- How (channel)
- Attention (Time)
- Alternatives (social graph)

Analysis

Figure 38 Implications for content consumption

The WHAT principle and the WHO effect

The WHAT principle

The focus of today's higher value services is personalisation – the making of your user experience, creating value from the reduction in churn and incremental service revenue, assuming that any incremental margin is not eroded by competitive pressures. The focus on personalisation is a focus on WHAT: what you as a user want to do; what service you want; what is needed now? The sole benefactor is the individual, but does this create any new value? The assumption is that personalisation provides focus, and that this focus leads to the ability to deliver engaging and personalised services, including advertising. This advertising being derived from the same advertising budgets, which is now redirected from other display channels. Therefore, does personalisation actually create any new value and will it actually grow the overall spend of the entire market?

Commentators, consultants and media sellers will provide convincing evidence to back their own propositions. The purpose of this viewpoint is not to debate the personalisation opportunity, but to introduce the WHO effect. Whilst personalisation will increase value for the provider [more effective marketing and efficient sales]; assuming that there is value for the user, it does not itself create new value for the entire converged industries. However, personalisation could create value, if the focus is on WHO and not WHAT.

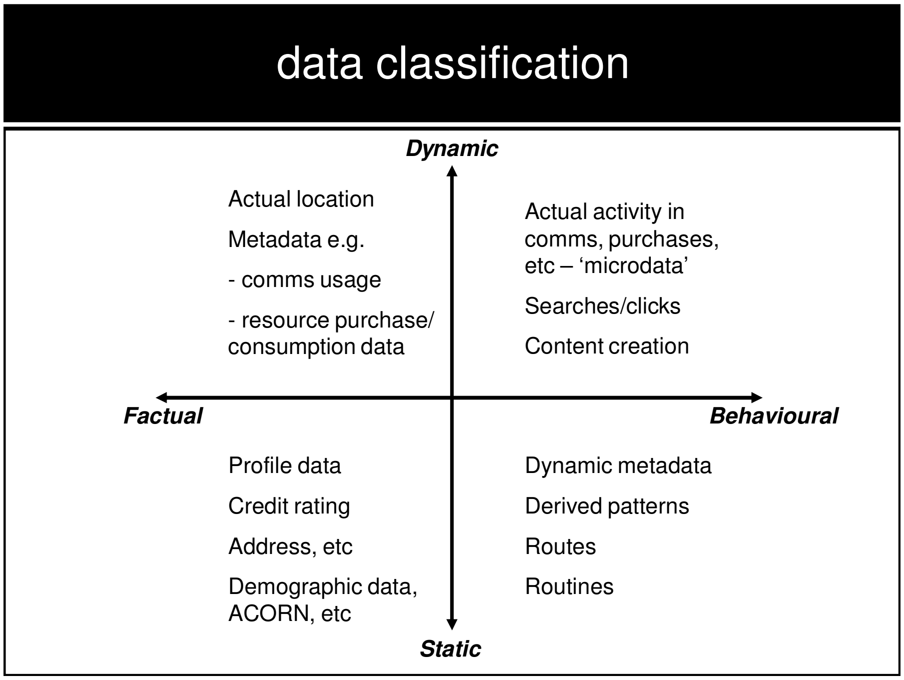
The WHO effect

Personalisation has been about the WHAT principle. This has focused on a single customer: 'you'. The WHO effect is the multiplier. The focus shifts from WHAT to orientate on WHO you are doing something with. In simple terms, when you go out for dinner, who are you with? When you are in a business meeting or seminar, who are you with? When you are at a concert, in school, or on holiday, who are you with? The opportunity is that these 'WHOs' are gravitating towards and enjoying the same experiences as 'you'. The additional profiles of those who you are 'with' can combine to create a new and incremental market value, assuming, as a company, you are able to reach these customers and deliver services that they want.

Consider the advertising issue created through personalisation, it reaches you – one person in two billion. The world is divided into two billion personalised worlds, only relevant to one person at any given time, and each person with an unequal bite of the advertising spend. The WHO effect would suggest that as you are enjoying something with others, even though it is outside of their personalised preference, it is possible that it would be worth providing information on products and services to the group. The WHO effect is the electronic 'word of mouth'. It assumes and depends on the fact that we adopt at different rates, and some not at all. These issues provide the limitation to personalisation and the WHAT principle, but also the opportunity to the WHO effect.

How WHO works

WHAT-based decisions are not using information in a sophisticated enough way. To move to the WHO-based system, customer data needs to be understood in a more nuanced way. We think of data on two main axes: how 'active the data is' (static to dynamic) and the 'type of data' (factual to behavioural) as per Figure 39.



data classification

Dynamic

Actual location

Metadata e.g.

- comms usage

- resource purchase/ consumption data

Actual activity in comms, purchases, etc – ‘microdata’

Searches/clicks

Content creation

Factual

Behavioural

Profile data

Credit rating

Address, etc

Demographic data, ACORN, etc

Dynamic metadata

Derived patterns

Routes

Routines

Static

Figure 39 Data classification

Factual vs. behavioural

- Factual data is just that – date of birth, where you live, etc.
- Behavioural data is what you do, your digital footprint (www.mydigitalfootprint.com) over time.

Dynamic vs. static

- Static data is data that doesn't change (e.g. your date of birth). Highly dynamic data is what is changing every minute – movement, current location, etc. Some data is dynamic, but over long time periods (e.g. place of work, home address, etc).
- Some dynamic data is repetitive (there are patterns in it, such as the daily commute) that allows one to predict behaviours and events.

Clearly, if a service provider has a grasp of this information, they would be able to make better decisions about the user context and thus serve up better information or services to the user – critical for the limited real estate on even the smartest of mobile devices.

Creating advantage

This WHO effect is not open to the traditional broadcast, TV and entertainment companies, although they are the traditional home of the display advertising budgets. This is because they have no real feedback loop. This service could be offered by web companies; however, as your profile and personalisation has a dependency on your web access time, it could be difficult. The major benefactor of the WHO effect will be mobile-orientated providers, as the mobile device becomes the platform to collect data, interrupt the connection and deliver the value (Mobile Web 2.0).

Caveats

It should be noted that there are a few caveats to this value growth model which include:

First: the opportunity to exploit the WHO effect is not open to companies who want to 'control' the user experience and developer environment, such as Apple, they can only enjoy the WHAT principle. Open mobile platforms, open access services and developers whose services work across all devices will be able to exploit the WHO effect. The multiplier value of mobile is not just knowing WHAT you are doing (location, time and attention), but WHO you are doing it with.

Second: user pre-acceptance and buy-in is critical. ‘Snooping’ behaviour has already blown up in a number of companies' faces, Facebook, Phorn, Google and MySpace are known examples. Users need to be certain that data will not be misused, sold on or otherwise exposed.

Third: trust is critical. Is your brand value one of trust, and what will the user trust you for?

Fourth: regulation and law. This is a black hole of debate currently, and differs by country, we would say the key here is to adapt – and actually practice – the mantra of ‘don't be evil’. If you wouldn't like people to do something to you, don't do it.

Fifth: this concept is deeply embedded in ‘web as a platform concept’ and AAS (As A Service). This is a change from domains, destinations and portals. A way to understand this is to view that owning a top level domain such as www.news.com is not important. Users will rarely visit this site, but rather interact with the data that comes from this source via Twitter, RSS feeds, readers, blogs and social sites. The stats on your site are not important, but rather where your content is consumed and how? Customer ownership, as per caveat one, is not important – owing the customers' data is.

Sixth: Although both the originator and user of **MY DIGITAL FOOTPRINT** is the user / consumer, where the WHO effect applies the user actually consumes data that relate not just to themselves but also to others with whom they interact. How we collectively address the privacy issues around this involvement of others in the key feedback loop, which is a rather important point, is somewhat overlooked.

The rainbow of trust

Marketing knows that segmentation works. Segmentation is usually based on a factual measureable component; however, the rainbow of trust that is described in this section is a concept to help determine and focus on customers who will aid the virtuous circle to start and grow. The bonds and bridges between trust, privacy, identity, risk and security have already been explored; the rainbow of trust explores one aspect of these bonds and outlines how **MY DIGITAL FOOTPRINT** can be a mechanism for purchasing commercial services based on a concept of 'trust'.

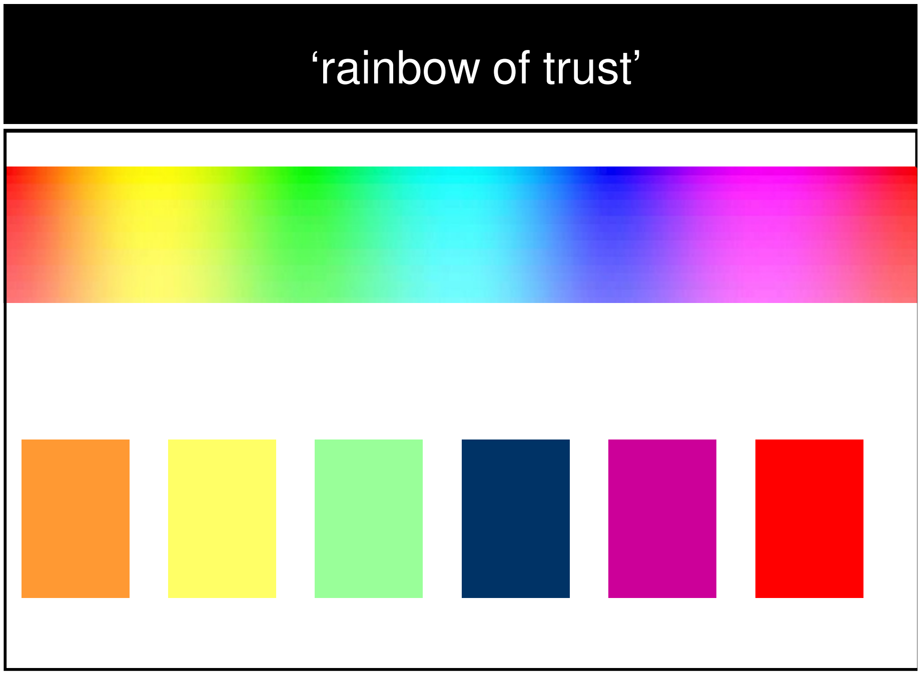
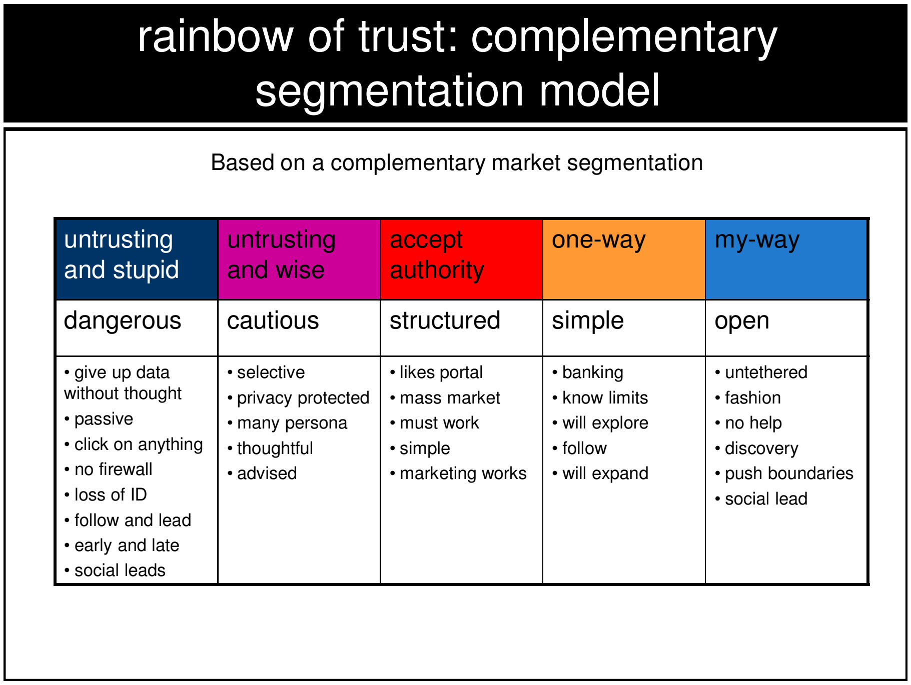


