

碩士學位論文

銀行의 DB마케팅을 위한
데이터웨어하우스 構築에 관한 研究

指導教授 千 瑛 秀



제주대학교 중앙도서관
JEJU NATIONAL UNIVERSITY LIBRARY

濟州大學校 經營大學院

經營情報學科 經營情報學專攻

成 昌 球

1999年度

銀行의 DB마케팅을 위한 데이터웨어하우스 構築에 관한 研究

指導教授 千 瑛 秀

이 論文을 經營學 碩士學位 論文으로 提出함.

1999年 12月 日

濟州大學校 經營大學院
經營情報學科 經營情報學專攻

成 昌 球

成昌球의 經營學 碩士學位 論文을 認准함.

1999年 12月 日

委 員 長

고 재 건



委 員

김 두 경



委 員

천 연 수



차 례

제1장 序 論	1
제1절 問題의 提起	1
제2절 研究의 目的과 方法	2
제3절 論文의 構成	3
제2장 銀行 데이터베이스와 DB마케팅	4
제1절 銀行의 데이터베이스	4
1. 銀行 데이터베이스의 發展段階	4
2. 銀行 데이터베이스의 問題點	7
제2절 DB마케팅의 定義와 特徵	11
1. DB마케팅의 定義	11
2. DB마케팅의 特徵	13
제3절 DB마케팅의 成長 背景과 適用 分野	16
1. DB마케팅의 成長 背景	16
2. DB마케팅의 適用 分野	19
제4절 銀行의 DB마케팅 導入 背景과 考慮 事項	22
1. 國內銀行의 DB마케팅 導入 背景	22
2. DB마케팅 導入의 先決 課題	26
3. 銀行 DB마케팅의 成功 要件	29

제3장 데이터 웨어하우스 34

제1절 데이터 웨어하우스의 定義 및 特性 34

제2절 데이터 웨어하우스의 登場 背景 37

제3절 데이터 웨어하우스 關聯 概念 40

1. 데이터 마트 40

2. 데이터 마이닝 43

제4절 데이터 웨어하우스 構築의 前提 要素 46

1. 데이터 웨어하우스에 대한 마인드 形成 46

2. 全社的인 協助 49

제5절 데이터 웨어하우스 構築 段階別 체크 포인트 51

제6절 데이터 웨어하우스 構築 事例 66

1. 國內 事例 : 三星카드(株) 66

2. 海外 事例 : 체이스맨하탄 銀行 74



제4장 DB마케팅을 위한 데이터 웨어하우스 模型設計 81

제1절 銀行 데이터 웨어하우스 具現 方式 81

제2절 銀行 데이터 웨어하우스 設計 83

제3절 데이터 웨어하우스 基盤의 DB마케팅 시스템 模型 87

제4절 데이터 웨어하우스 具現 過程 90

1. 데이터 모델링 90

2. 데이터 抽出, 變形 및 整齊 91

3. 데이터의 積載(Loading) 95

제5장 結 論 99

 제1절 研究의 要約 99

 제2절 研究의 限界 및 向後 研究 方向 100

參 考 文 獻 101

ABSTRACT 104



[표 차례]

<표 2-1> DB마케팅의 성공 및 실패분야의 일반적 구분	20
<표 2-2> DB마케팅이 활용된 사례	21
<표 2-3> DB마케팅의 구성 요소 및 세부 이슈	30
<표 3-1> 데이터 마이닝의 응용 분야	44
<표 3-2> 체이스에서 사용된 분석 도구들	76
<표 4-1> 은행 고객 분석보고서의 주요 구성 내용 예	83
<표 4-2> 시스템 구성 요소	87

[그림 차례]

<그림 2-1> 전통적 은행정보시스템 체계	8
<그림 3-1> 데이터 웨어하우징 프레임워크	39
<그림 3-2> 데이터마이닝의 흐름	45
<그림 3-3> 의사결정 피라미드	47
<그림 3-4> 역 의사결정 피라미드	48
<그림 3-5> 구축 단계별 체크 포인트	51
<그림 3-6> 삼성카드(주)의 시스템 처리 FLOW	70
<그림 3-7> 체이스의 데이터 웨어하우스/데이터 마트 아키텍처	76
<그림 4-1> 차원적 모형 설계의 예	85
<그림 4-2> 자료 테이블의 예	86
<그림 4-3> 차원 테이블의 예	86
<그림 4-4> 데이터 웨어하우스 모형	87
<그림 4-5> 은행 DB마케팅 시스템 모형	88
<그림 4-6> 은행 데이터 구조의 변환	94

제 1 장 序 論

제1절 問題의 提起

최근 금융기관간의 경쟁에 있어서는 시장 점유율보다 고객 점유율이 중요한 경쟁 포인트가 되고 있다. 금융시장의 경쟁에서 우위를 점하기 위해서는 기존고객의 지속적인 유지와 신규 고객의 창출이 필수적이며 이를 위해서는 고객에 대한 지식을 축적·분석하여 이를 기초로 한 적절한 상품의 개발 및 서비스의 제공이 이루어져야 한다는 것이다. 이런 점에서 데이터 웨어하우스(Data Warehouse) 구축을 통한 효율적인 마케팅 활동의 전개는 은행 경영 전략의 최우선 과제가 되고 있다.

그 동안 국내에서는 일부 신용카드 회사와 보험업계 등 제2금융권을 중심으로 부분적으로 데이터베이스(Database, DB) 마케팅 시스템을 구축·활용해 왔으나, 최근 IMF 체제에 들어서면서 불어닥친 금융 구조 조정 과정에서 생존 위기를 벗어난 은행들도 DB마케팅에 대한 관심이 더욱 높아지고 있다.

금융시장에서 경쟁력을 확보하여 살아남기 위해서는, 급변하는 고객의 금융 욕구(needs)에 신속히 부응할 수 있는 새로운 환경의 전산 인프라로 무장한, 보다 적극적인 영업 전략이 요구되고 있다. 특히 금융시장 개방으로 선진 마케팅 기법을 앞세운 외국 금융기관과의 고객 유치 경쟁이 치열해질 것으로 예상되면서, 고객에 대한 서비스 향상이 영업뿐만 아니라 생존의 관건으로 작용하고 있기 때문이다.

그러나 DB마케팅이 은행의 경영 전략에 시사하는 점은 은행에 따라 다르게 나타나고 있다 즉, 지방은행 및 후발은행에게는 DB마케팅이 은행의 차별화 전략의 일환으로 이용될 수 있다. 금융환경의 변화에 따른 은행간의 경쟁 심화로 은행의 차별성을 유지하는 것이 더욱 중요하게 되었기 때문이다. 그리고 은행의 차별화는 결국 은행 고객의 차별화에 의해 이루어질 수밖에 없다

특히 미국과 일본의 경우, 대형은행이나 도시은행보다는 지역은행이나 지방은행에서 이러한 고객 차별화가 적극적으로 추진되고 있다. 대형은행이나 도시은행의 대규모 점포 증설이나 전산 투자전략에 맞서서, 동일한 방법으로 경쟁하는 것이 쉽지 않

은 지방은행이나 후발은행들은 고객층을 차별화하여 개별 고객에게 특화된 서비스를 제공하는 방법으로 수익성 위주의 내실 경영을 추구하고 있다.

지난해 일부 지방은행들은 외환위기 이후 생존 전략의 일환으로 지역주민들의 향토애에 호소하여 '퇴출'이라는 최악의 상황을 넘긴 곳도 있지만, 앞으로는 더 이상 이와 같은 전략이 생존을 보장해 줄 수는 없을 것이다. 이러한 여러 가지 금융환경의 변화는 은행으로 하여금 정보기술을 활용한 효율적인 경영을 유도하고 있다.

제2절 研究의 目的과 方法

DB마케팅은 이미 오래 전에 제기된 개념으로 해외에서는 다양한 분야에 걸쳐 적용되어 왔으나, 국내에서는 90년대에 들어 소개되어 아직도 학계의 연구가 미미한 실정이며, 기업에서의 활용은 최근에서야 부분적으로 진행되고 있는 실정이다. 더구나 최근에는 경영기법과 정보기술이 전면적으로 결합하면서 새로운 단계로 접어들고 있다. 특히 전사적 데이터 웨어하우스나 부서별, 업무별 데이터마트 구축이 소개되면서 DB마케팅은 새로운 전기를 마련하고 있다.

물론, 데이터 웨어하우스만이 DB마케팅을 전개할 수 있는 기반은 아니다. 그러나, 중복 투자를 배제하고, 효과적인 DB마케팅을 전개하기 위해서는 데이터 웨어하우스가 필요하다. 또한 많은 문헌에서 데이터 웨어하우스의 대표적인 응용분야로 DB마케팅을 거론하고 있다.

최근에는 과감한 투자로 데이터 웨어하우스 기반의 DB마케팅을 시도하는 기업들이 증가하고 있고, 일부 시중 은행들도 외국의 업체들로부터 ISP(Information Strategy Planning : 정보 전략 계획) 컨설팅을 받은 후, DB마케팅을 통하여 은행의 수익성을 개선하고 고객과의 관계를 확고히 다지고자, 자사의 고객 데이터 베이스 및 자산관련 데이터를 통합적으로 관리할 수 있는 데이터 웨어하우스 구축을 본격적으로 추진하고 있다.

그러나 충분한 연구와 준비도 없이 막대한 투자가 뒤따르는 데이터 웨어하우스를 구축한다는 것은 DB마케팅의 성과를 낙관할 수 없게 만들뿐만 아니라 프로젝트의 책임문제를 일으킬 소지가 많다.

이러한 상황에서 본 연구는 다음과 같은 목적을 가지고 수행되었다
첫째, 새로운 정보기술의 조류인 데이터 웨어하우스의 필요성에 대한 이론적인 토대를 제시하고,
둘째, 은행에서의 DB마케팅 도입 필요성과 성공 요건을 제시하여 실질적인 활용에 길잡이가 될 수 있도록 하며,
셋째, 은행에서 효과적인 DB마케팅을 전개하기 위한 기반 환경으로서의 데이터 웨어하우스 구축 방안을 제시해 보고자 한다.
본 연구에서는 이를 위한 연구 방법으로서 문헌 연구와 사례 연구를 병행하였다. 먼저 문헌 연구에서는 국내외의 저서 및 논문 등을 참고로 하여 DB마케팅과 데이터 웨어하우스에 관한 이론적 고찰을 하였다. 사례 연구는 국내 은행에 대한 발표 사례가 없었기 때문에 제2금융권과 해외 은행의 사례를 중심으로 실시하였다.

제3절 論文의 構成

본 연구는 5개의 장으로 구성되어 있으며 그 내용은 다음과 같다.

제1장에서는 본 연구를 제기하게 된 상황 및 연구목적, 연구방법, 논문의 구성을 서술하였다

제2장과 제3장은 문헌연구를 통한 이론적 고찰을 하였는바, 제2장에서는 은행 데이터베이스의 현주소와 DB마케팅에 대한 선행 연구내용, 그리고 DB마케팅이 은행에 필요한 배경에 대하여 살펴보았으며, 제3장은 데이터 웨어하우스에 대한 선행 연구내용과 성공적인 데이터 웨어하우스를 구축하기 위한 요소들 그리고, 국내외의 데이터 웨어하우스 구축 사례를 살펴보았다.

제4장에서는 은행에서 데이터 웨어하우스를 기반으로 DB마케팅을 전개하기 위한 모형과 데이터 웨어하우스 구축 방안을 제시하였으며, 마지막으로 제5장에서는 본 연구의 요약과 연구의 한계 및 앞으로의 연구방향에 대해 기술하였다.

제 2 장 銀行 데이터베이스와 DB마케팅

제1절 銀行의 데이터베이스

은행은 오래 전부터 방대한 데이터베이스를 구축하여 사용하고 있지만, 데이터베이스로부터 마케팅 전개에 유용한 정보를 자유자재로 추출할 수 있게 된 것은 미국에서도 '90년대에 들어서이다. 국내 은행들도 전통적으로 계정(account)별로 데이터베이스를 구축해 왔기 때문에, DB마케팅을 전개하는 데에는 많은 문제가 있다.

따라서, 여기에서는 우선 은행의 데이터베이스 발전단계를 살펴보고¹⁾, 현재 은행 데이터베이스의 문제점을 논해 보고자 한다.

1. 銀行 데이터베이스의 發展段階

1) 업무별(application-specific) 데이터베이스

업무별 데이터베이스는 은행의 일상적인 업무들을 처리하면서 생성되는 데이터를 보관한 데이터베이스로 우리 나라와 일본에서 통용되고 있는 소위 계정계 데이터베이스가 이에 속한다. 즉, 은행이 고객과 일상적으로 거래를 처리하는 과정에서 발생하는 데이터들을 기록, 저장하는 곳이 계정계 데이터베이스이다.

은행업은 기본적으로 거래처리(transaction processing)를 위주로 하는 업종으로 매일 거래처리에서 발생하는 데이터의 양은 엄청나다. 일반 시중은행의 경우 하루 수 백만 건 이상의 거래가 발생하는 것으로 알려지고 있다. 그러나 은행으로서는 이러한 데이터로부터 원하는 정보를 뽑아 내고 분석하는 일이 쉽지 않다. 즉 계정계 시스템은 비정규적인(ad-hoc) 데이터 검색에는 적절한 시스템이 아니다. 이 시스템

1) 정철용·함유근, “고객정보시스템 구축 및 활용 전략”, 한국금융연구원 금융 PAPER 99-02, 1999, pp 4~7.

은 주로 데이터 처리 및 관리의 안전성과 효율성을 염두에 두고 설계된 시스템이다. 따라서 처리 업무별로 교차검색을 하기도 쉽지 않다

우리 나라 은행들의 경우 90년대 초에 완성된 소위 '제2차 온라인' 이후에야 여신과 수신 업무간의 연결 조치가 가능하게 되었다. 이러한 환경에서는 현업 부서의 특정요구가 주로 정보기술 부서에 의해 작성되는 보고서로 충족된다 그러나 이러한 방식은 비용과 시간 면에서 효율적이지 못하다.

2) 목적별(activity-specific) 데이터베이스

정기적으로 반복되거나 예측 가능한 정보요구를 충족시키기 위해 만들어진 데이터베이스로서 현재의 계정계 데이터베이스로는 해결이 어려운 데이터나 정보를 습득하기 위해 만들어졌다. 주로 한가지 목적을 위하여 데이터가 여러 업무로부터 추출, 종합되어 표준화된 후 목적별 데이터베이스로 입력된다. 위험관리와 수익관리와 같은 업무를 위한 데이터베이스가 가장 대표적인 경우다.

예를 들어 금리 위험관리를 위해 계정계 데이터베이스에서 원하는 데이터를 직접 추출할 수도 있으나, 그럴 경우 많은 데이터 조작 작업이 필요하고 결과적으로 메인 프레임 컴퓨터에 무리한 부하(load)가 걸릴 위험이 있다. 따라서 이러한 위험을 회피하기 위해서도 목적별 데이터베이스가 필요하게 되었다. 그리고 고정비용의 계산과 같이 특정한 식의 계산에 필요한 데이터 요소들은 한번 정해지면 이후로 변하지 않기 때문에 별도의 서버(server)에 모아서 계산하는 것이 효율적이며 안정적이다. 그리고 신속한 의사결정을 지원하기 위해서 정보의 최종사용자들이 비정기적인 조회를 보다 용이하게 할 수 있는 환경이 필요하게 된 것도 목적별 데이터베이스의 등장 배경이다.

이러한 목적별 데이터베이스가 수익관리나 위험관리 혹은 영업관리 등 각 기능별로 필요한 현실에서, 우리 나라 은행들이 그 동안 목적별 데이터베이스의 구축에 적극적이지 못했던 이유는, 우선 모든 업무에 대한 계정계 데이터베이스가 완성되지 않았으며, 투자의 효과를 단기간에 정당화하기 힘든 각종 관리시스템의 구축을 해당 부서만의 판단으로 추진하는 것이 쉽지 않았기 때문이다. 또한 정보기술 부서의 입장에서 각 부서별로 필요한 시스템을 개별적으로 추진하는 것보다는 정보기술 부

서가 총괄하여 종합적인 시스템을 구축하는 것이 데이터의 관리나 시스템의 안전성 측면에서 유리하다고 생각하고 있었던 점도 지금까지 이러한 목적별 데이터베이스의 추진을 어렵게 하였다

3) 종합 데이터베이스(enterprise-level warehouse)

미국 은행들의 경우 데이터베이스의 발전 과정을 보면, 실무 부서의 필요에 따른 목적별 데이터베이스가 먼저 정착된 다음 이를 통합하기 위한 종합 데이터베이스의 구축 필요성이 자연적으로 제기되었고 이러한 과정이 점진적으로 이루어져 왔다.

다시 말해 미국의 경우 종합적이며 통합된 정보계 시스템이 어느 한 순간에 만들어지기보다는 점진적으로 형성되어 왔다. 즉, 업무별 데이터베이스와 목적별 데이터베이스를 거쳐 세 번째 단계에 소위 데이터 웨어하우스의 개념이 소개되면서 이러한 종합적인 정보계 데이터베이스의 구축을 위해 많은 은행들이 막대한 비용과 오랜 시간을 투자하고 있다.

종합 정보계 데이터베이스가 등장하게 된 이유는 은행들이 하나의 목적DB에서 시작하여 이제 다수의 목적DB를 보유하게 되었고, 이에 따라 여러 가지 문제들이 발생하였기 때문이다. 우선 여러 목적을 위해 계정계 데이터베이스로부터 동일한 데이터를 중복하여 추출할 수밖에 없다. 예를 들어 여신 업무처리 시스템으로부터 생성된 데이터는 계리 시스템과 ALM(자산·부채 종합관리)시스템, 수익분석 보고서 작성 등 여러 분야에서 사용되므로 데이터의 중복 추출작업이 불가피하다. 그리고 여러 목적별 데이터베이스를 유지하면서 겪게 되는 또 한가지 현실적인 어려움은 동일한 데이터 항목에 대하여 목적별DB마다 다른 원칙이나 규정을 적용하여 데이터의 값을 유지하게 된다는 점이다.

예컨대 한가지 대출자산 항목에 대하여 어떤 데이터베이스에는 현재의 상환금리에 따른 값을 저장하는 반면, 또 다른 데이터베이스에서는 만기시의 잔존 현재가치를 저장한다. 미국의 은행들이 추진하고 있는 종합 정보계 데이터베이스 시스템인 데이터 웨어하우스는 자세한 수준의 데이터를 보관하여 관리하는 거대한 데이터베이스로서 이곳으로부터 여러 목적별 데이터베이스로 데이터가 공급된다.

2. 銀行 데이터베이스의 問題點

우리 나라 은행 정보화의 기반 구축은 '70년대 중반부터 시작되었다 은행들이 업무처리의 효율화를 위하여 보통예금, 저축예금 등의 결산업무에 배치(batch) 처리방식을 도입한 이후 온라인 처리방식으로 전환하기에 이르렀다 '80년대 중반에 이르러 종합적인 고객관리의 필요성에 의해 종합 온라인 체계를 구축하기 시작하였으며, '80년대 말부터 '90년대 초에 걸쳐 대부분의 은행들이 종합 온라인 체계 구축을 완료하였다 기존의 과목별 처리방식은 대고객 거래를 처리한다는 관점에서 본다면 효율적일 수 있었으나, 종합적인 고객정보의 파악이라는 관점에서 보면 필요한 고객정보를 추출하기 위해서 고객별로 별도의 작업을 수행하여야 하였다

따라서, 고객에 대한 종합 정보 추출을 쉽게 하기 위해서 종합 온라인 시스템의 도입이 필요하게 되었고, 이를 위하여 고객정보 파일(CIF : Customer Information File) 혹은 고객관리 파일(CMF : Customer Management File) 방식이 도입되었다.

각 은행이 종합 온라인 시스템으로 이행하면서 보인 가장 큰 특징 중의 하나는 거래처리를 위한 시스템(계정계)과 경영정보 관리를 위한 시스템(정보계)을 분리하여 구축하였다는 점이다. 종래의 온라인 시스템에서는 경영관리, 영업점 관리, 고객관리상 필요한 정보를 추출하기 위해서 온라인 가동이 끝난 뒤 배치(batch)처리를 통하여 이들 정보를 정리하고, 이를 다음날 온라인 시스템으로 제공하였다.²⁾

그러나, 종합 온라인 시스템에서는 계정계 시스템으로부터의 원시 데이터를 거래로그(log)로부터 디퍼드(deferred)방식으로 처리하여, 필요한 데이터를 즉시 추출해냄으로써 거의 실시간에 데이터를 정보계 데이터베이스에 저장하고, 이를 온라인 시스템으로 조회할 수 있게 되었다.(<그림 2-1>참조).

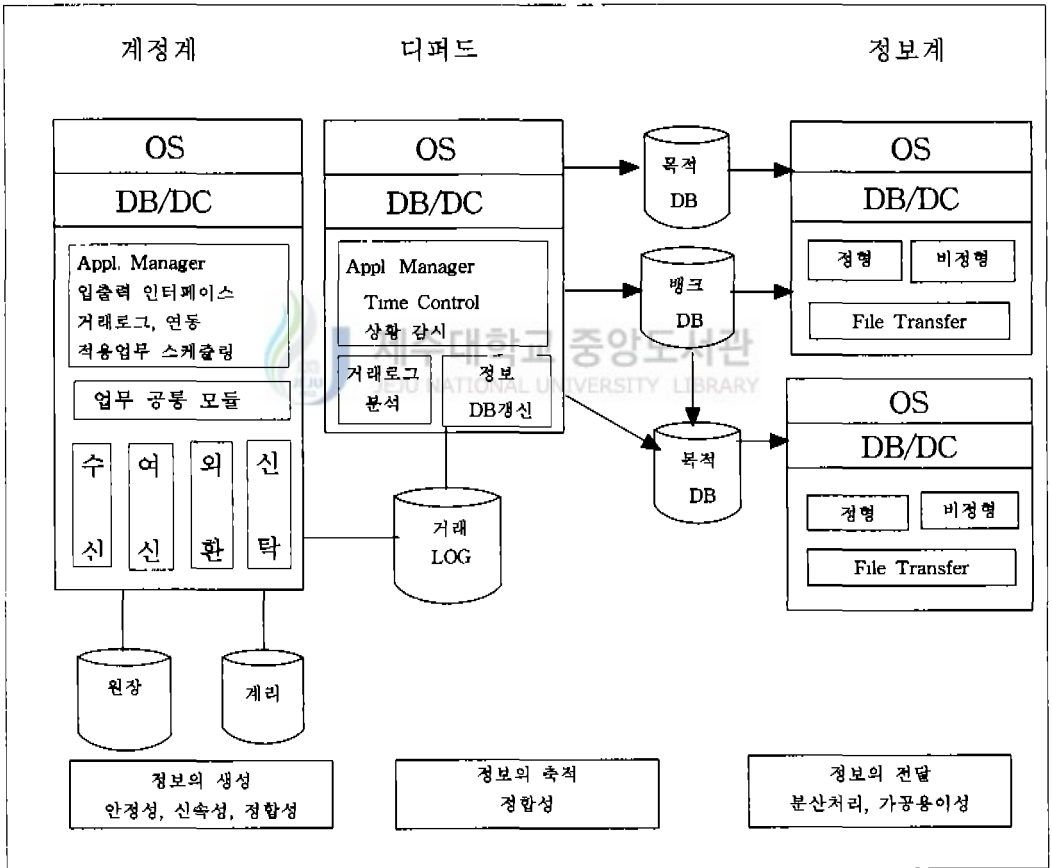
전통적으로 경영정보는 2계층 접근방식에 의해 전달되어 왔다 하위계층은 운영 응용시스템에 의해 사용되는 계층으로 실시간의 상세한 데이터를 갖고 현재의 업무 상황을 보여주고 업무를 수행하는 데 사용된다. 상위계층은 정보 응용시스템에 의해 사용되는 계층으로 하위계층으로부터 직접 복사하거나 적절한 계산을 통하여 생성된 파생(derived)데이터를 가지며, 업무 수행이 아니라 업무 관리를 위해 사용되는

2) 상계서, pp.53~54.

상세 혹은 요약수준의 데이터로 주로 읽기 전용으로 사용된다.

이렇게 분리된 방식으로 경영정보 데이터가 마련되는 데에는 두 가지 이유가 있다. 첫째로 상위계층 경영정보에는 실시간의 계정계 데이터베이스에는 없는 외부 데이터가 요구된다. 둘째로 정보계 데이터베이스 사용으로 인하여 계정계 거래처리를 위한 운영시스템의 응답시간이 허용 기준보다 늦어질 가능성이 있기 때문이다. 따라서 이러한 2계층 접근방법에 의한 데이터 추출 프로그램의 사용이 보편화되었다.

<그림 2-1> 전통적 은행정보시스템 체계



자료 : 상계서, p.54.

추출 프로그램은 처음에는 계정계 데이터베이스로부터 필요한 데이터를 뽑아 새로운 데이터베이스를 생성하였으나, 추출된 데이터베이스에도 다른 추출 프로그램이 적용되어 새로운 추출 데이터베이스를 만들게 되었고, 여기에 또 다시 다른 추출 프로그램이 적용되어 새로운 데이터베이스가 생성되게 되어 소위 자유방임적, “자생적” 정보계 데이터베이스 구조의 출현을 초래하게 되었다.

결과적으로 각 사용자부서로부터의 경영정보 요구를 개별적으로 추출 프로그램을 이용하여 해결한 결과 무수히 많은 개별적인 정보계 데이터베이스를 초래하였으며, 이는 다음과 같은 문제점을 가져오게 되었다.

- 데이터의 신뢰성 상실
- 생산성 하락
- 데이터로부터 정보로의 전환 불가능

데이터 신뢰성의 상실은 두 부서에서 서로 동일한 정보를 조회하였으나 서로 다른 정보 값을 갖게 되었을 때 발생하게 된다. 동일한 정보조회에 대해 서로 다른 두 값을 갖게 된 원인을 발견하여 이를 일치시킨다는 것은 현실적으로 거의 불가능한 일이다. 이러한 데이터 신뢰성의 상실은 다음과 같은 이유로 발생할 수 있다.

- 데이터 기준시간의 차이
- 데이터 계산방식의 차이
- 추출 수준의 상이
- 외부 데이터의 상이
- 기준 원천 데이터의 부재

예를 들어 어느 부서는 금요일 오전에 분석 데이터를 추출했으며, 다른 부서는 금요일 저녁에 추출했다면 그 분석 결과는 당연히 차이를 보일 것이다. 또한 어느 부서는 모든 계좌를 분석하기로 한 데 비해 다른 부서는 거액 계좌만 분석하였을 수도 있다. 또한 계정계 데이터로부터 분석에 이르기까지에는 여러 번의 추출과정이 있는 것이 일반적이며, 각 추출의 세부 수준의 차이에 따라서도 분석 결과가 다를

수 있다. 외부 데이터를 사용하였을 경우에도 데이터가 일단 데이터베이스로 입력되면, 그 데이터의 원천에 대한 정보는 상실되므로 서로 다른 외부 데이터를 사용하여 분석하였을 수도 있다. 마지막으로 분석 대상이 되는 데이터를 어느 데이터베이스를 시작점으로 하여 데이터를 추출해야 할지 항상 공통의 기준이 되는 원본 데이터베이스가 없다는 점이다.

데이터의 신뢰성 문제뿐만 아니라 정보계 데이터베이스로부터 새로운 정보를 검색해 내려고 할 때의 생산성도 커다란 문제이다. 예를 들어 새로운 보고서를 작성하려고 할 때, 우선 필요한 데이터가 어디에 저장되어 있는지를 알려면 수많은 파일 혹은 데이터베이스를 찾아보아야 하며 또한 각 데이터의 구조를 파악하여 이름, 정의, 계산 등 데이터의 의미를 분석해야 한다. 예를 들면, 동일한 이름이지만 서로 다른 의미로 사용되었는지 혹은 서로 다른 이름이지만 동일한 의미로 사용되었는지 등을 파악해야 한다. 데이터 위치가 파악된 이후에는 여러 곳에 흩어져 있는 데이터를 수집해야 하며 이를 위한 새로운 추출 프로그램을 작성해야 한다. 추출 프로그램 자체는 간단한 프로그램이지만 다양한 정보요구의 증가로 인하여 작성되어야 할 프로그램의 개수가 적지 않을 것이고, 원천 데이터가 있는 데이터베이스의 종류가 다양할 것이라는 점에서 보면 생각보다 간단한 작업은 아니다. 따라서 새로운 다양한 정보요구에 대하여 위와 같은 일련의 과정을 다시 거쳐야 할 것이므로 생산성 문제는 심각하다고 하겠다.

마지막으로 데이터에서 정보로 변환하는 과정이 원활하지 못하다는 점이다. 각각의 응용 시스템들이 통합을 전제로 개발되지 않았기 때문에 이들 응용 시스템들로부터 공통기준을 사용하여 정보를 수집한다는 것은 거의 불가능한 일이다. 뿐만 아니라, 의사결정 시스템에서 필요한 시계열 데이터가 충분히 저장되어 있지 않은 경우가 많다. 예를 들어, 대출의 경우에는 2년간의 데이터가, 저축예금의 경우에는 1년간의 데이터가, CD(양도성 예금증서)의 경우에는 18개월 분의 데이터가 저장되어 있다는 식이다. 또한 저장된 시계열 데이터의 시간적 간격도 일별 데이터, 주별 데이터, 혹은 월별 데이터 등으로 다를 수 있다. 이는 이들 응용 시스템이 의사결정 지원을 염두에 두고 설계된 시스템이 아니기 때문에 시계열 데이터에 대한 고려가 적었기 때문이다.

