White Paper on Enterprise Data Management (CDP/DMP)

CDP/DMP: Are We Using it the Right Way?

In the era of mobile Internet, digital media is getting more and more complex and changeable. Single-focused marketing strategy and slow-paced marketing campaign are no longer fit for the current marketing environment. At the same time, MarTech (Marketing Technology) is also rapidly evolving and iterating, and creating values in different scenarios such as advertising and communication, data flow management, e-commerce, offline retail, membership program operation, etc. Over the past two years, the construction of first-party data management platform (DMP) and customer data platform (CDP) has become a new trend, but this has not happened overnight. It is the product of the integration of multiple ecosystems including digital advertising, e-commerce, new retail sales and membership operation, triggered by the explosion of Internet applications in China over the past decade.

Now, let us first go back to the question in the title of this book – why do enterprises need CDP/DMP?

All companies will be data companies

There are both external factors that drive enterprises to build CDP/DMP, such as the improvement of data laws and regulations, the maturity of data collection technology, the change of consumer habits, etc.; and also internal factors, especially the inability of traditional marketing methods and means to keep up with the fast-changing consumer habits and fierce market competition, and the need of refining operation with data. In particular, as the value of data becomes increasingly prominent in the digital age and data plays a more important role in driving economic growth, data is the key to the digital transformation of enterprises. In this context, marketing data, which is closest to consumers, represents the core and foundation of brand digitalization. Data assets will be one of the most valuable assets of enterprises in the future.
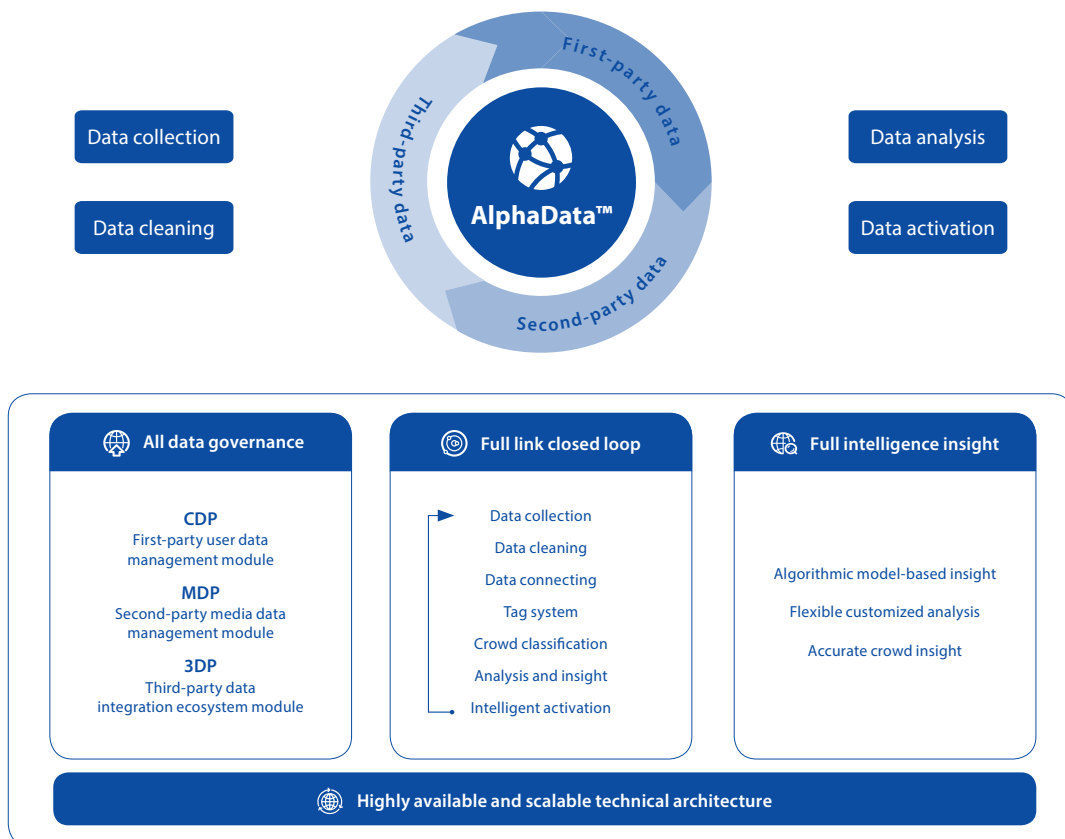
Ticket to the age of AI marketing, cornerstone of AI-enabled decision

The essence of AI is data empowerment. Only by mastering large amounts of big data can enterprises gain market initiative through algorithms, modeling, cloud computing and other capabilities. Therefore, we need to collect more data, activate and make good use of these data, teach the AI engine more knowledge of marketing patterns, learn to identify the unique features of marketing, find more marketing insights, make more scientific marketing decisions, and realize AI-enabled decisions. These capabilities are all dependent on data collection and management, which is the core function of CDP/DMP. Therefore, for brand owners, to build and make good use of CDP/DMP is the ticket to the age of AI marketing.

