Alan Mitchell

is a director of Ctrl-Shift and a
founder of Mydex Community
Interest Company.

The rise of Volunteered Personal Information

Alan Mitchell

Date Received (in revised form): 9th August 2010

Abstract

In its history, marketing has only had two forms of information to play with: market research (which is primarily qualitative and/or statistical), and transaction/interaction data (which are granular and personally identifiable). When transaction data first appeared on the scene around 50 years ago, they sent shockwaves through marketing as new insights and possibilities opened up. Direct and database marketing, customer relationship management and loyalty schemes are some of its progeny. Now a third form of market information — Volunteered Personal Information — is entering the fray. Its impact could be as or even more far reaching than the transaction data revolution that came before it.

Journal of Direct, Data and Digital Marketing Practice (2010) **12**, 154–164.
doi:10.1057/dddmp.2010.28

Keywords: CRM, social CRM,
personal data stores, consumer
empowerment, Volunteered
Personal Information (VPI), targeting

What is VPI?

Volunteered Personal Information (VPI) is information individuals *could* tell organizations, *if* they saw good reason to. The ‘could’ bit is mainly about technology. Until very recently, the technology and infrastructure for individuals to say ‘this is me and this is what I want’ on a mass, scalable way did not exist. Now that’s changing.

‘If they see good reason to’ is about incentives, trust and relationships. For organizations, VPI opens up new ways to cut costs *and* add value. From simple administrative updates such as ‘I’ve moved home’ through changing consumer goals and priorities that define markets, to the future purchasing plans and intentions that drive sales and marketing operations, VPI could bring organizations closer to the dream of the right information/offer to the right person at the right time. For individuals, VPI represents a leap forward in convenience, control and value. In other words, the win-win potential is huge. Can it be unleashed? If so, how?

Technology trends

There’s a good chance that some time recently you opened an email with a 5MB attachment. Back in 1956, when it first came to market, 5MB was the biggest hard drive available, and the necessary machinery for it filled a small room. By 1979, a hard drive capable of storing 250MB of data would fill a large supermarket trolley. Today, we nonchalantly slip iPhones and memory sticks into our pockets with ten times this capacity.¹

Information processing costs have tumbled in parallel. Back in 1960, \$1000 would buy you one computation per second. In the intervening

Alan Mitchell
Tel/Fax: +44 (0) 207 622 4987
E-mail: Alan.Mitchell@ctrl-shift.co.uk

From 'top down' to 'bottom up'

50 years we've witnessed a *billion*-fold productivity improvement: the costs of computation have fallen a thousand-fold in the last decade alone. That's why we are now surrounded by app-rich mobile devices, PCs groaning with super-sophisticated software, video games, data-rich comparison sites, and so on.

Another once-in-a-century sea change is unfolding in parallel. The information flows that define how our society works are being transformed. Until very recently, virtually all organized information flowed 'top down': from organizations to individuals. We're now in the midst of an explosion of communication from individuals to each other (peer-to-peer) and to organizations (consumer-to-business).

It started with the telephone. In 2007, the UK call centre industry handled 47 billion minutes' worth of in-bound calls. Now it's moved online. Last year, individuals sent an average 47 billion emails a day. Bloggers uploaded 270,000 words a minute. Tweepers tweeted 27.3 million times a day. YouTube users uploaded 20 hours of new content every minute. And that's not counting Google and Facebook, which have become ubiquitous.

In fact, the internet behemoths of today — Google, YouTube, Facebook, Twitter — are all driven by information from individuals, not from organizations. Google, for example, is driven by information input by individuals (the search term). So is Facebook (user profiles and messages). So is Twitter (user messages). Separately and together, these new information services are already having a profound impact on marketing. Yet the two underlying trends they represent — information as a tool in the hands of the individual, and 'bottom up' — are still in their infancy, evolving largely in separate, parallel streams.