제2절 DB마케팅의 定義와 特徵

1. DB마케팅의 定義

'80년대부터 시작된 DB마케팅에 대하여 많은 학자나 실무자들이 다양한 정의를 내리고 있다. 마케팅 컨설턴트인 Holtz(1992)는 “고객에 대한 접근법, 마케팅 전략, 방법론 등의 마케팅에 대한 제요소가 단순한 구매자 리스트가 아닌 잠재고객에 대한 풍부한 정보에 근거한 마케팅”이라고 정의하고 있고³⁾, Roberts(1992)는 “DB마케팅은 컴퓨터에 수록된 개별 고객 차원의 자료에 통계적 분석과 모델링 기법 등을 적용하는 것이다. DB마케팅은 현재고객 및 잠재고객과의 직접적인 커뮤니케이션을 위한 비용 효율적인 마케팅 프로그램의 개발을 지원해 주는데 이용된다. 또한 특정 마케팅 프로그램의 결과를 평가하는 데도 이용된다. 즉, DB마케팅은 현재고객 및 잠재고객이 장기간에 걸쳐 자사의 제품이나 서비스를 반복적으로 구매하게 하기 위해서 고안된 고객과의 사전에 준비된 커뮤니케이션을 의미한다.”라고 기술하고 있다.⁴⁾ DB마케팅 컨설턴트인 Hughes(1994)는 “기존고객과 잠재고객에 대한 적합한 정보가 수록되어 있는 컴퓨터화된 관계형 데이터베이스 시스템을 고객에게 보다 질 높은 서비스를 제공하고 이들과 장기적인 관계를 구축할 수단으로 운용하는 것”으로 정의하고 있다.⁵⁾

한편, '90년대 후반에는 국내에도 DB마케팅이 소개되면서 연구기관 및 학자들이 다양한 정의를 내리고 있다. 몇몇을 살펴보면, 한국능률협회컨설팅(1997)에서는 DB마케팅을 “기업의 전략적 경쟁우위를 창출하기 위하여 필요한 고객정보를 획득하고, 획득된 정보를 이용하여 데이터베이스 및 분석시스템을 이용하여 다양한 마케팅을 구현하는 것”이라고 정의하고 있고⁶⁾, LG경제연구원 경영컨설팅센터(1997)에서는

3) H. Holtz, *Databased Marketing*, John Wiley & Sons, Inc., 1992, p.5.

4) M. L. Roberts, "Expanding the Role of the Direct Marketing Database", *Journal of Direct Marketing*, 6(2) 1992, p.51

5) A. M. Hughes, *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*, Irwin, 1994, p.9.

6) 한국능률협회컨설팅, “데이터베이스 마케팅 시스템 구축과 활용”, 한국능률협회컨설팅 CS마케팅 OBU, 1997.

“데이터베이스에 축적되어 있는 내·외부 정보를 활용하여 고객과의 지속적인 관계를 구축해 나감으로써 마케팅 효율성을 높이고 기업의 경영성과를 향상시키고자 하는 과학적 마케팅 어프로치”라고 정의하고 있으며⁷⁾, 조재희(1997)는 DB마케팅을 “산재되어 있는 내·외부의 데이터로부터 고객중심의 데이터 웨어하우스를 구축하고, 이를 효율적으로 탐색하고, 과학적(통계적)방법을 이용하여 새로운 특성을 발견하여 이를 기반으로 한 마케팅 전략과 프로그램을 수립·실행하는 일련의 과정”이라고 정의하고 있다⁸⁾.

이처럼 DB마케팅은 관점이나 강조하는 측면에 따라 다양하게 정의될 수 있지만, 이 정의들의 공통적 특징은 컴퓨터에 수록된 고객DB를 바탕으로 고객과의 장기적인 관계 구축을 위한 마케팅 전략이 수립되고 실행된다는 점에 있다.



7) LG경제연구원 경영컨설팅센터, “마케팅 혁신 기법 : 데이터베이스 마케팅”, 81회 한국오라클(주) 정기세미나 자료집, 1997

8) 조재희, “데이터웨어하우징 기술과 DB마케팅 전략”, 한국산업정보학회 추계학술대회 발표논문집, 1997.11.

2. DB마케팅의 特徵

DB마케팅은 소비자의 욕구를 파악하고 그에 부응하기 위하여 기업이 사용할 수 있는 제 수단(흔히 마케팅믹스 변수라고 일컬어지는 제품, 가격, 유통 및 촉진)을 적절히 구사함으로써 시장에서의 경쟁력을 확보한다는 근본적인 목적에 있어서는 일반마케팅과 동일하다고 할 수 있지만, 경쟁력을 확보하는 구체적인 방법에서는 일반마케팅과 상당한 차이점이 있다.

DB마케팅의 특징으로는 고객과 일대일 관계의 구축, 쌍방향 의사소통, 고객 데이터베이스의 구축을 들 수 있다.⁹⁾

1) 고객과 일대일 관계의 구축

DB마케팅은 고객과의 접촉을 통하여 고객 개개인의 욕구를 파악하고 충족시키기 위한 일대일 커뮤니케이션 활동을 그 첫 번째 특징으로 한다. 즉, 잠재고객이나 기존고객의 욕구는 고객 데이터베이스를 통해 식별될 수 있으며, 따라서 개개인에 대해 가장 적절한 전략을 일관성 있게 실시하는 것이 가능하다. 시장세분화에 있어서도 총합데이터(aggregate data)와는 달리 세분시장의 구성원이 누구인가를 구체적으로 알 수 있으며, 이에 따라 이들을 대상으로 한 개별적인 마케팅활동의 전개가 가능해진다.

이처럼 DB마케팅은 고객 개개인의 욕구를 만족시켜주는 것을 목적으로 하기 때문에 다품종 소량생산을 근간으로 하여야 한다. 소품종 대량생산 방식을 흔히 취하게 되는 일반 마케팅에서는 도매점 및 소매점을 통한 전통적인 유통경로를 이용하기 때문에 경쟁업체간에 소매점에서 보다 넓고 보다 눈에 잘 띄는 진열공간을 차지하려는 경쟁이 치열하게 전개된다 따라서 자사제품이 소매상에게 일정 수익을 보장해 줄 수 있을 정도로 팔려나가지 않으면 그 제품은 시장에서 퇴장하게 되는데, 이러한 이유로 기업에게는 많이 팔든지 혹은 생산을 중단해야 하는 선택밖에 주어지지 않는다.

9) 박찬욱, 「데이터베이스 마케팅」, 연암사, 1996.

그러나 DB마케팅에서는 전통적인 유통경로 외에 고객과의 일대일 관계에 의한 '직접판매'라는 유통경로가 추가로 제공되기 때문에 특정계층에서만 환영받는 제품도 소매점에서의 진열공간과 관계없이 시장성을 가질 수 있다.

2) 쌍방향 의사소통

DB마케팅의 두 번째 특징은 쌍방향 의사소통(two-way communication)이라고 할 수 있다. DB마케팅에서는 기업이 제공한 정보에 대해 소비자들이 반응을 보이거나 의견을 피력할 수 있으며, 이러한 과정을 통해 소비자가 기업과 직접적인 커뮤니케이션을 하게 된다.

일반마케팅에서의 의사소통은 일방적이기 때문에 기업이 소비자에게 전달한 정보가 얼마나 성공적인가를 아는 것은 대단히 어렵다. 즉, 일반적인 이미지광고에서 보듯이 대부분의 기업과 소비자와의 의사소통은 일방적(one-way communication)이어서 광고의 시장성과(매출액 등)에 대해 소비자로부터 정확한 피드백(feed back)을 받기는 불가능하다 그래서 기업은 광고의 효과를 나타내는 지표로서 실제의 시장성과 보다는 상표 인지율이나 상표 회상률과 같은 커뮤니케이션 효과의 측정에 치중할 수밖에 없다. 그런데 인지율이나 회상률에 대한 측정은 광고가 나간 후 상당한 기간이 경과한 후에 현장조사 등을 통하여 이루어지고 따라서 많은 비용이 발생함은 물론 시장상황의 변화에 대한 시의 적절한 대응이 불가능하다.

이에 반하여 DB마케팅에서는 소비자가 기업과의 커뮤니케이션 과정에 직접 참여하기 때문에 기업은 일반광고에서처럼 일정기간을 기다릴 필요 없이 즉각적으로 소비자의 반응을 알 수 있게 된다. 즉, 쌍방향 의사소통을 통해 기업은 자신의 판촉활동에 대한 기록을 축적할 수 있기 때문에 각 제품라인의 목표시장이 누구이며, 또한 어떠한 마케팅믹스 변수에 보다 민감하게 반응하는가와 같은 정보를 알아낼 수 있고 이렇게 소비자와의 커뮤니케이션 과정에서 수집된 정보는 이후의 마케팅활동에 즉각 반영될 수 있다.

고객과 기업간의 쌍방향 의사소통에는 여러 가지 수단이 사용될 수 있다 흔히 생각될 수 있는 다이렉트 메일이나 카탈로그, 전화는 물론 직접반응 광고와 같은 경우에는 신문이나 TV와 같은 대중매체가 사용될 수도 있고, 멀티미디어가 발달되고

있는 오늘날에는 PC통신이나 인터넷도 쌍방향 의사소통에 사용될 수 있다.

쌍방향 의사소통을 위한 여러 가지 수단 가운데 고객이 기업에 주문이나 문의와 같은 자신의 의사를 전달하는 매체로 가장 많이 사용되고 있는 것은 전화라고 할 수 있다. 이러한 의미에서 텔레마케팅의 중요성이 강조되고 있으며 인터넷의 확산과 더불어 인터넷마케팅 또한 중요한 수단으로 부각되고 있다.

3) 고객 데이터베이스의 구축

DB마케팅은 컴퓨터에 의한 고객 데이터베이스의 구축을 전제로 하고 있다. 기업은 컴퓨터를 이용하여 고객의 지리적, 인구 통계적, 심리적 특성이나 구매기록을 수록해 두었다가 DB마케팅 캠페인의 적합한 대상을 추출해내는 데는 물론 시장세분화를 위한 기초자료로도 활용된다. 따라서 데이터베이스는 DB마케팅의 기획에서 실행 및 효과 측정에 이르기까지 DB마케팅 모든 분야의 근간이 된다.

데이터베이스의 활용이 DB마케팅의 요체가 되는데는 컴퓨터 산업의 발전이 결정적인 기여를 하고 있으며, 이러한 컴퓨터 기술의 향상으로 인하여 시장을 구성하고 있는 소비자들에 대한 정보를 저렴한 비용으로 시의 적절하게 수집, 축적하는 것이 가능하게 되었다. 이에 따라 소비자 전체를 대중 시장으로 간주하고 마케팅전략을 시행할 필요성이 감소하였으며, 개개인의 소비자를 대상으로 하는 전략의 구사가 보다 타당성을 갖게 되었다고 할 수 있다.

제3절 DB마케팅의 成長 背景과 適用 分野

1. DB마케팅의 成長 背景

DB마케팅의 발상지인 미국에서의 성장 배경을 살펴보면, '60~'70년대에 다이렉트 마케팅을 고객별로 실시하는 진전이 있었다. 컴퓨터 하드웨어와 소프트웨어의 발달, 신용카드의 보급, 수신자 요금부담 전화의 도입, 인구 센서스에 관한 정보의 이용, 통계 및 재무모델의 발달 등 수많은 변화가 빠른 속도로 과학적인 마케팅을 발달시켰다.

하지만 이러한 기술들은 '60~'70년대에는 실험적이고 단편적으로 이용되었다. '80년대에는 이러한 기술과 정보를 공유하는데 있어 더 많은 진전이 있었으며, 이를 통해서 기업들은 보다 정교하고 통합된 마케팅 프로그램을 실시할 수 있었다. DB마케팅의 성장 배경에는 다음과 같은 기술 진보가 있었다.

1) 데이터베이스 기술

DB마케팅에 가장 기여도가 높은 혁신은 관계형DB(relational database)의 개발이라고 할 수 있다. 관계형DB는 고객정보의 접근과 조작이 용이하고, 다른 기업과 고객정보의 공유를 가능하게 하는 장점이 있다. '80년대 초에 IBM은 DB2라는 메인프레임(mainframe) 시스템용 프로그램을 발표하였으며, 다른 소프트웨어 업체에서도 비슷한 프로그램을 메인프레임용과 퍼스널컴퓨터용으로 개발하여 발표하였다.

관계형DB의 도입이 중요한 이유는 이를 이용한 시스템이 갖는 탄력성이라 할 수 있다. 관계형DB는 외부 정보를 연결, 결합하고 저장과 분석을 용이하게 하는 장점을 가지고 있다. 더욱이 관계형DB는 마케팅 담당자에게 메뉴 방식의(menu-driven) 시스템을 쉽게 제공할 수 있어서, 고객에 대한 정보를 보다 신속하고 쉽게 접근할 수 있게 한다.

2) 컴퓨터의 발달

컴퓨터 기술은 '80년대에 이르러 매우 빠른 속도로 발달되어, 마케팅 데이터의 복잡한 조작과 분석이 용이하게 되었다. '70년대 후반, 용량은 작았지만 퍼스널 컴퓨터가 처음 출시되었다. '80년대 중반에 이르러 컴퓨터 기술의 발달로 마케팅 담당자들은 이전에는 메인프레임이 필요하였던 작업들을 그들의 책상 위의 컴퓨터로 할 수 있게 되었다. 메인프레임도 또한 성능이 개선되어 수많은 고객들에 대한 상세한 정보를 조작하고 저장할 수 있게 되었다.

점점 향상되는 컴퓨터의 성능은 기업들이 고객과의 거래를 통해 발생하는 많은 자료를 처리하고, 수많은 외부 정보를 기업의 데이터 파일(data file)에 추가시킬 수 있게 하였다. 또한 기업들은 그들의 성장에 매우 중요한 고객에 대해 수집된 정보, 예를 들면 국민통계 및 라이프 스타일과 같은 정보를 다룰 수 있게 되었다.

3) DB마케팅 개념의 출현

컴퓨터와 데이터베이스 기술의 발달로 DB마케팅이라는 새로운 개념이 나타났다. 여기서 DB마케팅이란 개념은 소비자 개인의 정보를 다이렉트 메일에 적용하는데 국한하지 않고 제조업자가 내구성 소비재 등의 제품을 소비자나 다른 기업에게 판매하는데도 적용되는 개념이 되었다.

4) DB마케팅의 시작

'80년대에 처음으로 전통적인 대중 마케팅(mass marketing) 담당자들이 DB마케팅의 개념을 그들의 사업에 이용하기 시작했다. 기업들은 고객의 성명과 주소뿐만 아니라 상세한 개인 정보와 구매에 관련된 정보들을 수집하기 시작한 것이다. 항공사들이 선두에 나섰다. '70년대 후반 미국 정부의 규제가 해제되자, 항공사들은 자주 항공기를 이용하는 고객들에게 충성도(loyalty)를 형성하기 위해 할인 프로그램을 마련하기 시작하였다. 소비자들은 처음에는 무료로 항공기를 탈 수 있는 쿠폰으로 보상받았다. 그러나 이 쿠폰들이 암시장에서 거래되어 문제가 발생하자 항공사들은

자주 항공기를 이용하는 고객들의 기록과 이들의 구매행위에 관한 기록을 보관하기 시작했다 이렇게 보관된 기록을 이용하여 항공사들은 마케팅 목적에 적합한 정보를 형성하여 이용할 수 있게 된 것이다.

항공사들의 뒤를 이어 다른 여행 관련 기업, 예를 들면 호텔이나 렌터카 회사들도 자주 이용하는 고객을 위한 프로그램을 시행하기 시작했다. 다른 기업들도 이러한 추세를 받아 들었다. 장거리 전화회사들도 거래기록을 사용하여 우수고객을 선별하고, 이를 유지하기 위한 차별적 마케팅을 하기 시작했다. 자동차 제조회사와 신용카드사들도 DB마케팅을 이용하기 시작했다.

5) '90년대와 앞으로의 DB마케팅

대중 마케팅에 의존하던 많은 분야에서 기업들은 DB마케팅을 활용하여, 고객별로 차별적인 수단을 이용함으로써, 고객과의 커뮤니케이션 증가를 통한 관계를 형성하고 있으며, 이러한 추세는 앞으로도 계속될 것이다.

한편, DB마케팅의 진전에 따라 앞으로 어떠한 영향이 나타날 것인가에 대해서 많은 의견이 나타나고 있다 예를 들면, 어떤 이들은 소비자의 프라이버시 문제에 대해 관심을 가지고 있다. 이들의 주장에 의하면 소비자의 프라이버시 문제 때문에 소비자와 개별적으로 커뮤니케이션 할 수 있는 마케팅 담당자의 능력은 앞으로 현저하게 감소하게 될 것이라는 것이다.

다른 주장은 소비자의 프라이버시 문제는 소비자가 얻는 만족에 비하면 상대적으로 작기 때문에 별 문제가 되지 않거나, 프라이버시 문제 때문에 마케팅 담당자들이 DB마케팅 활동의 초점을 현재 확보된 고객에 대해서만 효과적으로 실행할 것이라는 것이다.

또 다른 주장은 DB마케팅의 성장은 기업이 개별적이고 차별적인 기준에 따라 소비자에게 커뮤니케이션 할 수 있는 새로운 방법을 개발하고, 뿐만 아니라 소비자가 기업에게 커뮤니케이션 할 수 있는 방법을 제시할 수 있을 때만 가능하다는 것이다. 이러한 불확실성과 의견의 불일치와는 관계없이, DB마케팅 영역에 영향을 미칠 많은 변화가 수년 안에 발생할 것이다

2. DB마케팅의 適用 分野

DB마케팅의 주된 목적은 기업과 고객간의 관계형성(relationship)에 있다고 할 수 있다. 또한 DB마케팅의 성공은 관계형성의 성공여부에 있다. 따라서 DB마케팅의 성공분야를 파악하는 것은 릴레이션십 마케팅의 성공분야를 살펴보는 것이라 할 수 있으며, 이는 정보활용의 측면을 살펴봄으로써 가능할 것이다.

DB마케팅에는 서로 관련된 두 개의 독립된 형태가 존재한다.¹⁰⁾ 하나는 기존고객과의 관계형성이며, 또 하나는 가장 수익성이 있는 프로파일(profile)을 개발함으로써 잠재고객을 선별하는 것이라 할 수 있다. 두 번째 형태의 마케팅은 기술적으로 첫 번째 형태보다 실행이 용이하다고 할 수 있다. 잠재고객을 대상으로 하는 DB마케팅은 항상 실험(test)을 포함한다. 이러한 실험은 신경망(neural network)을 이용하여 반응을 이끌 수 있는 요소를 결정하여 이루어진다. 이 영역에는 별다른 문제가 없으며, 미리 정해진 과정에 따라 수행하면 실패할 확률이 거의 없다.

기존고객과의 관계형성을 위한 DB마케팅에는 많은 실패 요인이 존재한다. 왜냐하면, 이러한 마케팅 활동의 성공여부를 판별하는데 많은 시간이 소요되기 때문이다.

기존고객과의 관계를 유지하거나 형성하기 위한 DB마케팅은 어느 분야에서는 성공적이고, 어느 분야에서는 성공적이지 못하다. <표 2-1>에서는 DB마케팅의 성공 및 실패분야에 대한 일반적인 구분을 볼 수 있다.

그러나 이와 같은 구분을 확정하는 것은 매우 위험한 일이다. 이 표의 구분은 단지 일반적인 분류일 뿐이다. 경우에 따라서는 이러한 구분과 다른 결과를 얻을 수도 있을 것이다

실제로 DB마케팅은 기존고객과의 지속적인 관계를 유지함으로써 반복구매를 이끌어낼 수 있는 분야나 연관 제품에 대한 교차판매가 이루어질 수 있는 분야에서 가장 활발하게 이루어진다. 따라서 DB마케팅은 은행, 항공, 호텔 같은 서비스업이나 백화점, 슈퍼마켓 같은 소매유통업에서 가장 먼저 도입되어 지금도 활발하게 적용되고 있다. 이들 분야에서는 고객과의 지속적인 거래를 통해 고객에 대한 정보를 영업 활동 과정에서 꾸준히 축적할 수 있으며 동일고객에 대해 다양한 상품을 판매할 수

10) A. M Hughes, 전게서, pp.311~320.

있다.11)

<표 2-1> DB마케팅의 성공 및 실패분야의 일반적 구분

성공 분야	실패 분야
자동차서비스, 자동차 유아용품 은행 등의 재무서비스 음료 서적, 음반, 비디오 아동용 식품 통신용 기기, 컴퓨터 신용카드, 백화점 다이어트 또는 헬스센터 다이어트 식품 약국 오락, 연예, 스포츠 이벤트회사 화원, 정원용 설비 가사용 전기, 가스 주유소 주택수리 서비스, 잡화상 병원, 의료용 제품 보험회사 변호사 잡지 음식점 여행사	의상(상표가 없는 제품) 전자제품 건축자재 카페트 유리, 거울 가구 전기제품 자물쇠 제조공 이삿짐센터 사무용품 부동산업자 장난감 스포츠용품 벽지 페인트

자료 : 김정섭, “데이터베이스 마케팅에 관한 연구”, 석사학위논문, 서울대학교 대학원, 1995, p.29.

11) 박천승, “데이터 웨어하우스 환경에서의 전략적 데이터베이스 마케팅 시스템 구축”, 석사학위논문, 고려대학교 대학원, 1997, p.85.

다음의 <표 2-2>에서 DB마케팅이 활용된 사례를 살펴보면 앞에서 언급한 DB마케팅의 성공과 실패분야의 구분이 시사하는 바가 적지 않을 것이다.

<표 2-2> DB마케팅이 활용된 사례

업종	적용 예
소매 / 마케팅	<ul style="list-style-type: none"> • 고객분류 그룹별 특성발견 • 시간별 패턴분석 • 광고 및 판촉 효과분석 • 상품계획 • 표적(target) 마케팅 • 마케팅 효과 관리 • 장바구니 분석 • 예측모델 작성 • 범주(category) 분석 • 이탈고객 분석 • 대리점 및 고객관리
은행 / 금융	<ul style="list-style-type: none"> • 교차 판매 • 예측적 라이프사이클 관리 • 자산 부채 관리 • 해약 고객 관리 • 부정행위 적발(fraud detection) • 고객 분류(segmentation) • 주식거래 규칙 발견 • 상품 수익 분석 • 표적 마케팅
통신 / 언론	<ul style="list-style-type: none"> • 통화 상세기록 분석 • 고객 충성도 분석 • 표적 마케팅 • 무선 부정행위 적발 • 구독자 분석 • 교차 판매
보험	<ul style="list-style-type: none"> • 부정행위 적발 • 상품 기획 • 위험고객의 특징 발견 • 위험도 분석
기타	<ul style="list-style-type: none"> • 고객 분류 • 보증 • 자동차 생산 • 제조업-경쟁사의 입찰액 예측 • 화학공정반응에 대한 시뮬레이션 • 항공사-우량고객 우대 프로그램 • 제조업-공정요인검출

자료 : 박천승, 전개논문, p.86.

제4절 銀行의 DB마케팅 導入 背景과 考慮 事項

앞에서 DB마케팅의 적용 분야에 대하여 살펴 볼 때 언급되었지만, 은행은 고객 DB를 이미 보유하고 있다든지, 혹은 고객과의 일대일 관계형성의 필요성이 크든지 하는 점에서 보면, DB마케팅을 가장 먼저 도입하였을 법하지만 실제로는 미국의 경우에 있어서도 은행권은 매우 보수적인 자세로 새로운 마케팅 기법을 받아들이는데 있어서는 선두에 서지 않았다. 그러나 '80년대에 미국의 금융자유화는 금융기관 간의 경쟁을 격화시키고 이러한 과정에서 은행들은 경쟁력을 강화시키기 위한 수단으로 DB마케팅을 받아들이기 시작하였다.

1. 國內銀行의 DB마케팅 導入 背景

국내은행들이 고객 지향적인 마케팅 개념을 도입한 것은 '80년대 중반의 일이며, 더욱이 DB마케팅에 대한 관심이 생기기 시작한 것은 매우 최근의 현상이라고 할 수 있다. 따라서 국내은행들의 DB마케팅 도입은 '90년대 중반에 들어서 시작되었다고 할 수 있고, 서구 국가에 비해 약 10여년 뒤져 있다고 할 수 있다. 은행에서의 DB마케팅의 기술적 의미는 은행의 내부 및 외부 정보를 고객별로 데이터베이스에 통합, 정리하여 은행의 영업에 활용하는 것을 뜻한다. 그러나 현재 은행들이 DB마케팅을 필요로 하게 되기까지는 이러한 단순한 DB마케팅의 의미보다 다양한 배경이 존재한다.¹²⁾

1) 내실 위주의 은행 경영

금융사상 초유의 은행 퇴출로 이어진 IMF 체제하의 금융환경 변화에 따른 은행의 전략적인 변화가 DB마케팅의 필요성을 증대시키고 있다. 규제완화 및 금융개방과 함께 은행간은 물론 비은행 금융기관과의 경쟁이 심화되고 있고, 경기침체에 따른 은행들의 부실여신이 증가하면서 외형위주의 성장보다는 수익성 등 내실을 기하

12) 함유근, "DB마케팅의 의의와 은행의 도입", 금융 11월, 1997, pp.12~14.

는 은행 경영이 절실하게 되었으며 DB마케팅은 이러한 변화를 실현시키는 수단으로 주목받고 있다

DB마케팅은 은행의 영업활동이 기존의 상품 중심에서 고객 중심으로 변화하고 있음을 의미한다. 고객 중심 영업의 핵심은 고객과 은행간에 일대일의 밀착된 관계를 형성하여 이를 바탕으로 고객의 욕구 변화에 적극적으로 대처하는 소위 고객관계 관리(CRM, Customer Relationship Management)의 도입이다. 이는 은행의 영업활동이 불특정 다수를 대상으로 하는 대중마케팅에서 고객 개개인에 대한 개별적인 관계 형성을 통해 상품 및 서비스를 차별화하고, 고객의 충성도를 확보하여 고객 유지와 수익성 확보를 동시에 추구하는 방향으로 전환되고 있음을 의미한다.

2) 정보기술의 발달과 고객 수익성의 측정

DB마케팅의 중요성이 높아지고 있는 또 다른 배경 중의 하나는 발달된 정보기술을 활용하여 은행들이 고객별로 수익성을 측정하는 것이 가능하게 되었다는 점이다. 컴퓨터의 대중화로 DB마케팅의 필수요소인 고객정보 수집에 소요되는 비용이 대폭 하락하였으며 이를 분석하고 유지하는 데 소요되는 추가비용도 줄어들게 되었다. 또한 전화, TV, 멀티미디어, 인터넷 등과 관련된 정보기술의 발전으로 고객과 은행간의 직접 커뮤니케이션이 보다 저렴한 비용으로 실현될 수 있게 되어, 고객과의 정보 교환을 통한 마케팅 기회가 증가하게 되었다.

은행들은 거래처리시스템을 구축하여 기본적인 회계단위인 계좌(account)에 대한 데이터를 정보화하여 수집, 관리할 수 있게 되었으며, 더 나아가 고객정보파일 등이 구축되면서 각 개인에 대한 정보가 체계적으로 관리될 수 있게 되었고, 이러한 거래처리시스템으로부터 데이터를 고객별로 취합하여 고객DB가 구축될 수 있었다. 그러나 아직 대부분의 우리나라 은행들은 DB마케팅을 위한 고객DB로까지는 완전하게 구축하지는 못하고 있는 실정이다.

3) 고객 욕구의 가변성과 다양성 증가

고객의 욕구가 고정적이지 않고 시간과 상황에 따라 변화하는 정도가 심해지고

있으며 고객들의 금융서비스 수요도 다양해지고 있다는 사실이 DB마케팅의 필요성을 높이고 있다. 금리에 대한 민감도가 높아서 금융기관간 혹은 상품간의 금리차이가 발견되면 즉시 이를 감지하고 금리가 조금이라도 자신에게 유리한 곳을 찾는 고객이 있는 반면에, 금리보다는 서비스와 편의성을 추구할 만큼 경제적으로 여유가 있는 고객 계층도 증가하고 있다. 이들은 자신이 원하는 서비스를 받기 위해서 기꺼이 추가로 비용을 부담할 준비가 되어 있다. 고객의 욕구가 예전보다 더 세분화되고 다양화되면서 개별적으로 특별하게 대우받기를 원하는 고객들이 늘어나고 있는 것이다.

이러한 고객의 변화에 효과적으로 대처하기 위해 은행으로서는 고객의 현재 수익성뿐만 아니라 평생 동안의 수익성(lifetime profit)도 관리해야 하는 필요성이 높아지고 있다. 그리고 고객에 대한 이해를 높여 그들의 수요를 미리 예측할 필요성이 높아졌다. 이는 은행들이 이제는 자신의 고객이 어떤 특성을 가지고 있는지를 파악한 상태에서 영업을 해야 됨을 의미한다. 이와 같이 현재의 고객 중 과연 어떤 유형의 고객이 미래에 높은 수익을 은행에 가져다주는가를 예측하는 것도 DB마케팅의 중요한 기능이 되고 있다.