Do we need DMP or CDP:
Concept is not important, but essence is the key

According to the definition of Forrester, an international authoritative independent research institution, DMP includes three key points: first, integrate all data of the first, second and third parties; second, send all data to self-media or vertical media after standardization and subdivision management; third, establish insights on the target audience through data analysis, and then generate strategies. And the function of the CDP platform is to integrate existing customers based on the contact data of the enterprise, and help the enterprise to develop customer insight and conduct customer operation.

AlphaData™, the one-stop intelligent enterprise data management platform developed by DeepZero/iPinYou, has four key advantages, including full data governance, full link closed-loop, full intelligent insight, and high-availability and expandable technology architecture. It enables enterprises to comprehensively manage the data of users (first party), media (second party) and external partners (third party), and make full use of the data accumulated by DeepZero/iPinYou over the years to build technological advantages, so as to help enterprises integrate the data of both private and public domains, including users, members, fans on social network, media, etc., and generate recommended strategies and support marketing decisions through media value insight, crowd behavior insight and complete attribution model.

Data collection

Data cleaning

First-party data

Third-party data

AlphaData™

Second-party data

Data analysis

Data activation

All data governance

CDP
First-party user data management module

MDP
Second-party media data management module

3DP
Third-party data integration ecosystem module

Full link closed loop

Data collection
Data cleaning
Data connecting
Tag system
Crowd classification
Analysis and insight
Intelligent activation

Full intelligence insight

Algorithmic model-based insight

Flexible customized analysis

Accurate crowd insight

Highly available and scalable technical architecture

Be it CDP or DMP, from the beginning to the end, the key demand of enterprises is to have a complete data management solution that can help them process data collection, governance, mining and activation to achieve business growth. So how to quickly locate their essential need based on the available conditions, and come up with a solution? This book contains some of the problems we have encountered in the "application" of DMP and CDP, which we hope could serve as a reference for enterprises that choose to use CDP/DMP.

Does CDP/DMP have clear business objectives/application scenarios?

Why it matters

When we realize the importance of digital transformation for the enterprise and decide to take the first step, we should first set out a clear business objective for enterprise data management, or specify the application scenario of CDP/DMP, so as to define the kind of data capabilities that are required for CDP/DMP. To put it simply, find the destination first, and then see how to get there. In any case, for enterprises, data management platform is just a tool, and the quality of such a tool is to be judged by its ability to help enterprises achieve their goals. Every marketing activity needs a set of key indicators for performance assessment in order to make improvement and generate growth. Likewise, in order to better guide marketing decisions, we need to be clear about the business objectives and application scenarios, and use big data and AI technology to produce more effective and accurate insights.

How to help enterprises solve this problem

To define the business objectives or application scenarios of CDP/DMP, we will need a large collection of actual cases and rich practical experience. However, the reality is that many enterprises are struggling under digital transformation. Enterprises know that they now live in the age of digital transformation, and that they must keep up with the pace of the times. Every transformation is a new experiment with both risks and opportunities. Therefore, it is necessary to interact with different technology partners and test the system. Only those partners with rich experience in building data management platform, who are deeply involved in the industry, can help enterprises build CDP/DMP consistent with their business objectives and application scenarios.

Case

Philips one-stop intelligent data management platform

As one of the world's largest brands of electronic products, Philips owns the data of thousands of products in multiple categories and from sources both online and offline. In order to boost the efficiency of marketing management, seek better business performance, and build brand influence, Philips has decided to put all marketing activities under unified monitoring and management, with consistent marketing targets.
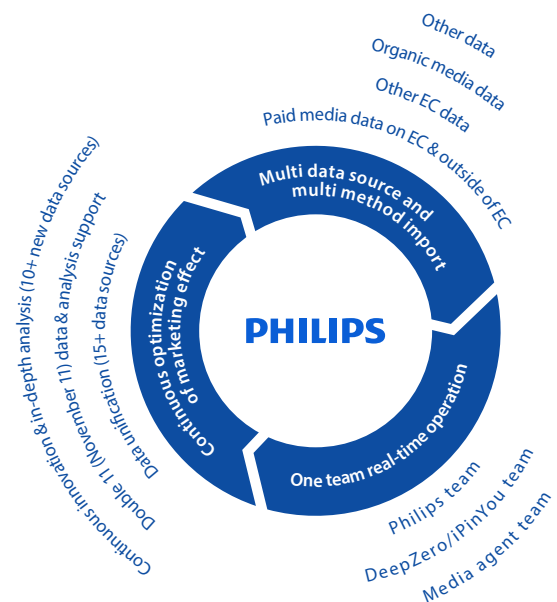
Is our data usable?

🎯 Why it matters

Data is the cornerstone of intelligent marketing and efficient customer operation, but it is also the massive data volume that makes the application process complicated. As we all know, the important goal of building first-party data management platform (DMP) and customer data platform (CDP) is to promote marketing growth. It forms customer insight through unified management of decentralized first-party data, combines the high-quality customer data owned by the enterprise itself with second and third-party data to create a 360° user profile and enables business decisions with more comprehensive customer knowledge and more accurate analysis. But the foundation to achieve this ultimate goal is the possession of truly "valuable" data. Locating valuable data is essential for building successful CDP/DMP and realizing enterprise data management.