Figure 40 Rainbow of trust

The assumption is that the rainbow of trust Figure 40 (which does not look good in black and white print) is about grades of colour (or grey). It is a continuum and shows that people have a different propensity of trust, from untrusting to very trusting, and what they trust (a brand, the government, people, things, friends, relationships) would also provide different dimensions. To highlight the types of trust that could be possible for service differentiation let's imagine a customer walking into a mobile phone shop and saying to the sales person: "Hello, I am Mr Blue". What the sales person understands from this is that this customer wants an unfettered device that has no constraints, and they will sort out all the IT problems themselves, including content that they download. In contrast, a customer called Mr Red only wants to do online and mobile banking and is concerned most about security. Mr Purple wants the walled garden (wants the provider to manage all their content and services. Mr Yellow wants voice and text and Mr Tickle wants a laugh for those following the Mr Men theme. These customers are differentiated based on their trust in both the device, the service provider and the services they demand. Enterprise customers will also vary by their attitude to risk. The analysis demands needs based on value by trust and risk, and not by product or lifestyle. An example of some possible profiles is shown in Figure 41.



rainbow of trust: complementary segmentation model

Based on a complementary market segmentation

untrusting and stupid	untrusting and wise	accept authority	one-way	my-way
dangerous	cautious	structured	simple	open
• give up data without thought • passive • click on anything • no firewall • loss of ID • follow and lead • early and late • social leads	• selective • privacy protected • many persona • thoughtful • advised	• likes portal • mass market • must work • simple • marketing works	• banking • know limits • will explore • follow • will expand	• untethered • fashion • no help • discovery • push boundaries • social lead

Figure 41 Complementary segmentation model

In this rainbow of trust-based segmentation, there is no market aligned to age, lifestyle, income, demographics, early adopters or followers. It moves the ideas from the young who explore, and the old who stay with what they know, even if it is not the best. It will be (is now) possible to determine trust as we now have access to the very data needed to determine it.

In this model, the untrusting and stupid are a segment who give up data without any thought, they always sign the terms. Largely they are passive in terms that they don't offer reviews or blogs. They will click on anything, adverts, banners, etc, and are easily swayed by marketing and advertising. They don't have a firewall, or it's out of date or switched off as it stops services that they want. They believe that the loss of ID is a hazard of the modern world. Due to their intrepid sprits of 'trying anything', they are often a social lead. However, for a brand and a service, this is a dangerous market as they are fickle, disloyal and will move with fashion.

Those in the untrusting and wise segment are the antipathy of untrusting and stupid. This trust-based segment is selective, concerned about privacy, understands about protection and its value, and have many persona to aid this concealment of true identity; they are well advised and considerate of outcomes. Whilst cautious, this is an attractive market as trust earned with this segment will keep them loyal.

Those who accept authority – that still have Yahoo, MSN or AOL as their home page – see the value in the portal or are unable to change it. It has to work with one click, it must work out of the box; it is broken if it's not simple. Marketing really works to this trust segment. They like structure and will follow.

Those in the one-way segment want to do one thing at a time. When they are happy with the one and that it is good and works (trust built) they will expand and enjoy anther service, one at a time, but, over time, building a wide range of services, slowly and steadily. They will only ever follow what is recommended and already trusted.

Those in the my-way segment – alas, like most people who will read this book – demand untethered devices, we don't want help, we will discover on our own and tweet, we will push boundaries, we trust no-one until we have done it ourselves. This trust group just want it open and transparent so that they can do it for themselves. As we will discover later, they are some of the few who could manage their own digital footprint and exploit it.

It would be fairly safe to assume that, based on the trust capital model, our trust in services and brands changes based on experience and others in our social group's experience. Therefore, Figure 42 shows the likely migrations routes between the suggested trust-based segmentation categories explored.

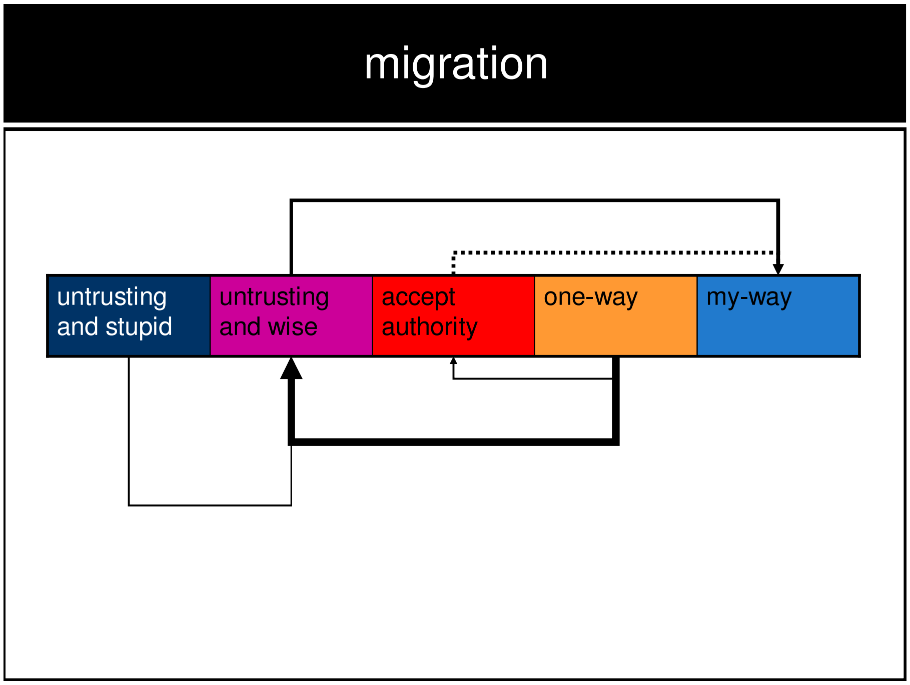


Figure 42 Migration through possible trust models

It is assumed from the research that the majority of the western population start in the untrusting and stupid, untrusting and wise and one-way segments, and from these there is a migration to other trust camps as users/consumers explore and discover the value. This raises an important question though, which was a point made at the opening of the book – some people will engage and some will not. However, what does this classic segmentation look like?

Figure 43 shows a set-based view on the total population broken down into various categories. From the entire population it is assumed that there are some who like Marmite and some who do not; some will engage and others will not; some will trust and some will not. The focus here is on who are the ones who will be prepared to engage. In this segment there are:

- the economically unviable: those who, even if they gave up all their data, have such a small influence on others, or spend so little, or are so unreceptive to the benefits;

- those who sit back and enjoy the benefits. A segment who will impart their data to whomever and, as long as they perceive that there is a fair trade for data in the form of services, are very happy to continue with the relationship; and
- the controlling. This is the segment who wants to take charge of their own digital footprint and exploit for themselves, as we will see in the possible business models.

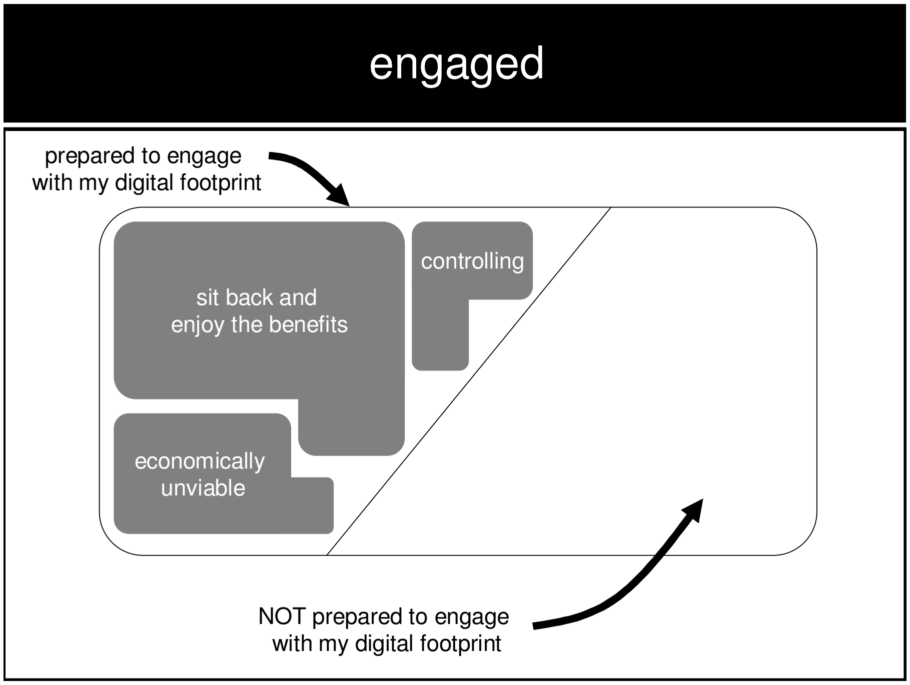

engaged
prepared to engage with my digital footprint
controlling
sit back and enjoy the benefits
economically unviable
NOT prepared to engage with my digital footprint

Figure 43 Who will engage from the overall population?

