Richard Webber

is the originator of the Mosaic classification system, Managing Director of OriginsInfo.Ltd and Visiting Professor at the Department of Geography, Kings College London.

Papers

The evolution of direct, data and digital marketing

Richard Webber

Abstract

Since the late 19th century when the concept first originated, Direct Marketing (DM) has evolved from a process that underpinned a distinct mode of transacting business with customers into an activity in which virtually every large organization is involved. This evolution has been most rapid since the founding of the Institute of Direct and Digital Marketing (IDM) in 1988. This paper describes the key trends that have shaped the evolution of DM over this period. Despite the introduction of new channels, the expansion of data processing capabilities and the proliferation of data analysis tools, the paper finds a surprising level of continuity in terms of the underlying business concepts and models associated with the industry. The period has been one of evolution rather than disruptive revolution. Many of the business concepts that underlie DM are as relevant today as they were 25 or even a hundred years ago. The paper represents the collective opinion of the Editorial Board of the IDM journal, led by Webber, to celebrate the 25th anniversary of the IDM.

Introduction

This year, the Institute of Direct and Digital Marketing (IDM) celebrates the 25th year of its existence. During this period, the practice of Direct Marketing (DM) has been radically changed not just by the emergence of new communication technologies, but also by fundamental changes in management thinking about the way businesses manage the relationship with their customers.

In many other fields of business, the impact of new technology has been disruptive. Although the impact of new technology on the practice of DM has been transformational, the changes have been evolutionary rather than revolutionary. Were it not so, the Institute would not have retained its position as the foremost professional training organization in this industry.

Most of the early innovations in DM practice continue to be used in some form or the other. Notwithstanding the proliferation of new

DM changes evolutionary, not revolutionary

Richard Webber
Managing Director
OriginsInfo.Ltd
16 Broadlands Road
London N6 4AN, UK
Tel: +44 (0) 20 8340 3034
E-mail: Richardwebber@originsinfo.eu

communications channels, many continue to underpin businesses' data management processes and communication strategies. However, the practice of DM has expanded and, in particular, has diffused into additional stages in the cycle of customer acquisition and retention.

As a result, practitioners in any one branch of DM now find it difficult to acquire knowledge and experience across the entire range of competences that make up the profession to which they belong. What at the time of the foundation of the IDM was a relatively self-contained 'industry' has in the past 25 years fragmented into a number of specialized branches. Many practitioners in these branches have limited grounding in DM as a whole, being content to learn new skills on a 'need to know' basis. Indeed, an increasing number of practitioners of DM would not necessarily recognize it as the primary description of the occupation in which they are engaged.

Established processes often taken for granted

With the passing of time, many of the techniques that, when first introduced, were subjects of great wonder and discussion have assumed so low a profile that it is easy for practitioners to take them for granted or even to be unaware of their existence. For example, how many of us properly understand the process that underlies the accurate identification of duplicate names and addresses?

The intention of this paper, which has benefitted from inputs from the members of the Journal's Editorial Board, is to provide a summary overview of the evolution of DM since its inception so that the practitioners who are engaged in some of the newer technologies can better understand how the industry has evolved.

The paper does not aim to chronicle the rise and fall of influential people and the businesses they established, nor is it intended as a strictly chronological account. Its purpose is to show how an industry that originated with the narrow aim of enabling manufacturers to by-pass retail intermediaries has evolved, step by step, into one that can orchestrate the complex variety of ways in which virtually all consumer-facing organizations now manage the communications with their customers.

Apologies are due to practitioners in business-to-business marketing for the fact that this paper restricts itself to business-to-consumer marketing — the evolution of that side of the industry can only be properly told in a separate paper.

The origins of DM in the American mid-West

The unifying concept that underpins DM practice is that there are many circumstances in which it makes better business sense to market goods and services directly to consumers, rather than rely on intermediaries located in retail outlets, or 'bricks and mortar' as they are now popularly described.

Selling directly, as distinct from face-to-face, implies the existence of media through which people can communicate with each other using channels other than face-to-face. Until 30 years ago, that communication relied almost exclusively on a communications channel developed over 400 years ago — the printing press.

Absence of established retail centres encourages DM

The earliest widespread use of print as a routine medium for recruiting new consumer business occurred in the mid-West of the United States in the late 19th century. Near-universal literacy combined with the absence of an established network of local retail centres made it an attractive proposition not just to promote goods via newspaper advertisements, but to include response mechanisms enabling readers to purchase these goods directly.

US Postal Service facilitates mailing of catalogues

Prerequisites for the success of such a mode of commerce included newspaper titles that reached a mass audience, a reliable and affordable postal system, customers having identifiable addresses to which goods could be sent and the development of an effective and affordable mechanism for making payments directly. The US Postal Service facilitated the growth of mail order by classifying mail order publications as aids in the dissemination of knowledge, thereby enabling catalogues to be posted at a rate of 1 cent per pound weight.

Photographs were critical in building desire among consumers for such products. This is why 'off the page advertising' did not really take off in Britain until the 1960s by when innovations in colour printing technology enabled the *Sunday Times* to publish its first full-colour supplement.

The practice of including direct response advertisements in newspapers naturally evolved into the promotion of manufacturers' catalogues. Before World War I, it would have been customary even for manufacturers selling through 'bricks and mortar' to produce illustrated catalogues to display their range of stock. William Morris first did this in England in 1862. Promoting these catalogues directly to end customers was a natural evolution.