💡 How to help enterprises solve this problem

1. Conducting full-link data cleaning to remove abnormal data flow.

Data cleaning is not just a simple procedure of format rearrangement and basic forms of data processing such as duplicate removal, but a process of multi-dimensional filtering and abnormal data flow removal based on intelligent algorithm model. At the same time, under the multi-dimensional, customized criteria for screening invalid arrival data, the built-in algorithm of the "deluxe" DMP version is able to intercept fraudulent data in real time, realize full-link data cleaning and guarantee the data quality of CDP/DMP at the source .

2. Establishing link between the data of multiple parties to solve the problem of data islands.

When conducting data management, many enterprises find they are trapped in "data islands". The data of different departments are not connected, and the data of different sales channels cannot be matched. The massive self-owned data of enterprises cannot be put under unified and efficient management. Therefore, connecting the data of various contacts and departments in a way consistent with the law, matching data at their source and building an exclusive label system is an important part of CDP/DMP application in intelligent marketing and also the key to ensure the success of customer/potential customer marketing.

3. Connecting with third-party data to meet diversified data analysis needs.

In the marketing industry, due to the "data garden" built by giant platform operators led by BAT (Baidu, Alibaba and Tencent), a large number of data have failed to be activated and realize their value. In the environment of data barrier, the ability to establish the ecosystem for third-party data cooperation and connect with third-party data will add wings to the enterprise. Integrating internal and external data to form a closed-loop for data feedback can activate data and help make informed decisions.

Case

Metro creates super ID to establish link between member data collected through multiple channels

Metro, the world's leading member-based e-commerce enterprise, is a typical company that collects data from multiple sources both online and offline. As a member-based supermarket, Metro's data comes from multiple marketing contacts, including brick-and-mortar shopping mall, WeChat, APP, online shopping mall, Tmall, official Metro website, e-news, e-mail, SMS and commercial media. Metro has established the link between member information collected through multiple channels, and built its own user label system "Super ID" to generate marketing insight and drive user growth.

1

360° customer profile

Based on the online and offline consumption habits of members, DMP can locate the target users, learn about their preferences through the portrait, and gain insight into their interests and consumption habits.

+

2

Pinpointing targeted customers and expanding potential customer groups

The DMP can make analysis on the target customers, generate the seed group, and expand the scope of potential members with the extended function of looklike provided by the third party, improve the effectiveness of member recruitment, and translate expanded membership into sales growth.

+

3

Media strategy optimization

DMP's marketing-to-sales funnel tool can make timely analysis and comparison of the marketing-to-sales data of different advertising media, evaluate the media value, optimize the media strategy, and improve the advertising effect.

Is our data safe to use?

🎯 Why it matters

Over the past decades, Chinese enterprises have mostly adopted IT application, and accumulated a large volume of high-value data, which provides a good foundation for the algorithm upgrade of AI technology and the expansion of application scenarios. At the same time, enterprises are also seeking to "activate" these valuable data through artificial intelligence, and find new areas of business values, business processes or customer demands from data.

One of the processes of "activation" includes the establishment of data management platform. To build first-party data management platform, enterprises need to make greater investment in capital and personnel, and further integrate with external platforms, as well as exchange data. This also makes data security critically important for enterprises. On the other hand, the data security specification system in the Chinese market is not mature enough, which makes enterprises feel less "secure". In the first research report on thought leadership of first-party data management platforms in China, From Data Management to Intelligent Decision-making (hereinafter referred to as "thought leadership research report"), we sent out 100 questionnaires to persons responsible for DMP management, decision-making and implementation in large and medium-sized Chinese enterprises, and found that 81% of the respondents recognized the importance of data security and customer privacy management capability, and the legal departments of 72% of the enterprises would participate in the screening process of the platform.

The global attention to data security is clear to all. The General Data Protection Regulation (GDPR) is under implementation in the EU, while China was enacted its Personal Information Security Code for Information Security Technology on May 1, 2018. Regulators and consumers are both gaining increased awareness of data privacy and security. This prompts enterprises to choose platform providers that can keep up with this trend while strengthening the compliance in their own data use.

💡 How to help enterprises solve this problem

Technology-wise, the life cycle of data can be divided into six stages: collection, transmission, storage, processing, exchange, destruction; data security is to make sure that data is secure in all stages.
The issue of data security should be considered in the context of full life cycle:
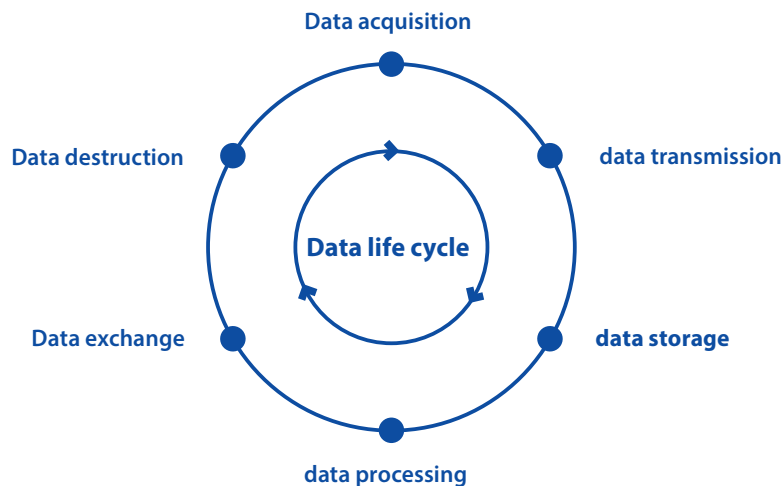
For example, data classification, cleaning, comparison and quality monitoring in the stage of data acquisition; storage security, access control, data copy, data archiving and data timeliness in the stage of data storage; distributed processing security, data encryption, data desensitization and data traceability in the stage of data processing and exchange; data import and export, sharing, release and exchange monitoring in the stage of data exchange; media use management, data destruction, media destruction and other security issues in the stage of data destruction.

