



### 저작자표시-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이 저작물을 영리 목적으로 이용할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

物流學碩士 學位論文

빅데이터를 이용한 부산항신항 ITT 효율적인  
운영방안에 관한 연구

A Feasibility Study on Effective Operation of ITT(Inter Terminal Transportation)  
Utilizing Big Data



韓國海洋大學校 海洋金融·物流大學院  
海運港灣物流學科  
白 龍 主

本 論 文 을 白 龍 主 의 物 流 學 碩 士 學 位 論 文 으 로 認 准 함 .

委 員 長 申 宰 榮 ㉠

委 員 辛 容 尊 ㉠

委 員 申 英 蘭 ㉠



2017年 06月

韓 國 海 洋 大 學 校 海 洋 金 融 · 物 流 大 學 院

## < 목 차 >

Abstract .....	i
<b>제1장 서 론 .....</b>	<b>1</b>
제1절 연구의 배경과 목적 .....	1
제2절 연구의 방법 및 구성 .....	2
<b>제2장 이론적 고찰 .....</b>	<b>4</b>
제1절 빅데이터의 개념 및 특징 .....	4
1. 빅데이터의 출현배경 .....	4
2. 빅데이터 정의 .....	6
3. 빅데이터의 특징 .....	8
제2절 빅데이터의 구성요소 .....	11
1. 자원 .....	11
2. 기술 .....	13
3. 인력 .....	15
제3절 빅데이터 분석 인프라 및 분석기법 .....	17
1. 빅데이터 분석 인프라 .....	17
2. 빅데이터 분석 기법 .....	20
제4절 빅데이터를 이용한 부산항 정보의 구축과 활용 사례 .....	29
1. 항만물류정보 역할 및 필요성 .....	29
2. 항만물류정보관리 문제점과 사례 .....	33
3. 최신 ICT 기술을 이용한 스마트 물류정보관리 사례 .....	35
<b>제3장 빅데이터를 이용한 부산항신항 ITT 이동패턴 분석 .....</b>	<b>47</b>
제1절 타부두 TS 현황분석 .....	47

1. 부산항 환적화물 현황 .....	47
2. 환적화물의 변화와 전망 .....	48
3. 타부두 TS 현황 .....	48
제2절 부산항신항 ITT 이동패턴 분석 .....	50
1. 분석 방법 및 타부두 ITT 패턴분석 .....	50
2. 터미널내 혼잡도 분석 .....	55
<b>제4장 부산항신항 ITT 운영방안 .....</b>	<b>63</b>
제1절 부산항신항 ITT 현황 및 문제점 .....	63
1. 부산항신항 터미널 운영현황 .....	63
2. 부산항신항 ITT 운영 현황 및 문제점 .....	65
제2절 ITT 플랫폼 설치를 위한 기본 방향 및 운영방안 .....	69
1. ITT플랫폼 설치를 위한 기본방향 .....	69
2. ITT 플랫폼 운영 및 구축 방향 .....	70
제3절 빅데이터를 이용한 ITT 운영 시스템 구축방안 .....	71
<b>제5장 결론 .....</b>	<b>73</b>
제1절 연구결과의 요약 및 시사점 .....	73
제2절 연구의 한계점과 과제 .....	74
<b>참고 문헌 .....</b>	<b>76</b>
<국내 문헌> .....	76
<외국 문헌> .....	77
<b>부    록 .....</b>	<b>78</b>
1. 주요업체(기업별) 활용사례 .....	78
2. 각국의 빅데이터 활용 사례 .....	81