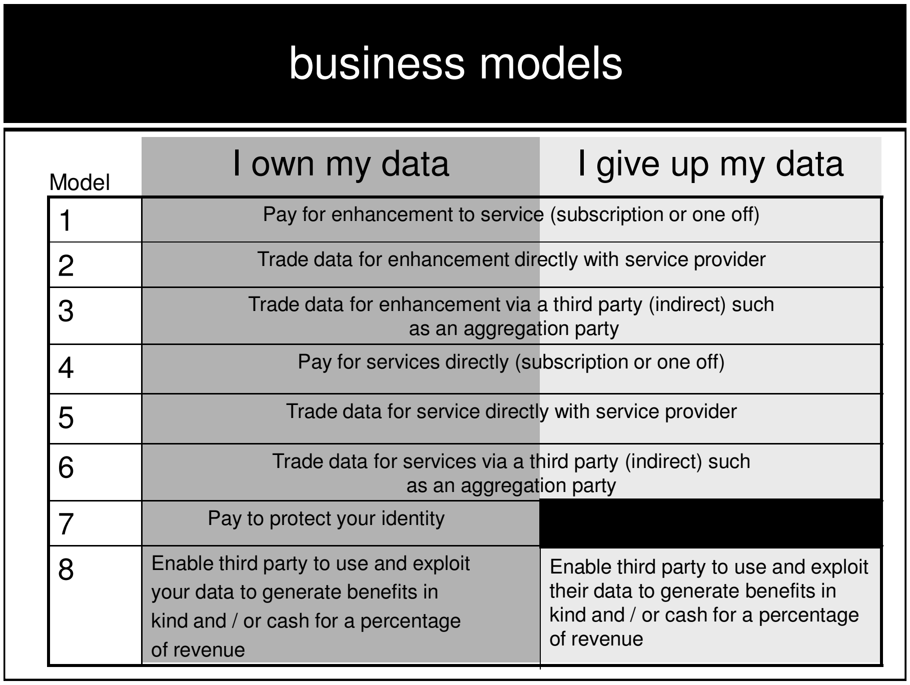
This idea of segmentation and trust brings us to the most basic of questions about the possible business model, which we explore in the next section. The fundamental question is: who owns my data?

BUSINESS MODELS

MY DIGITAL FOOTPRINT, as defined in this book, is about the system of collection, storage, analysis and value. The inputs to the system have focused on data types that can be collected as the user is willing to provide the data (explicit/active) and data that can be gathered by sensor.net (passive/automatic). It has already been stated that there is little value in the long run in collection (harvesting) and storing (regulated). There are possibly a few expectations to this, which are data types that are slow to replicate and can create a differential advantage by having/owning. There is a lot of value in the algorithm and good analysis tool. The understanding of value creation opportunities from analysis will create differential advantage. The outputs or value components are well understood in terms that they can be seen to create value. Additional value is created from the feedback loop as this provides a method to hone, focus and provide depth on responses to an individual based on their data inputs, and also the ability to add flavour, breadth and width based on the individual's social graph. It has been explored who will engage and participate, and how to create this virtuous cycle and keep it going by understanding the bonds and bridges between risk, privacy and trust. This section focuses on the fundamental question of who owns the data and what the business models are that MY DIGITAL FOOTPRINT creates.

The models to exploit MY DIGITAL FOOTPRINT are determined by who owns the data, the options being: 'I own my own data' or 'I give up my data'. Indeed, it is likely that, in many cases, both owning and sharing will be a healthy and amicable compromise, but it is worth focusing on the two separate models of ownership, as from this it is possible to draw clear intentions, value and models. Those that combine joint and shared ownership will, like the rainbow of trust, create and fill in the prime groups.

Figure 44 provides the outline of the models that will be explored.



business models

Model	I own my data	I give up my data
1	Pay for enhancement to service (subscription or one off)	
2	Trade data for enhancement directly with service provider	
3	Trade data for enhancement via a third party (indirect) such as an aggregation party	
4	Pay for services directly (subscription or one off)	
5	Trade data for service directly with service provider	
6	Trade data for services via a third party (indirect) such as an aggregation party	
7	Pay to protect your identity	
8	Enable third party to use and exploit your data to generate benefits in kind and / or cash for a percentage of revenue	Enable third party to use and exploit their data to generate benefits in kind and / or cash for a percentage of revenue

Figure 44 Business models based on data ownership

What data can I own?

I am in no doubt that owning data is difficult, ignoring the fact that even if I can get it, the analysis may be too difficult to create value. Some data (name, address, date of birth, certificates and other identity data as discussed at the outset) are in fact easy to own but hard to prove. Utility bills, bank and credit card statements are easy to collect and are easy to add to your database. Electronically collecting this data is easy and there are programmes that will allow you to build your own spend profile. Data from Amazon, eBay, PayPal, Yahoo, MSN, iTunes, SMS, email, AudioBoo, Palringo, Google, Facebook, Flickr, blog, Twitter, etc, is rather a more difficult case.

Yes, I replicate everything via a small widget on all my devices/screens so that I own a copy of my data (passive and active) and so does the service provider. It may be difficult to replicate some purchase information, especially cash and near field cashless. Currently the terms and conditions of many of these services determine that they have the rights over your data, they are currently prevented from sharing this to help you. How this small widget combines all the data streams is somewhat more difficult, as is how does my algorithm compare data from my social group to add colour and flavour to my services. Finally, how will my analysis output become available so that it can be fed into a service and, hence, I can enjoy the value?

Difficult and practical questions, and I am aware that some companies are working on them and why there is a lot of wealth to be created in this area and why the business model is wide open.

I purposely have not mentioned too many examples as the website http://ww.mydigitalfootprint.com allows readers to add their own examples and promote their own services. I have also avoided the cross-subsidies, two side and freemium categorisation. In many cases, most of the economic models are available; it really depends on the motivation of the service provider, end customer and other parties on how they trade value for data.

I am aiming to gather views on collective action. Could/Should we as creators organise a tribe via Twitter and Google to recover ownership? Google own the link, but didn't create it.

I own my data

Model 1: Pay for enhancement to service

In this simple and traditional business and economics model, data that is collected may provide access to some simple or basic features and functions. However, to gain access to enhanced, professional or premium services, a user needs to pay. The payment can be a one-off fee or a subscription-based model. The pricing of such services is becoming increasingly complex and there is a move from the traditional supply–demand economics to funding pricing that is based on 'bits are free'. In this case, I own my own data and there is a requirement for me to pay for the service of collection, store and analysis. I also need to pay for the presentation of my data to be able to receive the benefits that are offered.

Model 2: Trade data for enhancement directly with service provider

In this case, the user/customer still worries about collection, store, analysis and presentation of the data; however, a service provider takes the analysis and trades this for access to enhanced services as this customer has value. This is a simple barter.

Model 3: Trade data for enhancement via a third party (indirect), such as an aggregation party

Like the direct trade model, the user/customer has the same responsibilities; however, in this case, an aggregator of data collects many users together to create a proposition that enables the user to gain access to many service providers' enhanced services. The aggregator may well be performing a translation service if there are a number of presentation formats for **MY DIGITAL FOOTPRINT** generated from a fragmented market.

Model 4: Pay for services directly (subscription or one-off)

The simplest form don't bother with the data and just buy the service.

Model 5: Trade data for service directly with service provider

Unlike model 2, trading for enhancements, in this case the user/customer has to barter information to gain access for the simplest or basic level. This model assumes that there is a lot of demand for the basic service and there is value in the user/subscriber. BLYK is a good example.

Model 6: Trade data for services via a third party (indirect), such as an aggregation party

This is like model 3; however, in the same vein as model 5 (trade directly) this model depends on an aggregator of data to sit between the user and the provider, much like comparison shopping sites.

Model 7: Pay to protect your identity (digital footprint)

The model here is where a user/customer wants to pay to be protected. Ignoring any value in the data for an exchange and preferring to pay.

Model 8: Enable third parties to use and exploit your data to generate benefits in kind and/or cash for a percentage of revenue

This is a complex model; the user/customer has the same responsibilities as for all ‘I own my data’ models. However, in this case, the user opens the data presentation layer to a trusted third party who will use the analysis of the data to bring and offer services to the user. This offer could be in the form of free enhanced services or cash (or cash equivalents) given to the user/customer, as their data is seen as valuable to certain brands or marketing companies, remembering that **MY DIGITAL FOOTPRINT** covers all the screens of life across mobile, web and broadcast.

I give up my data

In each of these models there is a question about the user giving up control to a single service provider or many. If control is given to many, does each of the service providers have identical data or specific to their utility value? Does eBay only see your auction items, or can it see Amazon’s and Google’s data, or will single service providers, such as mobile phone operators, collect it all and distribute it? But will the user trust them in return, and how moveable will the data types need to be to prevent unfair competitive barriers being erected?

Model 1: Pay for enhancement to service

As with owing my own data, this is a simple model, data that is collected may provide access to some simple or basic features and functions. However, to gain access to enhanced, professional or premium services, a user needs to pay. The payment can be a one-off fee or a subscription-based model. The pricing of such services is becoming increasingly complex and there is a move from the traditional supply–demand economics to funding pricing that is based on ‘bits are free’. In this case, I give up my data to a service provider in exchange for collection, store and analysis.

Model 2: Trade data for enhancement directly with service provider

A service provider takes the collection, store and analysis and trades this for access to enhanced services, as this customer has value. This is a simple barter, but does demand enormous trust.

Model 3: Trade data for enhancement via a third party (indirect), such as an aggregation party

Like the direct trade model, the service provider has certain responsibilities; however, in this case, an aggregator of data collects many users together to create a proposition that enables the user to gain access to many service providers' enhanced services. The aggregator may well be performing a translation service if there are a number of presentation formats for **MY DIGITAL FOOTPRINT** generated from a fragmented market.

Model 4: Pay for services directly (subscription or one-off)

The simplest form don't bother with the data and just buy the service.

Model 5: Trade data for service directly with service provider

Unlike model 2, trading for enhancements, in this case the user/customer has to barter information to gain access for the simplest or basic level. This model assumes that there is a lot of demand for the basic service and there is value in the user/subscriber, and that the user has an economic value to the service provider.

Model 6: Trade data for services via a third party (indirect), such as an aggregation party

This is like model 3; however, in the same vein as model 5, trade directly, this model depends on an aggregator of data to sit between the user and the provider, much like comparison shopping sites. The issue being, as described at the beginning of this session, does the aggregator have the power or the service provider, and which one does the user trust?

Model 7: Pay to protect your identity (digital footprint)

Not possible if you give up your data, the model assumes that you keep it to yourself.

Model 8: Enable third parties to use and exploit your data to generate benefits in kind and/or cash for a percentage of revenue

This is a complex model; the service provider has the same responsibilities as for all 'give up my data' models. However, in this case, the service provider opens the data presentation layer to other trusted third parties who will use the analysis of the data to bring and offer services to the user. This offer could be in the form of free enhanced services or cash (or cash equivalents that could be bartered with the service provider for free access or basic services) given to the user/customer, as their data is seen as valuable to certain brands or marketing companies, remembering that **MY DIGITAL FOOTPRINT** covers all the screens of life across mobile, web and broadcast.

Business model 2.0

The reason for soliciting and collecting consumer data has fundamentally changed due to the web.