Enterprises should pay attention to the compliance and integrity of user authorization process when collecting and processing user data: different types of data have different risk levels. Enterprises often collect and process personal data from all kinds of channels at the same time. It is unrealistic and costly to use the method of user consent and authorization for the compliance control of all types of data. Therefore, when formulating compliance strategies according to GDPR and the Personal Information Security Code for Information Security Technology, the enterprises concerned should first put data into different categories according to the source, method and type of data collected by products or services, conduct classification and multi-tiered management, and establish different compliance processes.

For the data voluntarily provided by the users, the enterprise should collect data in the way of "explicit consent and authorization", inform the users the purpose, scope and function of data collection, and establish a fully-functional review mechanism for user data storage and utilization. At the same time, the enterprise should uphold the principle of "necessary and minimum collection", do not collect personal user information unrelated to products or services, and do not collect unnecessary information.

For the data shared with the third party, the enterprise should establish a security assessment and audit system for transaction and use, and classify the data provided by the third party and the enterprise's own data for management purposes; for the data provided by the enterprise to the third party, regular supervision and audit should be conducted; for the violations of the third party, the enterprise should issue a notice of rectification in time, or terminate the use of the data by the third party.

For the data automatically collected from public sources, the enterprise should establish a data legality review process to ensure that the automatically collected data indeed comes from public sources and that such information is not collected without authorization. At the same time, the enterprise should ensure that the data capture process does not violate any agreement or restrictive arrangement.

For the data information with special content, the enterprise should separately establish a special technical encryption and storage management system, put in place a more stringent compliance process on data use, and conduct de-identification or anonymization in data use to avoid leaking information and data.

Many enterprises are more inclined to adopt local deployment when building CDP/DMP due to the consideration of data security, and are not willing to put data in the cloud. Some enterprises will choose hybrid solutions to activate data value by encrypting first-party data and putting it in the cloud for integration with third-party data. This not only ensures the security of customer data, but also addresses the flexibility of algorithm and activates the value behind data. The whole implementation process is automatic and transparent. Finally, all the data are deposited as first-party user data. In terms of the application of AI technology, "federated learning", an encrypted paradigm of distributed machine learning, make it possible for all parties to build a model without disclosing the original data. The application of "federated learning" in CDP/DMP construction could also mitigate the problem of data security.

Do I Really Know My Consumers?

Why it matters

In the 1990s, the story of "beer" and "diaper" shocked the marketing world. Two seemingly unrelated products ended up in the same shopping basket of a young father. This story has also opened a new horizon for enterprises to learn about their consumers. Do you really know your consumers? This is actually a typical case of data application.

Enterprises are now paying more and more attention to collecting and sorting out consumer data of various kinds through various means and constructing their own CDP/DMP platform, with the aim of unlocking the commercial value of these structured and unstructured data. But how to achieve this is actually a technical problem. Under the premise of legal compliance, a very important step to take once data governance and connection is achieved is data labeling. Putting a label on the enterprise's data, and establishing the enterprise's own first-party data label system, so as to give a unique label to each ID, and, through management and processing, achieve real insight into the internal needs of consumers could guide the enterprises to make better marketing decisions.

How to help enterprises solve this problem

So here come the questions... :

1. How to design a fully-functional first-party user label system, how to put the label on and what types of label are needed?
2. After the establishment of the label system, how to realize its commercial value?

Female | Focus on quality

Like fashion | Love Yoga/jogging

Love to try new things | Petty bourgeois

Like wearable devices | Love to watch American TV series

Like to buy on overseas shopping sites | **Healthy life**

Love to wear make-ups

Like Lancome | Learning driving

**Diagram of Labels in the Context of Single User ID**

How have these labels been put?

Firstly, sort out and clarify the business scenarios. Enterprises need to build their own user label system according to their own actual situation. They need to make multi-dimensional designs based on their own business attributes, product features, market demand and other factors.