So far, for example, the explosion of 'bottom-up' communication in social media has mainly taken the form of unstructured information that's hard to gather, process and use efficiently on a mass scale. VPI will only really come into its own when 'bottom up' and 'information as a tool in the hands of the individual' merge in the form of new 'personal information management services' (PIMS) that help individuals gather, use and share the information they need to organize and manage their daily lives more efficiently.

PIMS are driven mainly by information from the individual: who I am, what I want, what I want to find out, etc. They generate, capture and use structured, commercially valuable personal data on a scalable basis. They turn individuals into 'VPI factories'.

Decision support services

To see the potential power of 'information as a tool in the hands of the individual' combined with 'bottom up', let's look at some PIMS that are likely to evolve over the next few years.

Our current definition of search is defined by Google — a search is a single, isolated search term. That's ridiculous. When I'm researching something such as 'my next holiday' or 'my next car', I conduct multiple different 'searches' all of which are small contributing elements

Next-gen search

to one, bigger research project. Next-gen search services will help users manage their searches on this ‘project’ basis (collecting, collating and sifting all the information that’s relevant to their end goal).

The mere fact of such sorting and defining involves these users in inputting information about the underlying purpose of their searches. By adding just a few further bits of information, such as ‘my budget’, individuals could create rich packages of data about what they are trying to achieve — data that they could share with suppliers if they wanted to.

Next-gen comparison

Today’s comparison services are product-centric. They compare the features of products. Tomorrow’s comparison services will become increasingly person-centric. They will help buyers build a specification of what they are looking for, and then collate comparisons around these specifications. These specifications — ‘the features and attributes that really matter to me’ — are created by the individual in the course of their research ... and are potential gold dust to sellers.

Next-gen advice giving

For the most part, today’s peer-to-peer advice services are chaotic and clunky. Over time, they will evolve to become far more structured, disciplined and easy to navigate: they will move from useless comments such as ‘It’s great!’ or ‘It sucks!’ to more carefully considered, collated and analysed perspectives on multiple product/service attributes such as price, ease-of-set up, ease-of-use, cost of maintenance, service reliability, quality and longevity, green issues, ethical issues. For such advice to be easily and efficiently accessible, however, users need to be able to specify the issues that matter to them: ‘I am interested in ease-of-use but not green/ethical issues’, or vice versa. Such ‘advice specifications’ are a potentially rich source of insight into what really matters to consumers when planning their purchases.

Personal infomatics

Personal data ‘mash-ups’

Loyalty schemes, GPS mobile devices, contactless payment mechanisms, behavioural targeting, smart energy meters: in every aspect of our day-to-day lives more information — information that previously evaporated the instant it was created — is now being captured digitally.

So far, virtually all of these developments have been driven by organizations capturing information about their customers’ behaviours. But there is no law stopping individuals from using the same technologies for their own purposes, to track and monitor aspects of their own lives. In the health sector, for example, multiple devices help sufferers of chronic diseases monitor key factors such as heart rate, blood pressure and blood sugar levels. Innovative services like Nike+ help individuals collect and analyse data about their fitness regimes. This is highly personal information that individuals can share with others if they see fit.

Personal data hand-backs

In parallel, there are growing incentives for organizations to ‘hand back’ data they have collected about their customers. Energy companies will provide individuals with digital data about their energy

usage to help drive energy saving. Banks will be asked to hand back financial transaction data in digital form, so that individuals can do their own financial accounts and analytics. In health, the concept of personal health records — where clinical and other data are held by individuals in their own health database — is gaining ground with the new UK Coalition Government's decision to 'put patients in charge of making decisions about their care, including control of their health records'. (The Conservative Party's election manifesto said: 'Wherever possible we believe personal data should be controlled by individual citizens themselves'.²)

Creating the customer's single view

Personal data repatriation is not just about transferring data from one party to another. It enables the creation of completely new data sets. If all the travel agents, airlines and hotels I buy tickets and rooms from enter these transaction details into my personal travel database, then, over time, I generate a full, comprehensive picture of my total travel behaviours. This 'single view' is potentially invaluable to suppliers for research and segmentation purposes, targeting, insight and so on. Once I have these data, I can share them with companies I want to do business with.