은행은 기존고객으로부터 창출될 수 있는 미래의 수익성을 고려한 영업전략을 필요로 하며, 이를 달성하기 위해서는 고객의 금융상품 이용추이는 물론 각종 채널의 이용도에 대한 데이터를 축적하고, 이를 바탕으로 고객의 행태를 분석하여 적절한 시기에 적절한 채널을 통해 고객이 필요로 하는 금융수요를 충족시킬 수 있어야 된다. 예를 들어 인척이나 친구들을 은행에 소개할 가능성이 높은 고객은 미래의 수익성이 높은 고객으로 분류하여 마케팅을 해야 할 것이다.

고객수요의 다양성은 고객이 선호하는 판촉 수단과 서비스 전달 채널에도 나타나고 있다. 은행은 일반적으로 새로운 상품을 다양한 수단을 통해 고객에게 소개한다. 고객의 집이나 직장으로 우편물을 발송하여 신상품에 관한 소개를 할 수도 있고 점포에 오는 고객들에게 텔러들이 직접 상품 판촉을 할 수도 있다. 그리고 전화를 이용하여 은행의 상담원이 고객에게 소개할 수도 있다. 물론 신문이나 잡지의 광고를 통한 방법도 있다.

이와 같이 다양한 판촉 형태에 고객이 어떻게 반응하는가에 대한 정보를 축적하여 고객의 선호도에 따라 교차판매(cross-selling)나 상향판매(up-selling)의 기회를

포착할 수 있다. 예를 들어, 판촉 형태에 따른 고객의 반응 데이터가 축적될 경우 각종 우편물에 대한 반응이 없는 고객에 대하여는 전화나 인적 접촉과 같은 보다 적극적인 방법으로 판촉을 행할 필요가 있을 것이다. 또한 전자금융의 발달로 은행의 서비스 전달채널이 PC, ATM, 전화, 무인점포 및 유인점포 등으로 다양화되고 있는데, 축적된 고객과의 거래 데이터를 활용하여 어떤 채널이 어떤 고객과 상품유형에 효과적인가를 파악할 수 있다

4) 기존고객의 유지 필요성

DB마케팅은 새로운 고객을 대상으로 하는 영업보다는 기존고객에 대한 영업을 극대화하기 위하여 개발되었다. 즉 새로운 고객을 발굴하여 이들을 대상으로 영업하는 것보다 기존고객에 대한 영업을 강화하는 것이 은행으로서는 더욱 용이할 뿐만 아니라 수익성도 높다는 사실이 DB마케팅을 촉진시키고 있다. DB마케팅은 고객에 대한 데이터를 지속적으로 축적함으로써 고객과 은행간의 일대일 관계를 발전시켜 궁극적으로는 고객당 수익성을 제고시키는 수단이 된다. 예를 들어 영국의 금융산업 전문 DB마케팅 회사인 하티 행크스(Harte Hanks)는 고객과의 관계지속을 위해 DB마케팅을 다음과 같은 여섯 가지 분야에 사용하고 있다.

- ① 고객을 분류하여 계좌를 폐지할 것으로 예상되는 고객들을 미리 파악한다.
- ② 고객가치를 계산하여 고객 중 최우선적인 영업 대상을 선정한다.
- ③ 벤치마킹에 의한 내부 경쟁을 유도한다.
- ④ 고객의 원리금 상환 패턴을 예측하여 체납을 방지한다.
- ⑤ 체납 예상 순위별로 고객의 명단을 작성한다.
- ⑥ 고객과의 지속적인 관계유지를 위한 적절한 상품의 판촉기회를 포착한다.

특히 기존의 고객들에 대해서는 은행과 거래하는 가운데 다양하고 풍부한 각종 정보들이 수집될 수 있기 때문에 은행으로서는 이러한 정보의 활용을 극대화하면 효과적인 영업활동을 할 수 있다. 더욱이 은행은 비은행 금융기관에 비해 많은 점포를 보유하고 있어 고객정보의 수집에서 상대적으로 유리하다는 점도 DB마케팅 도입의 필요성을 높이고 있다.

2. DB마케팅 導入의 先決 課題

우리 나라 은행들의 DB마케팅은 이제 시작 단계에 있다고 할 수 있다 최근 들어 각 은행에서 경쟁력을 강화하기 위한 수단으로서 고객 관계 관리, 체계적 고객관리, 데이터 웨어하우징, 데이터베이스의 통합, 텔레마케팅이나 직접 우편(Direct Mail; 이하 DM)을 이용한 직접 유통수단의 활용, 우량고객 우대 프로그램 등의 주제를 가지고 많은 논의가 이루어지고 있지만, 기본적으로 우리 나라 은행들은 고객의 욕구에 기반을 둔 기업문화가 정착되어 있지 않기 때문에 DB마케팅의 본격적인 도입과 성공적인 수행에는 앞으로 많은 시간이 필요할 것으로 예상된다. DB마케팅의 도입을 앞두고 먼저 해결해야 할 과제는 다음과 같다.

1) 마케팅 계획 수립과 조직 정비

DB마케팅의 도입은 은행이 실시하는 각종 마케팅 캠페인이 단발성의 이벤트로 끝나지 않고 그 결과를 차기의 계획 수립에 반영시킴으로써 점진적인 노하우의 축적과 보다 발전된 형태의 마케팅 프로그램의 개발을 가능케 해준다. 그 동안 우리나라 은행의 마케팅 활동은 그 규모 면에서 그다지 활발하지 못했을 뿐만 아니라 경쟁력 강화를 위한 전략적 의미를 크게 부여하지 못했던 것이 사실이다

따라서 마케팅 활동의 결과에 따른 책임소재의 문제나 성패원인의 분석을 통한 발전 지향적인 마케팅 계획 수립의 필요성이 크다 하겠다¹³⁾

또한, DB마케팅이 제대로 실현되기 위해서는 조직의 모든 면이 고객 중심으로 정비되어야 한다. 예를 들어 DB마케팅을 위해서는 고객에 관한 모든 정보를 집중시킨 고객DB가 필수적인데 고객 중심의 문화가 정착되지 못한 채 이러한 시스템이 은행에 도입될 경우 고객데이터의 활용이나 관리에 대하여 부서간의 혼란과 이해 대립이 초래될 가능성이 있다. 즉, 정보기술 부서에서는 시스템을 안정적이고 효율적으로 운영하는 입장에서 고객DB를 관리하기를 원할 것이며 고객지원부서 또는 마케팅 부서에서는 고객DB의 효율성보다는 이를 통해 얼마나 유용한 정보를 얻을 수

13) 박찬욱, "고객정보를 활용한 은행 데이터베이스 마케팅 전략에 관한 연구", 한국 금융연구원, 1998, p.14.

있느냐가 주요 관심 사항일 것이다. 그리고 종합기획부 등은 은행의 수익관리 측면에서 고객DB를 활용하기를 바랄 것이다

이와 같이 은행의 각 부서가 고객 중심이 아니라 부서의 편의성을 중심으로 움직인다면 효과적인 DB마케팅을 실현하기 어렵다. 즉 DB마케팅은 단순히 고객과 마케팅부서간의 문제가 아니라 은행 전체가 개별 고객과의 관계를 일대일 관계로 발전시키는 작업이다. 다시 말해 DB마케팅이 영업에 국한된 기능이 되어서는 아니 되며 고객 중심의 은행경영을 실현하는 수단으로 인식되고 활용되어야 한다.

2) DB마케팅 도입을 위한 준비

은행에서 효과적인 DB마케팅을 실현하기 위해서는 은행 내부의 고객정보에 대한 접근성(accessibility)과 활용성(usability)을 높여야 한다. 은행 내 누구든지 고객DB를 직접 접근하거나 혹은 정보기술 부서에 요청하여 원하는 고객정보를 손쉽게 신속하게 얻을 수 있어야 한다 특히 영업점에서 고객정보를 얼마나 용이하게 접근하여 활용할 수 있느냐에 따라 DB마케팅의 실효성이 나타난다고 할 수 있다. 궁극적으로는 모든 영업점에서 직접 온라인으로 고객정보를 입력하고 고객별로 즉시 데이터를 검색할 수 있는 고객정보시스템(CIS : Customer Information System)이 구축되어야 할 것이다.

그리고 DB마케팅에 대한 활용도를 높이기 위해서는 데이터베이스에 대한 투자 이외에 데이터의 수집과 통계적인 모델링(statistical modeling) 기술의 개발과 이에 대한 투자가 요구된다. 금융업 중에서 DB마케팅 분야에서 가장 앞선 외국 신용카드 업체의 경험에 따르면 전통적인 신용점수화(credit scoring)기법 이외에 각종 예측기법을 이용하여 카드 이용대금 체납률을 낮추는 한편 고객당 수익성을 높이는데 DB마케팅이 효과를 발휘하는 것으로 알려지고 있다

또한, DB마케팅이 효과적으로 활용되기 위해서는 이에 대한 은행 내부의 신뢰를 얻는 것이 매우 중요하다. DB마케팅은 단순히 마케팅 부서만의 노력으로 이루어지는 것이 아니라 이를 위해선 상품개발팀은 물론 수신과 여신 부서와의 의견 조율과 협력이 절대적이다. 이러한 부서들이 DB마케팅의 신뢰성이나 유용성에 대하여 회의적이라면 DB마케팅은 성공을 거둘 수 없다. 앞서서도 언급하였지만 DB마케팅을 성

공시하기 위한 제일 중요한 요인 중 하나는 바로 고객을 중심으로 사고하고 행동하는 조직 문화를 은행 내에 정착시키는 것이다. 그래야만 부서마다 동일한 목표에 따른 일관된 영업활동을 펼 수 있다. 경영층의 관심이 수신 중대에만 있고 수신 부서나 점포의 책임자도 여전히 수신에만 집착한다면 기존고객과의 관계 발전을 중시하며 고객당 수익성을 극대화하는 데 초점을 둔 DB마케팅은 성공을 거두기 어렵다¹⁴⁾



14) 정철용 · 함유근, 전게서, pp 98~99.

3. 銀行 DB마케팅의 成功 要件

은행의 DB마케팅이 성공하기 위해서는 극복해야 할 많은 장애물들을 사전에 파악하고 이에 적극적으로 대처해야 한다. 먼저 DB마케팅 담당자들은 DB마케팅의 제 요소들을 정확히 이해하고 있어야 한다. DB마케팅은 데이터베이스에 축적된 고객 관련 데이터를 활용하여 마케팅의 효율성을 높이는 비교적 단순한 개념에서 출발한 마케팅 기법이다. 그러나 DB마케팅은 마케팅 활동, 통계적 분석기법, 정보 시스템 등의 여러 부문에 뿌리를 둔 복합적인 시스템이며, DB마케팅의 성공을 위해서는 이 요소들 각각의 역할과 문제점을 명확히 인식하여야 한다.¹⁵⁾

DB마케팅의 구성요소는 DB마케팅의 구현 과정에 따라 데이터, 분석, 전략, 실행으로 구분할 수 있으며 각 요소는 나름대로의 세부 이슈와 구현 도구(tool)를 가지고 있다. DB마케팅의 구성요소별 세부 이슈와 관련 도구들은 <표 2-3>과 같다. DB마케팅의 각 이슈들은 은행이 추구하는 목적에 따라서 반드시 필요한 것도 있고, 당분간은 필요가 없는 것도 있다. 은행에서는 우선 자신이 도달하고자 하는 목표를 분명히 하고, 필요한 구성 요소별로 당면한 문제들을 파악하여 해결 방안을 찾아야 한다. 구체적으로 어떤 업무를 어떤 방식으로 지원하고, 어떤 성과를 얻겠다는 그림에서 출발해야 하는 DB마케팅 계획은 두 가지 상반된 입장이 있다.

그 첫 번째는 전체적인 전략의 방향과 DB마케팅의 주요 부문별로 목표를 설정하고, 구체적인 세부 실행계획을 사전에 명확히 수립하는 것이다. 두 번째는 전체적인 전략의 방향만을 명확히 하고, 구체적인 실행계획은 DB마케팅을 추진해 나가면서 사안별로 결정하는 방법이다. 첫 번째 방법의 경우는 미리 정해진 계획에 따라 추진됨에 따라 一絲不亂하게 진행되어 진행 도중의 혼란과 비용을 줄일 수 있다는 장점이 있다. 그러나 이 방법은 사전에 명확한 계획을 수립하기가 어렵다는 현실적인 한계와 함께, 초기의 계획이 잘못 수립되었을 경우 DB마케팅 전체의 구조를 완전히 새로 만들어야 한다는 위험이 따른다. 두 번째 방법은 DB마케팅에 관한 경험이 축적되지 않은 상황에선 실패의 위험이 적고 유연함을 가질 수 있지만, 구축기간이 길어지고 많은 시행착오를 겪을 위험이 있다.

15) 권오준, “데이터베이스 마케팅의 성공 포인트”, LG주간경제 제469호, 1998,
<http://www.lgeri.com/>

<표 2-3> DB마케팅의 구성 요소 및 세부 이슈

구성 요소	세부 이슈	관련 도구
데이터	데이터의 분석 목적 재정의 분산된 데이터의 통합 유용한 데이터의 추가	Data Transformer Data Warehouse Datamart
분 석	단순질의/리스트 추출 데이터간 비교/교차 분석 연관관계/인과관계 파악, 예측, 통계 분석	Query MOLAP Mining/Statcal Tool
전 략	고객 맞춤 전략 지속적인 고객 애호도 제고 전략	1:1 마케팅/대량고객화 프로그램 고객 관계 관리/로얄티 프로그램
마케팅활동	고객 정보에 기초한 효과적인 프로모션 고객접점의 가치증진 방안 및 지속적인 데이터 수집 방안	DM/TM 프로그램 Call Center/Web 운영

자료 : 상계자료, p.1.

은행의 입장에서는 이 두 가지 방법을 절충하여 전체적인 전략의 방향을 명확히 하고 각 구성 요소별로 목표를 분명히 하고, 각 구성 요소별 세부 계획은 실행 과정에서 제기되는 문제에 따라 결정하고 보완하는 방법이 가장 적절하다고 볼 수 있다.

DB마케팅 도입 과정에서 나타나는 문제점들을 극복하고 성공적인 DB마케팅을 구축하기 위해서는 다음의 사항들을 고려하여야 한다.¹⁶⁾

1) 선진 사례의 주체적 수용

대부분의 DB마케팅은 선진 사례에서 자극을 받아 시작하게 되며 그들의 경험에서 무언가 쓸모 있는 것을 얻고자 한다. 그러나 실제로 선진 사례를 우리의 현실에 바로 적용할 수는 없다. 그 이유는 벤치마킹의 대상인 미국이나 유럽의 금융 관행이 우리와는 많이 다르기 때문이다. 외국의 경우 고객 및 영업 관련 데이터가 매우 정확하고, 유용한 데이터를 수집할 수 있는 경로 또한 다양하다 그리고 계량적인 분

16) 상계자료, pp 3~6.

석을 경영에 활용하는데 익숙하며 데이터의 표준화도 잘 되어 있다. 이밖에 업무관행이나 조직 문화도 우리 은행과는 다른 점이 많다. DB마케팅을 도입하려는 우리 은행은 선진 사례에서 방법론이나 분석의 유용한 틀만을 습득하고, 세부적인 데이터 정의나 분석 결과의 활용은 우리 실정에 맞게 수행하는 자세가 필요하다

2) 창의적이고 지적인 사용자 육성

DB마케팅은 구축 단계에서부터 활용 및 유지 단계에 이르기까지 사용자들의 지적능력과 창의성에 영향을 크게 받는다. DB마케팅 구축 초기의 작업인 社的인 데이터의 재정의와 분석 기법의 개발을 비롯해서 실제 활용단계에서 문제를 정의하고, 분석 방법을 선택하며, 결과를 해석하는 일련의 과정은 통계적인 분석방법과 과학적인 마케팅에 대한 이해를 필요로 한다. 이러한 작업은 모범 답안이 있을 수 없으며, 거의 사용자의 지적능력과 창조적인 문제해결 능력에 따라 성과가 크게 좌우된다. 이런 점에서 DB마케팅은 사용자들에게 까다로운 도구로 받아들여질 수 있다. 이 문제를 해결하기 위해서는 필요한 분석 기법들을 사용자가 쉽게 사용할 수 있도록 사전에 시스템화 할 수도 있다.

그러나 다차원 분석이나 데이터마이닝과 같은 심도 있는 분석을 하기 위해서는 사용자 자신의 역량을 키우는 것 이외에 다른 방법은 없다. 현재의 사용자 수준에서 DB마케팅이 요구하는 다양한 분석 능력을 단기간에 갖추는 것은 쉽지 않은 일이지만, DB마케팅의 성공을 위해서는 지속적인 교육 프로그램을 통해 사용자의 분석 능력과 창조적인 활용 능력을 향상시켜야 한다 특히 데이터마이닝과 같은 고도의 분석기법을 활용하기 위해서는 통계 분석과 전산 지식, 마케팅 마인드를 두루 갖춘 전문인력을 은행 내부에서 양성하는 것이 바람직하다.

3) 기술과 기법에의 집착 제거

최신 기술이나 어려운 기법들의 도입이 DB마케팅의 목표는 아니다. 최신 기술과 기법을 위주로 DB마케팅을 추진하기보다는, 은행의 전략적인 목표에 부합하며 실질적인 성과를 낼 수 있는 단순하고 명확한 DB마케팅이 더 가치가 있다. 단순하고

명확한 기술과 기법은 구현하기 쉽고, 사용자들이 쉽게 적응하고 활용할 수 있다. 기술이나 분석 기법에 지나치게 집착할 경우, 시스템화되기도 어렵지만 현실적이고 유용한 아이디어가 제외되는 우를 범할 수 있다.

4) 외부 전문가와 내부 인력의 조화

DB마케팅과 관련된 다양한 이슈들에 대응하기 위해 필요한 지식들을 은행 내부 인력들이 단기간에 습득하기는 어렵다. 따라서 실제 DB마케팅을 구현한 경험이 많은 외부 전문가가 도움이 될 수 있다. 하지만 외부 컨설팅을 받더라도 은행 내부 인력들의 능동적인 역할은 DB마케팅의 성공을 위해 필수적이다. 현업 부서의 고유한 정서나 업무와 관련된 요구 사항들이 DB마케팅에 반영되지 않았을 경우, 실제로 DB마케팅 프로그램을 실행하게 될 사용자들이 거부감을 가질 수 있기 때문이다. 이렇게 되면 아무리 좋은 시스템과 기법을 도입하더라도 성과를 기대하기 힘들다.

5) 이상과 현실의 조화

새로운 경영 혁신 기법이 자주 실패하는 이유 가운데 하나는 지나치게 이론적인 모형에 집착하여 현실을 도외시하거나, 추진 도중에 현실을 지나치게 많이 수용하여 결국 자동화된 도구의 개발로 그치기 때문이다. DB마케팅도 예외는 아니다. 현재의 업무 요구와 제약 조건을 따르다 보면 기존 비즈니스의 잘못된 관행을 그대로 답습하게 되거나, 또는 장래의 비즈니스 환경에 맞지 않아 다시 DB마케팅을 추진하게 될 위험이 있다. 반대로 현재 상태를 무시하고 이상적인 DB마케팅만을 추구하다 보면, 사용자들이 수용하지 못하거나 생각지도 못했던 여러 가지 결함이 연이어 발견되는 경우도 있다. 이 때문에 전체적인 전략의 방향을 명확히 하고 현재의 업무 방식과 장래 비즈니스 환경을 함께 고려하는 유연한 자세가 필요하다.

6) 고객이 느낄 수 있는 가치의 제공

DB마케팅은 DB마케팅 자체만으로 성공할 수는 없다. 많은 DB마케팅 프로젝트들

의 경우 그 추진 방향이 지나치게 기법과 도구에 치중하는 경향이 있다. DB마케팅의 궁극적 목적은 고객의 만족이며, 만족한 고객과 우호적 관계를 지속적으로 유지하는데 있다. 이를 위해서는 단순히 분석을 과학화한다거나 단편적인 고객 맞춤 프로모션 활동에 그쳐서는 안된다. 고객의 입장에서 중요한 것은 은행이 제공하는 상품이나 서비스로부터 가치를 느낄 수 있어야 한다는 것이다. 성공한 DB마케팅의 결과는 단지 은행이 고객에 대해 많은 지식과 통찰을 얻는데서 그치는 것이 아니라, 그러한 지식과 통찰을 이용하여 고객에게 전달되는 상품과 서비스의 본원적인 가치를 향상시키는데 있다 아무리 고객의 성향을 잘 파악했다 하더라도 그러한 성향을 만족시킬 상품이나 서비스를 갖고 있지 못하다면, DB마케팅은 아무 소용이 없음을 명심해야 한다.

7) 지속적인 보완 발전

DB마케팅을 추진하는 많은 은행들이 DB마케팅을 단기간에 완성시키는 데만 급급하고 지속적으로 보완하고 발전시키는 내부에는 관심이 적다. 이것은 매우 위험한 태도이다. DB마케팅은 지속적으로 조직 안에 관련 지식과 활용 경험이 쌓여서 고도화될 때 비로소 경쟁은행이 따라올 수 없는 경쟁우위 요소가 된다. DB마케팅이 고도화된다는 것은 두 가지 의미가 있다. 하나는 시스템 측면에서 데이터가 충실해지고, 보다 의미 있게 재정의되며, 유용한 분석기법이 지속적으로 개발되어 시스템에 반영되는 것이다. 다른 하나는 활용 노하우(know-how)의 축적이다. 은행은 고객에 관한 데이터의 추출과 활용을 반복하면서 고객을 더욱 잘 이해하게 되며, 동시에 고객을 더욱 잘 만족시킬 수 있는 다양한 아이디어를 얻게 된다. 이 같은 DB마케팅의 고도화를 위해서는 지속적인 피드백(feed back) 과정이 제도화되어야 한다.

또한 영업전략의 변화와 정보·기술의 발전에 따라 분석기법도 지속적으로 보완되어야 한다 예를 들면 현재는 단순한 쿼리(query) 분석을 통해 고객에 관한 지식을 획득하는 수준에 머무르고 있으면, 보다 심도 있는 분석을 하기 위해 다차원 분석이나 마이닝 기법의 도입을 고려해야 한다 정보수집 경로도 보다 다양화시켜 웹(web)이나 콜센터(call center) 등의 새로운 채널을 통한 정보 수집과 고객에 대한 정보 서비스 제공을 고려해야 한다.

제 3 장 데이터 웨어하우스

제1절 데이터 웨어하우스의 定義 및 特性

데이터 웨어하우스는 '80년대 중반 처음 도입된 개념으로, 이후 이 개념은 많은 H/W, S/W 및 도구(tool) 공급업체들에 의해 발전되었으며, '80년대 후반 Inmon이 데이터 접근 전략으로 데이터 웨어하우스 개념을 사용함으로써 많은 관심과 집중을 받게 되었다.

데이터 웨어하우스에 대하여 가장 대표적인 정의로서 Inmon(1994)은 “기업의 의사결정 과정을 지원하기 위한 주제 중심적이고 통합적이며 시간성을 가지는 비휘발성 자료의 집합”으로 정의하였다.¹⁷⁾ 또한 Kelly(1994)는 全社的 데이터 웨어하우스를 “기업 내의 의사결정 지원 어플리케이션들을 위한 정보 기반을 제공하는 하나의 통합된 데이터 저장공간”으로 정의하였다.¹⁸⁾ 한편 Poe(1994)는 데이터 웨어하우스를 운영 시스템과 연계하여 “의사결정 지원에 효과적으로 사용될 수 있도록 다양한 운영 시스템으로부터 추출·변환·통합되고 요약된 읽기 전용 데이터베이스”로 정의하였고¹⁹⁾, Widom(1995)은 “데이터 웨어하우스는 분산 환경에서 다양한 정보 소스들이 가진 데이터를 빠른 질의와 분석을 목적으로 통합하여 저장한 것이다. 데이터 웨어하우스 환경에서의 정보 소스들은 자율적, 이질적인 특성을 가지고 있다.”라고 정의하였다.²⁰⁾

17) W. H. Inmon and R. D. Hackathorn, *Using the Data Warehouse*, John Wiley & Sons, 1994.

18) S. Kelly, *Data Warehousing: The Route to Mass Customization*, John Wiley & Sons, 1994.

19) V. Poe, "Guidelines for Warehouse Development", *Database Programming & Design*, September 1994.

20) J. Widom, "Research Problems in Data Warehousing.", *Proceeding of the 4th Int'l Conference on Information and Knowledge Management (CIKM)*, November 1995.

이상의 정의들에서 보면, 데이터 웨어하우스는 의사결정에 필요한 정보를 효율적으로 제공하기 위하여 조직 내·외부의 데이터들을 통합하여 저장한 데이터베이스라고 볼 수 있으며, 다음과 같은 몇 가지 특성을 가지고 있다²¹⁾

① 웨어하우스의 데이터는 사용자들의 의사결정 지원에 전적으로 이용된다. 즉, 웨어하우스가 존재하는 가장 일차적인 이유는 사용자의 의사결정을 지원하기 위한 것이다. 데이터 웨어하우스는 올바른 정보를 올바른 형태로 적시에 제공하기 위해, 대량의 원시 데이터를 유용한 정보로 변환하는 엔진이다.

② 기업의 운영 시스템과 분리되며, 운영 시스템으로부터 많은 데이터가 공급된다. 데이터 웨어하우스에는 여러 개의 개별적인 운영 시스템으로부터 데이터가 집중된다. 또한 데이터 웨어하우스의 기본적인 자료 구조는 운영 시스템의 그것들과 완전히 다르므로 데이터들이 데이터 웨어하우스로 이동되면서 재 구조화되어야 한다.

③ 데이터 웨어하우스는 전사적 모델에 기초하여 통합된다. 기존 운영 시스템의 대부분은 항상 많은 부분이 중복됨으로써 하나의 사실에 대해 다수의 버전이 존재하게 된다. 하나의 객체를 지칭하는 다양한 이름이 존재하거나 데이터가 가지는 의미가 서로 다르다. 데이터 웨어하우스에서 이러한 데이터는 全社的인 관점에서 통합된다.

④ 시간성 혹은 역사성을 가진다. 즉 일, 월, 년, 회계기간 등과 같은 정의된 기간과 관련되어 저장된다. 운영 시스템의 데이터는 사용자가 사용하는 매순간 정확한 값을 가진다. 즉 바로 지금의 데이터를 정확하게 가지고 있을 것이 요구된다. 반면 웨어하우스의 데이터는 특정 시점을 기준으로 정확하다.

⑤ 주제 중심적이다. 데이터 웨어하우스는 전통적인 데이터베이스와 근본적으로 구분된다. 전통적으로 데이터베이스는 대부분 어플리케이션의 일부분이었다. 이러한 환경에서 각 테이블과 필드는 어플리케이션에서 미리 정의된 화면상의 필드들과 논리적으로 밀접한 관계를 가진다. 운영 시스템은 재고관리, 영업관리 등과 같은 기업 운영에 필요한 특화된 기능을 지원하는 데 반해, 데이터 웨어하우스는 고객, 제품 등과 같은 주요한 주제를 중심으로 그 주제와 관련된 데이터들로 조직된다.

21) 조재희·박성진, 「데이터 웨어하우스정과 OLAP」, 대청정보시스템, 1996, pp.44~48.

⑥ 컴퓨터 시스템 혹은 자료 구조에 대한 지식이 없는 사용자들이 쉽게 접근할 수 있어야 한다. 조직의 관리자들과 분석가들은 그들의 PC로부터 데이터 웨어하우스에 연결될 수 있어야 한다. 이런 연결은 요구에 즉각적이어야 하고, 또한 신속성을 보여야 한다. 만약 접근을 다른 사람을 통해서 한다면, 느리거나 사용하기 어렵고 신뢰할 수 없다면 받아들여질 수 없다.

⑦ 데이터 웨어하우스는 읽기 전용 데이터베이스로서 갱신이 이루어지지 않는다. 운영 시스템 환경에서는 추가·삭제·변경과 같은 갱신 작업이 레코드 단위로 지속적으로 발생한다. 반면 웨어하우스 환경에서는 데이터 로드(load)와 활용만이 존재하며, 운영 시스템에서와 같은 의미의 데이터의 갱신은 발생하지 않는다.



제2절 데이터 웨어하우스의 登場 背景

조직의 정보시스템을 구성하는 근간은 거래처리 시스템이다. 은행의 온라인 시스템은 전형적인 거래처리 시스템으로써 대량의 고객 관련 거래를 실시간에 처리하는 방식으로 구성되어 있으며, 조직 운영에 있어서 핵심적인 부분이다. 그러나, 거래처리 시스템이 산출하는 단순 데이터는 의사결정을 내리는데 직접적인 도움이 되지 못하며 필요한 정보로 변환되어야 의사결정자의 요구를 충족시킬 수 있다.

오늘날의 경영자들은 과거에 비하여 더욱 짧은 시간에 훨씬 더 중요한 의사결정을 내려야 하는 상황에 처해 있다. 이러한 배경 아래에서 의사결정 과정을 지원하는 의사결정 지원시스템이 정보시스템 분야에서 개발되었다.