Secondly, establish the classification logic of the label system. In theory, data are generally classified into three categories:
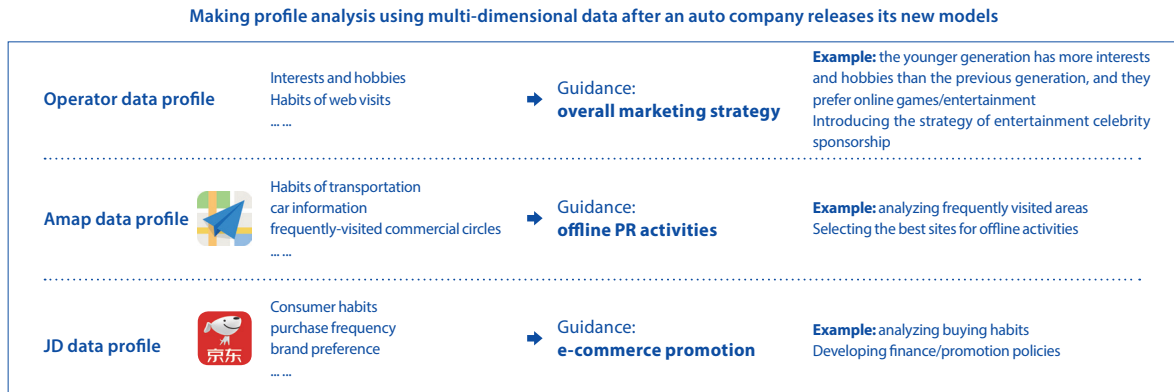
1. **Factual label.** That is, label based on the data processed by consumer behavior logs, such as consumer behavior records, advertising clicks and purchase records, etc.

2. **Model label.** That is, label based on customized model. Common models include AIPL model, RFM model and other algorithm models. There are also models tailored to customer needs, such as user rating model, additional purchase/replacement model, lead rating model, etc.

3. **Predictive label.** It makes a forecast of user preference and the next possible behavior based on existing factual data. Predictive label is an advanced label, which is a summary of the basic tags of individual ID numbers processed through the algorithm model. For example, for a fast-moving consumer product, its user has recently bought a bottle of shampoo. If, based on historical data, the cycle of re-purchase is 3 months, he may buy it again in 3 months' time.

Generally speaking, different industries and enterprises have different needs, which requires the construction of label system according to the actual situation of enterprises. Taking the automobile industry as an example, the dimensions of the first-party label system are as follows:
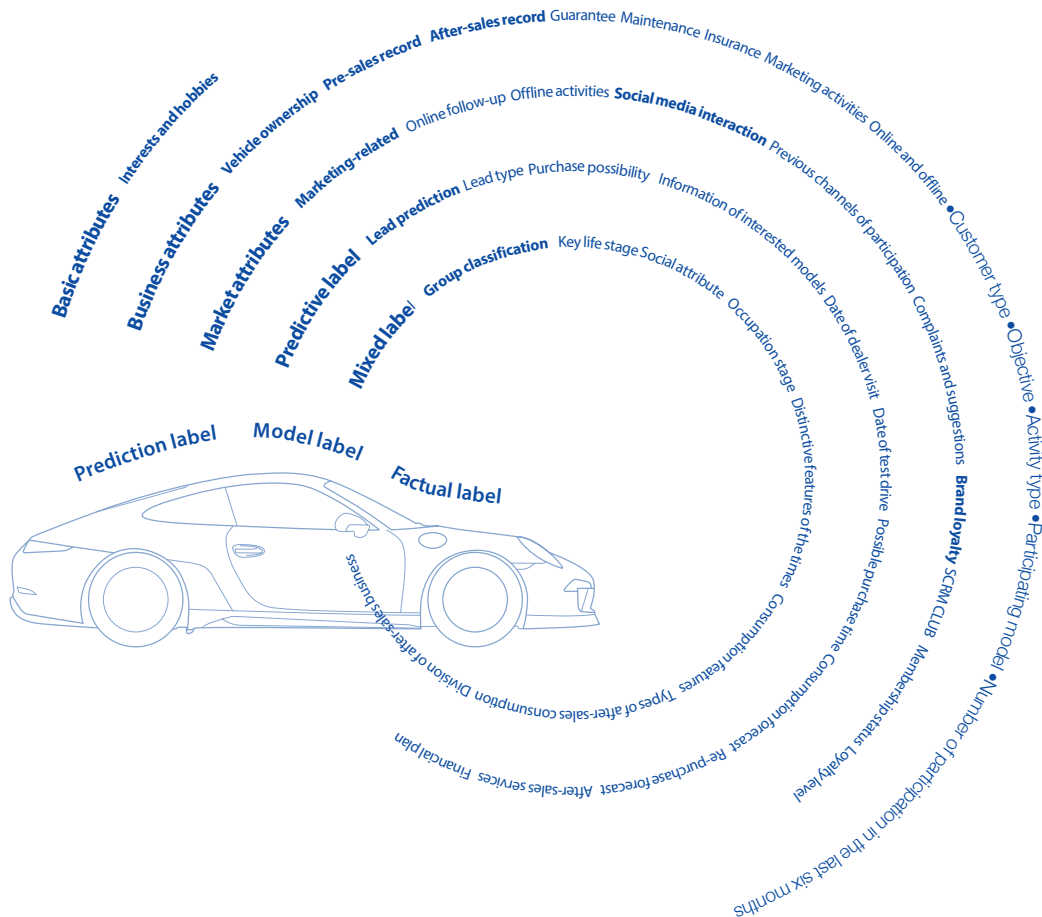
Thirdly, which is also the most important point, connecting with third-party data, continuously expanding the label system and improving the user profile.