Web 1.0 enabled successful companies to gain power by collecting, aggregating, and analysing the customer data they collected and offered something relevant back, this was a useful addition to the zero cost of distribution.

Web 2.0 brought about a new dimension to data, that of data created by the user. Users started to contribute explicit data about themselves, their social graph, how they consumed content or about purchases. This data extended click-and-search data that characterised Web 1.0, and Amazon reviews, using inherent knowledge to offer trusted recommendations, led the way. Online collaboration, such as Wikipedia, increased transparency leading to users being able to help themselves and their community. Web 2.0 is a model of creation, engagement and collaboration where success is built on rewarding users who contribute data, honest data now builds a reputation. Marketing strategies are evolving from how brands engage with consumers based on what data is collected and how they analyse it, to creating value. Consumers now expect and experience relationships and the ability to trade or barter (exchange).

The web has undergone three data revolutions, starting where companies implicitly captured consumer data, which they used to infer intent. This moved to consumers explicitly providing data about themselves and that embraced users/consumers expecting something in return for the data they provided. The third move involved the linkage of the person to social graphs and Web 2.0 was in full swing. The next phase of Web 2.0 will be sharing my digital footprint across platform (web, mobile and broadcast) and service provider. Thus, it will no longer be about what you consumed, but rather the social context of what you consumed. The semantic web (Web 3.0), where the web has intelligence, is a long way off.

Capitalism, it could be argued, has broken the trust between users and companies. For companies to thrive they need to rebuild (or build) trust through engagement, relationships, communication, participation, co-operation and collaboration with their customers, this could mean a shift from the economies of scale to a rainbow (economy) of trust.

When I die?

What happens to my digital data when I die? I have purposefully left this section open to collect views from the web site. Personally I am happy for it to remain, but if I believed it could be a hindrance for my family, maybe I should find a route to erase it? I can see that for the family of a celebrity is could be of great value and great pain, knowing what was done and when, but amazing information for a biography or life works. Divorce law and pre-nuptial agreements have just become more complex.

What should happen to www.tonyfish.com when I am gone?

A FAIRY TALE OR AN EPIC?

Are there any implementations?

While the principles of **MY DIGITAL FOOTPRINT** are beneficial to the customer, the question arises: are there any real implementations? The adoption of the model outlined in this book calls for a shift in mindset and needs the customer to be empowered. Are there initiatives that empower the customer, or is it all a fairy tale and wishful thinking?

In this section, we discuss instances where customer empowerment is being considered by service providers. The issues become much more severe in a converged environment driven by mobile devices. The issues themselves are not one-dimensional, for instance, youth don’t like to be patronised and are often happy with their relatively liberal privacy outlook. Legislators make laws on behalf of the youth – who are not eligible to vote for them. On the other hand, many people trust their providers and/or are not really concerned about loss of privacy. Overall, the mobile network operators have a good track record of managing customers’ data and protecting vulnerable citizens, such as children. Many of the threats may be simply overblown by the media, for instance, with regards to mobile advertising. The fear was, every time you walked near a coffee shop, you would get spammed by ‘location-based messages’ offering you 10% off a cup of coffee. Marketers have drooled over these supposed 'targeted audiences' and privacy advocates have raised the spectre of doom. The reality is very different. It is not very cheap or easy to send targeted mass messages to people asking for '10% off the cup of coffee' when you are near a coffee shop. In other words, the all pervasive 'spam' model of mass targeted advertising is not economically viable and may not ever be.

Lest we forget public services, as every time we interact we leave a record, a medical or benefits report. These records should be the lifeblood of our public services. Securely, if the records are shared and analysed, they can help services to improve the quality of their performance. They can also help to prevent problems: identifying, for instance, those at risk from diabetes or even child abuse. Politics seems to support the view that data sharing is essential for leaner, fitter government but data remains unused and public services remain unsafe and inefficient. Powerful civil liberties lobbies, are against data sharing of almost any kind. Ross Anderson recently co-authored an influential report, 'Database State', (Joseph Rowntree Reform Trust)[75] which argued that for the government to hold 'information on every aspect of our lives' is a real threat to civil liberties. Underneath this debate lies a clear tension between two competing social goods: the desire to defend civil liberties and the need for better public services. Anderson's strong arguments for the former risk undermining the latter. Yet the small risks of a government holding data on citizens are greatly outweighed by the potential benefits.

With this background, we now explore some of the different implications, issues and problems when customers want ownership of their digital footprint data and services. Unfortunately these are only given a brief outline, rather than the in depth analysis and insight that these issues should and can command.

Strategic issues for the service providers

There are four principal strategic issues that service providers face if they are looking to empower the user, deliver conversation, engage in co-operation, collaboration and relationships, and build trust. These four issues are summarised in Figure 45

There are four themes, and for each theme there are two options, one is to remain in the same mindset and the other is to change. If all four are changed, then there is a game changing scenario; changing none to three is not likely to bring about the trust and collaboration that the AAS service engagement model promises and access to the economic uplift.

Taking each of the themes in turn, value, focus, competition and Porter's model[76]:

Value: currently the focus is on cost, cost control, removing cost and generally using web tools and other available resources to focus on getting gross and net margin higher through efficiency. In itself this is extremely good and, as far as corporate governance is concerned, a positive attitude. The reason that game changing focuses on profit rather than cost is about the attitude of mind. In this case, cost focuses on the customer as a single unit for a single transaction. Profit focuses on the lifetime value of the customer. The mandate is not on the single transaction, but trust, relationship, brand and generally how that customer and their influence on a social graph will create profit and value in the long term. Yes, focus on process efficiency is important, but, assuming customers cannot afford to be loyal to many brands, what will build the long-term profit? Trying to attract churning customers in a saturated market is extremely hard.

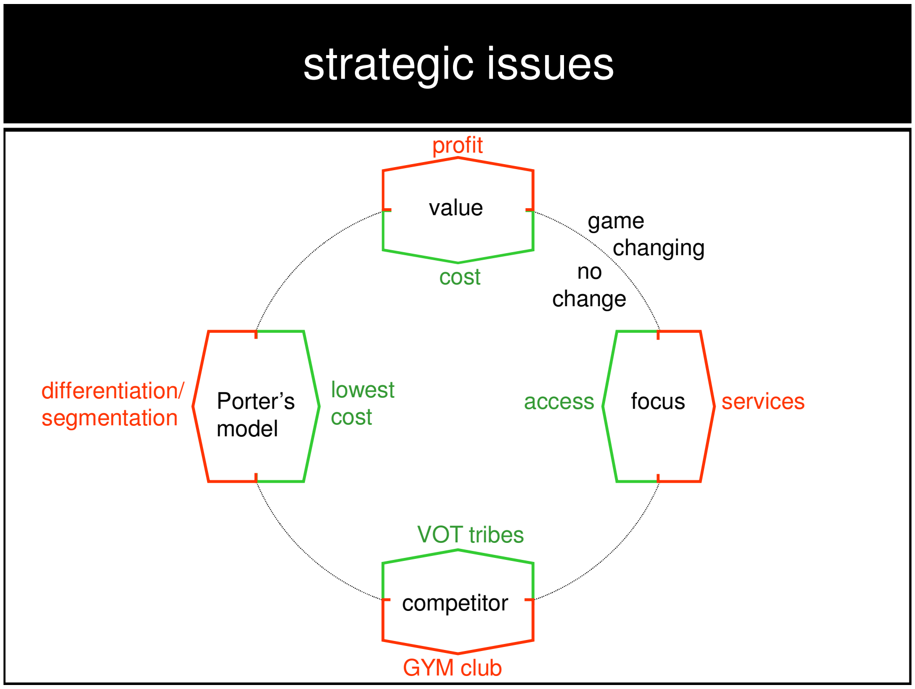


Figure 45 Strategic issues faced by service providers

Focus: this is the change between board attention on access or services. In the UK the Digital Britain report[77] focuses on getting everyone connected. This assumes everyone wants to be connected and have access. At some point this will happen and then, so what. I would propose that 'the digital divide will not always be about access, but those who engage and participate, and those who do not'. There is a requirement from those who want to lead to move away from access and focus on services. Like value, 'services' is about delivering on old-fashioned money, what the customers want. I am not keen on the idea of an access tax, 50p per month on my broadband, (which is different to universal service obligations) to build faster services, but I would pay the same for better spam, no viruses, a digital vault for **MY DIGITAL FOOTPRINT** and access to more valued services.

Competition. On one side is the traditional players in the service provider model, Vodafone, Verizon, Orange, O2, Telefonica, T-Mobile, Three, Sprint, Telstra, etc. These tribes focus on access, cost and competing with each other. The service providers emerging are Amazon, eBay, Google, Microsoft, Yahoo and Apple. Who is better positioned, a mobile operator who knows my location and call record or Apple/BlackBerry® who knows my location (iPhone), my music (iTunes), my contacts, my calendar or what about my applications (Facebook), my preferences (lastFM), Twitter or my blog? Who can gather and collect all the data? Who will the user trust? Who can create the most value? Competition is changing. It is becoming evident that two service providers can offer identical services, but build them from different datasets, ones they own. Competitive advantage, because you can collect data and not share or sell at high charges (e.g. location), is no longer a right to be able to create value.

Porter's model. In this case, this is his work on generic strategies for competitive advantage. He identified differentiation, segmentation and lowest cost providers. Again the whole focus remains on existing competition, access, cost control and, being the lowest cost provider, the operator is stuck on a treadmill to nowhere. Changing one or two of these blocks raises hopes, but others are attacking; with all four focused on the customer and services. Differentiation has so much value; segmentation endless possibilities. Whilst I am not sure about addressing every one as an individual and the two billion segment model, I am convinced about trust and social groups as a complementary model to support and aid insight into customer value and building trust from where they start.

Across these structural issues is an underlying problem insomuch that operators are more tightly regulated than, say, Google. The operators try to protect value for shareholders by not sharing location, which is trumped by Google who find a route round. The options are difficult as is the mindset change and cultural issues, let alone the corporate governance, somehow it is easier to see someone new on the block entering, than some of the existing players embracing the challenge of change.

Structural problems for mobile

As with the strategic issues that each service provider faces, there are four key structural problems. These are focused on the mobile industry; I believe that the mobile is part of solution as it is personal and will have access to more data than the other platforms. Further, in many parts of the world it will be the only device/screen. Essentially this is the summary of where Ajit Jaokar and I saw Mobile Web 2.0 when we wrote *Mobile Web 2.0* in 2006.

The four problems are shown in Figure 46, they are: network, applications and content, device, and platform. As with the strategic issues, remaining on the existing track will keep everything controlled and closed and will keep the mobile ecosystem closed. Reversing all of the problems delivers the open mobile ecosystem and opens the opportunities to compete. In many ways these are the implementation issues that the strategic problems throw up for a mobile operator.