전통적인 의사결정 지원 방식은 사용자의 요구에 의하여 정보기술 부서에서 추출 프로그램을 작성하여 메인프레임의 거래처리 데이터를 가공한 후 최종사용자가 손쉽게 접근할 수 있는 PC에 내려주는 방법이다. 이러한 방법은 데이터를 분석할 때 거래처리 시스템에 영향을 주지 않고 데이터가 사용자의 통제하에 오도록 해주는 이점이 있어 널리 사용되었다. 이른바 EUC(End-User Computing : 최종사용자 컴퓨팅) 방식으로서 全社의인 데이터 통합 노력 없이 필요할 때마다 분석활동이 이루어지도록 한 것이었다. 그러나 이러한 방식은 시간이 지나면서 추출된 데이터베이스에서 또 추출이 이루어질 뿐만 아니라 다음과 같은 여러 가지 문제들이 발생하게 되었다.²²⁾

① 각각의 시스템들은 데이터를 수집하는 기간에 차이가 있으며, 보고서 생성기간에서도 차이가 있다. 이로 인하여 서로 다른 시기의 데이터가 공존하게 된다. 또한 시스템간에 계산 방법이 상이하며, 데이터의 원천이 서로 다르고, 서로 다른 외부 데이터를 추가하게 된다. 이러한 결과로 추출 단계마다 문제점이 증폭되어 데이터의 통합 및 신뢰성에 문제가 생긴다.

② 분석 및 의사결정에 필요한 데이터의 위치를 파악하기 어렵고, 요구사항이 계속 변하여 추출 및 변환을 할 때마다 많은 프로그램을 매번 다시 작성해야 하며, 다

22) 박천승, 전계논문, pp.14~17.

양한 기술적 환경에 존재하는 데이터를 다루어야 하므로 정보제공의 생산성이 저하되고, 적시에 필요한 데이터를 지원하지 못한다.

③ 데이터 구조가 유연하지 못하여 다양한 사용자의 정보요구를 충족시킬 수 없다.

④ 수많은 응용 프로그램을 분석해야 하고, 과거 데이터가 저장되어 있지 않거나, 있더라도 쉽게 접근하기 어렵다는 문제점이 존재하여 정보생성과 관련된 문제가 생긴다.

⑤ 이러한 문제점을 해결하기 위해 통합 데이터 모델링을 하려고 해도 현실적으로 어려운 일이다.

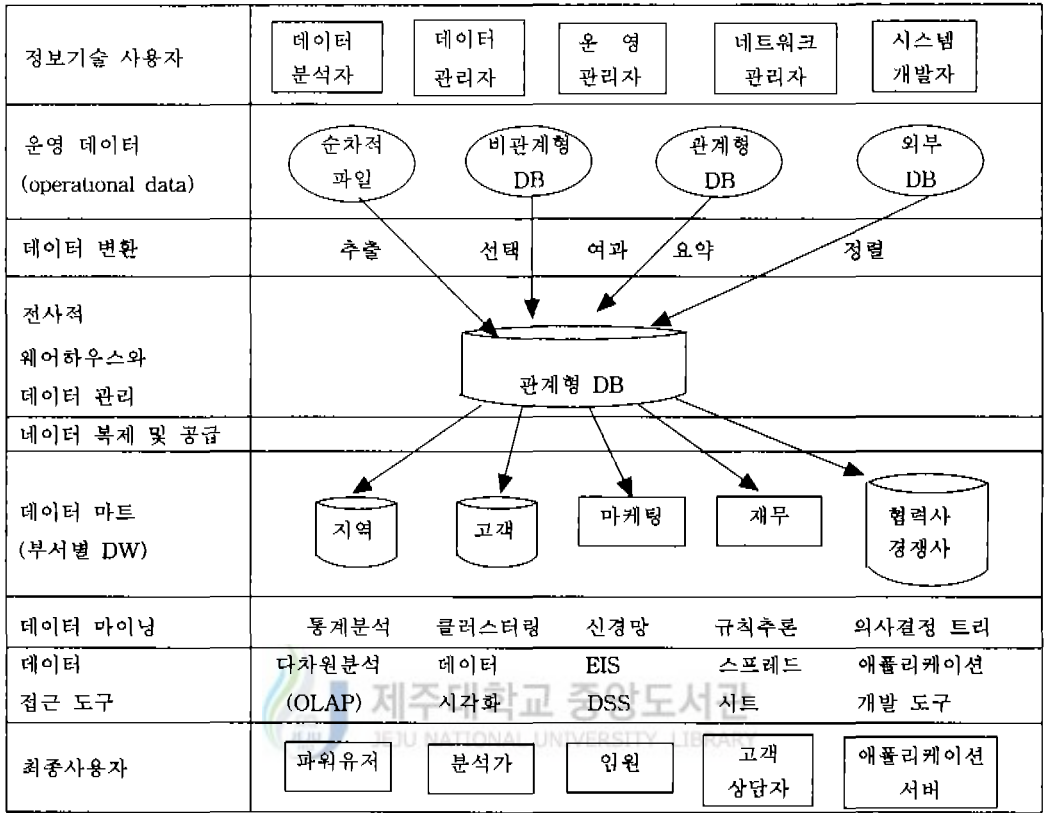
이상과 같은 문제점으로 인하여 의사결정 지원시스템은 새로운 정보 기반을 필요로 하게 된다. 경영환경의 변화에 따라 최종사용자들은 더욱 효과적인 의사결정을 해야하고 이를 위해 기업이 축적해온 대량의 데이터를 빠른 속도로, 全社적으로 통합된 형태로, 과거 이력 데이터까지 포함하여 분석함으로써 요약정보를 얻을 수 있어야 한다.

그러나 기존 시스템으로는 갈수록 복잡해지고 다양해지는 최종사용자의 정보요구를 수용하기에는 한계에 봉착할 수밖에 없다. 복잡한 데이터의 구성 및 분산된 시스템의 관리를 위하여 엄청난 인적·물적 부담을 질 수밖에 없기 때문이다. 이로 인하여 최종사용자들은 제때에 필요한 정보를 얻을 수 없게 되고, 이는 곧 정보기술 부서에 부담으로 남게되어 정보기술 부서 본연의 업무인 운영업무에 집중할 수 없게 된다.

이러한 상황이 데이터 웨어하우스의 등장 배경이 되었다. 다음의 <그림 3-1>은 가장 널리 인용되는 데이터 웨어하우징 프레임웍(framework)으로서 크게 원천 데이터, 데이터 웨어하우스, 데이터 마트, 최종사용자 층으로 이루어진다.

데이터 웨어하우스의 등장에는 물론 정보기술의 발전이 밑바탕이 되었다. 컴퓨터 하드웨어, 통신, DBMS의 발달로 병렬처리 컴퓨터, 병렬처리 관계형 DBMS, 미들웨어 기술, 데이터 분석 기술 등 정보사용자의 요구를 충족시킬 수 있는 기술적 기반이 성숙되었다.

<그림 3-1> 데이터 웨어하우징 프레임워크



자료 : Swift, Data Warehouse Report, 1995.

제3절 데이터 웨어하우스 關聯 概念

앞의 <그림 3-1>에서 보듯 데이터 웨어하우스는 全社的으로 통합된 관계형DB로서 최종사용자가 직접 접근하는 것은 현실적이지 못할 뿐 아니라, 전체 시스템 성능에 심각한 부하를 줄 수 있으며, 데이터 웨어하우스가 최종사용자들이 원하는 유용한 정보를 주지 못할 수 있다. 따라서 최종사용자가 데이터 웨어하우스의 정보를 효과적으로 이용하기 위해서는 다음에 설명하는 두 가지 개념, 즉 데이터 마트(data mart)와 데이터 마이닝(data mining)에 대한 이해가 필요하다.

1. 데이터 마트

데이터 웨어하우스로부터 여러 부서로 필요한 형태의 의사결정 지원에 사용할 목적으로 데이터가 흘러간다. 이러한 부서 기반의 의사결정 지원 데이터베이스를 데이터 마트라고 한다. 데이터 마트는 데이터 웨어하우스의 체계적 기반을 가지는 부서 내 의사결정 지원 데이터의 몸체이다.²³⁾

데이터 웨어하우스에는 세부 수준(granular level)의 데이터가 저장되어 있으며, 데이터 마트에는 정제된 수준(refined level)의 데이터가 존재한다. 데이터 웨어하우스에 기반을 둔 동일한 세부 데이터는 서로 다른 데이터 마트에서 각각 조회되거나 결합되어 있으며, 어떤 경우에는 서로 다르게 추가된다

이러한 데이터 마트의 개념은 중앙 집중적이고 全社的인 데이터 웨어하우스 방식이 여러 가지 문제점들을 내포하고 있음이 밝혀짐에 따라 등장하였다²⁴⁾ 데이터 웨어하우스는 데이터 자원의 사용에 대한 조직 내에서의 경쟁을 심화시키며, 데이터를 각 부서의 요구사항에 적합하게 변형하여 사용하기 어려우며, 처리에 많은 비용이 소요되며, 대용량의 데이터를 처리하기 위한 데이터 접근 및 분석용 소프트웨어가 흔하지 않다는 문제점을 갖고 있다.

23) W. H. Inmon, "What is Data Mart?", White Paper, D2K Inc, <http://www.d2k.com/d2k/library2.htm>, 1996.

24) W. W. Eckerson, "Building and Managing Data Marts", White Paper, <http://www.informatica.com/swp.html>, 1997.

또한, 여러 가지 이유로 데이터 웨어하우스는 급속히 대용량의 데이터플 가지게 된다. 데이터 웨어하우스가 커짐에 따라 데이터 마트의 필요성이 커지게 될 뿐 아니라, 다음과 같은 몇 가지 이유로 데이터 마트가 매력을 가진다.²⁵⁾

① 부서가 자신의 데이터 마트를 가지고 있을 때, 데이터 웨어하우스로부터 데이터 마트로의 데이터 흐름상에 데이터를 개별 조작(customizing)할 수 있다. 데이터 마트 내의 데이터가 전체 기업을 지원할 필요는 없다. 그러므로 부서는 다른 부서와의 관계를 생각할 필요 없이 핵심적인 데이터의 내용에 대하여 요약, 정렬, 선택, 구조화 등을 할 수 있다.

② 과거 데이터는 전사 차원이 아닌 부서 차원의 필요에 의한다. 대부분의 경우에 데이터 웨어하우스 내에 기초를 둔 데이터보다 훨씬 적은 양의 과거 데이터를 부서에서 이용한다.

③ 다른 부서와 관련되는 자원의 가용성 영향을 고려하지 않고 부서는 원하는 의사결정 지원처리의 어떤 것이든 언제든지 할 수 있다.

④ 부서는 원하는 분석 소프트웨어를 선택할 수 있다. 데이터 마트의 접근, 분석 소프트웨어들이 많이 소개되어 있다.

이러한 매력을 가진 데이터 마트는 데이터 웨어하우스를 대체하는 것이 아니라, 보완하는 개념으로 출발하였다. 데이터 마트가 데이터 웨어하우스의 일부분이라고는 하지만, 데이터 마트가 단지 소규모의 데이터 웨어하우스라고 할 수는 없다. 데이터 마트와 데이터 웨어하우스의 차이는 지원 대상의 규모뿐만 아니라, 기능에서도 존재하는 것이다. 물리적인 데이터량을 기준으로 한 데이터 웨어하우스와 데이터 마트간의 구분은 최근 데이터 마트가 보유하는 데이터의 양이 점차 커지면서 의미를 상실해가고 있다.²⁶⁾

오늘날 많은 기업들이 데이터 웨어하우스를 구축하기 이전에 데이터 마트를 먼저 구축한다. 이러한 현상의 가장 큰 이유 중의 하나는, 데이터 마트가 데이터 웨어하우스보다 구축하기 쉽다는 점이다. 데이터 웨어하우스를 구축하려면 비용, 기간 등

25) W. H. Inmon, 전계자료, 1996.

26) P. Brooks, "March of the Data Marts", DBMS Magazine, March 1997.

의 막대한 자원을 필요로 하며, 그래서 쉽게 실행으로 옮기기 힘들데 비하여, 데이터 마트는 초기 투자비용의 부담이 적고, 구축 기간도 상대적으로 짧기 때문에 위험 부담을 줄일 수 있으며, 비용 대비 효과가 큰 프로젝트를 우선 수행하여 이로부터 얻은 확신을 기반으로 안정적으로 확장해 갈 수 있다는 장점이 있다.

따라서, 대부분의 기업들은 데이터 웨어하우스의 문제점들을 해결 또는 보완하기 위한 수단으로 부분적으로 데이터 마트를 구축하고 있으며, 다수의 데이터 마트들의 연합체를 구성함으로써 데이터 웨어하우스를 사용하지 않으면서도 동일한 기능을 제공하고자 하는 접근 방식 등이 많이 사용되고 있다.²⁷⁾

그러나 전사적인 데이터 웨어하우스와 부서 단위의 데이터 마트간에는 어느 한쪽이 일방적인 우위에 놓여 있지는 않다. 그러므로, 기업의 정보시스템 담당자들은 각 기업의 특성에 맞추어 장기적인 관점에서 적절한 데이터 구조에 대한 결정을 내려야 한다.



27) C. Stedman, "Warehouse costs Scare up Changes", Computerworld, <http://www.computerworld.com>, 1996.

2. 데이터 마이닝

데이터 마이닝이라는 용어는 데이터로부터 유용한 정보 혹은 패턴을 발견하고자 하는 연구분야에서 다양한 명칭으로 사용되어져 왔으며, 이는 지식발견(knowledge discovery), 데이터 채취(data dredging), 지식추출(knowledge extraction), 정보수확(information harvesting) 등으로 불려졌다. 이 중에서 가장 보편적으로 사용되는 용어가 데이터 마이닝이다.²⁸⁾

즉, 데이터 마이닝은 “적극적으로 수집된 대용량의 데이터베이스 및 정보를 가공하여 기업경영에 결정적인 영향을 미칠 수 있는 지식을 발견하기 위한 일련의 작업 과정”이라 할 수 있다.²⁹⁾

데이터 마이닝을 위해서는 잘 정리된 많은 양의 자료가 필요하며, 따라서 데이터 웨어하우스는 최적이라고 말할 수 있다. 분석가는 데이터 웨어하우스가 구축되어 있으면, 데이터를 정제·통합·보충하는데 신경 쓸 필요 없이 마이닝 작업에 전념할 수 있다.³⁰⁾ 실제적으로 데이터 마이닝을 위하여 데이터베이스를 재정리하고 조작하는 작업은 데이터 웨어하우스에서 행하는 작업과 거의 유사하기 때문에 두 가지 프로젝트를 동시에 수행할 수도 있을 것이다.

데이터 마이닝을 통해 얻을 수 있는 정보의 형태는 매우 다양하며, 이에 따라 다양한 기법이 존재한다. 가장 대표적인 데이터 마이닝 기법으로는 사건들의 연관성(associations) 탐사, 연속성(sequence) 탐사, 분류(classifications) 규칙 탐사와 군집구분(clustering)을 들 수 있으며, 핵심 알고리즘으로 귀납적 추론(rule induction)이나 신경망(neural network)과 같은 인공지능(artificial intelligence)등을 응용한 방법이 소개되면서 데이터 마이닝이 더욱 활성화되기 시작하였다

28) 정보통신부, “데이터 웨어하우스 기반의 Data Mining 소프트웨어 개발”, 1997. 12월, p.48.

29) 이영철, “DB마케팅을 통한 의류브랜드의 데이터 분석”, ‘98 SPSS 사용자 사례 논문집 -데이터 마이닝을 중심으로-, p.57.

30) WH Inmon, “The Data Warehouse and Data Mining”, Communications of ACM, Vol. 39, No 11, November 1996.

데이터 마이닝은 다음의 <표 3-1>에 제시된 바와 같이 다양한 산업에서 적용되고 있으며, 특히 마케팅 분야에서 가장 활발하게 이용되고 있다. 데이터 마이닝을 통하여 자사의 서비스를 이용하는 고객의 특성을 분석하여, 개별 고객에 맞는 다양한 서비스를 제공할 수 있게 된다. 금융기관으로서는 상품간 혹은 고객 특성간의 상호 관련성을 발견하여 각종 위험을 예측한다든지, 고객의 기호에 맞는 다양한 금융 상품을 개발할 수 있다. 외국의 사례를 보면, 대형 금융기관만이 이러한 기법을 사용하고 있는 것이 아니라, 소형 금융기관도 직접 우편판매와 같은 마케팅 수단과 고객 특성들간의 상관관계를 밝혀 고객의 반응도를 높이는데 사용하고 있다.

<표 3-1> 데이터 마이닝의 응용 분야

응용분야	통신산업	유통업	은행업	보험업	기타
적용사례	통화기록분석	시장바구니 분석	고객집단분류 카드사기적발	사기색출	고객집단편성
	고객충성도 제고	시계열 패턴 분석	금리 및 추가예측	상품설계	자동차 제조
	라이프사이클 예측관리	예측모델의 개발	라이프사이클 예측관리	위험분석	항공사의 탑승객 인센티브

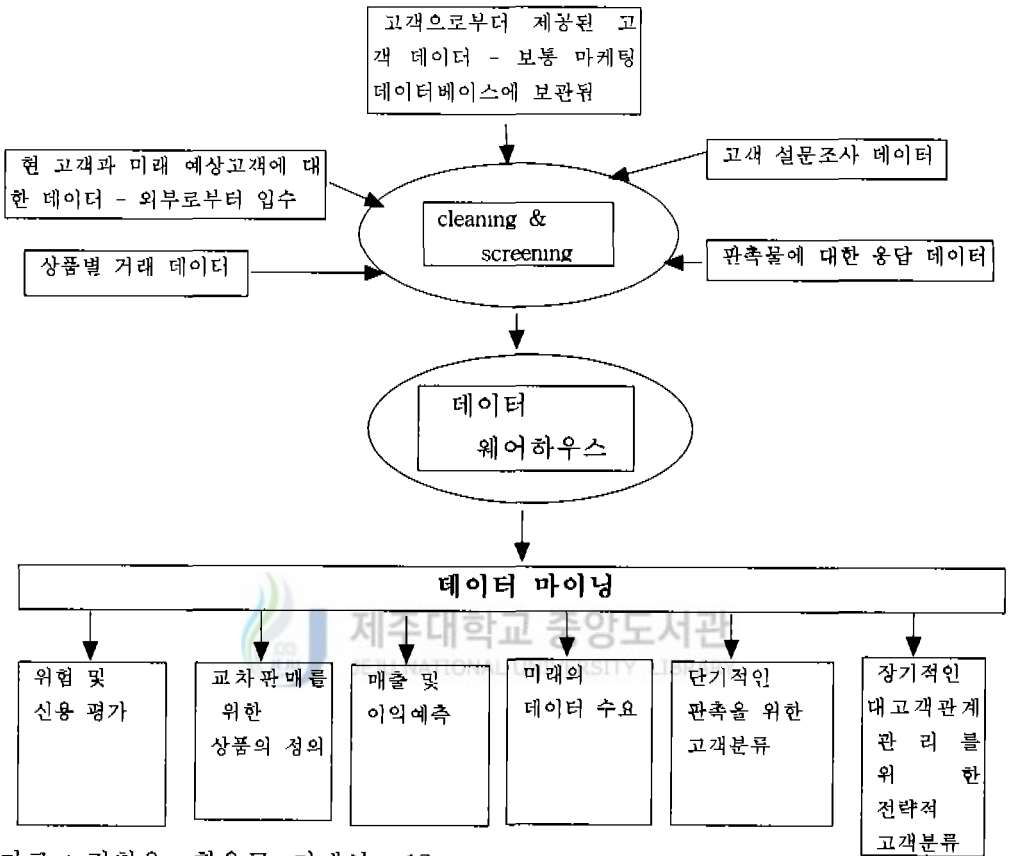
자료 이용처, "Data Mining을 이용한 리테일 뱅킹 전략에 관한 실증적 연구", 전국은행연합회 논문집, 1998, p14.

이러한 데이터 마이닝을 성공적으로 수행하려면 비즈니스 문제와 기술에 대해 깊이 있게 이해해야 한다.³¹⁾ 또한 사람에 의한 판단은 데이터 마이닝 과정에서 필수적인 부분이며, 사람만이 데이터 마이닝의 결과로 나타난 어떠한 패턴, 규칙 혹은 함수관계가 타당하고 기업에 유용한 것인지 판단할 수 있다.

다음의 <그림 3-2>는 데이터 마이닝의 일반적인 흐름을 나타낸 것이다.

31) H Edelstein, "Mining Data Warehouse", Information Week, Jan 1996.

<그림 3-2> 데이터마이닝의 흐름



자료 : 정철용 · 함유근, 전계서, p12.

제4절 데이터 웨어하우스 構築의 前提 要素

1. 데이터 웨어하우스에 대한 마인드 形成

현업 전문가가 데이터 웨어하우스를 활용하여 다차원적으로 분석된 의미 있는 자료를 추출해 놓고도 이를 활용하지 않거나, 최고 경영층에서 기업 경영에 반영하지 않는다면, 데이터 웨어하우스를 구축한 효과는 기대할 수 없게 된다. 결국 데이터 웨어하우스는 만들고 나서부터 시작이라는 인식이 全社的으로 공감대가 형성되어 있어야 한다.

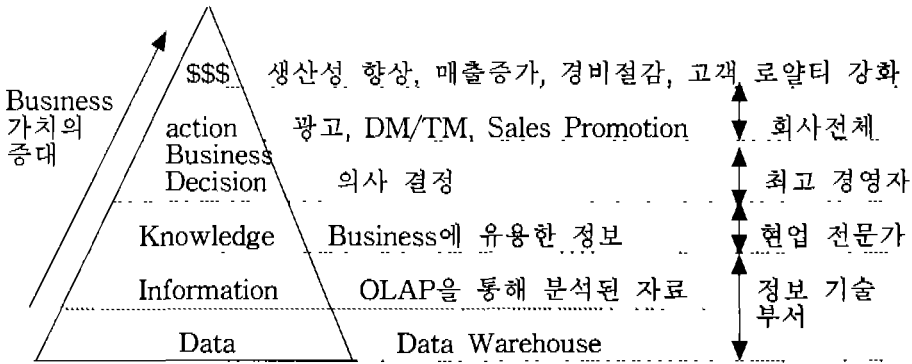
따라서 지금까지의 다분히 정보기술 중심의 데이터 웨어하우스 구축 방법을 지양하고, 현업이 충분히 참여하고, 최고 경영자가 적극적으로 후원하며, 全社的으로 데이터 웨어하우스 사용에 대한 의지를 가지고 데이터 웨어하우스를 구축하여야 한다.

그러나 아무리 데이터 웨어하우스를 잘 만들어 놓아도 현업 전문가가 이를 사용하지 않으면 데이터 웨어하우스 구축의 목적이 흔들리게 된다. 현업 전문가는 데이터 웨어하우스 이외에도 많은 곳에서 데이터, 또는 정보를 수집한다. 데이터 웨어하우스의 분석은 주로 OLAP(On-Line Analytical Processing : 온라인 분석 처리) 도구를 가지고 하는 데, 이 도구를 처음 사용할 때 저항감을 느낄 수도 있음은 물론, 같은 종류의 보고서라도 지금까지 만들던 방식을 그대로 고집할 수도 있다.

또한, 현업 전문가가 OLAP 도구를 활용하여 지금까지 생각해 왔던 것과는 다른, 비즈니스에 유용한 정보를 도출해 냈다고 하더라도, 최고 경영층에서 이를 받아들이지 않을 수도 있다. 만일 이러한 일들이 빈번히 반복된다면 현업 전문가는 데이터 웨어하우스를 여러 가지로 응용하여 분석하는 일을 하지 않을 것이고, 과거 방식대로 보고서를 만들어내는 것으로 만족하게 될 것이다.

이론적으로는 과학적 경영을 해야 한다고 하지만 의외성이 많은 것이 기업 경영이므로 최고 경영자가 판단하는 것은 또 다를 수가 있다.

<그림 3-3> 의사결정 피라미드



자료 : http://www.oracle.co.kr/magazine/win_cover1.html

결국, 데이터 웨어하우스 구축의 효과는 지금까지 데이터 웨어하우스 공급업체에서 약속한 것처럼 당연한 것이 아니며, 全社的인 의지 없이 구축 효과를 따지는 것은 무의미하다. 특히 데이터 웨어하우스의 ROI(Return On Investment)라는 것은 데이터 웨어하우스를 사용해 보지 않고는 생각할 수 없다. 따라서, ROI는 사용정도에 따라서 결정되는 것이지, 데이터 웨어하우스의 평균 ROI라는 것은 있을 수 없다. 심하게는 0%가 될 수도 있으며, 1000%에 육박할 수도 있다.³²⁾

데이터 웨어하우스를 구축한 후 사용하면서 실제로 데이터 웨어하우스가 시작되는 것이다. 지금까지는 데이터 웨어하우스 구축에만 매달려 있었기 때문에 데이터 웨어하우스를 어떻게 사용하는가에 대한 생각이 없었다. 이것은 데이터 웨어하우스를 구축할 때 현업 전문가가 참여해도 마찬가지이다. 결국 全社的인 지원과 협조 없이는 데이터 웨어하우스 구축 자체도 어렵지만 데이터 웨어하우스 구축 후에도 많은 문제를 남겨 놓게 된다 이러한 비즈니스적인 관점은 데이터 웨어하우스 구축 방법론에 대한 많은 반성을 하게 한다

기존의 구축 방법론에서는 데이터 웨어하우스의 구축 자체에 많은 중점을 두었다 그러나, 위의 <그림 3-3>에서 보면 실제 데이터 웨어하우스 사용으로 인한 의사결정 자체에 많은 비중을 두어야 한다는 것을 알 수 있다.

32) 장동인, “차세대 의사결정 지원 시스템-데이터 웨어하우스의 최신 동향”, 오라클 97년 겨울호, 1997.

그렇다면 어떻게 구축해야 올바른 데이터 웨어하우스가 될 것인가? 답은, 다음의 <그림 3-4>와 같이 의사결정 피라미드를 역으로 내려가면 된다는 것이다.

즉, 비즈니스 행동을 취하기 위한 최고 경영자의 의사 결정 과정을 분석해야 하며, 어떤 비즈니스 지식이 있어야 의사결정을 하는가를 알아야 한다. 중요한 것은 데이터 웨어하우스가 모든 비즈니스의 의사결정에 다 쓰이는 것은 아니라는 점이다. 데이터 웨어하우스는 몇 가지 특정한 비즈니스의 문제 해결을 위해서 필요한 것이다

다시 말해서 데이터 웨어하우스 구축에는 분명한 목적이 있어야 한다는 것이다.

<그림 3-4> 역 의사결정 피라미드



자료 : http://www.oracle.co.kr/magazine/win_cover1.html

예를 들어 은행의 경우에 있어서, '현재 우리 은행에 가장 수익을 많이 가져다주는 우수 고객 층을 찾아 효과적인 마케팅을 실시하여 더 많은 우수 고객을 확보하여 매년 10% 이상의 더 많은 수익을 창출 하고자 한다'는 구체적인 목적을 위해서 데이터 웨어하우스를 만들어야 데이터 웨어하우스 구축의 성공과 실패가 분명하고 ROI 계산도 분명해진다. 성공적인 데이터 웨어하우스 구축을 위해서는 데이터 웨어하우스 구축 목표를 확실히 해야 한다.

2. 全社的인 協助

데이터 웨어하우스는 반드시 정보기술 부서와 현업의 조화와 협조로 시작되어야 한다. 그렇지 않으면 데이터 웨어하우스가 단순히 보고서 제작용 시스템으로 전락하고 만다. 실제로 데이터 웨어하우스 구축을 두고 정보기술 부서는 현업에 대해서 다음과 같이 생각하는 경향이 있다.

- 현업은 자기들이 무엇을 원하는지 모른다.
- 현업은 자꾸 자기들의 생각을 바꾼다.
- 현업은 너무나 많은 것을 단시일 내에 해달라고 한다.
- 현업은 모든 데이터를 다 보여 주어야 원하는 것을 말하겠다고 한다.

반면에 현업은 정보기술 부서에 대해서 다음과 같이 불만을 이야기한다.

- 정보기술 부서는 시스템을 개발하는 데 시간이 너무 걸린다.
- 정보기술 부서는 아무리 내 요구사항을 설명해도 이해를 못한다.
- 정보기술 부서에 요청해도 내가 원하는 데이터를 가질 수 없다.

이러한 정보기술 부서와 현업의 의견차이는 사소한 것처럼 보이지만, 실제 데이터 웨어하우스를 구축한 후에 그 태도가 크게 달라진다. 정보기술 부서는 '원하는 것은 여기 다 있으니 알아서 써라' 라는 자세이고, 현업은 '그것이 우리가 원하는 전부는 아니다' 라고 하거나 '사용하기 불편하다', '안 되는 것이 너무 많다' 등 셀 수 없이 많은 문제들을 제기한다.

또한, 최고 경영층의 확고한 의지도 중요한 요소이다. 지금까지는 최고 경영층으로부터 데이터 웨어하우스 프로젝트를 재정적으로 후원해 주는 역할 정도만 기대하였으나, 성공적인 데이터 웨어하우스 구축과 활용을 위해서는 재정적인 후원 이외에도 할 일들이 많다. 즉,

- 데이터 웨어하우스 구축 목표를 확고하게 잡아 주어야 한다.
- 데이터 웨어하우스를 구축하는 과정에서 생기는 부서간의 많은 갈등을 창조적으로 조정해 주어야 한다.
- 全社的으로 데이터 웨어하우스에 바탕을 둔 과학적 경영을 할 의지를 선언해야 한다.
- 全社的으로 데이터 웨어하우스를 잘 활용할 수 있도록 분위기를 조성해야 한다.