After building their own first-party label system, enterprises need to jump out of their own ecosystem, connect with third-party data sources, and expand and improve their own first-party label system across platforms and dimensions, thereby realizing 360°consumer insight and management. Data sources complementing the regular label systems:

Making profile analysis using multi-dimensional data after an auto company releases its new models

|			
Operator data profile	Interests and hobbies Habits of web visits … …	→ Guidance: **overall marketing strategy**	**Example:** the younger generation has more interests and hobbies than the previous generation, and they prefer online games/entertainment Introducing the strategy of entertainment celebrity sponsorship
Amap data profile	Habits of transportation car information frequently-visited commercial circles … …	→ Guidance: **offline PR activities**	**Example:** analyzing frequently visited areas Selecting the best sites for offline activities
JD data profile	Consumer habits purchase frequency brand preference … …	→ Guidance: **e-commerce promotion**	**Example:** analyzing buying habits Developing finance/promotion policies

Finally, the application of data labels:

1. Create more detailed segmentation of target groups, with clear classification and layer-division of users

2. Establish a profile data system in line with the application scenario to support enterprises in quickly connecting big data technology and data services

3. Conduct user profile modeling, design marketing activities, improve marketing efficiency, optimize customer experience, etc. according to the established user label system

The marketing case of business value creation through data application:

Project overview: The customer is an auto company. By integrating the customer's own first-party data, as well as the data resources of iPinYou, operators and other third parties, a user rating model has been created to help locate the leads to maximum values.

Results: the rating model can automatically pick up high-value leads and distribute to potential customers, leading to a 46% increase in test drive rate.
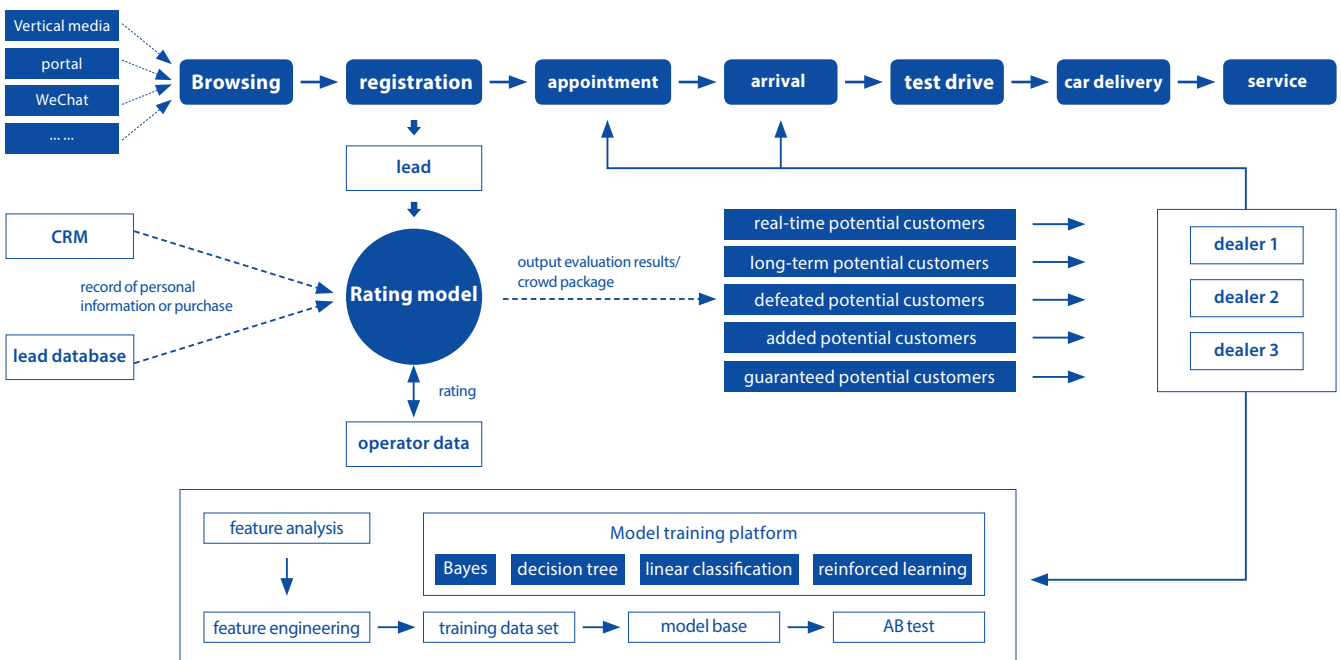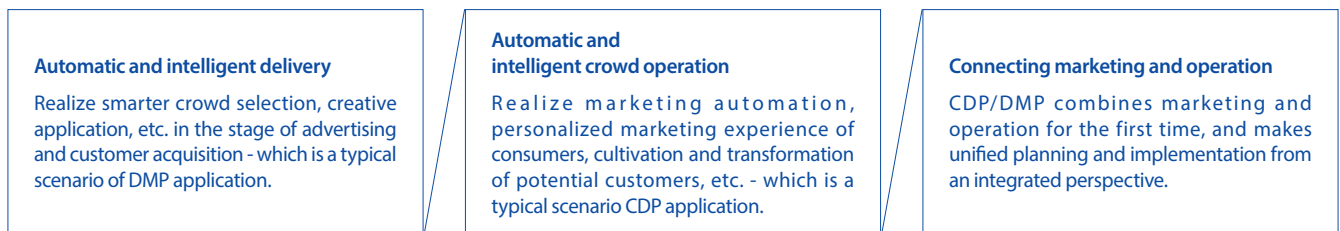
Can CDP/DMP really make intelligent decisions?