Personal data stores

Personal data stores (PDS) — databases built, updated and managed by individuals for their own purposes — are a necessary support for all the above services. A common reaction of many people coming across the PDS concept for the first time is negative: 'it sounds too much like hard work to me'. In fact, consumer take-up of PDS will be driven by the convenience they offer.

Take Mydex's 'subscribe to me' service.³ Under this service, organizations ask permission to subscribe to certain fields of an individual's data store. If the individual agrees, then every time the individual updates or changes a field in his own data store, subscribing organizations are instantly and automatically alerted to the update. So, for example, if the individual moves home and has 100 suppliers subscribing to his data store, all he has to do to inform them of his change of address is to change a field in his own data store, once. This saves the individual huge amounts of time, money and hassle while addressing organizations' endemic data synchronization headaches at very low cost.

Reaching the parts other data cannot reach

What these examples point to is an environment where huge amounts of rich, timely, structured and accurate personal data are generated 'naturally', as a by-product of things individuals do to organize and manage their lives more efficiently. It reaches the parts other types of data cannot reach primarily because it embraces:

- Future plans and intentions, as crystallized by specification-building and advice-seeking activities for example;

Information 'in my head'

Information 'only I can see'

- Stated goals, priorities and preferences, for example preferred product attributes, areas of concern and interest, preferred contact methods and times, etc.
- The creation of a *genuine* comprehensive view of the customer's total category activities (eg 'all my flights and hotels) as distinct to partial 'single' views silo by silo.⁴
- The identification of 'persona': me acting in my role as homeowner, as concerned citizen, as father, as professional career person, as investor, etc.
- Understanding context, as the data individuals collect and share about their lives gets richer and rounder.

Information 'as it happens'

Most organizations' customer data take the form of snapshots whose validity and accuracy begins to decay the minute they are collected. The forms of VPI discussed above generate and share information when it is being used: ie when it is important.

Valuing VPI

Just how big will the market for VPI be? Figure 1 lists the main components of the four broad categories of VPI (who I am; what my views, feelings and preferences are; what I want to find out; what I want). It doesn't cover the entire VPI landscape. It ignores most peer-to-peer communication, for example. Even so, it demonstrates two critical points:

- separately and together, various VPI feeds can act as an information input into every customer-facing activity from insight and innovation

Different types of VPI

Who I am

Factual updates: e.g. change of address, payment card, contact details

Lifestyle updates: e.g. I'm now married/divorced; got kids; living alone

Personal profiles 'mash-ups': combinations of data from a variety of different sources such as 'my financial circumstances and net worth', 'my personal health', 'my wardrobe', 'my total household possessions'

My views and feelings

Complaints

Suggestions

User reviews

Reasons why

What I want to find out

Factual queries: e.g. 'where is?' 'what is?' 'can I?'

Online search

Requests for information e.g. 'please tell me more about ...'

Advice seeking e.g. 'should I?' 'what if?'

What I want

Specifications: e.g. 'I want something that addresses these needs, behaviours and priorities'

Orders and advance orders; future plans and intentions

Purchase preferences: e.g. products, variants, flavours, brands etc

Operational preferences: e.g. preferred communication channels and times, delivery times etc

Figure 1: The next pivotal data resource? How VPI articulates all aspects of consumer demand. *Source:* The New Personal Communication Model: the rise of Volunteered Personal Information, Ctrl-Shift, 2009.

through sales and marketing operations to customer service and customer relationship management (CRM), including routine administration.