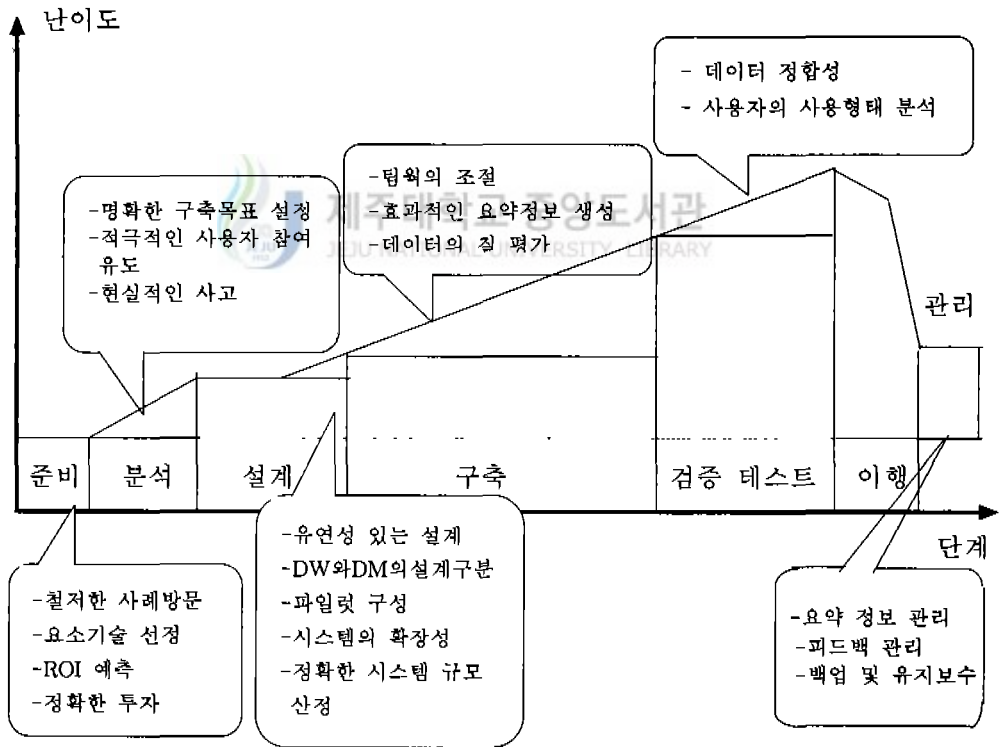
이와 같은 노력 없이는 데이터 웨어하우스의 구축 효과를 기대하기 어렵다.



제5절 데이터 웨어하우스 構築 段階別 체크 포인트

구현된 데이터 웨어하우스에 대하여 '성공'과 '실패'의 판단은 개발자가 내리는 것이 아니라 최종 사용자와 관리자가 내리는 것이다. 현실적으로는 데이터 웨어하우스를 구현 했느냐가 중요한 것이 아니라, 필요한 정보를 손쉽게 획득할 수 있는 시스템을 가지고 있는가 하는 것이 중요하다는 의미이다. 이제부터 성공적인 시스템을 갖추기 위해서 필요한 여러 가지 사항들을 각 프로젝트 수행 단계별로 나누어 살펴보고자 한다.³³⁾

<그림 3-5> 구축 단계별 체크 포인트



자료 : 상계자료, p.226

33) 전완기, "DW를 성공으로 이끌기 위한 전략", 컴퓨터월드 3월호, 1999.

1. 計劃 段階

1) 철저한 사전 준비와 자료 수집

데이터 웨어하우스의 구축은 메인 프레임, 오픈 시스템, OLAP, MINING, WEB 등 모든 정보기술의 총아라 해도 틀린 말이 아니다. 그러나 데이터 웨어하우스는 전사 데이터의 통합 시스템으로써 구현되는 반면, 다양한 사용자를 관리하여야 하므로 시스템 측면에서는 보안에 대한 철저한 대책이 필요하며, 이는 프로젝트의 성공에 치명적인 영향을 줄 수 있는 변수이기도 하다. 어떠한 프로젝트를 수행하는데 있어서 최적의 방법론은 프로젝트의 성공을 좌우하는 중요한 요소가 된다 많은 자료 수집 및 파일럿 시스템의 구현이나 프로토타이핑(Prototyping)을 통한 접근 방법론의 정립은 프로젝트의 열쇠가 될 수 있다.

또한, 이미 구축된 기업의 경험과 노하우를 습득하기 위한 사례 방문을 통하여 사용자들의 만족도를 조사하고, 개발 담당자의 조언을 듣고, 구축 상황 및 환경을 파악함으로써, 보다 현실적인 데이터 웨어하우스를 접할 수 있다. 이렇게 함으로써 실제 환경에서의 데이터 웨어하우스에 대한 장·단점을 평가할 수 있는 좋은 계기를 마련할 수 있다

2) 투자에 대한 회수 예측

기업이 투자를 하는데 있어 가장 중요한 목표는 고객 서비스 질의 향상과 이를 통한 이윤의 극대화이다. 고객과의 직접적인 접촉을 위한 새로운 마케팅 기회와 서비스의 개선, 보다 자동화된 마케팅 소프트웨어의 개발, 데이터 통합을 위한 데이터 웨어하우스 등이 기업이 최근 관심을 갖는 투자 대상이다. 투자에 대한 회수는 크게 두 가지로 나누어 볼 수 있는데, 정보기술 부서 및 사용자들의 보고서 작성을 위한 시간 절약과 기업의 이윤 극대화라는 측면이다.

1998년 8월에 Dataquest³⁴⁾에서 조사한 바에 의하면, 데이터 웨어하우스나 데이터

34) Gartner Group's Market Dynamics의 부서

마트의 활용 형태가 아직은 주로 보고서를 작성하는 수단으로 많이 사용되고 있음을 알 수가 있고, 그 다음은 마이닝(Mining), 위험 관리(Risk Management), 고객 관계 관리(CRM) 등의 순서로 조사되었다.

데이터 웨어하우스를 구축하기 전까지 대부분의 사용자들은 보고서 작성을 위해서 많은 시간을 소비해야만 했다 그러나 데이터 웨어하우스는 이 시간을 사용자의 시간으로 돌려주었다. 이는 데이터 웨어하우스가 통합 정보를 위한 창고로서의 역할을 충실히 하고 있다는 증거일 것이다.

또한, 데이터 웨어하우스를 이용한 고객 분석을 통하여 기업에 이익을 주는 우량 고객들을 파악해 냄으로써, 이들을 대상으로 보다 적극적인 마케팅을 전개하여 우량 고객의 유지는 물론 신규 고객의 유치로까지 확산시킴으로써 기업 이익의 극대화를 도모하여 투자에 대한 회수를 이끌어 낼 수 있게 된다.

3) 투자 순서와 목표 설정

기업들은 이제 데이터를 정보화하는 단계에 접어들었다. 앞으로의 과제는 새롭게 만들어지는 정보를 어떻게 활용할 것인가 하는 문제이다. 이제 DB마케팅과 데이터 웨어하우스를 어떻게 구분할 것인가 혹은 어떻게 활용할 것인가를 생각해보자.

많은 정보기술 종사자들은 데이터 웨어하우스를 구축해야 DB마케팅을 할 수 있다고 일반적으로 생각하고 있다. 그러나, 데이터 웨어하우스는 DB마케팅을 위한 기반으로 활용될 수는 있지만, 반드시 우선되어야 하는 과정은 아님을 알아야 한다.

DB마케팅은 독립적인 의미를 가지는 경우가 많다. 全社的 데이터의 통합으로서의 의미가 아니라, 고객관리 차원의 마케팅 영역을 중심으로 한 시스템의 구현이라는 면에서 데이터 웨어하우스와 다른 면을 보여 준다. 이러한 이유로 DB마케팅은 분리될 수 있는 하나의 독립적인 시스템으로서 새로운 마케팅 수단으로 활용될 것이다.

그러나 DB마케팅은 데이터 웨어하우스의 통합된 데이터를 기반으로 손쉽게 구현할 수 있다. 물론 쉽다는 의미가 모든 문제가 한꺼번에 해결되며 결과가 산출된다는 의미는 아니다. 단지, 全社的 차원으로 통합된 데이터를 활용함으로써 비교적 효과적으로 DB마케팅을 수행할 수 있다는 것이다. 따라서 투자의 순서는 데이터 웨어하우스를 기반으로 한 DB마케팅의 실현이라고 할 수 있다.

4) 적절한 요소 기술의 선정

최종 사용자는 프로젝트의 성패를 가늠할 수 있는 권한을 가진다. 왜냐하면 시스템의 최종 목적지는 최종 사용자들이기 때문이다. 도구 선정은 사용자들이 사용하기 위한 EUC 도구부터 개발에 필요한 여러 가지 도구들을 대상으로 한다. H/W, DBMS, OLAP 도구, 보고서용 도구, CASE 도구 등이다. 모든 도구들은 사용자의 요구를 100% 만족시키지는 못한다. 각각의 도구들은 자신만의 장·단점을 가지고 있다. 그러나 중요한 것은 가장 필요로 하는 기능을 제외한 나머지 기능은 사용자들에 의해 5% 미만으로 사용되거나, 대부분 사용되지 않는 기능이 많다는 것이다.

우선 적절한 요소기술을 선정하기 위해 가장 필요로 하는 기능부터 우선 순위를 정하여 도구를 평가하는 것이 필요하다. 사용자 도구를 선정하는 가장 좋은 방법은 사용자들로 하여금 직접 사용해 보게 하는 방법이다. 그리고 크게 관리자와 사용자의 관점에서 평가하는 것이 바람직하다. 개발 및 관리, 유지보수의 편리성과 사용과 기술 습득의 용이성 등을 고려하여야 한다.

5) 투자 규모의 결정



데이터 웨어하우스 구축을 위한 투자는, 적절한 기간과 비용, 그리고 철저한 준비에 의한 방법론, 적절한 기술 요소 선정, 경험 있는 인력의 투자 등 각 분야의 전문가와 자문 위원을 통하여 규모를 결정하는 것이 중요하다. 데이터 웨어하우스는 다른 시스템 보다 초기 투자가 많은 편이다. 단계적인 투자와 투자의 우선 순위 결정 등으로 초기의 무리한 투자를 방지할 수 있다.

대부분의 기업에서 초기 과잉 투자에 대한 우려와 투자 가치에 대한 판단을 제대로 하지 못하고 있다. 이것은 아직도 정보기술의 힘을 마케팅과 연계하여 생각하지 못한 이유일 것이다. 투자에 대한 두려움을 없애고, 정확한 판단으로 투자 규모를 결정하는 것이 기업 측면에서는 중요한 요소이다.

2. 分析 段階

1) 사용자의 적극적인 참여 유도

프로젝트의 성공요소 중 가장 중요한 것이 경영자와 현업 사용자들의 관심이다. 어떠한 시스템이라 할지라도 지금까지 대부분의 프로젝트에는 사용자들의 참여가 있었다. 하지만, 개발 자체를 이끌고 가는 주체는 정보기술 부서였고, 시작도 정보기술 부서에 의해서였다. 그러나, 이제 시스템의 개발은 최종 사용자로부터 시작이 되어야 한다는 것을 인식하여야 한다. 분석 단계는 전 단계 중에서 사용자의 참여를 가장 많이 유도해야 하는 단계이다. 분석가들은 사용자의 요구 사항을 최대한 반영할 수 있도록 해야 한다

시스템의 성패는 사용자에게 의해 결정된다. 사용자가 제시한 요건으로 만든 시스템을 사용자가 사용하지 않는다면 실패일 수밖에 없다. 사용자의 참여 하에 데이터 웨어하우스에서 제공되어야 할, 그리고 앞으로 제공하게 될 내용에 대해 분석하여야 한다. 사용자 환경의 변화에 따라서 사용자가 직접 데이터를 가공하고 분석하는 사고의 변화가 필요하다. 사용자들은 이제, 정보기술 부서에 의존해 왔던 지금까지의 작업들을 직접 처리하는 새로운 환경에 빠르게 적응하고, 발전된 사고를 받아들여야 한다

2) 구축 목적의 명확화

데이터 웨어하우스의 구축 목적이 분명해야 한다. 데이터 웨어하우스가 가져다주는 이점은 여러 가지이지만, 단순히 데이터 통합을 위해서인지, 진정한 의사결정을 위한 것인지를 확인해야 한다. 목적에 따라 시스템의 설계 방법이 달라질 수가 있기 때문이다. 즉, 시스템 구현의 근본이 바뀔 수 있는 문제이다. 데이터 웨어하우스를 확대 해석하는 경향이 많기 때문에, 근본 목적을 뚜렷하게 하여야 한다. 단계적으로 시스템을 구축하되, 각 단계별 구축 영역을 정의하고, 구축 목적과 효과를 많은 사람들에게 알리는 것이 좋다.

3) 환상의 제거

데이터 웨어하우스가 모든 것을 해결해 줄 것이라는 환상과, 단기적으로 데이터 웨어하우스를 이용하려는 생각을 버리도록 해야 한다. 모든 것을 데이터 웨어하우스가 해결할 것이라는 잘못된 판단은 투자자에게 실망을 안겨 준다. 장기적인 투자를 기본으로 데이터 웨어하우스 구축 효과를 기대해야 한다. 단기적으로 거둘 수 있는 효과는 사용자를 위한 보고서 작성 시간을 단축할 수 있다는 것이다.

투자한 만큼의 기대를 만족시킬 수 있는 시스템이 구현되기 위해서는 많은 노력과 시간이 필요하다는 점을 깨우치도록 해야 한다. 그리고 경영자나 현업 사용자들에게는 초기 투자만큼의 효과를 가져다주지 못할 수도 있다. 막연히 모든 것을 데이터 웨어하우스에 맡기려는 생각은 너무나 위험한 생각이다. 기대 효과를 너무 크게 홍보하는 것은 개발자에게 또 다른 위협으로 다가올 수 있는 문제이다. 데이터 웨어하우스의 가치와 미래를 정확히 이해시키는 것이 무엇보다도 중요하다.

4) 필요성에 대한 올바른 인식



데이터 웨어하우스가 구축될 필요성이 있다는 사실에 대해서는 모든 사람들이 공감하고 있다. 때문에 데이터 웨어하우스의 개념 및 구축 방법론에 대해 많은 사람들이 관심을 가지고 있고, 이에 대해 많은 경험을 쌓아오고 있다. 대부분의 기업에서 겪고 있는 공통적인 문제점을 들자면, 방대해진 데이터의 관리 문제와 치열해지는 경쟁에서의 승자가 되는 것이다. 새로운 개념에 대해 우리는 모든 것을 유행처럼 받아들여, 한번쯤은 경험하고 싶어하는 충동을 느낀다.

이러한 면에서 데이터 웨어하우스라는 개념은 이미 우리의 걸을 깊이 파고 들어왔다. 지금까지 수많은 용어들을 접하며, 새로운 용어들의 탄생과 소멸을 지켜보고 있다. 우리는 데이터 웨어하우스라는 용어를 유행으로 받아들이기보다는 당면한 문제를 해결하기 위한 좋은 해결책이 될 수 있는가를 생각하는 것이 중요하다. 그것은 데이터 웨어하우스가 될 수도 있을 것이며, 아니면 다른 해결책이 될 수도 있다. 중요한 것은 데이터 웨어하우스를 구축 했느냐가 아니라, 그것을 이용해서 무엇을 할 수 있는가 이다.

3. 設計 段階

1) 시스템 규모의 산정

시스템 용량의 산정(Capacity Planning)은 현재의 데이터를 기준으로 미래의 필요 자원을 수용할 수 있는 적절한 시스템의 크기를 예측하기 위한 것이다. 데이터의 크기, 동시 사용자의 수, CPU의 성능 및 개수, 메모리의 양, 디스크의 속도, 네트워크 속도, 데이터 증가량 등은 시스템의 용량을 산정하는 기본자료로 이용된다. 디스크의 크기는 일반적인 계산법으로 산정하면 낭패를 볼 수 있다. 요약 정보를 위한 공간, 임시 영역을 위한 공간, 작업 공간, 데이터 공간 등을 미리 계산해야 한다.

2) 설계 기법의 결정

데이터 웨어하우스는 다차원 분석을 위한 설계가 기본이라고 생각하고 있는 사람들이 많다. 데이터 웨어하우스, 데이터 마트를 분명히 구분하고, 다차원 분석을 위한 설계 기법이 과연 데이터 웨어하우스를 위해서 필요한 것인지, 데이터 마트를 위해서 필요한 것인지를 구분하여야 한다. ODS(Operational Data Store ; 운영 데이터 저장소), 데이터 웨어하우스, 데이터 마트 등 목적에 따라 약간씩 틀린 설계기법을 사용하게 된다. 일부 트랜잭션을 수용하기 위한 ODS, 장기적인 상세 정보를 저장 관리하게 될 데이터 웨어하우스, 요약 정보 형태의 데이터 마트 등은 제각기 다른 설계 기법을 적용하는 것이 바람직하다.

3) 프로젝트 진행 회의

프로젝트 진행을 위한 회의는 일 주일 단위가 좋다. 지난주에 대한 평가와 반성, 다가올 주에 대한 계획과 일정 조정으로 전체 일정은 채워져야 한다. 주 단위 평가를 위한 평가서를 만들고, 팀별로 달성율을 검사하여, 일정 조정에 반영하는 것이 바람직하다. 관리자는 주 단위, 월 단위의 상세한 개발 계획 및 운영 계획을 수립하여야 하고, 각 주별 변경 사항을 관리하여야 한다. 중요한 것은 개발 범위를 일정에

끼워 맞추는 것이 아니라, 효율적인 인원 및 일정 관리를 통해서 전체 일정을 효과적으로 운영하는 것이 중요하다.

4) 위험 요소의 확인

각 프로젝트 단계에서 필수적으로 진행되어야 할 사항들과 각 기술 요소별 지원 사항, 미지원 사항을 발견하고 대처 방안을 마련한다. 예를 들어, 데이터 추출 시에 발생할 수 있는 문제(소스 데이터의 문제, 도구의 기능 문제 등)들을 나열하고, 각각에 대한 해결 방안을 마련한다든지, H/W 장비나 기타 S/W의 설치 지연에 따른 대처 방안이나, 각 단계별 일정 및 기술 구현에 있어 문제가 될 수 있는 부분을 도출하여 미리 대처방안을 마련하는 것이 위험 요소를 찾는 목적이다. 그리고 주기적으로 사용자의 참여를 유도할 수 있도록 하고, 각 단계에서 중요한 부분을 간과해 버리는 것이 없는지 확인하는 것이 좋다.

5) 충분한 일정 계획의 수립

프로젝트는 시간과의 싸움이라고 말하는 이가 많다. 그러나 결코 시간과의 싸움이 아니라, 좋은 시스템을 만들기 위한 솔루션과의 싸움이다. 중요한 것은 시스템 자체가 아니라 시스템이 보유하고 있는 잠재적인 능력이다. 사용자를 위해서는 원하는 정보와 편리성을, 기업을 위해서는 경쟁력 강화를 바탕으로 한 이윤 창출이 그것이다. 이러한 면에서 무리한 일정으로 짜여진 시간에 맞추려는 의도는 잘못된 것이다. 충분한 준비 기간과 개발 시간, 검증 시간을 갖고, 여러 가지 요인들에 의해 변할 수 있는 요소들을 찾아 적절한 해결책을 찾을 수 있도록 하는 것이 중요하다. 일정에 쫓기다 보면, 모두들 지치게 되고 의욕을 상실하게 된다. 충분한 휴식시간을 활용하여 사용자, 개발자, 설계자의 사상이 충분히 반영될 수 있도록 하는 것이 프로젝트 성공의 열쇠이다.

6) 파일럿 시스템의 중요성

데이터 웨어하우스의 기술 검증과 기술 습득을 위한 가장 좋은 방법은 파일럿을

통한 소규모 시스템을 구현해보는 것이다. 이론으로 습득한 기술은 실전 앞에 무력해 질 수 있다. 파일럿 시스템은 설계 단계에서 일부 주제 영역을 선택하여 모든 기술 요소를 적용하고, 각 요소의 기능을 확인함으로써 각 단계에서 필요한 사항들을 실험한다. 파일럿을 통해, 사용자들은 데이터 웨어하우스에 대한 개념을 이해하게 되고, 프로젝트 관리자는 진행 일정 및 기술 요소 평가 결과에 대한 대비책을 마련한다. 프로젝트 참여자들은 이를 통해 프로젝트의 진행에 대한 기술적인 노하우를 습득하는 수단으로 삼아야 한다

7) 확장성의 고려

데이터 증가에 따른 시스템의 확장성, 지식관리 시스템과의 연계성, 전자상거래를 위한 시스템과의 연계 등 다양한 하부구조와의 연계에 대한 확장성을 고려하여야 한다. 데이터 웨어하우스의 역할은 그 자체로서 끝나지 않는다. 다양한 시스템으로부터 자원을 통합하는 역할과 이를 활용한 효과적인 의사결정 정보를 제공하는 정보 제공자로서의 역할을 하여야 한다. 데이터 웨어하우스는 각 기업의 중심에서 각 시스템에 필요한 정보를 생성하는 정보 생성자의 역할을 담당하여 어떠한 시스템 혹은 사용자의 요건에도 답할 수 있는 통합 정보 시스템으로서의 역할을 담당해야 한다.

4. 構築 段階

1) 팀위의 조절

구축 단계에 이르면 프로젝트 참여자들은 점점 다가오는 프로젝트 종료 시간과 생각하지 못했던 문제의 발생으로 인한 일정의 지연 등으로 쫓기게 되고, 야간작업 까지도 해야 하는 상황에 이르게 된다. 결국 팀원의 사기는 저하되고, 프로젝트에 대한 흥미를 잃어버리게 된다. 그러므로 적절한 시간 관리와 팀원간의 적절한 업무 분담을 위한 프로젝트 관리가 필요하다. 사기진작을 위한 파티나, 워크숍 등은 피로해진 팀원의 사기를 끌어올리는 원동력이 될 수 있다.

2) 철저한 데이터의 질 평가

좋은 시스템이 되기 위해서는 기본적으로 데이터의 질이 좋아야 한다. 데이터가 많다고 해서 좋은 정보를 생성할 수 있다고 말할 수는 없다. 데이터의 질 평가는 분석을 목적으로 구축되는 시스템이라면 더욱 중요하다. 데이터의 질이 충실하지 못한 부분을 찾아, 데이터를 보완할 방안을 마련하는 것이 중요하다.

데이터는 하루아침에 만들어지는 것이 아니다 꾸준한 노력과 준비 작업으로 가치 있는 정보를 생성하고, 수정해 가는 과정을 통해서 의미 있는 데이터가 만들어지는 것이다. DB마케팅의 효과를 분석하려면 적어도 2년 이상은 지나야 한다고 한다. 이것이 바로 데이터의 질적인 문제 때문에 발생하는 현상이다. 초기의 고객 데이터는 단순하게 최초 가입자들에게서 받은 신청서의 내용을 수집한 것이었다. 그러나 시간이 지남에 따른 고객의 신상 변화에 적극적으로 대처하지 못하였고, 그러한 노력을 하지 않았다. 이러한 문제로 기존고객 정보는 과거의 정보를 그대로 보관하고 있는 곳이 대부분이다. 이러한 정보를 수정할 수 있는 방안의 마련이 시급하다.

3) 추출을 위한 반복적인 테스트의 필요

추출은 전체 데이터 웨어하우스 구축 시간의 50~80%까지 소비해야 하는 가장 중요한 작업이다. 추출, 정제, 로드 등은 단순해 보이지만, 데이터의 질과 소스 데이터의 복잡성 및 소스 데이터의 개수, 추출 방법에 따라 난이도는 달라진다. 데이터 웨어하우스 초기에 Inmon은 전체 데이터 웨어하우스 구축의 80%를 추출에 소비해야 한다고 이야기했지만, 현재는 기술의 발전으로 상당히 낮아졌다.

그러나 아직까지도 구축 시간의 50% 정도는 추출에 소비해야 할 정도로 많은 시간과 노력을 필요로 한다. 여기에 추출을 위한 프로그램 개발 시간이 포함된다면, 더욱 많은 시간의 할애가 필요하다 추출이 이루어지면, 무엇보다도 중요하게 진행되어야 할 것이, 데이터에 대한 整合性 검증이다. 정확히 정제되지 못하고 추출된 데이터나, 로직의 오류로 인하여 잘못 생성된 데이터, 그리고 데이터의 질 때문에 발생하는 문제 등으로 데이터 웨어하우스는 잘못된 데이터(garbage) 웨어하우스가 되고 만다. 이러한 이유로 추출 단계는 반복적인 테스트를 통해 오류의 소지를 없애

야 한다.

4) 요약 정보를 위한 granularity(粒度)의 정의

데이터 양의 증가, 사용자 수의 증가, 질의의 형태에 따라 다양한 사용 형태를 발견하게 된다. 이에 따라 발생될 수 있는 문제점이 시스템의 성능이다. 성능 향상을 위해서 중요한 요소 중의 하나는 요약 정보를 생성하는 것이다. 그러나 요약 정보를 모든 사용자의 사용 형태에 따라 만들다 보면, 성능 향상에 도움이 되지 않는 정보들이 디스크 공간을 차지하게 되는 경우가 많다

이러한 요소를 제거하기 위한 좋은 방법은 적당한 granularity를 이용하는 방법이다. 과연 요약 정보가 많을수록 성능이 빠르냐 하는 의문을 가질 수 있는데, 예를 들어 1천 건의 레코드를 저장하고 있는 테이블과 1만 건의 레코드를 저장하고 있는 테이블에 같은 질의를 하였을 경우에 이론적으로 1천 건을 가지고 있는 테이블에 대한 질의가 10배 빨라야 한다. 그러나 현실적으로는 두 테이블에 대한 질의의 속도는 거의 비슷하다. 이유는 1천 건을 검색하는 속도나 1만 건을 검색하는 속도는 페이지나 블록 단위로 처리하는 시스템에서는 디스크를 읽는 속도가 거의 차이가 없다는 것이다. 이 때문에 사용자가 느끼는 성능의 차이는 거의 없다. 그리고 그만큼 시스템의 성능도 개선이 되었다. 그래서 요약 정보 생성에 대한 granularity를 적절한 수준에서 정의하는 것이 바람직하다.

한편으로 조심하여야 할 것으로는 데이터의 쓸림 현상이다. 요약할 하다 보면 특정 부분(특정 월이나 특정 상품 등)에 데이터가 많이 몰릴 경우가 있는데, 이러한 경우에 때로는 요약의 의미를 퇴색시키는 경우가 있으므로 조심해야 한다. 요약한 데이터와 그렇지 않은 데이터의 크기에 별 차이가 없을 경우가 그렇다.

5) 요약 정보의 관리 방안 마련

여러 사례를 통해서 보면, 데이터 웨어하우스 구축 이후, 문제시되는 부분이 요약 정보의 생성과 관리에 관한 내용이다. 요약 정보는 성능과 밀접한 관계를 가지고 있다. 요약 정보의 형태는 EIS나 의사결정을 지원하기 위한 정보를 생성하기 위한 수

단으로 대부분 사용이 되기 때문에 성능상의 문제가 필수적으로 발생한다. 이러한 상황에서 요약 정보를 관리하는 방안이 마련되지 않으면, 관리자는 상당한 고통에 시달리게 된다. 생성된 요약 정보의 내용이 잘못되었을 경우나, 요약 정보 생성이 관리자의 손에 의해 하나하나 수작업으로 되어야 한다면, 데이터 웨어하우스는 관리자에게 고통으로 다가온다. 모든 요약 정보의 생성 및 관리 작업이 도구에 의해 자동으로 이루어져야 하고, 메타 데이터를 통한 스케줄링이나, 각 작업에 대한 관리가 손쉽게 이루어질 수 있도록 하여야 한다.

6) WEB 활용 방안 마련

이제 인터넷은 우리 생활의 일부로 자리하고 있다. 컴퓨터와 네트워크의 발달로 전 세계는 글로벌화 되어 있고, 기업들은 이미 글로벌 마케팅을 위한 경쟁 체제로 방향을 잡아가고 있다. 위성통신 및 고속 네트워크의 발전으로 은행에서도 PC뱅킹, 텔레뱅킹 등의 홈뱅킹 이용이 확산되어 가고 있다. 더욱이 기존의 클라이언트/서버 환경에서 개발된 어플리케이션 형태의 시스템은 관리상의 큰 문제점을 안고 있다.

어플리케이션의 내용이 바뀌었을 때, 사용자들이 사용중인 어플리케이션을 수정해야 한다는 것이다. 물론 버전관리를 통해서 관리한다고 하지만 버그의 수정, 버전 업그레이드 등이 있을 때마다 매번 사용자들은 새로이 수정 및 업그레이드 작업을 해야 한다. 또한 관리자는 모든 사용자들을 관리하기 위해 많은 시간과 비용을 소비해야 한다. 이러한 문제를 간단히 해결할 수 있는 방법이 WEB을 통한 어플리케이션의 개발이다.

5. 檢證 및 테스트 段階

1) 통합 정보 시스템으로서의 역할 확인

데이터 웨어하우스의 본질을 이해하고, 전사 데이터의 통합 시스템으로서의 역할을 제대로 수행하고 있는지를 확인해야 한다. 데이터 웨어하우스가 구축되어 있음에도 불구하고 기존 방식을 고수하거나, 사용하기에 불편해 졌다면, 데이터 웨어하우

스의 역할을 제대로 하지 못하고 있는 것이다. 개발자는 사용자들이 기존 시스템과의 차별성을 인식하고, 개선된 프로세스를 통해서 작업이 이루어지고 있는지 파악해야 한다

2) 사용자 사용 패턴의 발견

데이터 웨어하우스의 장점 중의 하나는 정형이나 비정형 질의에 대한 대응 능력이 좋다는 것이다. 대부분 사용자들의 사용 패턴은 일정한 형태를 취하지만, 간혹 새로운 요건에 따라 패턴이 달라지는 경우가 있다. 이러한 경우는 미리 예측하기 힘든 부분이고, 성능에 가장 많은 영향을 미치는 부분이기도 하다. 대량의 데이터를 다루는 데이터 웨어하우스에서는 성능 향상을 위한 다양한 작업들이 이루어진다. 특히 사용 패턴에 따른 요약 정보의 생성이나 인덱스의 생성, 데이터 파티션 등으로 최대한의 성능을 보장할 수 있도록 한다. 이러한 이유로, 검증 단계에서 사용자들의 사용 형태를 발견하고, 미리 대처할 수 있는 성능 향상 방안을 마련하는 것이 좋다.