## 〈표 목차〉

<표 2-1> Cisco의 모바일 데이터 트래픽 예측(2013~2018) .....	5
<표 2-2> 빅데이터의 대표적인 3가지(3V) 구성요소 .....	10
<표 2-3> 데이터 자원 확보를 위한 단계적 방법 .....	12
<표 2-4> 정보화 시대 VS 스마트 시대의 데이터 관련 이슈 변화 .....	13
<표 2-5> 빅데이터 처리 프로세스별 기술 영역 .....	15
<표 2-6> 빅데이터 활용에 따른 분석 방법의 변화 .....	21
<표 2-7> 텍스트 마이닝 과정 .....	22
<표 2-8> 물류 프로세스별 주요 기술 .....	37
<표 2-9> CTA 장비 운영 현황 .....	39
<표 2-10> ZPMC 장비 운영 현황 .....	40
<표 3-1> 부산항 물동량 처리현황 .....	47
<표 3-2> 신항 ITT(타부두T/S) 발생 현황 .....	49
<표 3-3> 2016년 터미널별 타부두 TS 물동량(From/To) .....	49
<표 3-4> 1일 기준 복화율 .....	62
<표 4-1> 부산항신항 운영사 현황 .....	63
<표 4-2> 부산항 신항 추가 부두 건설 현황 .....	64
<표 4-3> 주요 시기별 선박 대형화 추세 .....	64
<표 4-4> 부산항 신항 화물 유형별 처리 비용 .....	66
<표 4-5> ITT 플랫폼 운영 주체별 역할 .....	70

## 〈그림 목차〉

<그림 1-1> 데이터 마이닝 및 시나리오 Rolling 과정 .....	3
<그림 2-1> IDC & EMC Digital Universe Study(2011) .....	5
<그림 2-2> 빅데이터 활용을 위한 3대 요소 .....	11
<그림 2-3> 빅데이터 처리 프로세스 .....	14
<그림 2-4> 데이터 중심 관리들은 어느 곳을 향하고 있는가? .....	16
<그림 2-5> 하둡의 구조와 분산처리기술 .....	18
<그림 2-6> 웹 사용 마이닝 시스템 .....	24
<그림 2-7> IT의 발달과 소통방식의 변화 .....	26
<그림 2-8> 소셜 프로세스 .....	27
<그림 2-9> 스피드 포트 .....	41
<그림 2-10> 다목적 물류정보시스템 .....	43
<그림 2-11> CoolGuardian 시스템 구성 .....	44
<그림 2-12> 고품질 택배시스템 구성도 .....	46
<그림 3-1> 연도별 얼라이언스 재편 현황 .....	50
<그림 3-2> 선박입출항에 따른 타부두 TS 및 ITT 개념도 .....	51
<그림 3-3> 터미널별 신항 ITT 발생도(From/To) .....	52
<그림 3-4> 전항지/차항지별 신항 ITT 발생도(From/To) .....	53
<그림 3-5> 수입환적 요일별 처리 물동량 .....	54
<그림 3-6> 수출환적 요일별 처리 물동량 .....	54
<그림 3-7> 신항 터미널별 야드장치시간 빅데이터 분석 화면 .....	55
<그림 3-8> 신항 터미널별 야드장치시간(요일/시간별) 혼잡도 분석 화면 .....	56
<그림 3-9> 터미널별 공'컨' 장치기간 및 적재율 분석 화면 .....	57
<그림 3-10> 터미널별 턴어라운드 타임 분석 화면 .....	59
<그림 3-11> 터미널별 턴어라운드 타임분석(요일/시간별) .....	60
<그림 3-12> 터미널별 턴어라운드 타임분석(요일/시간별) .....	61
<그림 4-1> 현행과 ITT 플랫폼 도입시 비용 비교 .....	67
<그림 4-2> 신항 ITT 효율화 개념도 .....	68
<그림 4-3> 공동배차 방식 개념도 .....	69
<그림 4-4> ITT Platform 기능 체계도 .....	71
<그림 4-5> 터미널내 혼잡도 최종 UI 화면 예시 .....	72
<그림 5-1> 출발지 및 목적지 입력에 따른 선대서비스(예시) .....	74

## Abstract

### A Feasibility Study on Effective Operation of ITT(Inter Terminal Transportation) Utilizing Big Data

Baek, Yong-Joo

Department of Shipping and Port Logistics  
Graduate School of Marine Finance and Logistics

The global maritime crisis and the financial slowdown are accelerating a fierce competition to induce container cargo volumes against other global ports.

In particular, this competition among them is more intensified by offering large-scale incentives or free port dues in order to attract global alliances of liners.

First above all, shipping lines constantly request that Busan Port Authority introduce the ITT, so that the subsidy of extra operating costs caused by ITT can secure transshipment container cargoes.