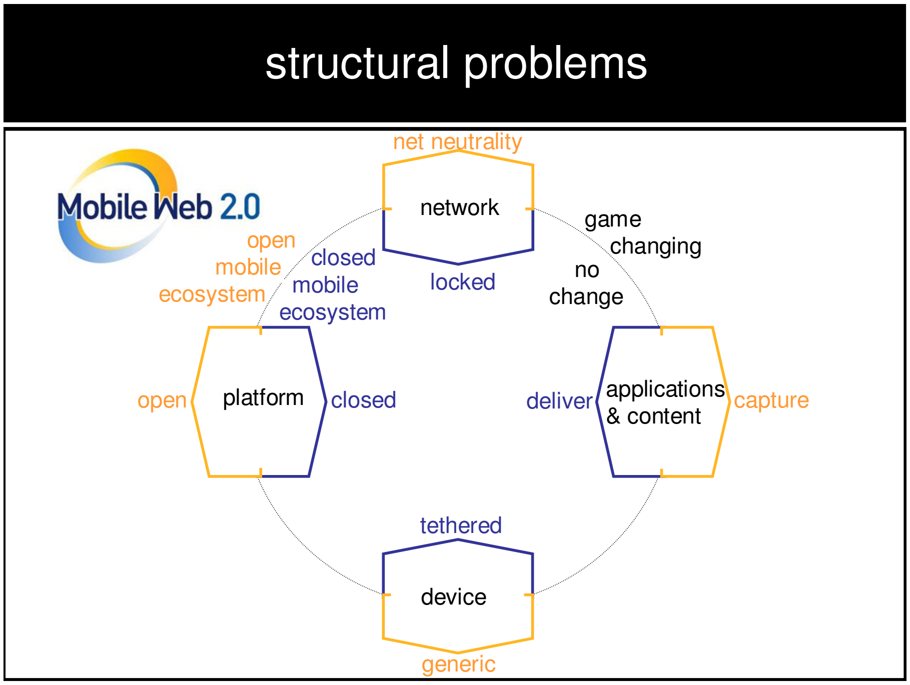


Figure 46 Structural problems in the mobile ecosystem

Network: there are books on this topic and a way to consider it simplistically is the move from a locked closed, controlled, proprietary network and API interfaces (at the network and signalling level) to provide network interoperability and billing to net neutrality. The issues are far wider and deeper covering economics, competition and technical. The takeaway here is about executive thinking of control, are relating back to the cost issues in the strategic discussion, but also the entire engagement model and thinking required for a focus on value and growth. Owning the network is not important, as collecting or storing data. Analysis and the algorithm are wealth creating, as is delivering the outputs. The operator can connect the feedback loop in real-time, but that is just data gathering (collection) not wealth creating.

Applications and content. There has been a long-term focus on the delivery of content, indeed one could argue that most of the 3G business cases centred on the user consuming content and applications on the phone. **MY DIGITAL FOOTPRINT** suggests that there is more value in getting data off a device than providing content and applications to a device. The focus on delivery and trying to be a content company has diluted earnings and growth of service providers for some time. Capturing data should be the focus, but remain mindful of where the wealth from this activity is.

Device. Fashion and shiny are absolutely drivers for purchase decisions, but the device argument is the difference between a device that is locked to a service provider (maybe because of the subsidiary business model) but it is tethered. The generic device is one that will work cross platforms; it will inter-operate with the web and broadcast. The idea about devices is not about one device but what the user can do with the device.

Platforms. Included in this generic bucket is both middleware platforms (BSS, OSS, SMS and MMS gateway, IM, security, WAP, Voicemail, etc) and edge-of-network platforms (location, real-time video, gaming, OTA, payment, phone book). These platforms, to a large extent, work on a single operator, but try porting your network-stored address book and it becomes evident that they are closed with the intention of keeping you loyal. When I move operator why should I lose my stored SMS or saved emails? Does this really make me more loyal, or is this like the banking industries intention of making it so hard and awkward to migrate, you lose the will to live? An open mobile ecosystem is about enabling users to use, create, engage and collaborate, these are the focus of an open mind to how value and wealth can be created. I expect that there will be a new layered architecture at some point, one based on services such as digital money and digital identity which will allow platform to inter-operate.

The current mobile and service provider ecosystem is closed. Networks are locked, devices are tethered, content is mostly consumption-based and platforms are closed. The options for the industry are to stay as it is (no change) or to adopt a more open, game changing approach. However, as networks evolve and converge, more opportunities will arise for new entrants to break the rules. One such example is Femtocell (bearer) aware web services. It is hard to sell capabilities of networks themselves (or for that matter to charge for networks). However, operators can sell services. Customers understand services; they are used to paying for them. The basic version of the service could be free, followed by some premium features.

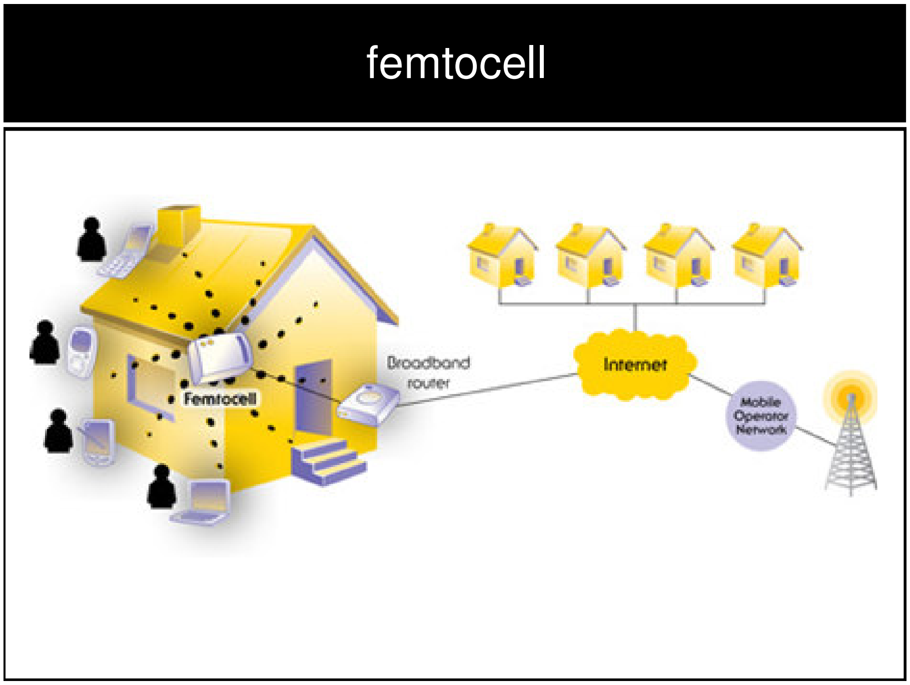
From a mobile perspective, services could be:

- long tail (completely decoupled from the device or the network – this is mainly in App Stores);
- coupled to the device, for example, deep integration of a web service to the device (e.g. address book integration of Facebook (INQ1) or Skype (N97)); or
- coupled to the network.

Learning from Amazon (and Web 2.0 in general), the more the customer interacts with the provider, the better the service could be because the provider captures insights and preferences from the customer and can use them to enhance the service.

While there are many examples of long tail services, and services coupled to devices, there are not many examples of services coupled to networks. In one sense, services should not be coupled to networks. However, Femtocells could provide an exception to this model by creating services that are useful to the customer. From an operator standpoint, they provide an opportunity for customers to stay on their network longer and to provide services that can be improved by usage.

As per the Femtocell forum[78], Femtocells are low-power wireless access points that operate in licensed spectrums to connect standard mobile devices to a mobile operator's network using residential DSL or cable broadband connections.



femtocell

Figure 47 Femtocell

Customers use mobile phones in the home, even when there's a fixed line available. Friends and family usually call a mobile number first, and it's where messages and contact lists are stored. Femtocells are important to the operator because mobile operators can get a share of the fixed call revenues. A study from ABI Research says that by 2011 there will be 102 million users of Femtocell products on 32 million access points worldwide.

Femtocells also fit in well with LTE (4G networks) by providing 'in building' coverage and offloading high traffic coverage to the Femtocell network.

Services based on Femtocells would need to be 'bearer aware' (detect that the network has switched to Femtocell from 3G or vice versa). This translates to the customer being 'at home' or 'at work' (depending on the Femtocell). What services are possible from bearer aware scenarios? Some examples from the from the Femtocells forum:

'Femtocells enable all kinds of new services to be created for your mobile phone when it's at home. Example Femtozone services could include:

- a virtual home number (rings all mobile phones currently in the home), allowing families to keep a home number even if they cancel their fixed line phone;

- get automatic SMS alerts when your kids arrive or depart the home, providing reassurance for working parents;
- automatic 'I'm at home' profile/presence update on social networking websites;
- automatic back-up of photos and videos from your phone to the web and/or your PC when you arrive at home; and
- automatic podcast reload on your phone when you get home, avoiding the hassle of having to manually synchronise with a PC.'

Future generations of Femtocells will connect your mobile phone to your home network, allowing you to do many things such as play a slideshow of photos from your phone on your TV, stream videos from your digital video recorder to view on your phone, and use your phone to control other devices in the home (e.g. to instruct the hi-fi system to play music stored on a home computer or media server).

We could extend this idea to any web service (specifically Twitter), so Femtocells could basically create a location version of any web service. For starters, we could start with two 'places' we all most spend time at: home and work. Then, when we roam into these cells – the phone could trigger messages from these services (just like the Facebook fridge reminders). Over time, we could extend this idea to other Femtocells in other locations (e.g. cinemas).

The web

Having reviewed the generic strategic issues and some specific mobile problems, we now consider the web.

The web is the most mature medium in terms of rights of the customer (television forces us to watch advertising). In contrast, the web is based on the idea of customer empowerment. The discussion about customer empowerment on the web will continue and, with the uptake of mobile devices, the discussion will become more complex. Below are some of the areas where discussion continues.

Full disclosure

Much of disclosure and transparency is related to trust and reputation which was discussed earlier in the book. It is worth noting at this point that disclosure and transparency (trust and reputation) will add to customer loyalty. Transparency is a complex issue especially as technology gets more complex. However, certainly an attempt needs to be made to educate the user how to manage their own data and the implications of revealing their data to providers.

Vendor Relationship Management

There are a number of emerging initiatives that empower the customer. One such initiative is Vendor Relationship Management (VRM), outlined at Harvard Law School[79]. According to their site: 'VRM, or Vendor Relationship Management, is the reciprocal of CRM or Customer Relationship Management. It provides customers with tools for engaging with vendors in ways that work for both parties.

CRM systems for the duration have borne the full burden of relating with customers. VRM will provide customers with the means to bear some of that weight, and to help make markets work for both vendors and customers — in ways that don't require the former to "lock in" the latter.