3) 철저한 데이터 整合性 및 아키텍처의 검증

시스템의 생명은 데이터의 整合性이다. 결과가 틀린 시스템은 가치를 잃기 마련이다. 사용자를 위해서 개발된 어플리케이션과 각종 사용자 도구들을 이용하여 데이터의 整合性을 검증하고, 필요하다면 소스 데이터와의 비교 검토를 통해 잘못된 곳을 찾아 해결 방안을 마련한다. 일반적으로 데이터의 整合性이 맞지 않은 이유는 각 테이블에서 보유하고 있는 코드들의 불일치로 생기는 경우가 많다.

예를 들어, 소스 시스템에서는 코드의 길이가 2바이트였는데, 데이터 웨어하우스에서는 1바이트로 정의되어 있다든지, 서로의 데이터 타입의 불일치로 인해 테이블 결합의 결과가 틀리게 나타나는 문제 등이다. 이와 더불어 시스템 아키텍처와 사용자를 위한 OLAP 도구의 기능 및 시스템 성능, 추출 도구의 성능 등 다양한 요소들에 대한 검증 또한 필요하다. 가능하면 사용자들이 편하게 사용할 수 있도록 시스템이 개발되는 것이 바람직하다. 물론 관리적인 면도 보완이 되어야 하겠지만, 우선은 사용자를 위한 시스템이 되어야 한다

6. 移行 段階

1) 구축 후 사용자 만족도 조사

시스템의 구축이 끝나면 사용자들의 관심과 사용 빈도 및 활용도를 측정하는 것이 좋다. 사용자들의 불만은 무엇이며, 어떠한 형태를 많이 사용하고 있는지, 이 시스템을 기반으로 향후 발전 방향을 어떻게 할 것인지 등을 파악하기 위해서는 설문 조사 형태의 사용자 만족도 조사를 실시하는 것이 좋다. 그 결과를 바탕으로 구축된 시스템의 평가 및 효과 분석과 시스템의 활성화 방안을 마련하여 필요하다면 재투자 및 시스템 개선을 위한 노력을 하여야 한다.

2) 철저한 사용자 교육

데이터 웨어하우스가 구축되기 전, 대부분의 사용자는 필요한 정보 생성을 위해서 정보기술 부서에 의뢰해야만 했고, 정보기술 부서에서는 프로그램 코딩을 통해서 원하는 정보를 생성하는 과정을 반복해야만 했다. 이러한 과정은 생산성을 떨어뜨리고, 개발자는 매번 많은 시간을 프로그램 개발에 투자해야 했다. 하지만, 데이터 웨어하우스를 기반으로 다차원 분석 도구와 보고서용 도구를 사용하게 되면, 개발자의 프로그램 개발 시간을 단축시킬 수 있을 뿐만 아니라, 사용자로 하여금 직접 원하는 데이터에 대한 분석을 용이하게 함으로써, 정보기술 부서의 역할을 다양화시킬 수 있다. 결국 사용자들이 어려움 없이 모든 도구들을 사용할 수 있도록 하는 것은 정보기술 부서의 부담을 한층 더는 것이 된다.

3) 지속적인 피드백

한번 구축된 시스템은 수명이 다하는 날까지 구축된 그대로 존재하는 것이 아니다. 사용자 부서의 지속적인 문제점 발견과 이에 대한 해결, 새로운 기술의 접목, 새로운 요구사항의 수렴 등을 통해 계속적으로 성숙하고 발전하는 단계를 거쳐야 한다. 변화하는 환경을 수용하지 못하고 새로움을 추구할 수 없다면 이는 생명력이 없

는 시스템이나 다름없기 때문이다 시스템에 새로운 생명력을 불어넣기 위해서는 지속적인 피드백을 통한 보완 작업이 계속되어야 한다



제6절 데이터 웨어하우스 構築 事例

1. 國內 事例 · 三星카드(株)

국내 신용카드 회사 중에서 DB마케팅을 위한 데이터 웨어하우스를 최초로 구축한 사례로 알려져 있으며, 금융기관을 대상으로 하는 각종 데이터 웨어하우스 구축 사례 발표 시에 대표적으로 소개되고 있는 삼성카드(주)의 사례를 분석한다.³⁵⁾

1) 구축 배경

삼성카드(주)는 국내 카드 매출 2위의 대규모 신용카드 회사이다. 1988년 3월 창사 이래로 꾸준한 고객 유치와 서비스를 지속적으로 추구해 온 결과, 1997년 7월말 현재고객 수 약 500만 명, 카드 발급 수는 700여만 매에 이르고 있다. 다른 카드회사 보다는 차별적이고 효과적인 마케팅 전략을 수립하는데 총력을 기울여 왔던 동사는 최근에 들어서는 신용카드 시장 개방과 심화된 경쟁, 다양화하는 고객들의 요구 등의 금융 환경 변화뿐만 아니라, 정보 처리 기술과 통신 기술의 급속한 발달 등의 기술 환경 변화와 기존 정보계 시스템의 문제점 및 한계를 극복해야 할 필요성이 대두되었다.

이러한 과정에서 강력한 전략 정보 시스템으로 데이터 웨어하우스가 유일한 대안으로 선정되었으며, 1996년 6월 내부적인 검토 작업과 파일럿을 거쳐 1996년 12월부터 全社的인 데이터 웨어하우스 프로젝트에 착수하게 되었다.

2) 대상 업무

대부분의 회사와 마찬가지로 동사도 호스트에 정보계 시스템을 가지고 있다. 정보계 시스템은 주로 정기 보고서 형태의 실적을 조회할 수 있는 帳表 위주로 구성이

35) 김희진, “삼성카드 데이터 웨어하우스”,
<http://dwserver.hit.co.kr/kolee/구축사례/삼성카드.html>

되어 있기 때문에 고객과 매출의 급증에 연계한 다양한 마케팅을 실시하기 위한 적절한 대응을 하기에는 미흡한 면이 많았다. 따라서 현업 사용자나 전산실 모두 기존 시스템에서 다음과 같은 어려움을 겪고 있었다

- 어떠한 보고서가 전략적인 정보로서의 가치가 있을 지에 대한 정보 시스템 대상 기준이 불명확하다.
- 전산실에 요청한 후 짧게는 1일에서 10일까지 소요되는 등 적기에 적절한 보고서를 얻기가 어렵다.
- 개인별, 팀별로 동일한 보고서에 대한 데이터를 중복 작성함으로써 데이터 불일치가 발생한다.
- 보고서마다 한 개 혹은 여러 개의 새로운 프로그램을 작성해야 하므로 자원 및 인력 낭비가 초래된다.
- 고객, 매출, 가맹점의 데이터를 개별적으로 관리함으로써 이들 간의 교차 분석이 불가능하다.
- 회사의 매출에 기여하는 회원 군, 이탈 가능성이 있는 회원 군 등 회원 특성별 적절한 회원 관리가 용이하지 않다.
- DM/TM 대상이 되는 회원별 이익 기준을 설정하기 어렵다.
- 인구 통계 정보, 보험 정보 등 다양한 외부 자료와 연계한 정보 분석이 어렵다.
- 시계열 분석이 어렵다.
- 동일한 보고서에 대해서 담당자별 분석 기준이 다른 경우가 많아 새로운 보고서를 다시 작성해야 한다
- 위와 같은 이유로 산재한 보고서, 채도하는 요청에 대응할 수 있는 정보계 시스템의 유지, 관리가 어렵다

이와 같은 어려움은 정보 시스템을 구축한 회사라면 대부분 체감하고 있는 문제점들이며, 데이터 웨어하우스는 이와 같은 어려움을 해소할 수 있는 전략적 정보 시스템이어야 한다. 데이터 웨어하우스는 절대로 모든 것을 해결해 줄 수 있는 마법의 상자는 아니지만 평범한 들을 일순간에 황금으로 변화시킬 수 있는 가능성을 지니고 있다

데이터 웨어하우스에서는 추상적인 정보에 대한 요구 사항을 구체화하는 것이므로 업무 흐름에 의한 업무계 시스템을 구축하는 것과는 다른 형태의 접근을 요한다. 예를 들면 계정계 시스템을 구축하는 경우 시스템의 흐름은 업무의 흐름과 동일하므로 그것에 맞춘 적절한 프로그래밍을 하면 된다. 물론 응답 시간을 최소화함으로써 생산성을 극대화하는 것이 궁극적인 목표이다.

하지만 데이터 웨어하우스는 앞으로 발생할 수 있는 잠재적인 정보 요구 사항을 수렴하여 잠재적인 데이터베이스를 구축해야 하는데 그 어려움이 있다.

3) 범위 설정

데이터 웨어하우스 구축의 첫 번째 단계는 요구 사항 분석, 즉 현업과의 인터뷰였다. 앞서서도 밝혔듯이 이미 全社의인 데이터 웨어하우스라는 大命題가 전제되어 있었기 때문에 회사 전체의 정보계 시스템으로서 대단한 기대를 받고 있었다. 하지만 데이터 웨어하우스는 '데이터 헛간'이 아니다. '데이터 창고'에 적재할 만한 중요한 데이터별로 우선 순위를 결정하는 것이 가장 중요한 일이었다. 全社의 데이터 웨어하우스라고 해서 한꺼번에 모든 정보를 담을 수는 없다. 자원의 한계는 예전이나 지금이나 마찬가지로 제약 사항으로 존재하고 불필요한 데이터 때문에 정말 중요한 정보를 분석하는데 영향을 받지 않기 위해서이다.

全社의 데이터 웨어하우스의 해법은 전사의 정보 요구 사항을 모두 분석하고 모델링하되 데이터 웨어하우스에는 우선 순위가 높은 것부터 단계적으로 구축하는 것이다. 따라서 동사에서는 데이터 웨어하우스에서 우선적으로 포함되어야 할 정보의 범위를 다음과 같이 설정하였다.

- 정정보고 자료 ; 고정적으로 반복되는 주요 지표, 일선 점포별 영업 실적 제공
- 비정정보고 자료 ; 회원 및 가맹점 교차 분석 자료, 임의의 회원 속성별 요구 분석 자료
- 전략적 마케팅의 활용 ; 고객 지향적 전략 수립 시행, 고객의 사용 실적 분석을 통한 이탈 최소화
- 일선 점포 활용 ; 일선 점포에 마케팅 틀 지원, 우량·이탈·신규 회원 정보를 제공함으로써 점포 영업의 독자성 확보

4) 구축시의 문제점 및 해결 방안

동사의 데이터 웨어하우스 구축에 있어서 당면했던 문제점들은 다음의 다섯 가지로 정리되었으며, 각각에 대한 해결 방안을 도출한 후 개발에 임하게 되었다

① 데이터의 整合性 확인

- 업무계 시스템에서 입력되는 코드 중 미입력된 부분 또는 오류의 정비
- 코드체계의 재정립
- 코드체계의 통합에 따른 새로운 업무 기준 수립

② 개발 대상 업무범위 확정

- 효율성 및 실현가능성, 코스트, 퍼포먼스 등 고려
- 업무별 우선 순위 및 중요도에 따른 범위 결정
- 단계별 업무 반영

③ 대용량의 데이터 처리 시 적정 응답속도 확보

- 적절한 H/W, DBMS, DISK Architecture 채택
- 자주 활용되는 분석 자료는 요약 테이블로 해결
- Ad-hoc 조회 사용자수 엄격한 제한
- 대용량 데이터 처리 작업 - Job schedule 기능 이용
- 향후 부문별 주제별 특화된 데이터 마트 시스템 구축

④ 각종 영입 및 분석 데이터의 기밀 보호

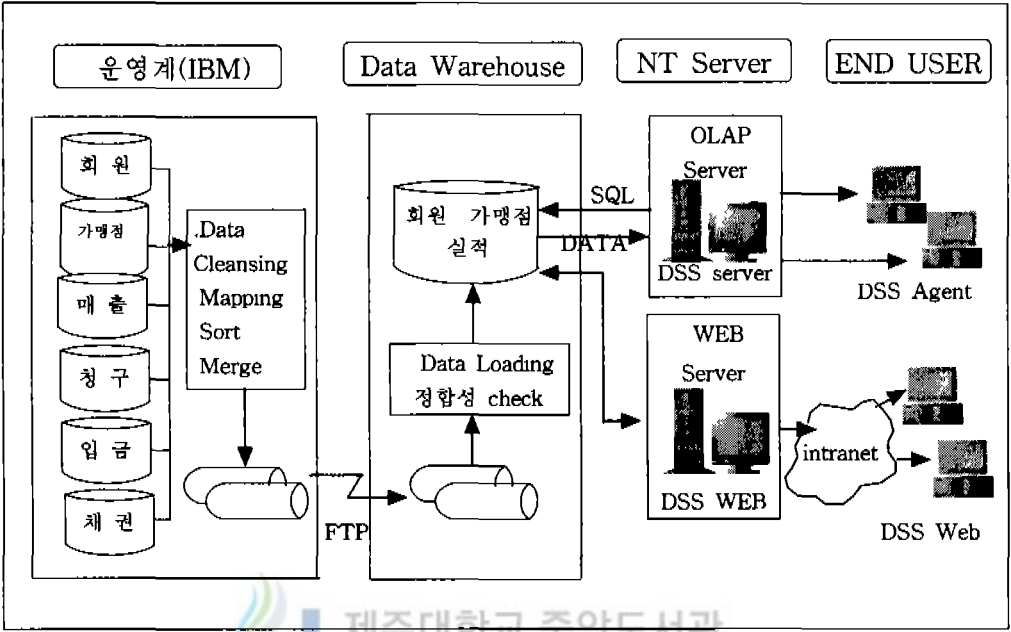
- 각 사용자별 사용권한 레벨에 따른 ID 부여

⑤ 시스템 사용 숙련도 향상

- Pilot 시스템을 이용한 사전 교육
- 시스템 오픈전 기본장표 생성시 개발자와 공동 작업

5) 시스템 개요

<그림 3-6> 삼성카드(주)의 시스템 처리 FLOW



자료 · 이창열, "Database Marketing 구축사례 삼성카드(주)", 서강정보공학연구회, 1999, p.45.

데이터웨어하우스의 테이블을 생성하기 위한 원시 데이터는 운영계로부터 제공된다. 운영계의 COBOL 프로그램으로 IBM 데이터베이스를 읽어서 데이터웨어하우스의 테이블을 위한 마스터 파일을 생성한다. 마스터 파일은 FTP를 사용하여 데이터웨어하우스 서버에 전송(Transportation)되며 오라클의 Loader를 사용하여 코드 변환(Transformation) 및 정제(Cleansing)를 거친 후 테이블로 로드된다. 테이블은 SQL Loader를 사용하여 로드한다.

데이터웨어하우스는 OLAP 툴인 DSS Agent와 웹 서버인 DSS Web을 사용하여 분석된다. DSS Agent는 비정형 보고서와 정보 분석가를 위해 제공되어지며 데이터웨어하우스 서버, DSS Server, DSS Agent의 전형적인 3-Tier 구조로 구성되어 있다.

6) 적용 사례

① 우수 고객 중에서 이탈한 고객에 대한 원인 분석

우수 회원 중에서 이탈한 회원이 누구인지를 파악하여 성별, 연령별, 직업별, 한도 등급별, 입회 년도별 등의 속성 분석을 통해 이탈하는 회원의 특성 분석을 실시하였다

② 추이 분석

특정 업종에 대한 매출액 현황의 추이 분석을 실시하여 수수료율 인하를 결정할 때에 손익에 미치는 영향을 파악하고, 가맹점의 수수료율 인하 요청을 거절한 후 매출액의 변화를 분석 해 봄으로써 손익을 감안한 대응 방안을 강구할 수 있었다.

③ 특정 고객에 대한 DM 행사실시

특정 양복을 구매한 고객 22,000명을 추출하여 할인 쿠폰과 함께 DM을 발송한 결과, 기존의 DM 실시 시의 고객 반응률 0.5%에 비하여 획기적으로 증가한 27.3%의 고객 반응률을 얻었다.



④ 영업활동 자료 제공

각 지역별 직군 또는 직업별 회원 분포를 파악한 후 각 지점의 시장 규모와 비교 분석 후 취약 지역을 선정하여 상담사의 유치 활동을 효과적으로 도모하였다.

7) 구축 효과

동사는 신규·우량·이탈 회원 등 다양한 회원 정보를 가맹점과 매출 정보와 교차하여 분석한 결과 마케팅에서 거둔 효과가 상당하다. 현재까지 동사에서 자체 분석한 데이터 웨어하우스 구축 효과는 다음과 같다.

① 정보 Infra의 구축

기존의 운영계에서 각각 독립적으로 관리하던 각종 정보(회원, 가맹점, 매출, 청

구, 입금 등)를 데이터 웨어하우스에 통합하여 관리함으로써 상호 연관 분석이 가능하게 되는 등 정보 활용을 위한 기반이 마련되었으며, 데이터 웨어하우스에서는 데이터의 무결성을 위하여 단일 코드 체계를 채택함으로써 일관된 정보의 관리가 가능하게 되었다.

② 정보처리 생산성의 증대

통상 7~10일이 소요되던 정보의 확보 시간이 1~2일, 혹은 몇 분내로 단축되었으며, 기존에는 기초 데이터를 이용하여 스프레드시트 등에서 분석 보고서 작성을 위한 재가공 업무를 해야 했으나 OLAP 도구를 이용하여 원하는 보고서를 바로 얻을 수 있게 되었을 뿐만 아니라, 현업 전문가가 데이터 웨어하우스를 직접 검색할 수 있게 되었으며, 정보 활용의 기회가 확대되었기 때문에, 보고서를 위한 정보기술 부서의 인력 낭비를 최소화하여 업무 생산성이 증대되었다.

③ DB마케팅의 기반 구축

고객 정보를 체계적으로 데이터 웨어하우스에서 관리함으로써 원하는 정보를 검색·가공·분석하여 목표 고객을 출력·활용할 수 있게 되었으며, 일선 지점에 우량·이탈·신규 회원 등의 다양한 정보를 제공함으로써 영업에서 유리한 고지를 점할 수 있고, 추세 분석을 비롯한 다차원 분석이 가능하게 되어 회원, 가맹점, 매출, 입금간의 교차 분석이 가능하게 되었다.

8) 향후의 과제

데이터 웨어하우스는 정보가 저장된 데이터 혹은 테이블의 집합이며 이것을 이용하여 결과는 알 수 있지만 원인을 분석할 수는 없다. 이러한 내용을 분석하기 위한 통계학적 혹은 데이터 마이닝적인 접근은 앞으로 해야 할 숙제로 남아 있다. 앞으로 데이터 웨어하우스에 추가되거나 보완되어야 할 사항이나 데이터 웨어하우스를 기반으로 구축해야 할 시스템을 살펴보면 다음과 같다.

① 마케팅 수익 분석

- 회원, 가맹점, 상품별 기여도의 평가
- 마케팅 결과에 대한 반응을 분석
- 영업실적 분석, 스코어링 시스템의 도입

② 업무별 데이터 마트의 구축

- 정형화 업무(경영 지표 등) 다차원 분석
- 부서별 목적 데이터 베이스(수익기여도, 반응률 분석, 세분화 정보 등)

③ 효율적인 시스템 관리

- 통합 메타 데이터 관리
- 추가 시스템의 데이터 통합 관리
- 비정형 보고서(ad-hoc query)의 성능 향상

9) 결론



동사의 데이터 웨어하우스가 성공적이라는 평가를 받게 된 것은 무엇보다도 데이터 웨어하우스가 회사의 사활과 관련된 전략적으로 중요한 시스템임을 인식한 최고 경영자의 전폭적 지지로부터 시작한 全社의인 공감대의 형성에 의한 현업, 개발팀의 공동 작업의 결과라고 볼 수 있다. 프로젝트 초기만 하더라도 데이터 웨어하우스에 대한 오해, 환상이 팽배해 있었고 데이터 웨어하우스의 장점을 최대화하고 단점을 극복하는 여러 사람의 노력에 의해 데이터 웨어하우스가 구축되었다.

성공적인 데이터 웨어하우스의 구축은 동사에서와 같이 데이터 웨어하우스가 회사의 전략과 더불어 발전되며 단기간의 개발로 효과를 검증 받을 수 없고 최소한 1~3년의 유예 기간을 두고 지속적인 경험의 축적과 투자를 해야만 비로소 결실을 볼 수 있을 것이라는 인식의 전환이 되어야만 가능하다.

2. 海外 事例 : 체이스맨하탄 銀行

국내 은행은 현재 데이터 웨어하우스를 추진 중에 있거나 검토 단계에 있기 때문에 구축 사례로 발표된 예가 없다. 따라서, 해외의 사례 중에서 미국의 체이스맨하탄 은행의 사례를 분석하여 은행의 데이터 웨어하우스 구축에 참고하고자 한다.³⁶⁾

1) 구축 배경

체이스맨하탄 은행은 총 자산 규모가 3,655억 달러에 달하는 미국 내에서 가장 큰 은행 지주회사이다. '90년대 중반 이후, 동행은 대규모의 인수합병(매뉴팩처러 하노버, 케미컬 은행 합병)을 통하여 현재의 규모로 성장할 수 있었고, 미국 내에 3,200만명의 가구를 고객으로 확보하고 있다.

동행이 데이터 웨어하우스 프로젝트를 추진하게 된 이유는, 기존에 많은 문제점들을 안고 있었기 때문이었다. '90년대 초반부터의 인수합병으로 각기 다른 고객들을 광범위하게 볼 수 있는 시스템을 개발하는 것은 상품 판매뿐 아니라, 서비스 제공과 기존고객에 대한 유지 마케팅(Retention Marketing)을 위해 절박한 상황이었다.

또한 합병되기 전의 각 고객의 거래 이력을 이해하는 것은 동행이 유지 마케팅을 하기 위해서는 필수적인 사항이었다. 동시에 사업 확장과 신규 고객확보에도 많은 관심을 가지고 있었기 때문에 데이터 웨어하우스는 잠재고객을 식별하고 효과적으로 권유하는데 좋은 도구가 될 수 있었으며, 대상이 된 고객에 대해 교차판매를 용이하게 해준다는 사실을 확신하고 있었다

한편으로는 데이터 웨어하우스와 많은 분석 도구를 사용함으로써, 근본적으로 업무처리 방식이 바뀌고, 단순한 영업적 관점뿐만 아니라 서비스의 질적 개선까지도 고려한 방향이었다. 동행의 의도는 거의 모든 데이터에 대한 고객 정보를 데이터 웨어하우스를 통해서 액세스할 수 있게 하는데, 웨어하우스는 다시 30여개의 데이터 마트로 업무별로 세분화되어 있다.

36) 이진권, “체이스맨하탄은행의 DW 구축사례”, 시사컴퓨터 11월호, 1998.

2) 프로젝트 계획

동행은 기존 고객 및 잠재 고객에 대한 서비스와 판매 증대 방법을 전환시키기 위해 데이터 웨어하우스 프로젝트를 시작하게 되었는데 이에 대한 추진 전략으로, 여러 단계로 세분화된 작업 공정을 통해 프로젝트의 리스크를 최소화했으며, 단계적으로 10개의 공정을 1996년 10월부터 1999년 6월까지 설정했다.

다단계 개발공정 방법으로 추진함으로써 동행은 전체 프로젝트를 효율적으로 관리할 수 있게 될 뿐 아니라 규칙적인 추진일정의 설정으로 한 단계에서 구축 성과물을 조정해도 프로젝트 전체 진행에 위협을 끼치지 않을 수 있다.

3) 데이터 웨어하우스 아키텍처

가장 흥미 있는 데이터 웨어하우스 아키텍처는 데이터 마트의 활용 부분이다. 동행의 데이터 마트는 특정 부서에서 사용할 데이터를 구체적인 업무에 국한시켜서 데이터 웨어하우스로부터 가져와서 구축된다. 현재 계획된 데이터 마트는 30여개 정도로 부서당 2~3개의 마트를 가진다.

◆ 데이터 마트

동행의 궁극적인 목표는 이렇게 세분화된 데이터 마트를 많은 현업 사용자가 자유자재로 쓸 수 있는 환경을 구축해 주는 것이기 때문에 웨어하우스 아키텍처의 유연성은 무엇보다도 중요한 사항이다. <그림 3-7>은 데이터 웨어하우스와 데이터 마트 아키텍처의 개요를 나타내고 있다.

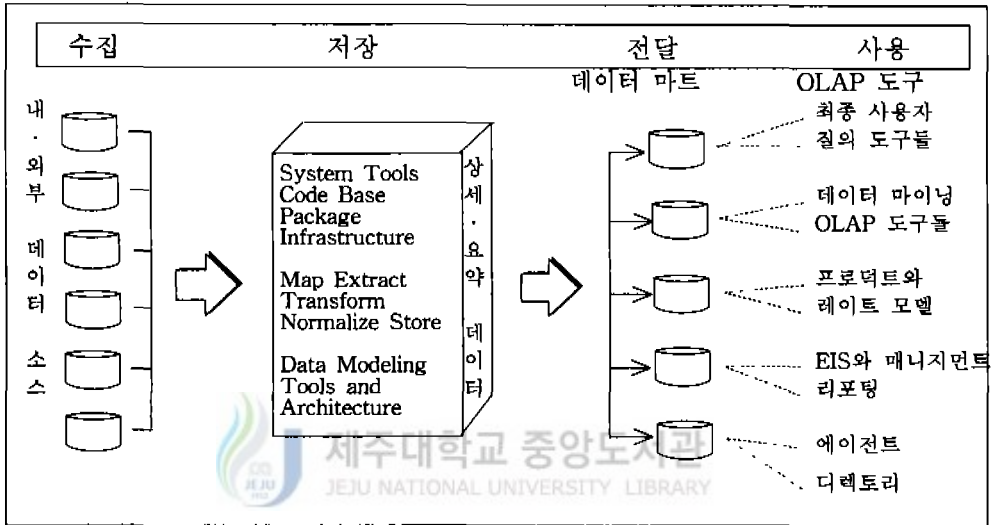
◆ 분석 도구

데이터 마트와 이를 사용하는 최종 사용자간의 인터페이스는 매우 특이하다. 여러 종류의 데이터 마트를 이용목적과 데이터 타입에 맞는 다양한 분석도구를 제공하여 준다. 이를 이용하는 이용자 그룹도 상품 개발팀, 가격 분석팀, 지역별 영업관리팀, 콜센터 등 다양하고, 분석하는 마트 정보도 여러 가지 종류이다. <표 3-2>는 현재 사용하고 있는 분석 도구와 최종 사용자 그룹 개요를 나타내고 있다.

◆ 하드웨어와 소프트웨어

개념적 설계 시 2개의 물리적 데이터 웨어하우스를 하나의 통합 데이터 웨어하우스로 구축하는 방향으로 구성되었기 때문에 여러 개의 데이터 마트에 대한 표준 공급 매커니즘이 필요했고, 이런 원칙에 의해 하드웨어와 소프트웨어가 선정되었다.

<그림 3-7> 체이스의 데이터 웨어하우스/데이터 마트 아키텍처



자료 : 체이스맨하탄 은행

<표 3-2> 체이스에서 사용된 분석 도구들

최종 사용자 그룹	사용자 수	도구들
Management group (Regional Sales)	10 여명	Lotus Notes, Power Play Cubes
Power Analyst (Product Pricing)	10 여명	Oracle Express Cubes, Power Play Cubes Knowledge Seeker, SGI Mineset Information, Advantage SAS
Market Analyst (Product Manager)	100 여명	Oracle Express Cubes, Power Play Cubes MS Access & Excel, Business Objects
Sales and Service Personnel	1000 여명	Oracle Express Cubes, Power Play Cubes MS Access & Excel

자료 : 체이스맨하탄 은행

하드웨어와 소프트웨어를 선정하는 과정에서 데이터 정합성과 상호 연계성은 생각보다 적게 고려해도 되었는데, 이는 전체 설계 과정에서 데이터 웨어하우스의 신속성과 확장성, 안정성, 가격적인 측면을 고려해서 구성했기 때문이었다. 데이터 마트 구성은 현업 부서와 협의해서 결정하도록 했다.

4) 프로젝트의 특징

동행의 프로젝트 특징은 다음의 다섯 가지로 요약할 수 있다.

① 3,200만 가구의 정보를 데이터 웨어하우스로 구현하는 것은 크기나 규모, 업무 범위에서 많은 위험을 가지고 있다는 사실을 개발팀은 인지하고서 프로젝트를 시작했다.

② 프로젝트의 리스크를 최소화하기 위하여 개발팀은 두 개의 핵심 전략을 수립했는데 첫째로, 물리적으로 두 개의 데이터 웨어하우스 시스템으로 분리해서 구축하고 개념적으로는 하나의 시스템으로 묶어주는 것이고, 둘째로는 데이터 웨어하우스 개발단계를 계획하고 추진하는 과정에 있어서 'Chunk and Conquer'(Divide and Conquer의 반대 개념)에 의해 추진하였다. 이는 작게 시작하고, 각각의 단계를 분리해서 추진하는 것으로서 연속적으로 데이터를 정제하고 데이터 웨어하우스 시스템 조각을 붙여나갈 수 있고 또한 프로젝트의 리스크를 최소화시킬 수 있었다.