Why it matters

Enterprises have realized that data is the most important asset in the future. CDP/DMP works like the brain, tasked to drive the collaboration of external tools. Marketing is also becoming more intelligent and operational. Therefore, CDP/DMP can help enterprises make intelligent marketing decisions.

How to help enterprises solve this problem

1. Identifying the areas of application

Automatic and intelligent delivery

Realize smarter crowd selection, creative application, etc. in the stage of advertising and customer acquisition - which is a typical scenario of DMP application.

Automatic and intelligent crowd operation

Realize marketing automation, personalized marketing experience of consumers, cultivation and transformation of potential customers, etc. - which is a typical scenario CDP application.

Connecting marketing and operation

CDP/DMP combines marketing and operation for the first time, and makes unified planning and implementation from an integrated perspective.
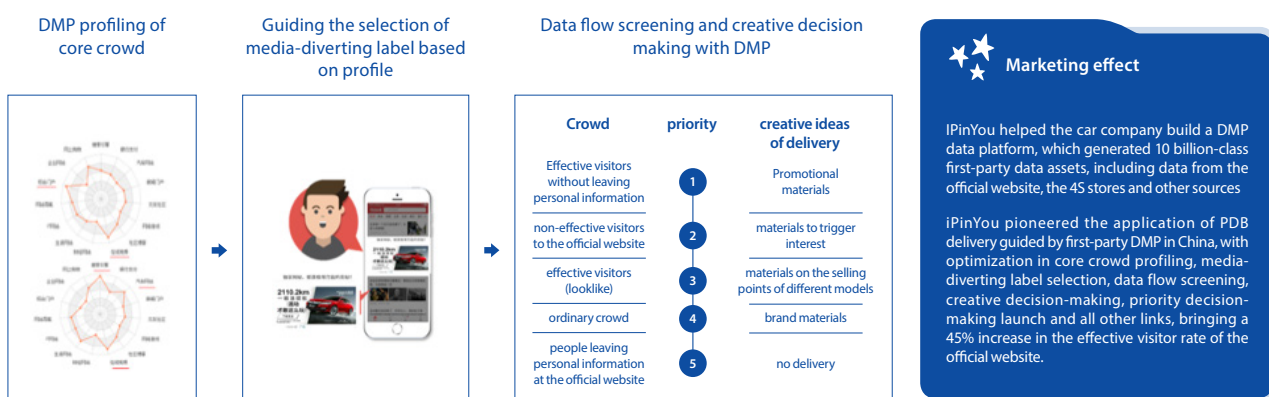
2. The fundamental value of CDP/DMP lies in its ability to connect data and decision, form a closed loop of marketing data, and promote model iteration and optimization.

The key purpose of DMP and CDP should be to help enterprises form their own closed loop of marketing data, so as to make marketing decisions more intelligent through repeated feedback and iteration. In this closed-loop, each campaign can select TA from CDP/DMP to reach, accumulate first-party data assets, make TA more accurate, and obtain better transformation. Post-delivery data monitoring is a very important feedback link. The key step of intelligent marketing is to continuously analyze, optimize and study the monitored feedback data. When choosing first-party DMP supplier, the enterprise should not only examine its ability related to DMP construction, but also examine whether the supplier has the sense and experience to build a marketing closed-loop from the overall marketing perspective.

Case

Constructing DMP to guide PDB delivery, resulting in a 45% increase in the effective visitor rate of the official website of a well-known car company.

DMP profiling of core crowd

Guiding the selection of media-diverting label based on profile

Data flow screening and creative decision making with DMP

Crowd	priority	creative ideas of delivery
Effective visitors without leaving personal information	1	Promotional materials
non-effective visitors to the official website	2	materials to trigger interest
effective visitors (looklike)	3	materials on the selling points of different models
ordinary crowd	4	brand materials
people leaving personal information at the official website	5	no delivery

Marketing effect

IPinYou helped the car company build a DMP data platform, which generated 10 billion-class first-party data assets, including data from the official website, the 4S stores and other sources

iPinYou pioneered the application of PDB delivery guided by first-party DMP in China, with optimization in core crowd profiling, media-diverting label selection, data flow screening, creative decision-making, priority decision-making launch and all other links, bringing a 45% increase in the effective visitor rate of the official website.