The goal of VRM is to improve the relationship between Demand and Supply by providing new and better ways for the former to relate to the latter. In a larger sense, VRM immodestly intends to improve markets and their mechanisms by equipping customers to be independent leaders and not just captive followers in their relationships with vendors and other parties on the supply side of the marketplace.

For VRM to work, vendors must have reason to value it, and customers must have reasons to invest the necessary time, effort and attention to making it work. Providing those reasons to both sides is the primary challenge for VRM.

VRM principles:

- Relationships are voluntary.
- Customers are born free and independent of vendors.
- Customers control their own data. They can share data selectively and control the terms of its use.
- Customers are points of integration and origination for their own data.
- Customers can assert their own terms of engagement and service.
- Customers are free to express their demands and intentions outside any company's control.'

We expect that there will be other initiatives that will arise similar to VRM with the goal of empowering the end user.

User-managed anonymity of data

Increasingly, we will see many mechanisms that anonymise data at the source with the user’s permission. User-managed anonymity could provide both safety, but also a business model, based on trust.

Web 2.0 has taught us the concept of harnessing collective intelligence. Companies, such as Google with PageRank, Amazon with Amazon reviews, and others have benefited from the idea of harnessing collective intelligence. So, the business model for the provider in harnessing collective intelligence is proven. Creators of data own the copyright to the individual data elements (reviews of books) but the providers own the value gained from harnessing that granular data. Providers of services would postulate that the granular data elements don't hold commercial value – it is only the aggregated elements (harnessing collective intelligence) that has value; or does it?

In other words, is there any value in the granular data as opposed to the aggregated data? Let us put this into perspective.

Currently, providers can (largely) provide personalised services and some form of targeted advertising and also segmentation that does not need customers to 'own' their own data. The question arises: Is there a model which would enable the providers and customers to both benefits if data is owned and managed by the customers themselves?

To explain this issue, we have to understand the problem of k-anonymity. The problem and solution of k-anonymity relates to re-identifying individuals from multiple datasets even if the data is (supposedly) anonymised. As we become creators of data, with Web 2.0 and especially Mobile Web 2.0, the problem becomes significant because data is collected by providers at a phenomenal rate. It is then possible to potentially re-identify people from datasets.

This discussion explores the possibility of making the problem of anonymisation into a business opportunity. Essentially, if data is anonymised at the source and is under the control of the customer, the customer will trust the provider who anonymises their data. In return for that trust, the customer could volunteer to reveal attributes about themselves which would enable the provider to create personalised advertising campaigns and also to be used in segmentation. This benefits both the providers (protection from legal action, personalised advertising, segmentation) and also the customers (anonymised data, personalised services, etc).

To elaborate this idea further, we consider the example of k-anonymity. The concept of k-anonymity is summarised in a paper by Latanya Sweeney (k-anonymity: a model for protecting privacy) School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA as per the paper abstract[80]. 'Consider a data holder, such as a hospital or a bank, which has a privately held collection of person-specific, field-structured data. Suppose the data holder wants to share a version of the data with researchers. How can a data holder release a version of its private data with scientific guarantees that the individuals who are the subjects of the data cannot be re-identified while the data remain practically useful? The solution provided in this paper includes a formal protection model named k-anonymity and a set of accompanying policies for deployment. A release provides k-anonymity protection if the information for each person contained in the release cannot be distinguished from at least k-1 individuals, whose information also appears in the release. This paper also examines re-identification attacks that can be realised on releases that adhere to k-anonymity unless accompanying policies are respected. The k-anonymity protection model is important because it forms the basis on which the real-world systems known as Datafly.' Argus and K provide guarantees of privacy protection.

The basic problem applies to all data. As we become creators of data with Web 2.0 and especially Mobile Web 2.0, the anonymising of data becomes a problem. Historically, data has been anonymised by removing explicit identifiers such as name, address, telephone number, etc. Such data looks anonymised, but it may not be when co-related with another dataset which may help to uniquely identify people.

For example, as per the paper, in Massachusetts, the Group Insurance Commission (GIC) is responsible for purchasing health insurance for state employees. Because the data were believed to be anonymous, GIC gave a copy of the data to researchers and sold a copy to industry. It was then possible to co-relate it with normal voter registration. This information can be linked using ZIP code, birth date and gender to the medical information, thereby linking diagnosis, procedures, and medications to particularly named individuals. For example, William Weld was governor of Massachusetts at that time and his medical records were in the GIC data. Governor Weld lived in Cambridge, Massachusetts.

According to the Cambridge Voter list, six people had his particular birth date; only three of them were men; and, he was the only one in his 5-digit ZIP code.

The solution for this problem is potentially a policy manager managed by the user which anonymises data (perhaps at the source itself, i.e. the client). From a mobile perspective, the policy manager should be designed to:

- be controlled by the user (user sets the policies); and
- manage all data – not just location

This is a potentially win-win situation because essentially, if data is anonymised at the source and is under the control of the customer, the customer will trust the provider who anonymises their data (and in turn protects them). In return for that trust, the customer could volunteer to reveal attributes about themselves which would enable the provider to create personalised advertising campaigns and also to be used in segmentation. This benefits both the providers (protection from legal action, personalised advertising, and segmentation) and also the customers (anonymised data, personalised services, etc).

The approach potentially provides a compelling argument for both the provider and the customer. In doing so, it is different from current approaches because, at the moment, advertising, segmentation, etc, can be implemented on a best case basis (without knowing the exact data from the customer) – but a trust-based approach will benefit all parties through accurate personalisation that is managed by the user.

Revocation

Conventional privacy models lean towards a closed, digital fortress. These can take many forms – Linkedin introductions, signed applications, third party trust endorsers, etc. In turn, conventional revocation models are provider-driven (e.g. certificate removal of malicious applications). The current methods don't fit the current open web ecosystem and, more importantly, a future web-based ecosystem, where there is a tendency to give up privacy with a younger generation. It could be possible to consider a user-driven revocation model.

Social networks are increasingly going to be the primary form of interface to the web for many of us. Unlike the open web, the social network has some form of structure (profiles, messages, etc). In this scenario (within the social web), privacy and revocation could go side by side, leading to a new privacy model. In other words, the user will be open to contact but, in return, she will choose to exercise the right to terminate that contact if I need her to. This model is based on ‘innocent until proven guilty’ as opposed to the existing digital fortress ecosystem (guilty until proven innocent).

Admittedly, the revocation engine may not work in the context of the whole web, but it may well work in the context of a social network. Currently, the spam features of Gmail work in a similar way (except Google does the revocation implicitly on our behalf).

The model is a switch on existing privacy models (strengthen the revocation not the moat bridge; let people cross freely at the moat, but always have the revocation engine as a defence mechanism).

The data portability implementation

Increasingly, social networks are becoming the primary point of interaction for many people, especially the young. In that context, social networks could be viewed as the highest layer of a unified mobile stack which spans both the web and the mobile domains.

The rise of social networks has taken the world by surprise with companies like Facebook and MySpace becoming household names. However, as social networks mature, users have realised that they can be closed. Further, privacy advocates have realised the implications of personal information being available on social networks. In response to these developments, the *Bill of Rights for the Social Web*[81] was released by Joseph Smarr, Marc Canter, Robert Scoble and Michael Arrington.

This document states: ‘We publicly assert that all users of the social web are entitled to certain fundamental rights, specifically:

Ownership of their own personal information, including:

- their own profile data;
- the list of people they are connected to;
- the activity stream of content they create;
  control of whether and how such personal information is shared with others; and
  freedom to grant persistent access to their personal information to trusted external sites.

Sites supporting these rights shall:

- Allow their users to syndicate their own profile data, their friends list, and the data that's shared with them via the service, using a persistent URL or API token and open data formats.
- Allow their users to syndicate their own stream of activity outside the site.
- Allow their users to link from their profile pages to external identifiers in a public way.
- Allow their users to discover who else they know is also on their site, using the same external identifiers made available for lookup within the service.

The overall goals are concise and clear. In practise, this means as a user you have freedoms like:

- If you are on MySpace and your friend is on Facebook, you should be able to contact them, share links, find out their friends.
- If you leave a social networking site, you should be able to 'take' all your data with you; including your contacts, friends and the content you have created.

In practise of course, this is not easy to implement and there are a number of initiatives created to address this problem, because user data is not portable across social networks. The data portability project[82] works towards the creation of open standards for the ability of data to be reused across interoperable applications. According to their website, the vision of the data portability project is: 'Data portability enables a borderless experience, where people can move easily between network services, reusing data they provide while controlling their privacy and respecting the privacy of others.'

The idea of data portability is based on a number of open technologies such as OpenId[83], Microformats[84], RDF (Resource Description Framework)[85], XMPP[86] (Extensible Messaging and Presence Protocol), FOAF[87], SparQL[88], OAuth[89], RSS[90], and OPML[91]. Besides these, we have friends connect from Google[92], and Facebook connects from Facebook[93] and MySpace data availability[94] all of which are getting traction.

Thus, data and privacy go together. For marketers, the temptation to treat social media as a 'channel' is strong along with the desire to retrofit the new world of communication to the familiar world of brands, traffic, audiences, growth, etc. However, this is not always in the consumers' interests. Social networks and mobile, due to their unique, personalised nature will have to go beyond 'opt-in' and may need higher standards beyond statutory regulation, based on moral and ethical integrity with a view to protect consumer interests. The future lies in empowering the customer and building relationships based on genuine trust, openness and transparency. Ultimately, such relationships will be profitable to the providers and the brands.

Data and privacy form the bedrock of this multi-way conversation between the marketer and the participants. Ultimately, we see the participants and the marketer enter into a trusted relationship based on transparency where the participants share data about themselves and entrust the marketer with their data in return for better and personalised services.

WHAT FUTURE WILL WE CHOOSE?

We have seen many facets of digital footprints. At the time of writing (Summer 2009), we have a choice – we can choose a dark vision based on an intrusive future, or we can choose an enlightened future based on a mutually beneficial relationship between service providers and customers. In reality this stark choice is not possible as most users will not choose their own benefit over a corporate offering to do it for them and there will be a balance. However, let's first review the dark side (the intrusive future).

The dark side

The digital signage

Interactivity could take an intrusive dimension as 'ads have eyes'. This idea appeals to advertisers – since it could 'personalise' advertising. But customers may not like the concept for its potential intrusion of privacy[95]. In its ultimate form, mechanisms such as facial recognition cameras and other invasive technology could be used. Cameras could record the race and gender of passers-by. Bluetooth and RFID tags could be used to solicit interaction and gain sign-up. Data could be collected on customers and analysed. Information on customers could be gathered without explicit notification. Interactivity and measurement makes the advertising valuable for the provider. Advertising could even 'change' depending on who is interacting with it. Advertisers could create profiles without seeking permission.