③ 두 개의 데이터 웨어하우스는 은행의 각 현업 부서(business unit)와 아주 긴밀하게 제휴 관계를 맺고 있는데, 하나는 고객 대출정보(customer credit) 데이터 웨어하우스이고 또 하나는 수신 및 투자 정보(deposit & investment) 데이터 웨어하우스이다. 두 개의 데이터 웨어하우스에 대한 개발비용은 현업 부서에서 직접 부담하였고, 개발팀은 명확히 현업 스폰서의 요구사항과 기대하는 바가 무엇인지 알고 있었다.

④ 동행은 개발팀을 감독하기 위한 일련의 위원회를 결성하였는데, 프로젝트 각

개발 단계를 계획하고 업무 요구사항에 대해 허가를 해주는 업무 정보 소위원회 (business intelligence subcommittee)와 전체 웨어하우스 사용자가 사용할 데이터와 포맷을 결정하는 액세스 소위원회(access subcommittee)로 구성되어 있다. 또한 각 소위원회의 활동을 정보기술 부서와 마케팅부서 책임자가 후원하도록 되어 있다.

⑤ 동행은 현재 수백 명이 데이터 웨어하우스(아주 세분화된 데이터 마트를 통해서) 시스템을 사용하고 있다. 동행은 광범위하고 손쉬운 데이터 활용을 통해서 조직을 보다 효율적이고 효과적인 방향으로 경쟁력을 갖추어 갈 수 있도록 하고 있다.

5) 구축 효과

동행이 데이터 웨어하우스 프로젝트를 시작하게 된 근본적인 이유는 '원천적인 사업 방식의 개혁'에 있었다. 또한 단위 부서의 업무 지원뿐 아니라 각 조직간의 원활한 정보제공도 최고 경영자의 데이터 웨어하우스 추진 목표 중의 하나이다. 현재 동행은 이미 데이터 웨어하우스를 사용하면서 많은 효과를 보고 있고, 이러한 효과들은 수치적으로 보여지는 것보다 프로세스상의 개선효과가 더 크다 하겠다.

마케팅부서의 예를 들자면 과거엔 엑스레이로 진단했다면, 이젠 MRI와 같은 도구를 제공해 준다고 할 수 있겠다. 즉, "우리 고객이 누구이며 무엇을 원하는가?"라는 질문에 보다 심층적으로 분석해서 처방을 내릴 수 있게 되었고, 이 처방에 대한 치료도 명확하고 신속하게 이루어질 수 있게 되었다. 다음의 내용들은 동행에서 평가하는 일반적인 구축 효과를 나타낸 것이다

- 과거 50여 개의 상이한 데이터 소스를 단일 고객정보로 통합함
- 데이터 웨어하우스 데이터의 整合성을 모든 뷰에서 일치시키는데, 일치된 뷰를 소스 시스템과 은행 회계 계수와 조정하여 계산 원칙을 표준화함
- 인트라넷을 통하여 데이터 디렉너리에 등록된 정보를 공유해서 조회할 수 있게 함
- 월 마감 작업을 15일로 단축시킴
- 전월 데이터부터 데이터 웨어하우스 시스템에서 이용할 수 있는데, 과거엔 3개

월 전 정보만을 볼 수 있었음

- 일별, 주별, 월별, 분기별 주기적으로 보관을 해야 하는 데이터는 현업 부서에서 요청을 하고 요청된 내용을 프로젝트 위원회에서 검토하여 결정하게 됨

현업에 대한 만족도를 가장 잘 나타내는 것은 현업 부서가 계속해서 데이터 웨어하우스 프로젝트에 비용을 지원하고 있는 사실일 것이다. 각 업무 부서는 데이터 웨어하우스 시스템의 투자 효과를 빠르고 확실한 결과로 도출하기 위해 지금도 다양한 정보를 요구하고 있다.

6) 결론

동행의 정보기술 부서는 매뉴팩처러 하노버와 케미컬 은행의 인수 합병으로 인하여 무엇보다 정확하고 통합된 데이터를 핵심 시스템으로 가지고 있어야만 했고, 각각의 데이터 요소에 대한 문서화는 필수적인 것이었다.

또한 동행의 크기나 규모를 고려해 볼 때 데이터 웨어하우스 프로젝트의 리스크는 매우 높았으며, 이를 최소화하기 위해 물리적 데이터 웨어하우스를 두 개로 분리해서 구축했고, 분리된 두 개의 업무영역(Chunk)은 각각의 업무에 맞게 프로젝트 진행이 되었으며 비즈니스 인텔리전스 소위원회와 액세스 소위원회에서 두 개의 시스템에 대한 통합 뷰를 컨트롤 해주는 역할을 했다.

마지막으로, 동행의 데이터 웨어하우스는 최고 경영자의 전폭적인 지지와 현업 부서와의 긴밀한 업무 협조가 매우 효과적이었고, 이는 현업 부서가 가시적인 업무개선 효과 없이 전체 데이터 웨어하우스가 완료될 때까지 기다려 주지 않을 것이라는 것을 개발팀은 잘 알고 있었고, 이를 대비하여 단계별 데이터 마트를 제공해 줌으로써 지속적인 업무 부서의 참여를 유도할 수 있었다.

7) 국내 은행에의 시사점

국내의 여건과 비교해 볼 때 우리가 고민하고 있는 것을 선진 금융기법을 구사하는 그들도 똑같이 고민하고 있다는 것을 동행의 사례를 통해 알 수 있다. 하지만 접

근 방법에 있어서는 많은 차이점을 발견할 수 있다.

첫째, 데이터 웨어하우스에 대한 개념의 차이이다. 사례에서 보듯이 동행에서는 全社의인 엔터프라이즈 데이터 웨어하우스 개념에서부터 시작해서 단계별로 전사 데이터를 하나의 뷰에 의해 모으는 작업을 했고, 이를 통해 마트 형태로 최종 사용자가 이용할 정보를 순차적으로 제공해주고 있다.

둘째, 추진 방법이다. 특징에서 언급한 것처럼 ‘Chunk & Conquer’ 방법의 핵심은 위험의 분산과 현업의 적극적인 참여 유도에 있다. 특히 현업의 무관심과 주도적인 리드없이 정보기술 부서의 의욕만이 앞설 경우 실패로 가는 지름길이라는 사실을 인식해야 한다. 이러한 점에서 점진적 접근 형태의 동행의 추진 방법은 좋은 사례가 될 것이다.

마지막으로, “우리 은행 데이터 웨어하우스의 핵심은 무엇인가?” 라는 원론적인 질문에 대한 대답은 데이터 웨어하우스 시스템을 사용할 현업 사용자 부서에서 내려주어야 하며, 그 해답은 당연히 해당 부서 업무의 핵심일 것이다. 동행의 사례에서 보듯이 무엇을 분석할 것인가에 따라 분석 도구나 소프트웨어는 얼마든지 바뀔 수 있고, 그에 대한 결정에는 이용할 사람의 뜻이 충분히 반영되어야 할 것이다.



제 4 장 DB마케팅을 위한 데이터 웨어하우스 模型 設計

본 장에서는 지금까지 살펴본 DB마케팅과 데이터 웨어하우스에 관한 선행 연구 내용과 실제 사례를 토대로 하여 은행에 적합한 시스템 구축 방안을 제시해 보고자 한다.

제1절 銀行 데이터 웨어하우스 具現 方式

데이터 웨어하우스의 구현 방식으로는 첫째, 미국형 접근 방식으로서 필요 부문별 데이터 마트를 먼저 구현하고 점차적으로 이들 데이터 마트를 은행 데이터 웨어하우스로 통합해 나가는 방식인 상향식 접근법(Bottom-Up)과 둘째, 유럽형 접근 방식으로서 은행 수준에서의 데이터 웨어하우스를 먼저 구현하고 나중에 데이터 마트를 구현하는 방식인 하향식 접근법(Top-Down)³⁷⁾이 있다.

앞의 사례 연구에서 살펴보았지만, 신용카드 회사와 같이 업무 범위가 단순하고 전체 업무를 통합적으로 분석해야 하는 경우(통신회사, 백화점 등)에는 전사 수준의 데이터 웨어하우스를 먼저 구현하는 하향식 접근법이 절대적으로 필요하다. 현업 부서별 업무별 정보 요구가 있을 때에는 적절한 규모의 데이터 마트를 만들어 줌으로써 현업 사용자들이 직접 접근하도록 할 수 있다.

그러나 이러한 구축방식은 비용과 시간이 많이 들고, 구축 성공에 대한 보장이 없기 때문에 최고 경영층의 전적인 지원을 받기 어렵다는 단점을 가지고 있다. 이를 해소하기 위한 방안이 주제별로 단계적으로 데이터 웨어하우스를 구축하는 것이다. 즉, 단계적으로 구축함으로써 각 단계에 대한 구축의 성공여부를 조기에 결정할 수 있게 되며, 따라서 최고 경영층에서 의지를 갖고 지원할 수 있게 되는 것이다.

체이스맨하탄 은행의 사례가 여기에 해당된다. 먼저 은행 전체적인 데이터 웨어하우스의 큰 밑그림을 그려 놓고 모자이크 그림을 맞추어 나가듯이 통합 데이터 웨어하우스의 형태로 진행하고 있는 것이다. 또한 사용자 부서의 참여도를 높이기 위하

37) B. Devlin, "From data model to data warehouse, in bite-sized chunks!" InfoDB, Vol. 9, No 5, 1995.

여 데이터 웨어하우스 구축과 함께 단계별로 데이터 마트를 제공해주고 있다.

한편, 데이터 마트를 먼저 구축하면 실제적으로 비용과 시간을 단축시킬 수 있다. 가트너 그룹에 따르면, 데이터 마트는 데이터 웨어하우스보다 최고 80%의 비용이 절감될 수 있다고 한다. 그러나 일반적으로 특정 부서의 정보 욕구를 만족시키기 위하여 구축되는 데이터 마트는 회사의 나머지 부분이 공유할 만한 매력을 제공하지 못하며, 서로 호환되지 않는 기술, 데이터의 중복, 엄청난 시스템 유지보수 노력 등과 같은 혼란을 가져올 수 있다.

데이터 웨어하우스 소프트웨어 업체인 파인 콘 시스템 社에 따르면 데이터 마트를 구축할 경우 기업은 문제가 커진 후에야 문제가 있다는 알게 된다고 한다. 즉 세 번째나 네 번째 데이터 마트를 구축한 후에 데이터를 통합하거나 유지보수를 할 때에 문제가 드러난다는 것이다. 많은 예산이 기술적 세부사항에 투자되기 때문에 이 같은 문제가 드러날 경우, 일은 각 현업 부서가 했지만 결국 비난과 책임 및 정리작업은 정보기술 부서가 져야만 한다.³⁸⁾

이상의 내용을 종합해 볼 때, 은행 데이터 웨어하우스의 구축은 체이스맨하탄 은행의 사례를 따르는 것이 효과적이라고 보여진다. 즉, 은행 전산인력의 기술 경험이 미흡하고, 데이터 웨어하우스 프로젝트의 국내 공급업체 지원 능력도 검증되지 아니한 상황이므로, 체이스맨하탄 은행에서와 같이 단계별로 데이터 웨어하우스를 구축하고, DB마케팅을 위한 데이터 마트를 활용할 수 있도록 하는 것이다.

또한, 데이터베이스의 선택에 있어서는 대용량의 데이터를 저장해야 할 경우에는 관계형 데이터베이스를, 비교적 독립적인 분야에서 단기간에 응용시스템을 개발하여 활용할 필요가 있거나 사용자가 빠른 응답을 요구하는 경우에는 다차원 데이터베이스를 선택하는 것이 바람직한 것으로 여겨진다.

38) Special Report, 데이터 웨어하우스 50%가 실패작 어떻게 극복할 것인가
<http://www.cio.seoul.kr/990401/main14.html>

제2절 銀行 데이터 웨어하우스 設計

은행의 데이터 웨어하우스는 개념적으로는 모든 데이터가 통합된 하나의 데이터 베이스이지만 물리적으로는 여러 개로 분할되어 설계되어 진다.

즉, 고객 데이터 웨어하우스는 고객 구조분석을 통하여 고객을 세분화하고, 소매 금융 마케팅 및 영업 전략을 수립하기 위한 기본자료를 제공하는 것이다. 은행의 고객 분석보고서에 포함되는 주요 내용을 다음의 <표 4-1>과 같이 들 수 있다.

<표 4-1> 은행 고객 분석보고서의 주요 구성 내용 예

구분	가계 금융	기업 금융
내용	<ul style="list-style-type: none"> - 수신 상품별 고객 수 및 수신 금액 - 고객 성별, 연령별 고객 수 및 수신 금액 - 고객 성별, 연령별 거래 기간, 거래 규모, 거래계좌 수, 부대서비스 사용 실적 - 거래기간별 고객 수 및 수신 금액 - 거래기간별 거래 규모, 거래 계좌 수, 부대서비스 사용실적 - 거래규모별 고객 수 및 수신 금액 - 거래규모별 연령구성, 부대서비스 사용실적 - 연동계좌 보유 현황 	<ul style="list-style-type: none"> - 연도별 기업 고객 수와 수신 및 여신 - 업종별 기업 고객 수와 수신 및 여신 - 거래기간별 기업 고객 수와 수신 및 여신 - 지점별 기업 고객 수와 수신 및 여신 - 수신 및 여신 규모별 기업고객 - 당좌거래 기업고객 - 할인어음거래 기업고객

고객 데이터 웨어하우스는 위의 표에 나열된 다양한 정보들을 정확한 시계열 데이터로 저장하고 있어야 한다는 점이 매우 중요하다. 과거 어느 한 시점의 데이터 값은 동일하며 불변이다. 정확한 시계열 데이터에 대한 요구가 새로운 요구는 아니다. 과거에도 정기적, 월별 보고서 형식으로 제공되었으며, 이러한 자료를 정리하여 모아두었다가 필요시에 사용하였다

이러한 요구에 맞는 데이터 웨어하우스를 설계하기 위하여 다음의 두 가지 방법을 고려하여야 한다. 개체 데이터 웨어하우스 설계와 차원적 데이터 웨어하우스 설계가 그것이다

개체모형은 정규화 된 관계형 데이터베이스 모델링 방법으로 설계시의 데이터 중복을 최소화한다. 이 모형은 완전하고 유지가 용이한 데이터베이스를 제공하기 때문에 한 시점의 데이터를 기록하고 이들을 모아 저장하는 것을 목적으로 하는 데이터 웨어하우스에 적합하다. 즉, 개체모형 접근으로 원천 데이터를 저장하는 개체 데이터 웨어하우스를 작성하게 되는데, 이는 수정이 거의 없는 데이터를 저장한다는 점에서 거래처리 시스템의 운영 데이터베이스와는 차이가 있다.

개체 데이터 웨어하우스는 아카이브(archive) 데이터를 저장하기 위해서 구축되며, 가장 중요한 목적은 데이터의 상실을 방지하는 데 있다. 또한, 원천 데이터의 변형, 정제 혹은 통합 등이 많이 일어나는 경우에도 아카이브 데이터는 매우 중요하다. 사용자들은 아카이브 데이터로부터 정보를 검색할 수 있지만 분석기능 및 보고서 작성은 매우 제한적이다.

차원적 모형에 의한 접근방식은 데이터 웨어하우스에 접근하여 데이터를 다양한 방식으로 검색해 보려고 하는 사용자들의 검색 경로와 일치하는 데이터베이스 설계를 제공한다. 자주 요청되는 총계치 혹은 계산 값은 미리 계산되어 데이터베이스에 저장되기 때문에 데이터 중복을 허용하게 되나 보고서가 작성될 때마다 계산들이 다시 수행될 필요가 없기 때문에 이들 데이터를 즉시 제공할 수 있게 된다. 데이터 마트는 부문별 사용자들에게 데이터를 보다 효율적으로 제공하기 위해 중앙 집중적 데이터 웨어하우스로부터 데이터를 공급받아 사용자에게 제공하는 데이터 배분센터로 생각할 수 있으며 따라서 데이터 마트도 차원적 데이터베이스로 설계된다.

개체모형이 레코드별 데이터 접근이라고 한다면 차원적 모형은 OLAP에서 요구하는 필드별 데이터 접근을 제공하여 OLAP 응용에서 요구하는 유연성, 재사용성, 그리고 성능 요건을 만족시킨다. 데이터 분석을 위해서는 데이터가 거래발생 수준에서 아니라 분석 차원의 단위에서 저장되고 나누어질 수 있어야 한다.

은행의 경우 기간별, 상품별, 지역별 그리고 고객별 평잔과 계좌 수의 추세에 많은 관심이 있다. 이러한 경우 상품, 고객, 지역, 기간이 각각 데이터 분석에 있어서의 서로 다른 차원으로 이해된다.

다음의 <그림 4-1>은 차원적 모형 설계의 예를 나타내었다 기간, 상품, 지역, 고객 차원에 대한 각각의 차원 테이블이 정의되며, 각 차원 테이블은 차원에 대한 설명과 차원 내의 그룹화를 나타내는 차원 구조에 대한 속성, 기타 여러 속성 등을 포함한다

자료 테이블은 표시된 각 차원수준에서의 시계열 자료를 포함한다. 예를 들어 지역차원 테이블이 지역본부, 지점, 지역특성을 나타내는 구조를 갖고 있으며 따라서 자료 테이블도 지역본부, 지점, 지역특성 수준에서의 평잔과 계좌 수 데이터를 모두 포함한다. 즉, 최하위 수준의 데이터가 거래발생 수준에서가 아니라 분석시간 단위로 나누어져서 요약되어 각각의 분석차원 수준별로 데이터가 저장되는 것이다.

<그림 4-1> 차원적 모형 설계의 예



자료 : 정철용 · 함유근, 전계서, p.75.

다음의 <그림 4-2>는 위의 차원적 모형에 대한 자료 테이블을 나타낸 것이다. 일반적으로 숫자 형태로 표시되는 분석차원 키로 색인된 시계열 데이터가 테이블로 표현된다. 자료 테이블은 기간, 상품, 지역, 고객의 분석 차원의 기본키들로 구성되고 색인된 복합 기본키를 갖고 있다.

<그림 4-2> 자료 테이블의 예

복합기본키

기간키	상품키	지역키	고객키	평잔	계좌수
199911	1001	201	301	₩100,000	2
199912	1001	201	301	₩110,000	2
200001	1001	201	301	₩200,000	3
200002	1001	201	301	₩210,000	2
200003	1001	201	301	₩240,000	4

<그림 4-3> 차원 테이블의 예

[기간 차원]

기간키	기간 설명	연도	분기	월
199911	1999년 11월	1999	4/4분기	11월
199912	1999년 12월	1999	4/4분기	12월
200001	2000년 1월	2000	1/4분기	1월
200002	2000년 2월	2000	1/4분기	2월
200003	2000년 3월	2000	1/4분기	3월
19993	1999년 3/4분기	1999	3/4분기	NA
19994	1999년 4/4분기	1999	4/4분기	NA
20001	2000년 1/4분기	2000	1/4분기	NA

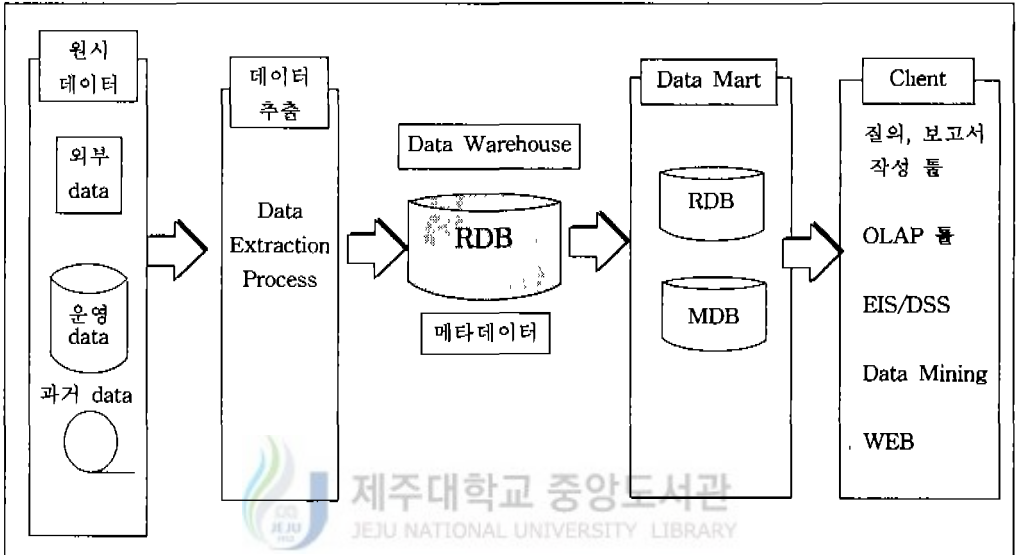
[상품 차원]

상품키	대분류	중분류	계정과목	수준
1001	수신	요구불예금	보통예금	계정과목
1002	수신	요구불예금	저축예금	계정과목
1101	수신	저축성예금	정기예금	계정과목
1102	수신	저축성예금	정기적금	계정과목
1099	수신	요구불예금	NA	중분류
1199	수신	저축성예금	NA	중분류
1999	수신	NA	NA	대분류

제3절 데이터 웨어하우스 基盤의 DB마케팅 시스템 模型

지금까지의 연구를 종합하여 은행의 데이터 웨어하우스는 다음의 <그림 4-4>와 같이 데이터 마트를 활용하는 일반적인 모형을 이용하도록 한다

<그림 4-4> 데이터 웨어하우스 모형



자료 조재희·박성진, 전개서, p38.

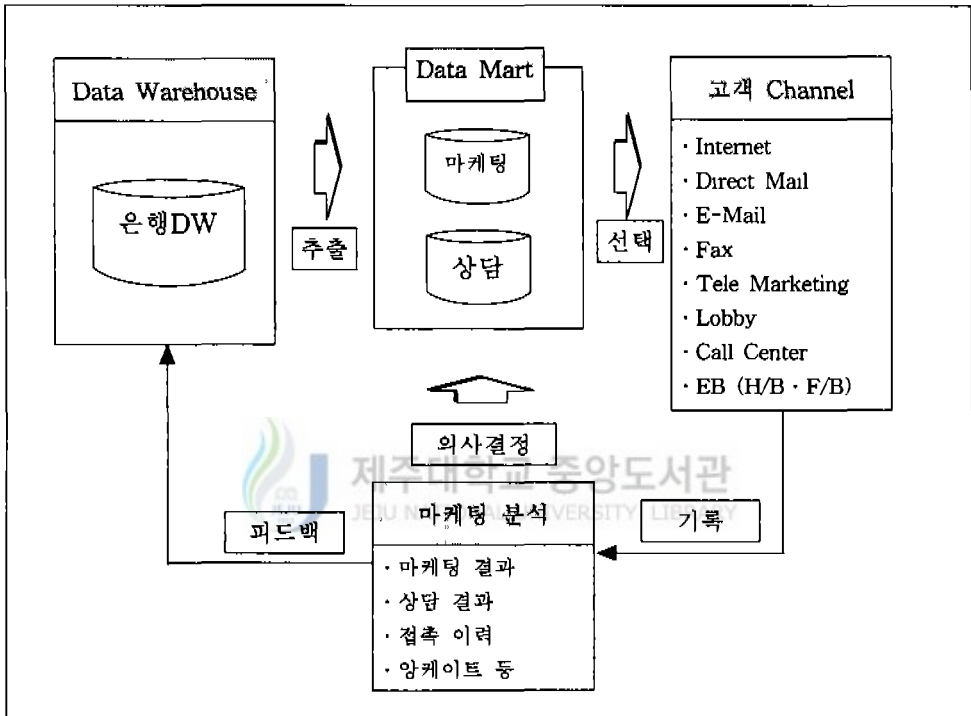
다음의 <표 4-2>는 위의 시스템 구성 요소들에 대한 보충 설명이다.

<표 4-2> 시스템 구성 요소

Component	내 용
Data Extraction	원시 데이터로부터 데이터를 추출, 정제, 로드하여 데이터 웨어하우스에 수록하는 작업을 수행
메타 데이터 관리	데이터 구조, 데이터 요약 알고리즘, 운영 DB와 데이터 웨어하우스와의 매핑 정보 등을 관리
EIS	사용이 용이하면서 다양한 분석 정보를 이용할 수 있도록 지원
OLAP	정규적인 다차원 분석
질의, 보고서	사용자별 개인 업무 적용을 위한 정보 추출과 보고서 작성

앞의 <그림 4-4>에서 보여준 데이터 웨어하우스 모형을 기반으로 하여 본 연구의 최종 목표인 데이터 웨어하우스 환경하에서의 은행 DB마케팅 시스템을 다음의 <그림 4-5>과 같이 제시한다.

<그림 4-5> 은행 DB마케팅 시스템 모형



자료 : 연구자 작성

은행은 데이터 웨어하우스로부터 각종 보고서를 작성할 수 있게 될 뿐만 아니라 DB마케팅을 위한 정보는 데이터 마트로 구축하게 된다.

데이터 마트는 다양한 분석 도구를 활용하여 고객의 세분화, 신용 평가, 행위분석 및 고객 수익성 분석 등 각종 고객분석 자료를 추출해 냄으로써, 목표고객과 고객의 행동을 예측하여 고객 각각의 특성에 맞는 상품 및 서비스를 다양한 접점을 이용하여 마케팅을 전개할 수 있게 되는 것이다.

고객 접점은 인터넷, 전화, 팩스, 메일 등 다양하며 고객과 은행이 만나는 곳으로서 중요한 의미를 갖고 있다. 은행은 다양한 접점의 활용으로 다양한 고객계층과의 만남을 가능하게 하여야 하며 고객의 요구에 부응하고 시대에 뒤떨어지지 않는 새로운 채널 기술에 대하여 지속적으로 투자할 필요가 있다.

고객의 접점을 관리할 때 접촉 시점과 접촉했던 채널, 접촉을 시도한 이유 등을 데이터로 관리하는 것이 중요하다. 한 고객이 얼마나 자주, 어떤 매체를 통하여, 어떤 시간대에, 왜 접촉했는가 하는 사실은 마케팅의 중요한 기본 자료로 쓰일 수 있다. 전화를 통하여 은행과 자주 접촉하는 고객에게는 전화로 하는 홍보가, PC를 통하여 자주 접촉하는 고객에게는 PC로 하는 홍보가 가장 적합할 것이기 때문이다

마케팅 결과의 분석을 통하여 홍보 및 마케팅 전략을 수립하는데 필요한 의사결정 정보를 제공하고, 새롭게 획득한 정보는 데이터 웨어하우스에 반영하는 작업을 반복함으로써 고객 세분화 작업을 더욱 강화할 수 있게 된다.



제4절 데이터 웨어하우스 具現 過程

DB마케팅을 위한 데이터 웨어하우스의 설계가 이루어지고 나면 이제는 어떠한 과정을 거쳐서 어떠한 방법으로 은행 내·외부의 실제 데이터를 데이터 웨어하우스의 데이터로 만들어 갈 것인가를 연구하여야 한다.

데이터 웨어하우스의 구축 과정은 데이터 모델링-데이터 추출, 변형 및 정제-데이터의 적재(Loading)로 구분할 수 있다.

1. 데이터 모델링

데이터 모델을 개발하는 것, 즉 은행 전체에 퍼져 있는 데이터 자원의 위치에 대한 지도를 작성하는 것은 데이터 웨어하우스 구축에 있어 아주 중요한 첫 단계이다. 데이터 모델이란 그래픽 표현 방식으로 조직의 데이터 자원을 정의하고 분류하며, 이들 데이터 자원 사이의 경로를 보여주는 것이다. 데이터 모델은 각 항목과 이들 항목이 구성하고 있는 데이터를 정의한다. 그리고 데이터 마이닝 기법을 통하여, 이들 항목 사이의 연결 관계를 파악하여, 어떤 상품이 어떤 고객에게 가장 많이 판매되었는지를 알 수 있게 된다.

데이터 모델링을 위해서는 다음의 몇 단계 과정을 거쳐야 한다

- ① 사용자의 요구를 확인하기 위하여 이들 사용자를 조사하여야 한다.
- ② 조사 자료를 분석하고 통합하여야 한다.
- ③ 원형 데이터 모델을 만들어 사용자에게 보여주고 확인을 받아야 한다.
- ④ 사용자의 사용 경험과 의견을 토대로 모델을 개선한다.

그러나, 이러한 일련의 과정을 거치더라도 한번에 조직의 모든 데이터 모델을 만든다는 것은 거의 불가능하다. 실제로 대규모 데이터 모델을 구축하려면 수년의 시간이 걸리며, 이 과정에서 엄청난 어려움을 겪게 된다.

따라서, 은행에서는 데이터 모델을 만들기 위하여 점진적 접근법을 택하는 것이 바람직할 것이다. 점진적 접근법은 정보기술 부서에서 각 단계에 대한 사용자의 의견을 수집한 후 다음 단계를 향상시키기 위한 자료로 활용할 수 있는 시간적 이점을 얻을 수 있기 때문이다

2. 데이터 抽出, 變形 및 整齊

1) 데이터의 추출, 변형, 전송(ETT; Extraction, Transformation, Transportation)

운영계 시스템(OLTP)에서 다양한 형태(DB2, VSAM, Tape, RDBMS 등)로 저장해 두고 있거나 새롭게 발생하는 소스에 인터페이스를 제공함으로써 데이터를 읽어 들여 정제, 변형, 통합 과정을 거쳐 데이터 웨어하우스로 적재하는 즉, 물리적으로 데이터 웨어하우스를 생성하는 전 과정으로서, ETT는 데이터 웨어하우스 구축의 성패를 좌우할 정도로 중요하다.