In due course, it became clear that, rather than each manufacturer developing its own bespoke catalogue, there was an opportunity for an intermediary to source products from a variety of manufacturers and consolidate them into a single, larger catalogue. The most successful of these catalogues was that of Sears, first published in 1888. The company's headquarters were in Chicago, its prime market mid-Western farmers. Many of America's foremost data compilers and data analytics companies continue to be based in this city.

Catalogue companies become retail intermediaries

This consolidation provided a number of opportunities for saving costs. Like a department store, the mail order catalogue provides the consumer with the option to purchase across a range of different products. It could be kept handy until the time when these products were needed. Purchasing from a single supplier was more convenient and involved less risk than buying from multiple suppliers. Quality of service and timeliness of delivery could be assured. Moreover, in many cases, the catalogue company could use its purchasing power to negotiate better terms from suppliers and therefore supply merchandise at more attractive prices.

The business model which this marketing proposition was based on is little different from that used by Britain's most successful

catalogue retailer, Great Universal Stores, founded in Manchester in 1900, and, albeit using different channels, more recently by Argos and by Amazon.

Centralization of point of contact improves understanding of customers

The centralization of the interaction with individual customers made DM a much easier environment in which to examine the impact on the business of micro-decisions. Each individual assistant behind the counter of a retail multiple may have known more about the preferences of individual customers. However, this knowledge could not be aggregated to build an understanding of how best to promote which products to which groups of customers using a process of controlled experimentation.

It is not clear exactly to what extent the early catalogue companies managed to link individual customer orders together. Before World War II, punched cards were the means by which customer behaviour was analysed and target customers selected. However, it is likely that even before the advent of the computer, most orders were filed at the level of the individual customer and it was not impossible in the pre-digital age to test the impact on business performance of variations in copy, pricing or terms, and to learn how best to make use of previous response history when communicating with repeat customers. Renowned for its unparalleled expertise in these techniques was the *Reader's Digest*, formed in 1922. In its heyday, it was the best-selling magazine in the United States, and after the launch of its UK edition in 1938 the company became the most internationally admired exponent of DM practice.

The arrival of computers

The world's first general purpose electronic computer started work in 1951. The business justification for the introduction of computers to the operations of direct response and catalogue companies was not initially to allow new methods of distribution. It was to automate existing clerical processes and thereby to reduce costs. The most obvious source of these savings was the automation of the billing process.

Automation applied to name and address deduplication

One key requirement of this process was the ability to reliably distinguish one customer from another and to identify instances of potential ambiguity. This requirement led to substantial investment in algorithms for parsing the various elements of a name and address, for converting them to a common format and for identifying duplicates. One of the earliest examples of such a system was that operated by the UK division of *Reader's Digest* in 1964.

Computers made it possible to improve the effectiveness of customer communications, not just through the streamlining of the process of generating mail, but also by permitting segmentation of the customer file, making use of whatever attributes about a customer were captured during the order processing and billing processes.

The rise of the mail order business

The willingness of organizations to make available to third party organizations computerized lists of the names and addresses of customers led to the emergence of another new form of organization, a type of business promoting its products principally or even

exclusively by mail. Businesses that had previously been reliant on newspaper or magazine advertisements to generate direct customers could now bypass off-the-page by appealing to the consumer directly via the post. This exchange of prospect names was facilitated by the growth of specialist intermediaries, list brokers.

The use of direct mail to win new customers provided an additional incentive for Direct Marketers to recruit specialists in statistics who could install and manage procedures for monitoring the response rates from names on different lists and work out how response varied according to the timing and frequency of offers, the copy and visual appearance of the mailing and the use of different incentives.[1]

Recency, frequency, monetary value becomes framework for evaluation

Many of these new mail order businesses promoted a narrower range of products than the catalogue companies and could trade profitably only through repeat business with existing customers. As they became more experienced, the focus of their analysis teams extended beyond their initial interest in the maximization of Response (R) to the understanding of factors that contributed to Frequency (F) of repeat purchase and the Monetary value (M) of orders using a form of optimization often referred to as RFM analysis.

The emergence of new numerical disciplines such as those involved in the maximization of RFM led practitioners in DM to become centrally involved in financial decisions associated with the concept of 'investment' in new customer acquisition and in the financial value of a loyal base of repeat customers. It also precipitated interest in the use of segmentation criteria as a means of identifying and targeting consumers who would be most profitable to trade with.

Mass prospecting via direct mail

Statisticians involved in the computation of these analyses were among the first marketers to make use of software packages such as SPSS (first released in 1968) and SAS (first created in 1966), both of which are now standard tools for direct and digital marketing analytics. The earlier use of single selection criteria was replaced by an environment in which multivariate scorecards placed each individual customer on a continuum from high to low in terms of predicted behaviour.

With the growth of direct mail as an alternative to the earlier direct response mode of acquiring new customers, the larger British companies that sold direct to customers — catalogue mail order companies in particular — began to experience difficulty in sourcing from list brokers sufficient volumes of names to meet their needs.

Insufficient number of mailable addresses constrains high volume mailers

This problem was more acute in Britain than in the United States because in America a very large proportion of rentable names were the names and addresses of newspaper subscribers. In Britain, where most newspapers are delivered by local newsagents, the names and addresses of their readers are not maintained by newspaper publishers.