Despite that there are many container terminal operators which have been splitted for a long time and disintegrated terminals have caused unnecessarily extra operating costs, however, an ITT operative problem in Busan Port can be solved by integrating all terminal operators into a unified single terminal.

In fact, due to our current difficulty of the terminal integration, therefore, only establishing an ITT platform can be the key to solve ITT issues in the

first place.

This paper shows that an ITT platform can adopt an ITT platform design and configuration algorithm which should be first simulated utilizing Big Data tool.

Big Data is one of the latest IT technologies in the 4th Industrial Revolution. It is a way to re-create valuable data through correlation analysis.

In order that the data are used for Big Data, I'd like to propose a scenario to optimize the ITT platform construction by confining the optimal design requiring actual data after data mining through analysis tool (R).

Particularly, optimized in-and-out container cargoes can be calculated by analysis of loading and unloading rates of transshipment cargoes between nearby two terminals and the turnaround time. Besides, an ITT platform can be optimized to the establishment of an ITT platform via analysis of the best route of terminal-to-terminal TS on a basis of day and time and the transportation pattern of cargoes.

Thus, it is possible to reduce the failure rate of the ITT platform by performing the simulation once using the big data before execution, and it could be reflected in the future ITT platform design.

# 제1장 서론

## 제1절 연구의 배경과 목적

부산항은 '16년 말 기준으로 처리물량 19,700,00TEU를 처리하여 매년 6%의 물동량 증가세를 보이고 있으나 글로벌 경쟁항만의 급격한 성장으로 과거보다 처리물동량 순위가 다소 떨어진 세계 6위항만으로 자리잡고 있다. 부산항을 글로벌 허브항만 및 세계 2위의 환적중심항만으로 육성하기 위해 정부와 항만공사가 정책 목표를 수립하고 여러 가지 전략으로 모색하고 있다.

특히, 환적화물 거점 지정 및 환적화물 유치는 대부분 선사가 결정권이 있어 영업손익 등 각종 비용을 계산한 후 환적 거점 항만을 지정하고 있다. 최근 선사와의 간담회에서 부산항 환적화물 유치를 위해 ITT(Inter Terminal Transportation)를 요청하는 선사가 매우 많아졌다.

ITT는 터미널 내 YT(Yard Tractor)를 이용하여 타부두 TS(Transshipment Cargo)를 운송하는 방법을 말한다. 부산항은 경쟁선진 항만인 상해항, 싱가포르 항만과는 달리 항만내 터미널이 쪼개져 있어 자부두 TS 비율이 상대적으로 경쟁항만보다 낮다. 자부두 TS의 경우 동일한 터미널에서 입항한 선박 화물을 다른 선박으로 이동시켜 출항한 경우로서 동일한 터미널내에서는 육상 이적(Movement)가 YT로 운송되기 때문에 이적비용이 발생하지 않는 반면, 타부두 TS는 자부두 TS와 동일한 방법이지만 다른 터미널내 발생함으로 육상 이적(Movement)에 대한 추가적인 이적비용이 발생한다.

글로벌 대형선사는 환적화물 주요 거점의 지정은 여러 가지 요인을 들 수 있겠지만 그 중 하나는 환적화물 비용절감 요인을 들 수가 있다. 앞서 설명한 선사의 입장에서 타부두 TS 육상 이적비용은 터미널이 쪼개진 이유로 불필요한 부담으로

작용할 수 있다.

가장 현실적인 방법은 터미널 운영사를 통합하여 타부두 TS를 줄이는 방법이 대안이라 할 수 있으나 터미널 운영사 통합은 그리 쉽지만 않을 것이다. 그러나 타부두 TS를 줄이기 위한 항만당국 노력은 터미널간 경계 울타리 허물어 ITT를 실행할 수 있는 플랫폼을 구축하려는 노력을 하고 있다.

본 연구에서는 터미널내 실데이터를 활용한 빅데이터 툴을 이용하여 플랫폼의 설계를 위한 데이터 분석을 하고자 한다. 따라서 ITT 플랫폼 설계이전에 TS화물 복화율, 타부두 TS 이동 동선, 각 터미널의 턴어라운드 타임 등을 분석하여 ITT 운영상 개선 사항을 도출하였다.