Should enterprises choose customized projects or standardized products?
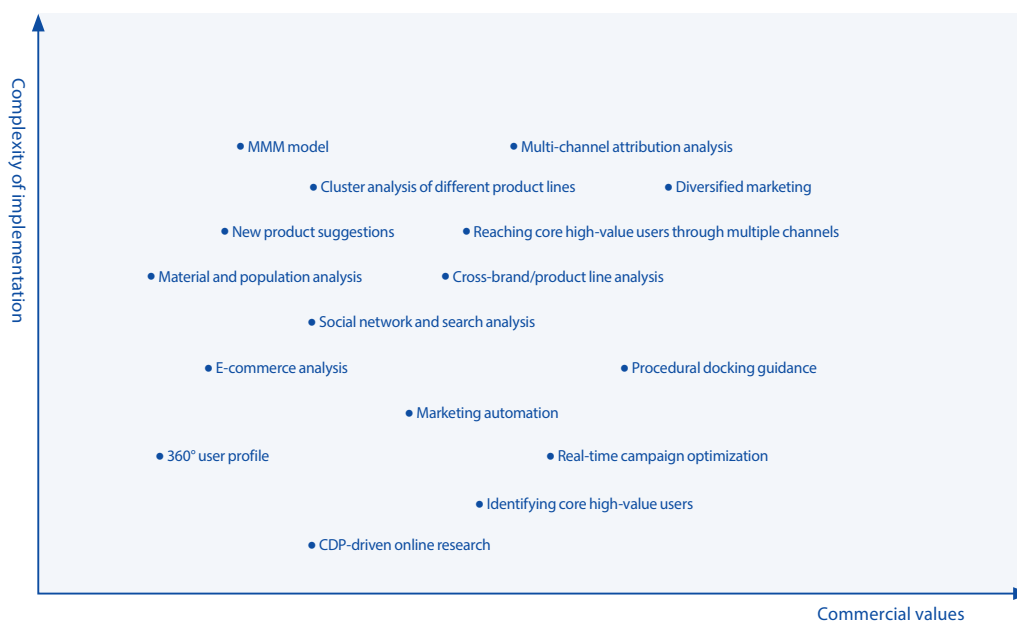
🎯 Why it matters

With the advent of CDP/DMP boom, more and more advertisers now want to construct their own CDP/DMP for the purpose of data management and activation. When it comes to the platform, a big problem for them is to choose between customized and standardized products. From the cost point of view, customized projects are relatively more expensive compared with standardized products. At the same time, the cycle of implementation will be very long, and it is difficult to produce quick benefits. It may take more than a year for the project to see actual output. However, standardized products are often unable to meet personalized product demands. In marketing, people talk about different marketing strategy for different subjects, and in terms of CDP/DMP, the demands of enterprises are also highly diversified.

💡 How to help enterprises solve this problem

Of course, this problem can be solved. Currently, many enterprises choose the mix of standardized and customized projects and seek to improve the product functions of CDP/DMP stage by stage, and ultimately realize the business goals of the enterprise.

Customized projects + Standardized products
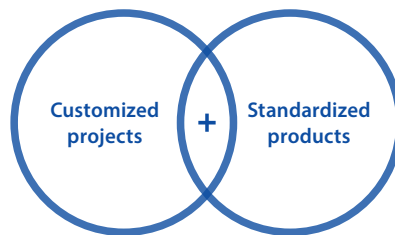
In most cases, we would recommend that customers establish their own business application scenarios according to the complexity and commercial prospect of implementation, and use the scenarios as the reference for choosing products. For example, for business scenarios like media evaluation/crowd analysis, which are not complex in implementation but high in commercial value, some powerful suppliers do have mature product modules, which can be directly deployed; demands that are based on specific business scenarios can be satisfied with offline operation support or phased-in functions. Service providers with a large customer base will be able to develop a host of customized functions and combine them into standardized modules.

Complexity of implementation

- MMM model
- Multi-channel attribution analysis
- Cluster analysis of different product lines
- Diversified marketing
- New product suggestions
- Reaching core high-value users through multiple channels
- Material and population analysis
- Cross-brand/product line analysis
- Social network and search analysis
- E-commerce analysis
- Procedural docking guidance
- Marketing automation
- 360° user profile
- Real-time campaign optimization
- Identifying core high-value users
- CDP-driven online research

Commercial values

Based on its extensive experience of customer service, the standard product functions of DeepZero/iPinYou can now cover 80% of the demands of the majority of companies. Business scenarios such as data collection, label management, campaign analysis, crowd analysis, customized analysis and one-button push can all be processed through DeepZero/iPinYou's AlphadataTM platform. Given the strong implementation ability of DeepZero/iPinYou, those special, customized demands can also be timely addressed and satisfied. For example, for car enterprise customers, we will help with the designing of effective visitors, rating models and algorithms of customer leads, real-time business display boards, etc. For FMCG enterprises, we will create RFM, forecast analysis and other models as well as develop the docking system for different business departments within the enterprise based on the attributes of different products and factors such as purchase cycle, frequency and quantity.

Reference

Forrester Consulting: From Data Management to Intelligent Decision-making -- Building a Successful First-party Data Management Platform (DMP) to Meet the Marketing Challenges in the Era of Big Data

Song Xing, Fenxi Data: White Paper - Private Domain Data and DMP/CDP of Chinese Enterprises in 2019

The Forrester Now Tech: Data Management Platforms In Asia Pacific, Q2 2019

China Academy of Information and Communications Technology: 2018 White Paper on Big Data Security

Morketing: Huang Xiaonan - CDP is the New Black in Town, What's Its Relationship with DMP?