During the 1970s, Britain's two highest-volume originators of outbound direct mail, the catalogue mail order groups GUS and Littlewoods, responded to this problem each by independently embarking on the data capture of the country's Electoral Register (ER).

DM enables the centralization of election campaigning

The sheer number of mailable names and addresses on the Register made direct mail a viable customer recruitment model for a number of non-traditional business sectors such as timeshare operators who became the most notable new users of the ER as a prospect file. A previous history of purchasing by mail was not an important attribute of prospects in this market.

As a result of financial deregulation, the 1970s and 1980s were characterized by a rapid expansion of consumer credit. This resulted in credit card mailing becoming a very high proportion of unsolicited mail. The advent of more powerful computers and newly acquired access to a near-universal file of mailable names and addresses made it practical for the first time to combine an organization's customer database, its prospect universe and its mailing history into a single integrated marketing database.

Leading this practice were the Conservative and Labour Parties that, during this period, invested heavily in the development of national voter databases, which, in a period when centralized mailing and telemarketing centres began to replace locally organized canvassers, were seen as critical sources of advantage in targeting swing voters in key marginal seats.[2] The integration of information on supporters and prospects in a single database continues to be a key source of advantage in political campaigning, even if the data sources contributing to this database now include social media data, as is evidenced in Mark Pack's review on pages 349–353 of Sascha Isenburg's account of Barack Obama's 2012 presidential re-election campaign.

Discovering the value of segmentation

Although the ER greatly expanded the volume of mailable addresses, relatively little was known about any elector other than perhaps his or her gender, as may be inferred from the personal name, and how many other electors were resident at his or her address. Over time, information could be derived on how many years an elector had lived at that address, a useful proxy for the likely age of that voter. However, ER data in its raw form provided little opportunity for the targeting of particular demographic groups.

Linking Census and ER data

The UK ER was first linked to the geography of the Census in 1972, enabling demographic data about a person's neighbourhood to be appended to each ER record. This was the year in which regression analysis was first used to create a multivariate model for selecting ER names and addresses for a prospect mailing.

Geodemographic classification systems were introduced 2 years later. Their use was predicated on the assumption that where you live is often a better predictor of your consumer preferences than the job you do and on the basis that the type of neighbourhood where you live can be inferred to a very high level of precision from your postcode. These classifications sought to group the country's 1.3 million or so postcodes into a manageable set of around 60 clusters, each distinctive in terms of its age, housing and socio-economic characteristics.[3]

Linking the classification to the ER via the postcode made it possible to refine the 30 million or so ER records so as to be able to target, for example, people living in ultra-affluent areas, rural retirement areas, military bases or areas of smart, highly educated inner city singles, the propensity of which to buy particular products could be established either from research surveys or by profiling the postcodes of existing customers.

The introduction of postcodes and the Postal Address File (PAF)

The use of geodemographics presupposes the existence of the postcode system. Such a system was introduced by Royal Mail in the United Kingdom during the early 1970s with the object of facilitating the sorting of mail. To enable high-volume users of postal services to code their files, the post office allowed commercial organizations access to what had previously been a file for internal use, the PAF. This incorporated in a standardized form the addresses of each of the country's 20 million delivery points.

PAF facilitates customer address recognition

The addresses on this file provided a valuable base against which the various formats in which a customer might record his or her address could be matched. In time, the requirement to match addresses to this file became seen to be a benefit rather than a burden. If each customer address was held in the form in which it appeared on the PAF, then it became much easier to identify multiple records relating to the same individual than if each record was held in whichever format the customer had written it down.

The link between geodemographics and the postcode system added further sophistication to market segmentation. It became possible to process a file of postcoded customers and to directly infer from their addresses the types of neighbourhood that generated the highest level of business, whether in terms of the number of customers, the amount they spent or their propensity to buy from particular product categories.

Profiling prospects by where they live

The link between Census and postcode geography also played a major role in the upgrading of the services offered by door-to-door distributors. Before the late 1970s, this form of distribution was used principally by retailers who needed to target customers within the vicinity of a particular new or existing store. With distribution now targetable by market segment, advertisers could achieve greater cost-effectiveness by targeting distribution at households known to be of a type that bought their products.

In due course, Direct Marketers also came to realize that the postcode field in the PAF could also be used to provide an effective index field for real-time access. This indexation was a key factor in enabling organizations to trace the history of previous transactions with customers who contacted them in real time over the telephone.

Viewing the customer in the round

In the early days of business automation, it was often the business processes of individual product divisions that were automated. Even

then, different processes in the customer cycle were often held independently on separate databases. Once these efficiencies had been achieved, there were many benefits to be gained — and not just by marketers — by linking together the information held on these different files in such a way as to form a single customer view. This integrated customer view often replaced the diverse set of views held by the divisions of the business that managed the promotion of individual products.[4]

For the recruitment of new customers and for the processing of orders, it had been quite sufficient until now to rely on batch processing and updating, often using magnetic tapes on which information was recorded in flat file format, a collection of records with no structured relationships. Information was read, updated and stored offline in an electronic archive.