## 제2절 연구의 방법 및 구성

본 연구는 4차 산업혁명의 최신 IT 기술인 빅데이터인 분석 툴(R)을 통해 데이터 마이닝한 후, ITT 플랫폼 시행 전 실제 데이터를 이용한 최적의 설계를 제한함으로써 ITT 플랫폼 구축을 최적화 시나리오를 산출하여 제안하고자 한다.

특히 타부두 TS 복화율 분석, 턴어라운드 타임 분석을 통한 최적의 반출입시간 산출, 타부두 TS 외부 최적의 경로 및 타부두 TS 요일 및 시간대별 이동패턴 등 여러 가지 시나리오를 산출하여 실제 데이터를 기초로 ITT 플랫폼 구축에 최적화를 제시하고자 한다.

시나리오 산출 방법은 관련기관(PA) 및 유관기관(선사, 운영사, 포딩, 화주 등)의 인터뷰를 통해 ITT 시행에 플랫폼 방안에 필요한 관련 데이터를 수집하여 데이터 마이닝을 통한 가치있는 데이터를 산출하는 방법이다. 동 방법론으로 구체적인 자료는 부산항만공사의 Port-MIS와 터미널 운영사의 TOS(Terminal Operating System) 및 선사와 터미널간의 COARRI(Container discharge/loading report message), CODECO(Container gate-in/gate-out report message)의 자료에서 추출

한 것으로 하였으며 구체적인 인터뷰 및 환경 분석을 통해 실행과제를 도출하였다.



<그림 1-1> 데이터 마이닝 및 시나리오 Rolling 과정

시나리오에 대한 구체적인 과정은 첫째, 최종 목적 데이터가 무엇인지 인지하고 무슨 효과가 있는지를 살펴야한다. 두 번째로 원천데이터를 찾고 데이터 분석 및 마이닝을 통해 시나리오 정제 작업을 거친다. 세 번째, 정제 후 시나리오 목적물은 최초 시나리오 설계와 비교해서 많이 벗어나면 다시 피드백을 한 후 적합성 여부를 판단 및 협업 인터뷰를 통해 원천데이터에 추가적인 데이터를 확보하는 지속적인 Rolling 작업을 거친 후 1개의 시나리오가 완성된다. 마지막으로 시나리오는 지속적인 Rolling 과 피드백과정이 필요하다. 이처럼 시나리오를 산출하는 과정을 반복해서 ITT 플랫폼 설계에 필요한 시나리오를 산출하는 과정을 반복하게 된다.

## 제2장 이론적 고찰

### 제1절 빅데이터의 개념 및 특징

#### 1. 빅데이터의 출현배경

요즘 4차 산업혁명 시대의 단어가 무척이나 많이 나타났다 특히, ICT의 발달로 빅데이터, IoT, 인공지능, AI 등 세상이 급변하고 있다. 특히, 빅데이터는 2011년부터 출현하여 산업계 및 학계전반으로부터 각광을 받아서 최근에는 4차 산업혁명분야에 빼놓을 수 없는 IT분야의 큰 축으로 성장하였다. 특히, SNS(Social Network Service)의 등장으로 데이터의 양이 폭발적으로 증가하고 있으며, 전 세계적으로 생산, 유통되는 디지털 정보량이 2011년 1.8 제타바이트(ZB) 약 2조 기가바이트에 달하고 있다.<sup>1)</sup>

최근 구글, 페이스북, 아마존 등의 소셜미디어나 하루 평균 1억 5,500만 건의 데이터가 트위터상에 발생하고 유튜브의 하루 평균 동영상 재생건수는 40억 회에 이르고 있다. 일상적으로 사용하는 단어 하나하나가 데이터가 되고, 기계적 물리적 상태를 데이터화 시키고 사람의 위치 등 상호작용 역시 데이터화를 시켜 모바일 폰, 온라인 쇼핑 소셜 네트워크, GPS등 현재 우리는 걸어 다니는 데이터 생성의 원인이 된다.<sup>2)</sup>

미국의 네트워크 통신사인 Cisco는 2018년까지 전 세계 모바일 데이터 트래픽이 매달 15.9 엑사 바이트(약 1.6조 기가 바이트)까지 성장할 것이며 2013년에서 2018

---

1) 천민경(2014), “빅데이터 역량 평가를 위한 참조모델 및 수준진단시스템 개발”, 한양대 석사 학위논문, p.2.