2) ETT 방식



ETT 방식은 데이터의 전송 방법에 따라 오프라인(Off Line)과 온라인(On-Line) 방식으로 구분하며, 은행에서는 필요에 따라 두 방식을 적절히 선택하여 운영하여야 할 것이다.

① 오프라인 방식

이 방식은 소스데이터(source data)에서 추출하여 데이터 웨어하우스에 로딩하는 프로세스가 끊어지는 것을 말하며, 많은 양의 데이터를 데이터 웨어하우스 레퍼지토리(repository)에 로드해야 하는 초기 적재시에 사용되나 몇 가지 주의해야 할 사항이 있다.

첫째, 운영계 시스템에서 소스데이터가 SAM파일로 변환되어야 하므로 운영계 시스템의 디스크를 많이 사용해야 하며, 또한 전송되어진 후에도 SAM파일 형태로 변

환, 정제가 이루어지므로 웨어하우스 서버의 디스크 사용에 부담을 준다.

둘째, 데이터 타입의 변환을 위한 프로그래밍이 필요하다. 소스 시스템이 EBCDIC 코드 메인프레임이고 타겟 시스템이 UNIX 또는 NT일 때에는 파일 전송을 위하여 ASCII 코드로 변환하여야 한다.

셋째, M/T(Magnetic Tape)를 이용하여 데이터를 전송하려면 웨어하우스 서버에 별도의 카트리지를 디바이스를 도입해야 한다. 이에 대한 비용의 부담도 만만치 않다.

넷째, FTP를 이용한 데이터의 전송은 간편하나 네트워크에 장애가 생겨 전송이 끊어지는 경우가 종종 있다. 이럴 경우 아무런 에러 메시지가 없으므로 정상적인 전송 종료임을 알 수 있는 방법이 없다. 전송이 끝난 후 보낸 파일의 크기와 받은 파일의 크기가 같은지 반드시 확인한 후 다음 과정을 진행해야 한다. 이 방식의 ETT는 비용이 적게 소요될 것이라고 생각하기 쉬우나, 실제로는 수 작업 처리가 많고, 프로그램의 하드 코딩을 해야하므로 많은 인력의 낭비와 생산성의 저하를 초래할 수 있다.

② 온라인 방식

온라인 추출 방식은 운영계 시스템에서 소스 데이터를 직접 내려 데이터웨어하우스에 로딩하는 전 프로세스가 통합적으로 이루어지는 추출 방식이다. 이는 추출 틀의 기본 방식이기도 하다. ETT 전 과정을 하나의 프로그램 안에 스텝별로 엮어 이전 스텝의 정상적인 작업 종료를 확인한 후 다음 스텝이 자동적으로 실행된다. ETT 프로그램의 관리가 용이하며 스케줄링에 의해 작업이 자동적으로 실행되기 때문에 수 작업으로 인한 번거로움을 배제할 수 있다.

3) ETT 과정

ETT 과정은 운영계 시스템에서 소스 데이터를 추출하여 데이터 웨어하우스 서버로 전송한 후 그 데이터를 변환 설제서에 맞게 변환, 정제하여 데이터 웨어하우스 레퍼지토리에 로드하는 것으로 이루어져 있다.

데이터의 변형 및 정제 과정을 운영계 시스템에서 하느냐, 데이터 웨어하우스 서버에서 하느냐의 문제가 발생할 수 있으나, 운영계 시스템의 부하를 적게 한다는 입

장에서 생각한다면 데이터 웨어하우스 서버의 시스템 자원을 사용하는 것이 바람직하다

① 소스 데이터(Source Data)

운영계 시스템에서 생성되고 저장되는 데이터로서 데이터 웨어하우스의 근원 데이터를 소스 데이터라고 한다. OLTP 데이터, 과거 데이터, 외부 데이터 등이 소스 데이터로 분류된다. 소스 데이터는 다양한 OS상에서 여러 형태로 저장된다.

② 데이터 추출(Extraction)

추출은 추출도구를 사용하는 방법과 프로그래밍하는 방법이 있다. 추출시에는 다양한 저장 형태에 따른 데이터 인터페이스뿐만 아니라, 소스 데이터 필드의 다양한 데이터 타입을 어떻게 변환하여야 하는가를 고려해야 한다.

③ 데이터 전송(Transportation)

데이터 웨어하우스 서버는 주로 소스 시스템과 다른 별도의 시스템에 존재한다. 그러므로 대량의 데이터를 타겟 시스템에 전송해야 한다.

데이터 전송 방법은 FTP, M/T, 게이트웨이(Gateway) 등이 있다. 게이트웨이는 주로 이기종간 특히, 메인프레임의 데이터를 추출하여 온라인 방식으로 타겟 시스템인 웨어하우스 서버에 넘겨준다. 게이트웨이는 TCP/IP, APPC, DECNET, HLLAPI 등의 네트워크 프로토콜을 사용한다.

④ 데이터 변환(Transformation)

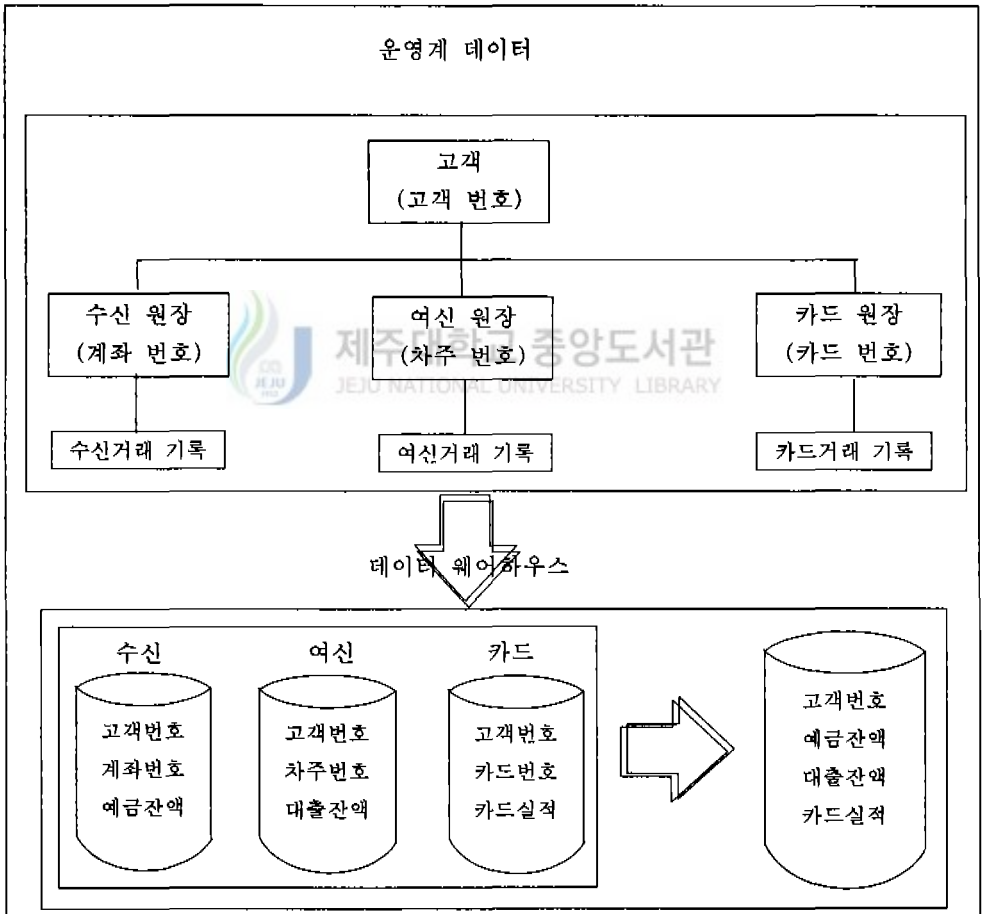
운영계의 소스 데이터는 다각도의 데이터 분석에 적합한 형태로 존재하지 않는다. 따라서 데이터의 변형작업이 필요하다. 데이터 타입의 변환, 데이터 구조의 변환, 코드의 재구성, 새로운 필드를 파생하는 작업 등이 포함된다. <그림 4-6>은 은행의 데이터 구조 변환을 나타낸 것이다.

⑤ 데이터 정제(Cleansing)

운영계 시스템의 데이터는 거래 프로세싱에 의해 생성되므로 고객과의 거래에 필

요한 필드의 유효성이 중요하다. 예를 들면 은행의 수신원장 데이터의 전화번호, 우편번호, 성별 구분 등의 필드는 중요하지 않다. 운영계 시스템에는 이러한 필드에 저장되어 있는 데이터 자체가 없거나, 유효하지 않은 값들이 저장 될 수도 있다. 그러나 데이터 웨어하우스에는 이러한 필드의 값이 매우 중요할 수도 있다. 그러므로 데이터 정제 작업은 웨어하우스를 사용하여 현업 사용자에게 정확한 정보를 제공하기 위해 로드하기 전에 해야 하는 반드시 필요한 작업이다.

<그림 4-6> 은행 데이터 구조의 변환



자료 : 연구자 작성

3. 데이터의 積載(Loading)

데이터의 적재는 변환, 정제과정을 거친 데이터를 리포지토리에 로드하는 것이다. 적재하는 기법은 크게 두 가지로 구분된다. 첫 번째는 전체 데이터를 비우고 초기 적재와 같이 새로이 적재하는 리프레쉬(Refresh) 방법이고, 두 번째는 소스 시스템에서 변경된 데이터만 추출하여 기존의 데이터에 변경된 내용만 반영시키는 변경 데이터 반영(Changed Data Capture : CDC) 방법이다.

초기 적재란 일반적으로 M/T에 수록·보관하고 있는 과거 수년간의 초기 데이터를 데이터 웨어하우스에 로딩하는 것을 말한다.

1) 리프레쉬 기법

데이터 웨어하우스의 데이터는 시간이 감에 따라 계속 성장하게 된다. 즉, 운영계 시스템에서 생성된 최신의 데이터를 반영하여야 한다. 이때 전체 데이터를 비우고, 다시 데이터를 적재하는 방법이 리프레쉬 기법이다. 이 기법은 최신의 운영계 데이터를 데이터 웨어하우스에 반영하는 가장 확실한 방법이다. 그러나 데이터 웨어하우스의 데이터 량은 굉장히 방대하기 때문에 다음과 같은 사항을 고려해야 한다.

첫째, 방대한 양의 데이터를 추출하는데 한계가 있을 수 있다 이를 위해서는 데이터의 추출시간을 고려해야 한다. 또한 데이터베이스에 저장된 데이터를 SAM 파일로 전환하여 FTP를 이용한다면 추출에 대한 디스크 스페이스를 고려해야 한다.

둘째, 데이터 전송시간을 고려해야 한다. 방대한 양의 데이터를 네트워크를 이용해 전송한다면 네트워크 성능을 확장해야 한다. 네트워크 성능을 확장하기 위해서는 네트워크 인프라를 확장해야 하는데 이에선 많은 비용이 소요될 것이다.

셋째, 전송된 데이터를 변형하여 적재하는 시간을 고려해야 한다. 많은 양의 데이터를 변형하기 위해서는 많은 양의 메모리와 임시 디스크 스페이스가 확보되어야 하며 데이터 양이 크면 이의 처리시간은 당연히 늦어질 것이다.

리프레쉬 기법은 데이터 웨어하우스의 데이터 정합성을 보장할 수 있는 가장 확실한 방법이나, 위와 같은 고려 사항들 때문에 운영계 시스템에서 변경된 데이터만

추출하여 데이터 웨어하우스에 반영하는 변경 데이터 적재에 대하여 많은 운영자들이 관심을 가지고 있다

2) CDC 기법

운영계 시스템에서 변경된 데이터만 추출하여 데이터 웨어하우스에 반영하는 변경 데이터 적재 기법으로는 다음과 같은 방법이 있다.

① Time Stamp 이용

운영계 시스템의 데이터베이스에 시간을 구분할 수 있는 필드가 있다면 이를 이용하는 방법이다. 즉, 추출하고자 하는 시간으로 데이터를 서브 셋팅하여 추출한다 가장 간단한 방법이나, 운영계 시스템의 데이터베이스에 이러한 필드를 관리하고 있어야 하는데 일반적으로 관리되지 않고 있다. 또한 삭제되는 레코드에 대한 정보를 관리하기 힘들다.

② Delta 파일 이용

델타 파일이란 변경되는 레코드를 별도의 테이블에 모아두는 것을 말한다. 델타 파일을 생성하는 방법은 DBMS에서 제공하는 기능을 이용하는 방법과, 트랜잭션을 처리하는 애플리케이션 프로그램에서 갱신되는 레코드를 별도의 다른 테이블에 쓰여지도록 하는 방법이 있다. 그러나 이 델타 파일을 만드는 것은 운영계 프로그램을 갱신해야 하거나, 트랜잭션 처리시간에 부하를 준다.

③ DBMS 로그 이용

이 방법에는 액티브(Active) 로그와 아카이브(Archive) 로그를 이용하는 방법이 있다. 액티브 로그를 읽어서 변경 데이터를 캡처하는 방법은 국내 금융권이 차세대 정보계 시스템을 오픈 시스템으로 구축하기를 원하면서 관심도가 높았다. 그러나 DBMS의 액티브 로그를 읽는다는 것은 OLTP에 부하를 줄 수도 있고, 기술적으로 매우 까다롭기 때문에 아직까지는 국내에서 성공한 사례가 없다.

이로 인해 아카이브 로그를 읽어서 변경데이터를 캡처하여 데이터 웨어하우스에

그 변경 내용을 반영하는 솔루션을 추출도구 벤더들이 제공하고 있다

④ 사용자 로그 이용

사용자 로그는 애플리케이션 프로그램으로 만들어지는 로그 파일이다. 국내 대다수 은행들의 운영계 시스템은 이 사용자 로그를 가지고 있다.

현재 은행에서의 CDC 기법은 사용자 프로그램에서 생성되는 로그를 이용하는 방법으로, 이러한 방법을 이용하여 구축한 CDC를 지연처리(deferred) 솔루션이라고 부르고 있다. 은행에서 지연처리를 이용하는 목적은 24시간 운영되는 운영계 시스템의 데이터를 목적 시스템으로 전송하여 데이터 웨어하우스를 구축하고, 구축된 데이터 웨어하우스의 데이터 적기화(適期化)를 이루는데 있다.

은행의 운영계 시스템은 조금의 부하 증가도 치명적일 수 있기 때문에 운영계의 OLTP 프로그램들이 생성하는 로그를 이용하는 방법이 보편화되어 있는 것이다.

지연처리는 운영계 시스템에서 발생한 데이터의 로그를 분석·가공하여 정보계 시스템에서 필요로 하는 각종 DB를 구축하는 과정으로 로그를 분석하는 과정과 DB를 갱신하는 과정으로 구분하여 구축한다. 지연처리를 이용한 CDC 솔루션을 구현하는 데는 많은 고려 사항이 따른다.

첫째, 운영계 시스템의 로그에 대한 정확한 분석이 필요하다. OLTP 프로그램을 작성할 때는 대부분의 프로그래머들이 운영계 만을 고려할 뿐 정보계는 잘 고려하지 않는다. 작성된 OLTP 프로그램을 구동하여 보면 때때로 사용자 로그의 상당 부분이 규약과 다른 점을 발견할 수 있다.

둘째, 정보계 시스템의 테이블을 어떻게 구축할 것인가, 갱신 주기는 최소 어느 정도여야 하는가 등의 목적 테이블에 대한 요건 정의가 정확히 이루어져야 한다.

지연처리 시스템은 사용자 로그를 읽어 분석을 수행하게 된다. 여기서 분석은 목적 테이블의 요건에 맞게 로그를 읽어 내는 것을 의미하므로 목적 테이블의 구성에 대한 분석이 없으면 로그에 대한 제대로 된 분석을 수행할 수 없다.

셋째, 기본적인 시스템과 네트워크의 인프라가 이루어져 있어야 한다는 것이다. 일반적인 CDC 솔루션과 달리 지연처리는 작업 도중에 사용자의 간섭을 허용하지 않

아야 한다. 따라서 시스템 내의 안정성 및 DB의 안정성은 기본적으로 시스템에서 제공해 주어야 하는 것이다.

운영계와 정보계 사이의 네트워크 단절이 심하고 정보계의 시스템이 자주 다운되며 DB의 부하 증가로 타겟 테이블의 갱신이 이루어지지 않는다면 지연처리 시스템의 구축은 생각하지 말아야 하는 것이다.

넷째, 데이터의 정합성을 보장할 수 있어야 한다 운영계에서 생성해 내는 로그의 정합성이 보장된다면 로그의 분석 및 목적 테이블로의 적재과정에서 하나의 데이터의 유실도 경영분석 및 판단을 내리는데 큰 영향을 줄 수 있으므로 데이터의 정합성을 보장하기 위한 정확한 관리 방안 및 정합성 검증을 위한 알고리즘을 개발하여 적용하여야 한다.

지연처리 시스템을 구축하다 보면 위와 같은 네 가지 고려사항 이외에도 여러 가지 문제점 및 제약사항이 나타난다. 하지만 개발을 하는 과정에서 미리 사용자의 요구를 확실히 파악하고 프로젝트에 임한다면 이러한 문제점들의 대부분은 사용자의 요구를 파악하는 단계에서 대부분 해결될 것이다.



제 5 장 結 論

제1절 研究의 要約

지금까지 은행의 데이터베이스와 DB마케팅 그리고 데이터 웨어하우스에 대하여 각종 문헌들을 살펴보고, 이를 토대로 은행의 DB마케팅을 위한 데이터 웨어하우스 구축 방안을 제시하였다.

현재, 은행에서 활용하고 있는 데이터베이스는 운영업무를 중심으로 구축하여 왔었기 때문에 IMF 체제 이후의 급격한 금융환경 변화에 적절히 대응하기 어려운 상황이다. 즉, 금융시장의 개방으로 외국 금융기관과의 경쟁뿐만 아니라, 업무영역의 철폐로 제2금융권과도 생존을 걸고 치열한 전쟁을 치러야 하는 현실에 직면한 것이다 어느 기업에서도 마찬가지이지만 은행도 이제 고객을 두려워할 줄 알고, 고객에 대한 서비스 여하에 따라서 생존이 결정되는 시대에 살고있는 것이다

따라서, 고객에 대한 정보를 자원으로 전개하는 DB마케팅은 은행에서도 거스를 수 없는 대세로 받아들이고 있다. DB마케팅에는 정보기술 뿐만 아니라 현업 부서의 지식습득도 중요한 문제이다. 아무리 최신의 정보기술을 활용하여 DB마케팅 시스템을 구축해 놓더라도 이를 이용하여 효과적인 마케팅을 전개하지 못한다면 시스템 구축 효과는 반감될 것이다

정보기술의 측면에서 보았을 때 DB마케팅 시스템은 데이터 웨어하우스를 기반으로 구축하는 것이 효과적이라고 많은 문헌에서 밝혀졌다. 여기에는 현업 부서에서 다양하게 활용할 수 있는 데이터 마이닝 도구의 출현이 그 배경이 되고 있다. 데이터 웨어하우스의 출현은 정보기술 부서에서도 적지 않은 이점을 준다. 즉, 지금까지 현업 부서에서의 정보제공 요구를 정보기술 부서에서 일일이 하드 코딩에 의해 처리하였던 방식으로부터 벗어나, 현업 부서에서 직접 데이터 웨어하우스에 접근하여 필요한 정보를 추출할 수 있게 된 것이다. 또한, 데이터 웨어하우스로부터 마케팅용 데이터 마트를 구축함으로써, 데이터 마이닝 기법을 활용하여 은행 고객의 특성을 분석하고, 개별 고객에 맞는 다양한 서비스를 제공할 수 있게 될 뿐만 아니라, 은행

상품간 혹은 고객 특성간의 상호 관련성을 발견하여 각종 위험을 예측하거나, 고객의 기호에 맞는 다양한 금융상품을 개발할 수 있게 된다.

새로운 정보기술을 이용한 은행 경영전략의 수립이 절실한 시점이다 이는 정보기술 부서만의 노력만으로는 불가능하며, 은행의 최고 경영진에서부터 DB마케팅 실무가에 이르기까지 DB마케팅에 대한 공감대의 형성이 무엇보다도 중요하다. 데이터 웨어하우스를 기반으로 하는 DB마케팅 시스템의 구축에는 지속적인 관심과 투자가 이루어져야 하기 때문이다. 그러나 투자에 대한 효과는 단기간이 아닌 중장기적으로 나타나기 때문에 단기 평가에 익숙한 우리 나라의 조직에서는 실천하기가 매우 어려운 면도 있다. 따라서 최고 경영진의 적극적인 이해와 지원이 필요하다. 최근에는 정보기술 부문에 대한 투자 없이는 앞으로의 경쟁환경에서 생존을 보장받기 어렵다는 사실을 최고 경영진을 비롯한 대부분의 조직 구성원들이 공감하고 있기 때문에 과감한 결정을 내릴 수 있는 여건이 되어있다

성공적인 데이터 웨어하우스의 구축을 통한 DB마케팅을 실천에 옮기기 위해서는 시스템 구축의 전과정에 걸쳐 정보기술 부서와 현업 부서의 협조와 노력이 필요하며 조직의 의사결정 과정에 대한 변화가 일어나야 할 것이다.



제2절 研究의 限界 및 向後 研究 方向

본 연구의 한계점으로는 첫째, 국내 은행에서 발표되는 DB마케팅 활용 사례가 부족하여 은행에서 전개할 수 있는 다양한 마케팅 전략을 제시해 볼 수 없었다는 점과 둘째, 국내 은행들의 데이터 웨어하우스 구축 사례를 살펴볼 수 없었기 때문에 본 논문에서 제시하는 데이터 웨어하우스 기반에서의 DB마케팅 시스템 모형과의 비교 연구가 되지 못하였다는 점이다. 따라서 본 논문에서 제시하는 내용들이 실제 사례를 통해 검증받지 못하고 이론적인 제시에 그칠 수밖에 없었으며, 실제 데이터 웨어하우스의 구축은 기술적인 문제를 포함하여 물리적으로 구현이 되어야 하는 앞으로의 과제로 남겨두게 되었다.

參 考 文 獻

<國內 文獻>

- 권오준, “데이터베이스 마케팅의 성공 포인트”, LG주간경제 제469호, 1998.
- 김정섭, “데이터베이스 마케팅에 관한 연구”, 석사학위논문, 서울대학교 대학원, 1995.
- 김희진, “삼성카드 데이터 웨어하우스”,
[<http://dwserver.hit.co.kr/kolee/구축사례/삼성카드.html>]
- 박찬욱, 「데이터베이스 마케팅」, 연암사, 1996.
- _____, “고객정보를 활용한 은행 데이터베이스 마케팅 전략에 관한 연구”, 한국금융연구원, 1998.
- 박천승, “데이터 웨어하우스 환경에서의 전략적 데이터베이스 마케팅 시스템 구축”, 석사학위논문, 고려대학교 대학원, 1997.
- 이영철, “DB마케팅을 통한 의류브랜드의 데이터 분석”, SPSS 사용자 사례 논문집-데이터 마이닝을 중심으로-, 1998.
- 이용희, “Data Mining을 이용한 리테일 बैं킹 전략에 관한 실증적 연구”, 전국은행연합회 논문집, 1998.
- 이진권, “체이스맨하탄 은행의 DW 구축사례”, 시사컴퓨터 11월호, 1998.
- 이창열, “Database Marketing 구축사례 삼성카드(주)”, 서강정보공학연구회, 1999.
- 장동인, “차세대 의사결정 지원 시스템-데이터 웨어하우스의 최신 동향”, 한국오라클, 1997 겨울호. [http://www.oracle.co.kr/magazine/win_cover1.html]
- 전완기, “데이터 웨어하우스를 성공으로 이끌기 위한 전략”, 컴퓨터월드 3월호, 1999
- 정보통신부, “데이터 웨어하우스 기반의 Data Mining 소프트웨어 개발”, 1997 12월.
- 정철용·함유근, “고객정보시스템 구축 및 활용 전략”, 한국금융연구원, 1999.
- 조재희, “데이터웨어하우징 기술과 DB마케팅 전략”, 한국산업정보학회 추계학술대회 발표논문집, 1997 11월
- 조재희·박성진, 「데이터 웨어하우징과 OLAP」, 대청정보시스템, 1996.

한국능률협회컨설팅, “데이터베이스 마케팅 시스템 구축과 활용”, 한국능률협회컨설팅 CS마케팅 OBU, 1997.

함유근, “DB마케팅의 의의와 은행의 도입”, 금융 11월호, 1997.

LG경제연구원 경영컨설팅센터, “마케팅 혁신기법: 데이터베이스 마케팅”, 81회 한국오라클(주) 정기세미나 자료집, 1997.

[<http://www.cio.seoul.kr/990401/main14.html>], “Special Report, 데이터 웨어하우스 50%가 실패작 어떻게 극복할 것인가”,

<海外 文獻>

Brooks, P., “March of the Data Marts”, *DBMS Magazine*, March 1997.

Devlin, B., “From data model to data warehouse in bite-sized chunks!” *InfoDB*, Vol. 9, No 5, 1995.

Eckerson, W. W., “Building and Managing Data Marts”, *White Paper*, 1997.

[<http://www.informatica.com/swp.html>]

Edelstein, H., “Mining Data Warehouse”, *Information Week*, Jan 1996.

Holtz, H., *Databased Marketing*, John Wiley & Sons Inc., 1992.

Hughes, A. M., *Strategic Database Marketing: The Masterplan for Starting and Managing a Profitable, Customer-Based Marketing Program*, Irwin, 1994.

Inmon, W. H., and R. D. Hackathorn, *Using the Data Warehouse*, John Wiley & Sons, 1994.

Inmon, W. H., “The Data Warehouse and Data Mining”, *Communications of ACM*, Vol. 39, No. 11, November 1996.

_____, “What is Data Mart?”, *White Paper*, D2K Inc, 1996.

[<http://www.d2k.com/d2k/library2.htm>]

Kelly, S., *Data Warehousing: The Route to Mass Customization*, John Wiley & Sons, 1994.

- Poe, V., "Guidelines for Warehouse Development", *Database Programming & Design*, September 1994.
- Roberts, M L., "Expanding the Role of the Direct Marketing Database", *Journal of Direct Marketing*, 6(2) 1992.
- Stedman, C , "Warehouse costs Scare up Changes", *Computerworld*, 1996.
[<http://www.computerworld.com>]
- Swift, R., "A Maturity Model for Decision Support and Data Warehousing", *Data Warehouse Report*, Issue 4, Winter 1995.
- Widom, J., "Research Problems in Data Warehousing.", *Proceeding of the 4th Int'l Conference on Information and Knowledge Management (CIKM)*, November 1995.



ABSTRACT

A STUDY ON ESTABLISHING THE DATA WAREHOUSE FOR BANK'S DB MARKETING

Chang-Koo Seong

Department of Management Information

Graduate School of Business Administration

Cheju National University

Supervised by Professor Young-Soo Cheon

Database marketing is a new marketing concept which is employed by increasingly more companies these years. Taking place of CS(Consumer Satisfaction) and BPR(Business Process Reengineering), it is becoming a competitive strategy critically important for a company's survival and prosperity. DB marketing is now passing a turning point by the support of the state-of-art information technology.

Especially, data warehouse and data mining are major technologies in the area of DB marketing. Through relevant data analysis, these techniques enable us to identify potential customers who may possibly render high profit to the company, and to make them loyal to the company by focusing marketing effort on them.

Banks are experiencing drastic changes in the financial environment in recent years. They have realized that it is hard to survive without relying on the strategy making maximum use of the information technology. Therefore, DB marketing based on the data warehouse became a very important issue in the management of a bank.

This study was conducted with the purpose to present a model to establish a data warehouse for bank's DB marketing. For this purpose, the factors that

make DB marketing a must for financial firms are examined in depth through literature survey, and prerequisites for the successful establishment of a data warehouse are analyzed. Case studies are added on the banks which have already established data warehouses and are successfully practicing DB marketing. Major findings are as follows;

Firstly, top-down approach is recommendable to implement a data warehouse. Domestic banks do not need to follow the same track taken by the banks of foreign countries. The latter have already established various data mart and are using them while the former have just started in this filed. For domestic banks, it is regarded as an effective approach to establish data warehouse step by step based on a long-term master plan and, at the same time, establish data marts which satisfy users' needs.

Secondly, refresh technique was known to be effective for the initial loading in the implementation stage of data warehouse. And deferred process technique using users' log was found to be the most relevant among CDC(Changes Data Capture) techniques. This technique can prevent confusion caused by drastic changes by utilizing bank's experience accumulated in the process of past operation.

DB marketing based on data warehouse is an inevitable task for banks. But it should be emphasized that sufficient preparation and considerate planning is a must to make successful the data warehouse system which demands a substantial amount of investment. It is also important to form a consensus among all the people including top-management and those who are engaged in DB marketing

This study have some limitations. It just presented a rather general approach to establish a data warehouse system based on the result of case studies. It is necessary to prepare a concrete and detailed implementation plan which takes into account various contingency factors that may be confronted in real application.