- VPI has multiple uses including:
 - as a trigger to action (right information/offer to the right person at the right time)
 - trend analysis (aggregated VPI data)
 - insight (from questions asked, concerns expressed and trends analysed)

Figure 2 shows our estimates of the actual UK market for VPI by 2020.⁵ These estimates are based on the likely frequency of data use, the proportion of households engaged in the activity and the current/likely future value of different types of data to organizations (for example, sales leads for cars and mortgages are worth more than address change data). The total figure of £20bn includes £3bn contributed by the already maturing market for online search, which, of course, is ‘powered’ by volunteered data.

The analysis suggests that two still-fledgling VPI markets could grow bigger than search. The first is ‘requests for information’. Essentially, this is reverse direct marketing, where individuals announce they are in the market for a certain type of product or service, and solicit relevant offers and communications from suppliers. The two key components of these services (we predict) will be the individual’s ability to use anonymity shields to avoid opening themselves up to spam, and the ability to turn the communication tap on and off (eg ‘all communications and offers to be received by 1 June’). The processes and mechanisms necessary for such services already exist (the hard part is the relationships, the user interfaces and the distribution of the value that’s created).

The second, closely connected market is for ‘specification building’ — for example, all the information generated as individuals go through a decision-making process such as:

- ‘I am planning to buy a new car’
- What sort of car? (SUV, saloon, four-wheel drive, etc)

Estimated size of the UK market for VPI	£m
Who I am	1850
My views and feelings	1750
What I want to find out	10,000
What I want	6750
TOTAL	20,350

Figure 2: Bigger than seven Googles — projected market for VPI, by 2020.
Source: The New Personal Communication Model: the rise of Volunteered Personal Information, Ctrl-Shift, 2009

- To fit what needs and purposes? (family, town chores, holidays, hobbies, dogs, etc)
- What price bracket?
- What core specifications? ('green' fuel, sun roof, sat nav, etc)
- Any brand preference?
- What are the financing alternatives?
- Model preference/fixtures, fittings, trim, colour, etc

Each stage of the decision-making process, from initial decision to final detailed specification, generates its own opportunities for new volunteered information and for new communications and responses from suppliers. For this reason, we predict that 'specification building' will become the biggest of all VPI markets over time.

Some areas of household spending, such as mortgages and cars, are likely to involve high levels of VPI input. Others, such as furniture and furnishing, food, alcoholic beverages and recreation, are not likely to involve high inputs of VPI. A quick back-of-the-envelope assessment of the relevance of VPI to different aspects of household expenditure suggests that by 2020 VPI could play a significant role in around third, or £270bn worth of consumer purchases.

Implications for CRM

We also looked at the likely impact of VPI on existing CRM processes and systems. There are some straightforward savings in data maintenance, external data purchase and internal data gathering (as an increasing volume of timely information comes direct from 'the horse's mouth'), but these cost-savings are offset by increased costs elsewhere — new systems and processes to elicit VPI, paying for VPI elicited by third parties, extra analysis and integration of VPI data. In addition, there are costs associated with running 'dual systems' as VPI matures and goes mainstream.

The net effect (for a financial services organization) looks to a reduction in the *operational* costs of CRM of between 10 and 20 per cent. However, this massively understates the real financial impact of VPI that lies elsewhere — in the knock-on ability to access and use richer, more relevant information to:

- Reduce the costs of wasteful actions based on guesswork and/or inaccurate data.
- Using richer, more timely insight to generate more efficient, effective innovation, segmentation, targeting, etc.

The biggest impact of all, however, is on the role of CRM itself. Some types of VPI are 'naturally' created *within* a company/customer relationship — for example, complaints and contact preferences. However, many other types of VPI are 'naturally' created by services, helping individuals interact with the market as a whole. Online search and personal data mash-ups, which combine

transaction and other data from a range of different sources, are obvious examples.