Towards a single view of the customer

Initially, the decision to migrate towards a single customer view could be achieved using flat files stored on magnetic tapes. However, so long as one relied on this format, it was necessary to pre-accumulate event and transaction details and for these to be summarized at account level. Summary data was typically recorded at points in time that corresponded to the billing cycle so that the value of spend, for example, would be calculated for each monthly cycle, and information on products bought would be stored at an aggregate level rather than at individual transaction level.

The 1980s and 1990s therefore saw the development of what came to be known as Customer Relationship Management (CRM) systems for storing, manipulating and accessing customer data in such a way so as to support a unified, integrated strategy for customer communications applied across all product divisions and all channels.[4]

Communications refocus on the customer, rather than the product

CRM systems were also essential to support the move towards more customer-centric management strategies adopted at the start of the 1990s. During the downturn of 1990–1992, companies began to realize that their existing customer base held more value than their prospect pool. It was also realized that the use of cross- and up-selling to existing customers yielded a higher return on investment than seeking to achieve higher market share by aggressive new customer acquisition.

The cost of making a sale to a new customer was estimated to be between five and seven times as much as the cost of making a sale to an existing customer. Increasing customer retention rates by 5 per cent a year can increase customer value by between 25 and 100 per cent, depending on the industry sector, according to Reichheld.[5] This is the consequence of maintaining a commercial relationship with each customer over a longer period of time at a lower cost, while also increasing the number, range and value of products or services they purchase.

The immediate impact of the financial pressures of the recession of the early 1990s and these findings about customer value was the widespread introduction of customer loyalty marketing programmes

that relied on creating a single view of the customer to aggregate their purchasing activity and CRM systems to deliver loyalty benefits and rewards.

Direct Marketers forced to work with others

The requirement for business to consolidate all information held about a customer within a single customer view was much more easily delivered using new relational database technologies that began to be promoted from 1970 onwards, not least by Oracle. However, this new technology and the fact that it linked together all different aspects of customer information often resulted in Direct Marketers losing control of operations through the centralization of computer processing, analytics capability and dedicated IT personnel.

Direct Marketers lose their independence

As a result of this change in technology, the decisions that Direct Marketers were previously able to take independently now started to be taken within a more complex environment in which other factors had to be considered. Decisions that were previously justified on the grounds of predicted mailing response now had to give due consideration to process handling capacity, stock levels, competitor activity, advertising, the creditworthiness of customers and so on.

Closer cooperation between DM and other corporate departments led to a re-evaluation of the criteria used to drive to optimization processes in computer modelling. Thus, the use of simple RFM models was challenged by more complex predictive models. At a purely DM level, these were used to forecast the expected payback from individual campaigns, based on the likelihood to respond and predicted purchase value. More sophisticated practitioners sought to predict the likely lifetime contribution to profits of recruiting a new customer, reflecting the discounted value of a new customer's future flow of transactions and resulting costs incurred over the lifetime of his or her account. This was an ambition that few marketers successfully realized.

Predictive modelling becomes more sophisticated

RFM models were easy for managers to understand as they did not require statistical software or expertise and were straightforward to apply to customer data. Unlike the RFM approach, which describes past purchasing behaviour and assumes that customers continue to behave in the same way, predictive techniques attempted to forecast future behaviour. By definition, RFM only incorporates three variables, whereas the inputs to predictive models are only limited by the number of variables recorded in the data and the processing power of the analytical software.

The increasing use of predictive models led to improved statistical techniques being introduced for modelling responses to DM campaigns. The Chi-Squared Automatic Interaction Detection (CHAID) method was launched in the late 1980s[6] and has become widely used, as CHAID combined the advantages of RFM and predictive modelling. When practitioners started to apply predictive analytics widely, either to predict response propensities or the purchase values of responders, they realized that the complexities went beyond the choice of modelling technique.

The advent of data mining

A best practice process was required for developing and implementing models. The activity became known as data mining, and the Cross Industry Standard Process guide[7] for data mining was launched in 1999. Nowadays, users have access to a wide range of systems and tools for data analysis and mining, including automated methods for building and deploying literally hundreds of models on a customer database. A review of the data mining process and the techniques currently available is provided by Leventhal.[8]

The shift from wave to event-driven communications

The advent of relational databases also extended the scope of analysis and segmentation from the use of data aggregated to account level to data derived from individual transactions. This spawned much greater interest in the analysis of product purchase below the category level, for instance, using as predictive variables the colour, the country of origin and the grape variety of the wine that was purchased rather than just the fact that it was wine that a customer bought. Such analysis could only be undertaken using data mining tools that can search for patterns among a very large number of different variables and reduce a diverse set of customers to a limited set of intelligible market segments.

Clustering applied to transactional data

Business analysts often found it beneficial, when tackling the analysis of a seemingly infinite number of product categories, to make use of tools for data reduction, in particular cluster analysis. This technique, originally used to build geodemographic classifications, could in this context be used to organize customers into broad-brush behavioural groupings, according to the products that they purchased.

One example of the use of cluster analysis to create customer groups was the FRuitS segmentation developed in the mid-1990s by Berry Consulting using NOP's Financial Research Survey. This segmentation went beyond a simple set of groups — it was designed to segment consumers on their use of financial services and therefore was based on dimensions of known importance in this market: life stage, financial strength (ie ability to purchase) and product portfolio (ie breadth of products used).