Deep packet inspection

A more ominous method is ‘deep packet inspection’ – implemented in the UK by companies such as Phorm[96]. Phorm is an advertising company whose model involves paying your ISP to handover information on which websites you visit. Some large Internet providers, such as BT plc, have run trials with Phorm. In 2009, the European Commission started legal action against Phorm[97].

The problem with network level tracking of user behaviour (even if you did get permission) is this – most people do not understand networks. And people do not trust what they do not understand. Personally, I believe that this should not be on the dark side as it is all to do with positioning and trust as any powerful technology can be used for good and evil. By contrast, Phorm are probably better at protecting and managing user privacy than Facebook and Google who do much more dangerous things with customer information resulting, for example, in the recent concerns about social networks generally raised by Article 29 Working Party[98]. Facebook and Google clearly have the resources and business clout to manage such criticisms. Such balancing will continue on virtually every privacy issue.

Considering the same issue from the viewpoint of a mail postal service (physical), the issue is less to do with knowing who a letter is sent to (any mail service knows this) and knowing the contents of the letter (opening it); the difference being keeping the contents of the letter and storing it, or throwing away the data and just keeping the analysis. If the output of turning the inspection into value is created, everyone is happy, as long as the contents are immediately thrown away. The disposal of data/content and keeping analysis is socially acceptable, keeping data/content is socially unacceptable, but both are legal. Whatever your opinion, the judgement should be based on the usefulness of the analysis; as we have seen, the focus should be on collecting valuable datasets and not just all data, as a lot of data is rubbish.

The worried well

This is scare headline material from any daily tabloid; it is the dark side as it is uninformed, one-sided and sells newspapers. Unfortunately it may prevent people from trying Marmite before they had the chance and therefore will miss out on the value.

Yes, your ISP could jump through some hoops to figure out that someone in your house is blogging about Enter Shikari[99] and falsely assume that it's you. Your son isn't on the electoral register because he's 14 and your wife isn't on it because she's American. Having snooped on you, the ISP then makes a load of false assumptions and you get a load of rubbish targeted at you. That would be just annoying and an invasion.

You then receive threatening e-mails, purporting to come from a foreign gang, because you posted on a message board about credit cards and they backed your name through Google to trace you down. Therefore you no longer use your real name on some of the message boards that you enjoyed interacting with and your footprint becomes dispersed across the four winds. You can no longer enjoy the value that could have been created as someone else has prevented holistic harvesting. In reality it could be the Department for Health tracking down people posting unhealthy recipes or boasting about how much they had drunk and passing it to insurance companies for increased premiums. It is just too easy to make up scary stories, but I can see how the next version of anti-virus software is going to add value to me and make me pay a premium.

The enlightened view

Many companies – especially telecoms network operators persist in the quaint notion of owning the customer. This obsession causes a blind spot because they miss out on the opportunities of knowing the customer (leaving aside the fact that the customer does not want to be ‘owned’). While operators continue to try to 'own' the customer – the web players continue to 'know' the customer better. Since most people will want to centralise their personal information with a few players, the longer telecoms delay in knowing the customer, the more they will lose out. The familiar argument goes, we [telecoms] have a lot of details about our customers, knowing the transactional information in CDRs, we can tell you a lot about the customer. But so can a bank.

If my bank attempts to 'bundle' insurance every time I book a flight on the web through my bank card most people will (rightly) be annoyed and sue them for lack of privacy. So, the notion of customer intimacy on the basis of CDRs (Call Data Records) alone is ambitious at best and litigious at worst. But what do we mean by knowing the customer, and how can we do it? Three things are needed:

For starters, the customer needs some free incentive to part with their personal information. This assumes that all privacy guidelines and permissions are adhered to.

Second, we need an open ecosystem/platform and factors like data portability are important (leverage the network effect and get others, third party developers, to do work for you and grow the ecosystem both in terms of content and new users). Again, Google/Android and Facebook exploit this strategy best.

Finally, we need touchpoints (places where the operator can interact directly with the customer and can deliver the advertisement). Here again, open systems are important as a place where advertisements can be placed. The call centre is a customer touchpoint, but it is an expensive one. Hence, operators need to make efforts to work with social networks and develop the mobile web.

ENDING THE JOURNEY AND THE TAKE AWAY

Like Marmite, some people like the idea of digital footprints and some do not, but, irrespective of personal preference, we all leave digital footprints and as we have discovered they are much more than identity. Digital footprints are about where we have been, for how long, how often, with whom and the inter-relationships, they are memories and moments. Digital footprints are not about your identity, your passport, bank account or social security number. Digital footprints come from your mobile, web and TV interactions and comprise of the digital data and also the Metadata of who we are, the true value and why the ownership of this data class is the web's next battleground.

Many interactions, such as creating a social networking profile or commenting on a picture on Flickr, leave a digital footprint. In a mobile context, CDRs (Call Data Records) are the transactional data that constitute the user's digital footprint. But the mere availability of transactional data alone is not enough since privacy and data protection rules will apply to the usage of data, and rightly so. It is the ability to store, analyse and create value from the digital footprint that differentiates the study of digital footprints. In other words, if we all left digital footprints – and nothing happened to those footprints – then there are no concerns and no benefits.

The concept, however, of collecting user consumption data has helped with the improvement, measurement and accuracy of marketing for the last 15 years, but there is a subtle, almost unnoticeable change happening which is advancing the assimilation and value from the collection of user data. This change is moving the orientation to who you are consuming with in addition to what you are consuming, the measurement and new value is being created from understanding the process of creation and consumption.

The web has already undergone three data revolutions, starting where companies implicitly captured consumer data, which they used to infer intent. This moved to consumers explicitly providing data about themselves and that embraced users/consumers expecting something in return for the data they provided. The third move involved the linkage of the person to social graphs and Web 2.0 was in full swing. The next phase of Web 2.0 will be sharing **MY DIGITAL FOOTPRINT** across platform (web, mobile and broadcast) and service provider. Thus, it will no longer be about what you consumed, but rather the social context of what you consumed and with whom.

MY DIGITAL FOOTPRINT extends the idea of raw data to the wider concept of capture, store, analysis and value created from data generated through digital engagement. This process is based on a structured approach incorporating inputs and outputs, and a feedback loop that governs the whole process. This feedback loop progressively enriches and refines the outputs (value) over time. The analysis phase is able to take raw data from various sources (which I refer to as the digital footprint) and generate value in the form of services such as personalisation, reputation or discovery – this analysis output of this process I call the ‘behavioural DNA’. The value derived from the process is **MY DIGITAL FOOTPRINT.**

The algorithm is the component that creates the value; the outputs are how that value is realised. The algorithm that computes, combines, compares and analyses the digital footprint is the differentiator for a service provider. A good algorithm can produce success, a poor one can bring a company down. Whilst a company can implement the same algorithm, the way it is presented to the community will also lead to success or failure. This provides the bridges and bonds to risk, trust and privacy and how governance and the culture of the company, led by the CEO, will bring some brands down and others to new heights. Considering the algorithm is important. It is a very complex component and the part of the process that will bring differentiation.

The two central ideas which underpin value in **MY DIGITAL FOOTPRINT** are: the feedback loop which enriches the digital footprint and the role of the mobile device in enriching the value from **MY DIGITAL FOOTPRINT**. The ability to get data out of or off a mobile device lends itself to the unique advantage a mobile device has. As presented, the mobile device once prevailed for the consumption of content and has evolved to enable the capturing of data on how, what you consume and with whom.

The inputs to **MY DIGITAL FOOTPRINT** have focused on data types that can be collected as the user is willing to provide the data (explicit/active) and data that can be gathered by sensor.net (passive/automatic). I believe that there is little value in the long run in collection (harvesting) and storing (regulated). There are possibly a few exceptions to this, which are data types that are slow to replicate and can create a differential advantage by having/owning. There is a lot of value in the algorithm and good analysis tools. The understanding of value creation opportunities from analysis will create differential advantage. The outputs or value components are well understood in terms that they can be seen to create value. Additional value is created from the feedback loop as this provides a method to hone, focus and provide depth on responses to an individual based on their data inputs, and also the ability to add flavour, breadth and width based on the individual's social graph. It has been explored who will engage and participate, and how to create this virtuous cycle and keep it going by understanding the bonds and bridges between risk, privacy and trust.

Is digital footprint not just advanced CRM and behavioural marketing? There is a focus on products and services and embedded in this approach are great works such as Kotler's '5 P's of Marketing', Porters '5 Forces', Andersons 'The Long Tail', 'Crossing the Chasm', and the 'Boston Matrix' these still stand and will continue to do so along with many others. Therefore, what will change and why could it be radical? I outline, based on two-dot-zero (2.0) thinking, that there is a move from product to process, from service to method. Marketing metrics pick up data on product/service/channel/marketing performance and this data is used to hone the product and channel and its appeal. This classic data collection for CRM is important and will continue, what I am suggesting is that a new class of data will become available on who you consumed the product/service with, where you consumed it, which channel influenced your decision, who you influenced and the most incredible aspect is that you will not even know that you have provided, given, contributed the underlying data. No user interaction! An outcome from **MY DIGITAL FOOTPRINT** and subsequent generation of your behavioural DNA is that, as a consumer, your entire 'Six Screens of Life' experience, the screens that form the ways you consumer digital media will be uniquely controlled and designed for you and your social media crowded by the control of your digital footprint.

The data that is collected will form your digital footprint and provides one side of the business model. This side includes the analysis of the data to generate your behavioural DNA, from which value can be extracted. This side of the business model will be governed by user control, trust and hiding/preventing true identity from being discovered and exploited.

The second side of the business model is the extraction of value which will come from many different avenues, some direct (e.g. selling a service for cash) and some indirect, recommendation, context or the ability to barter. Digital businesses will become dependent on locating itself between the user as the provider of data and the consumer of data! However, there are a number of possible controllers of the digital footprint including: the user, a trusted friend, a commercial locked platform, a commercial open platform, government or a QUANGO! I hope I have explained the unique role of the mobile in the collection of data for your digital footprint and explained that the original value is changing from what you are doing to who you are doing it with; from product/service knowledge to process and method insight.

Business strategy and the link to valuation is predicated on the NPV (Net Present Value) of future spend of your customers. By this I mean; win a customer, keep them loyal and I can depend on their future spend to create value. Advertising and marketing has focused on this principle and through the improvements in measurement and metrics can now demonstrate how successful this approach is. However, I believe and suggest that this fundamental ethos is changing, evolving, morphing developing or being re-written depending on your view, at worst you may see it as obvious and still based on NPV of future spend, but there is going to be a change and there will be new value created from **YOUR DIGITAL FOOTPRINT.**

EATING MY OWN ANALYSIS

This book will now try and live out the story of rejecting the old model where production was cheap (time) and distribution was expensive (print), and move to the new model where production is cheap (time) and distribution is free (online). The book is being offered online for free; bits cost nothing. There is a paid-for model of the book, this is the physical version; atoms cost money. The free, online book can be printed; however, it has been purposefully designed to make it hard to print the entire book. The reason for this is to make people consider the environmental effect of printing at home or in the office, but if someone wanted the PDF for local reference and research, it is free to download. The proper print version from Amazon is both more economical and better for the environment as it is printed on demand locally.