Other types of VPI are likely to be a mixture of both. Many car companies already have their own specification building tools, for example. But while these tools help the individual navigate to appropriate value *within* the brand, they don't help the individual navigate to appropriate value *between* brands. For this reason, the majority of specification building services are likely to migrate outside of direct company/customer relationships.

The upshot of this is that, according to our estimates, *two-thirds of all available VPI will be generated outside the direct one-to-one customer company relationship*. Traditional CRM data (mostly data relating to/an individual's transaction and interactions with that particular organization) will decline from its current role as 'the crown jewels of customer data' to become just one supporting element in a much richer data ecosystem. This suggests that organizations seeing 'more of the same CRM, only bigger and better' as the future may need to rethink their customer data/CRM strategy.

A new information contract with customers

VPI is information individuals could volunteer *if* they wanted to. The dark side of the technology trends that are making VPI possible is their ability to trigger an ugly adversarial battle between individuals and organizations over who 'owns' the data, and who should benefit from its use.

Recent history doesn't help here. Today's data *status quo* has its roots in a time when organizations were the only entities gathering and using customer data; when individuals were the passive 'subject' of data gathering and 'target' of communication; when information was a monopoly tool in the hands of organizations. This environment created a 'default mode' in organizations — that customer data is like fish in the sea, yours to do with as you like, if you can catch it (and stay within the law).

Applying this same default mode to the emerging environment is a recipe for conflict and declining trust. We've already seen two high-profile flashpoints: consumers' mistrust of behavioural targeting, and increasing public disquiet over Facebook's privacy policies.

The defining feature of the emerging era is that individuals will only volunteer information if they *want* to volunteer it. Organizations need to treat individuals as mutual data sharing partners, not 'targets'; the relationship has to be 'win-win'. To this end, we will need to develop new types of 'information sharing contract'.

Figure 3 outlines some of the issues these new information contracts will need to address: whose data are they, who controls how they should be used, who benefits from their use, and what processes are used to determine and monitor agreements relating to these questions. The two core features are a default mode that 'personal data is the person's', and selective disclosure — it is up to individuals to decide what information they want to share with who, for what purposes, under what terms and conditions.

The new company/customer 'information contract'		
	Yesterday	Tomorrow
Data 'ownership'	Default: Beneficial ownership goes to who gathers/collects the data	Default: Personal data is the person's. It is up to the person what s/he does with his/her data
Permission	Only needed where the law requires it. Once given, lasts forever.	Permission is required at every step: gathering, using, sharing. The individual can rescind permission at any time.
Control	The organisations decides what to do with customer data (within the law).	Individuals are able to specify what data is to be used for what purpose, and can change these specifications
Value	Customer data is a corporate asset, used to improve corporate ROI	Customer rewarded (directly by payment or indirectly 'in kind') for contributing a valuable resource
Transparency	Who does what with customer data is largely hidden from the customer	Transparency over who is doing what with customer data is a foundation of trust

Figure 3: A new information sharing market requires new rules and understandings.

Given the huge variety of different forms of VPI, and the different circumstances under which it will be shared, the details of such 'information contracts' will vary enormously. There is a vast spectrum of possibilities, which means the right answer will always be 'horses for courses'. Nevertheless, organizations and individuals will need to place themselves somewhere on the spectrum.

'Data ownership'

Legally, the question of ownership is a nightmare that's almost certainly best avoided. But 'morally' the stance that personal data is the person's sets the right tone for win-win information sharing. There is a very broad spectrum here. By definition, for example, all transaction data are effectively 'owned' by both sides of the transaction though, until recently, effective 'ownership' went to the organization. Tesco, not the customer, is the effective 'owner' of Clubcard data, for example. The waters are even muddier with behavioural data. The underlying assumption behind behavioural targeting, for example, is that the organization collecting data about the consumer's online behaviour is the effective 'owner' of those data.

As individuals realize the commercial value of the information they generate, pressure to gain more access and control over these data is likely to grow.