Cluster analysis was applied in two different ways. At the lowest level of the segmentation, clustering was used to create groups of consumers with different patterns of financial product ownership — this was the product portfolio dimension of FRuitS and identified nine types of financial services holders. The life stage and financial strength dimensions were specified by demographics such as age, marital status, presence of children, household income and total value of savings and investments. A fuller account of FRuitS development is given by Leventhal.[9]

Events used to trigger communications

As it became easier for analysts to access transactional records, time-specific events such as the anniversary of the date on which a person booked a holiday became a much more important source of information for use in analysis. As a result of these analyses, businesses began to appreciate the advantages of using specific events

as triggers for direct communications to a customer rather than, as previously, organizing all communications on a monthly cycle of selections. This shift from cyclical to event or trigger-based communications was facilitated by the move from maintaining data on flat files to the use of relational databases from which queries, as well as selections, could be conducted on an *ad hoc* basis.

As the amount of data that Direct Marketers could access grew, the focus of segmentation shifted from the selection of the customers who should or should not be mailed to the optimization of the timing and content of customer communications. As sales targets still applied at a product level, this resulted in the growing need for models for prioritizing and timetabling the different campaign mailings that should be sent to each individual customer during any particular period of time.[10]

Reduced reliance on 'wave' campaigns

As it became easier to access transactional and response data, Direct Marketers found themselves redesigning their communications strategies to focus on event-based trigger mailings rather than relying exclusive on (typically) monthly mailing campaigns based on data summarized at product or account level. This trend coincided with developments in printing technology that made it easier and much less expensive to create multiple selections of copy and imagery in a single run.

Previously prospects for different campaigns had to be separated out into separate files and the production of mailing material had to be undertaken segment by segment. Changes in print technology and data management resulted in the development of what became known as mass customization or one-to-one marketing.

Some believed that the shift from mass marketing to targeted marketing would eventually lead to the redundancy of segmentation and a regime in which every individual would be serviced in a manner reflecting his or her own unique characteristics. It may well be that there are some business areas where this prediction has come true — online marketing is theoretically capable of infinite variations — but achieving this goal has proved far harder than first imagined.

Recognizing the financial value of customer-contributed information

The demand for ever greater personalization at this time spawned the development of a new class of business, such as CMT's National Shopper Survey and ICD's Consumer Lifestyle Survey, the purpose of which was to generate qualified names and addresses for list rental and customized couponing.

Lifestyle surveys generate qualified prospect pools

Unlike market research surveys, which were carefully stratified and assured the anonymity of those who responded, these 'lifestyle' questionnaires sought to create large databases of names and addresses that could be segmented by a hundred or more different product, lifestyle and media preferences. Some of these businesses acted as direct response media in their own right. Mailings to survey respondents would carry a variety of coupons, the value of which was driven by their earlier responses. For example, if you reported that you owned a large dog, not

only would you receive a promotion for dog food, but the value on the voucher you received would be of a higher value than if you reported owning a small dog.

With increased consumer familiarity response rates declined to the point where the original business model was no longer profitable. However, the underlying concept of promotional coupons, personalized according to individual product preferences, continues to be used to great effect by companies selling a wide range of products to their consumers, as, for instance, with Tesco's Clubcard reward system.[11,12] Today, it is previous transactional data rather than survey response that drives the personalization.

An even more prescient insight of these businesses was the monetary value of information volunteered by individual consumers, a notion that underlies in different form the business models underpinning the success of Facebook, launched in 2004, and Twitter, 2 years later. Before the advent of online communication, this content could not be collected and processed at a low enough cost for the business model to be effective.

The shift from single to multi-channel operation

During its early phase, DM was often perceived as a new channel enabling a new type of business to develop using a distinct form of distribution. Some people, it was supposed, preferred to shop for certain products by mail order, some preferred telephone banking to high street banking, others preferred to buy insurance door-to-door, others liked using mail order agents. These changing preferences reflected changing work and leisure patterns. This thinking led to the launch in 1999 of Direct Line and First Direct, both providing a 24 hours telephone-only service.

Banks move to multi-channel

As the practice of DM became demystified, it began to find favour with an increasing number of large 'bricks and mortar' organizations. While some financial service providers set up their own telephone-based subsidiaries, others supplemented their high street presence by expanding their direct communications with customers from the mailing of monthly statements to that of product propositions. Recipients were invited to respond either by telephone or via the branch.

Banks' use of direct mail was facilitated by their newfound ability to use credit reference data to restrict mailshots to prospects with a good credit history. In the case of credit mailings, DM had the advantage over retail marketing in that applications for loans could be pre-approved. Segmentation could also be extended to the interest rate at which people could borrow or the upper limit of the value of the loan. The bank branch was restricted in the variety of tariffs that could be displayed visibly in the branch. That made it difficult to price products on a customer-by-customer basis taking regard of likely differences in customer profitability. The opportunity to customize terms and prices based on information about individual customers has been greatly increased by the advent of online self-servicing systems.

Therefore, banks were among the first to offer services on a multi-channel basis and increasingly sought to segment customers according to the channel preferred by the customer and the channel most likely to maximize profitability to the bank, reflecting the very different costs of servicing customers via different channels. An increasingly important part of banks' activity was to encourage low lifetime-value customers to migrate from high-cost to low-cost channels, and vice versa.