The website means that I can update as I go, recognising that this area is in continuous flux.

I believe that there is some value in this book, but there is more value in the comments and community. To read the comments a reader has to go by a round-robin email registration and confirmation. This is a small price of time to pay to allow you to read comments that others are paying to leave. In registering, the terms are set to allow me to analyse what you are looking at and to be able to use this data to advertise, as you are getting the contents for free.

The final part of the current model requires readers to pay to leave comments. This serves several purposes, it moves the abundance of time (wasters) into the realms of scarcity (money), it allows tracking for anything illegal and makes readers think about the value of their comments and contribution. As an added value to those who pay and comment, they are added to the creative commons publishing license to be able to replicate the work further including the rights to alter, transform, and build upon this work. Further, knowing attention spans are short, for those who want a speed read, the summary is set at a premium price.

From those who read or comment on the book or the summary, I now have a database of interested parties with terms and conditions to allow me to offer additional services and comments – this allows me to market services such as comment updates, debate, podcasts, new books, focused issue commentary, conferences and other mechanisms to a pre-selected and interested group who will pay for value. Further, if a company wishes to pay me to educate key staff or support board level strategy and thinking, I will charge a healthy rate. The comments will provide a feedback, as with other web data, to produce a second updated version at some point.

Thanks for joining the debate and now let’s see where it takes us.

AUTHOR'S PROFILE

Tony Fish is an experienced and qualified board level executive with professional experience crossing web, mobile and TV. He is an acknowledged leader, strategic thinker, creator and public speaker. Tony has spent over 20 years in technology, media and telecoms and has worked for leading blue chip, early stage companies and in venture capital.

Tony has many public acknowledgements as a leader in 2.0 thinking, through independent awards such as top 10 in *The Observer* and *Guardian* newspapers 'The future 500 rising stars', and from global recognition from his peer group.

Tony holds an MBA (1993) from Bradford School of Management and a B-Eng (1990) from Reading University in Electronic Engineering. Tony holds the following professional achievements: Chartered Engineer (C-ENG), Fellow of Institution of Engineering and Technology (FIET) and Fellow of the Chartered Institute of Marketing (FCIM).

Tony has previously co-authored two books on mobile and innovation with Ajit Jaokar: *Mobile Web 2.0: the innovators guide to developing and marketing next generation wireless/mobile applications*. August 2006; and *OpenGardens, the innovators guide to mobile data industry*, December 2004.

MY READING LIST (JUNE 2008 – JUNE 2009)

Grown Up Digital: How the Net Generation is Changing Your World by Don Tapscott

Free: The Future of a Radical Price: The Economics of Abundance and Why Zero Pricing Is Changing the Face of Business by Chris Anderson

The Future of Reputation: Gossip, Rumor, and Privacy on the Internet by DJ Solove

Nudge: Improving decisions about health, wealth and happiness by Thaler and Sunstein

The Future of the Internet: And How to Stop it by Jonathan Zittrain

We-think: Mass innovation, not mass production: The Power of Mass Creativity by Charles Leadbeater

Who Controls the Internet?: Illusions of a Borderless World by Jack Goldsmith, Tim Wu

Tribes by Seth Godin

Consumer Kids: How Big Business Is Grooming Our Children for Profit by Ed Mayo, Agnes Nairn

Convergence Culture: Where Old and New Media Collide by Henry Jenkins

The Digital Identity Reader 2008 by Dave Birch

Social Media Marketing by Jaokar, Jacobs, Moore and Ahvenainen

[1] http://en.wikipedia.org/wiki/Metadata July 2009

[2] http://battellemedia.com/archives/000647.php

[3] http://www.pewinternet.org/Reports/2007/Digital-Footprints.aspx

[4] http://www.wired.com/politics/law/news/1999/01/17538

[5] http://www.identityblog.com/

[6] http://www.idcorner.org/

[7] www.spock.com

[8] http://www.spock.com/do/pages/help#claim-search-result

[9] http://www.forbes.com/2009/01/12/mobile-marketing-privacy-tech-security-cx_ag_0113mobilemarket.html

[10] http://en.wikipedia.org/wiki/Data_Protection_Act July 2009

[11] http://www.telegraph.co.uk/news/newstopics/politics/4339771/Threat-to-privacy-under-data-law-campaigners-warn.html

[12] http://adage.com/digital/article?article_id=134036

[13] http://news.bbc.co.uk/1/hi/health/7760413.stm

[14] http://www.guardian.co.uk/media/2009/jan/26/marketing-online-children-kids-underage-regulation

[15] http://www.searchenginejournal.com/google-advertising-patents-for-behavioral-targeting-personalization-and-profiling/2311/

[16] http://www.out-law.com/page-6483

[17] http://pewglobal.org/reports/display.php?ReportID=247

[18] http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1

[19] http://en.wikipedia.org/wiki/Blind_men_and_an_elephant July 2009 and www.naturalchild.org

[20] Picture/ graphics from © Paul Davenport and © Jason Hunt

[21] http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html

[22] http://www.doubleclick.com/us/

[23] http://en.wikipedia.org/wiki/Tag_(Metadata) July 2009

[24] http://www.flickr.com

[25] http://del.icio.us/

[26] http://www.benkler.org

[27] http://www.benkler.org/CoasesPenguin.html

[28] http://www.randomhouse.com/features/wisdomofcrowds

[29] http://en.wikipedia.org/wiki/Rss July 2009

[30] http://www.maps.google.com

[31] http://www.navteq.com/

[32] http://www.digitalglobe.com/

[33] http://en.wikipedia.org/wiki/Mashup_(digital) July 2009

[34] http://www.housingmaps.com/

[35] http://www.craigslist.com/

[36] http://en.wikipedia.org/wiki/SOAP July 2009

[37] http://www.adaptivepath.com/publications/essays/archives/000385.php

[38] http://en.wikipedia.org/wiki/Wisdom_of_crowds July 2009. Also see http://www.randomhouse.com/features/wisdomofcrowds/

[39] http://en.wikipedia.org/wiki/System/390 July 2009

[40] http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html?page=1

[41] http://www.littlespringsdesign.com/blog/2005/09/14/the-carry-principle/

[42] http://en.wikipedia.org/wiki/Semantic_Web July 2009

[43] http://anthropology.si.edu/humanorigins/ha/laetoli.htm

[44] http://ulik.typepad.com/leafar/2006/10/ulik_unleash_id.html

[45] http://www.fredcavazza.net/index.php?2006/10/22/1310-qu-est-ce-que-l-identite-numerique (in French)

[46] www.identity.futuretext.com

[47] http://openid.net/

[48] http://www.projectliberty.org/

[49] http://en.wikipedia.org/wiki/IP_Multimedia_Subsystem July 2009

[50] http://en.wikipedia.org/wiki/Session_Initiation_Protocol July 2009

[51] Adapted from Wikipedia

[52] http://weblog.cenriqueortiz.com/mobile-context/

[53] Yao Wang, Julita Vassileva; Trust and Reputation Model in Peer-to-Peer Networks

[54] zhen zhang, xiao-ming wang, yun-xiao wang: PeerTrust: Supporting Reputation-Based Trust for Peer-to-Peer Electronic Communities

[55] zhen zhang, xiao-ming wang, yun-xiao wang: A p2p global trust model based on recommendation

[56] http://en.wikipedia.org/wiki/Public_key_infrastructure July 2009

[57] http://en.wikipedia.org/wiki/PageRank July 2009

[58] http://en.wikipedia.org/wiki/Recommendation_system July 2009

[59] http://hyveup.blogspot.com/2008/10/3-different-approaches-to-automated.html

[60] http://www.pandora.com/corporate/

[61] http://corp.strands.com/

[62] http://www.readwriteweb.com/archives/recommender_systems.php

[63] http://www.readwriteweb.com/archives/recommendation_engines.php

[64] http://www.msearchgroove.com/2008/10/23/judging-recommender-start-ups-in-switzerland-will-recommendation-engines-come-through-where-mobile-search-falls-short/

[65] http://en.wikipedia.org/wiki/Mosaic_(geodemography) July 2009 and http://en.wikipedia.org/wiki/A_Classification_Of_Residential_Neighbourhoods

[66] http://www.google.com/ads/preferences/html/about.html

[67] http://feeva.com/index.html

[68] http://thenowfactory.com/

[69] www.blyk.com

[70] http://en.wikipedia.org/wiki/Maslow's_hierarchy_of_needs July 2009. Additional good article on the history of human psychology http://psychclassics.yorku.ca/Maslow/motivation.htm

[71] http://www.amazon.com/Being-Digital-Nicholas-Negroponte/dp/0679762906

[72] http://www.mercurynews.com/mld/mercurynews/news/13567880.htm

[73] http://www.intel.com/products/viiv/index.htm

[74] http://en.wikipedia.org/wiki/Freemium_business_model July 2009

[75] http://www.jrrt.org.uk/uploads/database-state.pdf

[76] http://en.wikipedia.org/wiki/Porter_generic_strategies July 2009

[77] http://www.culture.gov.uk/what_we_do/broadcasting/5631.aspx/

[78] http://www.femtoforum.org/

[79] http://cyber.law.harvard.edu/projectvrm/Main_Page and http://blogs.law.harvard.edu/vrm/

[80] http://privacy.cs.cmu.edu/people/sweeney/kanonymity.pdf

[81] http://opensocialweb.org/2007/09/05/bill-of-rights/

[82] http://www.dataportability.org/

[83] http://openid.net/

[84] www.microformats.org

[85] http://en.wikipedia.org/wiki/Resource_Description_Framework July 2009

[86] http://www.xmpp.org/

[87] http://www.foaf-project.org/

[88] http://en.wikipedia.org/wiki/SPARQL July 2009

[89] http://oauth.net/

[90] http://en.wikipedia.org/wiki/RSS July 2009

[91] http://en.wikipedia.org/wiki/OPML July 2009

[92] http://www.google.com/friendconnect

[93] http://developers.facebook.com/connect.php

[94] http://developer.myspace.com/community/myspace/dataAvailability.aspx

[95] http://blog.cdt.org/2009/02/02/ads-with-eyes/

[96] http://www.theregister.co.uk/2008/02/29/phorm_roundup/

[97] http://news.bbc.co.uk/2/hi/technology/7998009.stm

[98] (http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2009/wp163_en.pdf) Facebook by the Canadian Privacy Commission (http://ec.europa.eu/justice_home/fsj/privacy/docs/wpdocs/2009/wp163_en.pdf).

[99] http://en.wikipedia.org/wiki/Enter_Shikari July 2009