Permission

The default assumption in a VPI-driven world is that organizations cannot gather, collect or use these data unless they have prior permission. There are, of course, many degrees of permission, from formal agreements to implied understandings. The trend is likely to be towards explicit opt-ins rather than vague opt-outs.

Control

The principle of selective disclosure lets the individual decide what information to share with who, and under what terms. Is it to be

personally identifiable or anonymized, for example? Under the Mydex model, before an organization can subscribe to the individual's personal data store, it has to sign the *individual's* privacy/data sharing agreement — a mirror image of the *status quo* online where, to complete a transaction, individuals have to sign the organization's terms and conditions.

Value

Individuals have no incentive to volunteer information if they get no rewards for doing so. In some cases, individuals will 'pay' for services they value with their information (as with Google). In other cases, organizations will pay individuals for access to their data. Exactly how much individuals are paid for their data will be a matter of negotiation.

Transparency

Under the current system what organizations do with customer data — what data they collect, how they use them — remains largely opaque. In a VPI world, this undermines trust: the trend will be towards increasing transparency over data collection and use.

Conclusion

We are witnessing the beginnings of a massive new market, which is going to transform the ways individuals and organizations relate and interact. The potential scope and range of VPI could make it the most valuable of all organizations' customer data. But getting the relationships, processes, technologies and business models right will involve climbing a steep learning curve.

The first thing organizations need to do is to recognize the change: there is a new and different opportunity and challenge.

A next step is to conduct a VPI audit: what information are individuals currently volunteering, via which touchpoints, and why? What is the organization doing with this information? What else could it do to a) elicit more volunteered information via existing processes and b) use these data to best effect?

Getting the 'information contract' with customers right will be crucial. Most of today's privacy policies, terms and conditions, etc were written to fit a different environment. What changes are needed to encourage individuals to volunteer information?

Finally, there is the development of a VPI strategy. Which of the emerging VPI 'feeds' are most valuable for our organization? How and where will we elicit them?

Notes

1. For some fascinating details about the history of data storage, see <http://royal.pingdom.com/2010/02/18/amazing-facts-and-figures-about-the-evolution-of-hard-disk-drives/>
2. This is taken from the invitation to join the Government of Britain; Conservative manifesto, 2010, p 79.
3. Equality and Excellence: Liberating the NHS, The NHS White Paper, July 2010. Disclosure: Alan Mitchell is a founder and director of the Mydex Community Interest Company. The first live test of the Mydex technology and service got under way as this article was going to press. Participants accepting 'live' VPI from individuals include a UK Government consortium including the Department of Work and Pensions, the Cabinet Office, the TellUsOnce initiative and DirectGov, three local councils (Brent, Croydon and the Royal Borough of Windsor and Maidenhead), BT, Experian and Netmums. Observers include

the Information Commissioner's Office, Olswang the City lawyers and privacy experts, and a number of smaller private sector organisations.

4. Marketers talk a lot about generating a single customer view, but this 'single customer view' is actually a very restricted, partial view of what the customer is actually up to. It is *the organization's* view of that customer's dealings *with that organization*. It tells us nothing about what the customer is doing with other organizations. Therefore, for example, Amazon has a good profile of my book-purchasing habits ... with Amazon. But it doesn't have a clue about my book purchases from Waterstones, Borders, Blackwells, Foyles and W H Smith. That's why the next purchase recommendations it generates for me are so often wrong. For example, it keeps on telling me I should buy *Nudge* by Richard Thaler and Cass Sunstein. But I bought it from Waterstone's. Amazon's recommendations are not working on the basis of real knowledge of my book-purchasing activities. They are guesses made on the basis of a partial view of my book purchases, and these guesses are often wrong. Personal data mash-ups transcend the intrinsic flaws of this organization-centric approach to a single customer view.
5. This is taken from Ctrl-Shift's research report: *The New Personal Communication Model: the rise of Volunteered Personal Information*, Ctrl-Shift, 2009.