The introduction of self-service

During the 1970s, the introduction of the Automated Telling Machine (ATM) provided an addition to the mix of channels offered, even if it was used for self-servicing rather than for direct selling. Initially, this channel was only accessible at the 'bricks and mortar' location of the bank operating the service. However, with interchangeability of cards and ATMs, and with the introduction of an increasing number of non-cash services (eg mobile phone top-ups), the ATM has become one of the most important touch points for financial service customers.

Improved consumer confidence enables shift to self-service

ATMs were important in that they developed consumers' confidence in their ability to use self-servicing devices. This enabled businesses to reduce costs and prices, as low-cost airlines did, by introducing automation to the customer ordering process. Although petrol retailers for some reason have been largely unsuccessful in persuading customers to pay at the pump rather than the kiosk, shoppers' greater familiarity with self-ordering systems has enabled supermarket retailers to persuade many customers to use automated self-service checkouts in place of manned tills.

This has now extended even further into the 'click and collect' cross-channel ordering and fulfilment used by, among others, Marks and Spencer, Argos and Tesco, which allows the consumer to purchase online or via a mobile phone app and then collect their purchase at a convenient store.

Finding an appropriate role for the telephone

With the growing use of telephones, particularly mobiles, and the declining cost of telephone charges relative to post, banks increasingly began to extend their use of telephone call centres to act as an additional channel for transactions, rather than just as a device for resolving customer service issues. The growth of the call centre as a channel for transactions itself precipitated a series of technological innovations. These included systems for managing queues, the pre-recording of customer information and systems for directing calls to appropriate handlers using the telephone as a keypad. Computer-assisted scripting systems enabled the way calls were handled to be more effectively and consistently controlled, scripting the different responses that could be given in different circumstances, enabling the recording of content for auditing purposes and providing real-time statistics on call centre performance. To this list in recent years has been added voice recognition systems and software that can infer callers' sentiment by reference to call content or to stresses in the voice.

Call centre management focuses on operations rather than segmentation

The use of the call centre had added hugely to the volume of information that can be used for DM purposes. However, the analytics needed to make use of the data are very different from those with which DM analysts are familiar. It has not proved easy to integrate event-based data recorded at a call centre with information derived from other sources. Applying customer segmentation in real time is also a major analytical challenge.

As a general rule, the management of telephone call centres has focused less than other DM channels on improving business performance via better customer targeting and much more on optimizing the use of the operational staff or, in the case of outsourcing operations to other countries, reducing the average cost per call.

Call centres handle the most difficult-to-service enquiries

In their early days, call centres were often used for handling direct response from TV advertisements. The effectiveness of this strategy has been undermined by difficulties in managing fluctuations in peak volumes and their consequent impact on costs. Sharing the use of larger operations servicing multiple campaigns has proved logistically difficult. By contrast, fluctuations in peak volumes is not a problem for self-servicing channels, which is why an internet URL is taking over from a telephone number in the call to action. As a result, these centres have become consigned either to outbound calling or to dealing with the residue of difficult-to-service enquiries that cannot be addressed using self-servicing channels.

The rise of 'clicks and mortar' operations

The introduction of telemarketing as a communications channel and, more recently, the internet, provoked discussion of whether the supposition that consumers had generic preferences between different channels was really correct. If they did, then commercial opportunities existed for different forms of direct organization based on direct mail, the telephone and the internet. The prevailing view to the contrary was that that most consumers prefer to deal with organizations that offer a range of different channels for self-servicing, information and handling customer issues and that consumer preferences for different channels varies more according to circumstances than to any innate preference.[13]

Retailers adopt online strategies

Such a conclusion implied that even traditional 'bricks and mortar' non-financial service organizations needed to develop an internet presence. Most grocery retailers have recognized the necessity of offering an internet-sourced delivery service.[14]

DM becomes business as usual

As a result of the rise of multi-channel operations, it has become increasingly difficult for DM to separate itself out as a distinct business philosophy. Its role has evolved to become a source of expertise in multiple business processes spanning the processing of certain data types, the use of certain analytical techniques and in the application of certain types of segmentation strategy based on customer needs. It is also the source of the 'test and learn' ethos, placing a strong

emphasis on measurement of outcomes to define next actions — this underpins many digital marketing strategies, such as multivariate testing of display ad copy through to dynamic pricing and ranging.

Measurement — A traditional DM strength

One consequence of this integration of DM into business as usual is that campaigns, whose effectiveness in a single-channel organization could be evaluated using a simple set of parameters, now need to be evaluated using more complicated metrics. In the early days of DM, one of its advantages was that it was an ideal environment in which to build statistical evidence of what worked and what did not.

Today, in order to evaluate the true impact of a retailer's marketing communications, it is necessary to measure not just the volume of direct response, but also the incremental value of additional purchases from the store, information that then needs to be linked back to the identity of the customers who were mailed. This has resulted in the drive to capture information on the clickstream of existing customers and to relate their web visiting behaviour with information on their subsequent order purchases.

Revisiting segmentation

The introduction of the internet as a communications and sales channel was disruptive of established methods in that it offered a much lower cost communications channel than any previously used by Direct Marketers. If it has virtually zero cost, one might suppose that there is no penalty for its wasteful use, in which case there is little need for the application of segmentation strategies required to ensure profitability from the use of higher-cost channels.

The internet — A zero-cost channel?

In practice, the 'zero cost' characteristic of the internet has created such a high volume of traffic that, for the recipient, a key consideration is the prioritization of messages that are relevant. Therefore, there is a role for segmentation in identifying and ensuring the relevance of communications to each target prospect or customer. Segmentation can, of course, be applied at a very fine level of detail in the purchase of keyword-based online display advertising and also in the manipulation of page-based search rankings, as well as to the use of social media.

Whereas in a pre-digital age the paucity of information about customers made it imperative to maximize the use of 'external' and inferred data, the digital marketer has access to so much behavioural information that there is a strong temptation to overlook the predictive value of contextual data, such as the identity of a person's social media friends and the demographics of the neighbours with whom they socialize.

Although we can now treat each consumer as a digital island, unconnected to social and geographical neighbours, analytics is increasingly recognizing the predictive value of understanding the influence of both forms of network on attitudes and purchasing decisions. Thus, in addition to the content of social media being of predictive importance in its own right, social media provides an additional important source of analytic information in the form of the character of the contacts that a person communicates with.

Linking DM to new sources of customer information

In a similar manner, the shift from fixed to mobile digital also provides opportunities for accessing information on consumers' locations, present and past, which themselves provide a new and important basis for segmentation, for example, using texts to inform people of offers in local outlets, and a further opportunity to integrate digital marketing with service through 'bricks and mortar' operations.

Enriching customer records with locational information

The internet also provides an alternative way of undertaking market research surveys. Of increasing importance in large organizations is the linkage of information from customer surveys with information held on a typical CRM database. Linking attitudinal information to behavioural data provides a much more rounded qualitative base for strategy development than the inference of attitudinal preferences based on transactional data or analysing survey data that is not linked to customer behaviour.

Examination of fast moving trends is also facilitated by the use of text mining to examine exchanges on Twitter. This has been used to understand public order issues (see Indianapolis case study on pages 321–328 of this issue), forecast the outcome of elections and even to predict the outbreak of diseases.

Marketing the 'long tail'

One of the positive features of the internet as a DM channel, its low cost of marketing low-visibility brands, has rejuvenated the marketing of brands with a 'long tail' of niche products. This facilitates the application of 'mass customization' to products and services, as well as to communications content.

However, one of the problems of the internet as a channel for large consumer-facing organizations is that they do not 'own' the data generated in the way that they do in relation to mail and telephony. Social networks can be mined for any data that is published publicly, that is where the information is available to all by default (as the ten core variables of the Facebook Graph are) or where the user has set privacy controls so that they are available to all. To link other data elements to the CRM systems will involve the creation of APIs and often a commercial contract with the third party in question.

The skills required to access, manipulate and interpret digital footprints from these new channels is a specialist activity currently largely disconnected from the experience of DM professionals. As a result, 'data scientists' who understand how to mine Big Data effectively are now often recruited on the basis of their technology-oriented skills base, rather than experience within marketing. This risks putting the solution ahead of the purpose and forgetting DM's core discipline of 'test and learn' and demonstrating a positive return on investment.

The web as an enabler of direct business

The web has transformed DM in numerous ways. The principal one of these is that it provides a much lower-cost channel both for the communication of content and for order processing than previous communications channels. Compared with previous forms of outbound communication, the web appears to offer access to information at

virtually zero cost — and email the opportunity to deliver messages almost free — and thus further intensifies the shift from marketing and selling through intermediaries to direct sales.[15]

Low-cost airlines take advantage of the web to introduce self-servicing

This acceleration is not just in the communication of product and marketing information, but also in self-ordering. Without the web, the airline industry would probably not be dominated by companies operating the low fare models of EasyJet and Ryan Air.

However, the consequence of the virtually zero communication cost of using email and the low cost of capturing data via the web is a high level of competitive activity. The result is that the marketer must now incur very heavy expenditure in order to achieve a sufficient level of visibility to generate an acceptable level of sales. This results in prices for key Google AdWords being subject to intense inflation, with generic keywords such as 'car insurance price quotes' and 'loans' commanding the highest prices. This investment is a particular requirement for products that form part of the mass market, in other words that are not the niche variants that form part of a long tail. Both large, high-profile brands with significant marketing budgets and niche products that buy less competed-for words are therefore advantaged over those in a squeezed middle.

The challenge posed by web analytics

The need to achieve acceptable returns from this investment motivates Direct Marketers to invest in specialist skills for maximizing page rankings, customer stickiness and click-through conversion. Each highly specialized skill is relevant to highly specific and often relatively narrow digital marketing activity. This form of knowledge does not borrow from the accumulated knowledge developed over the history of the DM profession. It constitutes a field that most practitioners would not naturally associate with the broader practice of DM. Despite this, their expertise is not hugely different in kind from the very specialist knowledge built up by catalogue mail order companies, regarding the layout of offers on their pages or indeed the media buying skills developed by advertising agencies that supported off-the-page Direct Marketers. The technologies have changed, but the underlying marketing imperatives remain the same.

Data manipulation at the heart of digital marketing

A second set of skills required for an effective web presence is the experience in the manipulation of the data. Traditionally, DM analysts had access to well-organized account summary and transactional data. Much of the data that can be accessed from the web and that is characterized by the term Big Data arrives in an unstructured or semi-structured format, or in a differently structured form that requires investment of time and ingenuity to re-order in such a way that it is suitable for conventional analysis.

New forms of analysis

Access to social media also confronts the analyst with the requirement to undertake wholly new forms of enquiry. Where previously a model

was fed by account and transaction level data, now it may be based on the customer's own network of communications.

On the other hand, the improvements in computer power and software tools designed for the purpose of creating and applying self-learning systems now make it possible to achieve very significant improvements in the speed and level of detail with which segmentation strategies can be implemented. These technologies have also made it far easier to apply these strategies in real time.

Segmentation updating cycles shortened

Whereas geodemographic segmentation systems would typically be updated every 5 years and customized response models perhaps every 2–3 years, with customers being reassigned between segments perhaps on a monthly update cycle web-based segmentation systems can operate in real time, based on access to a much richer base of real-time indicators with the variables and the segmentation being applied virtually in real time. Models can be updated more rapidly too and the variety of tracking tools from which it is possible to adapt practice to take account of evidence is far more sophisticated than that which older DM analysts would be familiar with.[16] These developments still clearly fall within the tradition of DM.

Skills — A key constraint

This explosion of analytic opportunity associated with Big Data is not necessarily easy to convert into improved targeting effectiveness. Business managers complain even more than they used to of an acute shortage of analytical skills. Where people with the requisite skills are recruited, it has often proved difficult for top mathematicians and computer scientists to understand the business context within which their models have to operate or to learn how to communicate their findings with people who have to make practical decisions regarding the use of segmentation in specific campaigns.

Conclusion

This paper has traced the evolution of DM from a process that underpinned a distinct mode of transacting business with customers into a key activity of virtually every large consumer-facing organization, government as well as commercial and not for profit. We have demonstrated that this process has been one of evolution rather than disruptive revolution. It is evident that most business concepts associated with DM are as relevant today as when they were first introduced, however differently they may now be implemented.

Despite the proliferation of technical skills now associated with different branches of the industry, there remains a strong case for a professional institute and for a professional journal that educates and stimulates awareness of these fundamental business concepts. Notwithstanding the plethora of new channels, the exponential growth of data and data processing capability and the associated proliferation of data analysis tools, it is important that practitioners should be reminded of the concepts, practices and intellectual heritage that they have in common. This requirement is likely to be as strong over the next quarter century as it has been in the past one.

References

1. Chen, D., Sain, S.L. and Guo, K. (2012) 'Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining', *Journal of Database Marketing & Customer Strategy Management*, Vol. 19, No. 3, pp. 197–208.

2. Webber, R. (2006) 'How parties used segmentation in the 2005 General Election Campaign', *Interactive Marketing*, Vol. 7, No. 3, pp. 239–252.

3. Webber, R. (2004) 'Designing geodemographic classifications to meet contemporary business needs', *Interactive Marketing*, Vol. 5, No. 3, pp. 219–237.

4. Nitsche, M. (2002) 'Developing a truly customer-centric CRM system', *Interactive Marketing*, Vol. 3, No. 3, 207–217.

5. Reichheld, F. (1996) 'The loyalty effect', *Harvard Business Review Press*.

6. Magidson, J. (1988) 'Improved statistical techniques for response modelling: Progression beyond regression', *Journal of Direct Marketing*, Vol. 2, No. 4, pp. 6–18.

7. CRISP-DM 1.0. (2000) 'Step-by-step data mining guide', *SPSS*, ftp://ftp.software.ibm.com/software/analytics/spss/support/Modeler/Documentation/14/UserManual/CRISP-DM.pdf.

8. Leventhal, B. (2010) 'An introduction to data mining and other techniques for advanced analytics', *Journal of Direct, Data and Digital Marketing Practice*, Vol. 12, No. 2, pp. 137–153.

9. Leventhal, B. (1997) 'An approach to fusing market research with database marketing', *Journal of Market Research Society*, Vol. 39, No. 4, pp. 545–558.

10. Berry, J. (2009) 'How should the goals for "contact optimisation" be set, and how should contact optimisation be managed in a multi-channel inbound and outbound environment?' *Journal of Database Marketing & Customer Strategy Management*, Vol. 16, No. 4, pp. 241–245.

11. Sleight, P. (2001) 'Tesco.com', *Interactive Marketing*, Vol. 2, No. 4, pp. 373–383.

12. Fairlie, R. (2004) 'Tesco.com: Defining the online shopping experience for the UK consumer', *Interactive Marketing*, Vol. 5, No. 4, pp. 373–377.

13. Peterson, M., Gröne, F., Kammer, K. and Kirscheneder, J. (2010) 'Multi-channel customer management: Delighting consumers, driving efficiency', *Journal of Direct, Data and Digital Marketing Practice*, Vol. 12, No. 1, pp. 10–15.

14. Starkey, S. (2010) 'E-retail — Using home delivery as a service differentiator and strategic marketing tool', *Journal of Direct, Data and Digital Marketing Practice*, Vol. 12, No. 2, pp. 165–173.

15. McCarthy, J. (2006) 'New technology briefings: E-catalogues and e-brochures: Their part in record e-retail figures', *Journal of Direct Data and Digital Marketing Practice*, Vol. 8, No. 2, pp. 151–161.

16. Hymas, J. (2001) 'Online marketing: Segmentation and targeted customer strategies for the web', *Journal of Financial Services Marketing*, Vol. 5, No. 4, pp. 326–331.