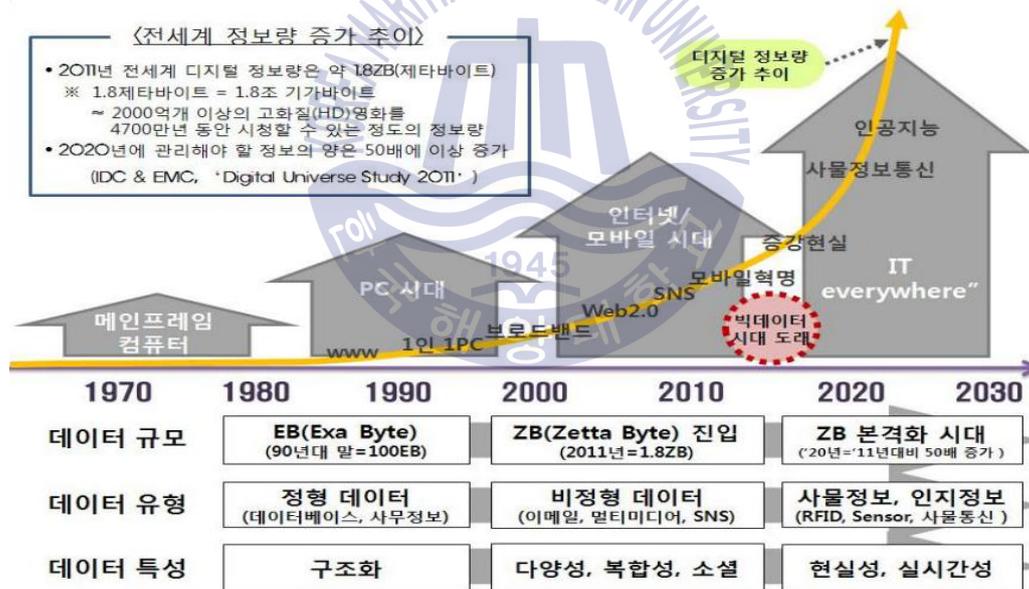
2) Andrew McAfee, Erik Brynjolfsson(2012), “Big Data : The Management Revolution”, *Harvard Business Review*, pp.60-68.

년까지 연평균 61% 정도로 모바일 데이터 트래픽이 증가할 것이라 예측하고 있다. 실제로 일상생활에서 생성되는 데이터들은 일정한 형태가 정해져 있지 않은 비정형데이터이며 거의 아무런 경제적 가치도 없는 데이터이다. 그러나 취합된 집단적인 데이터는 분석여하에 따라 소비자의 미래행동을 예측할 수 있으므로 기업에게는 위기가자 새로운 기회가 되어 막대한 부가가치 창출 효과를 낼 수 있다<sup>3)</sup>

<표 2-1> Cisco의 모바일 데이터 트래픽 예측(2013~2018)

구분	2013	2014	2015	2016	2017	2018
트래픽 양	1.5EB	2.6EB	4.4EB	7.0EB	10.8EB	15.9EB

출처 : Cisco VNI Mobile(2014)



<그림 2-1> IDC & EMC Digital Universe Study(2011)

3) BIR RESEARCH GROUP(2014), “Big Data(빅데이터) 분석 Global 혁신기술 동향”, B.I.R, pp.15-16.

빅데이터는 물류와 연관성이 깊다. 빅데이터라는 틀을 이용하여 새로운 비즈니스 모델을 사전에 모델링 하고 실제 사업의 위험을 줄이며 고객의 경험을 개선하고 새로운 사업모델을 만들 수 있는 잠재력을 가질 수 있다. 특히 물류에서 불필요한 비용을 줄이는 것에 효과적인 기술로 꼽힌다. 빅데이터는 고객사별 물류의 이동 경로와 재고량 증감 추이 등에 대한 실시간 분석을 가능케 하며 고객 성향 조사를 통해 화주기업이 정확한 수요예측을 진행할 수 있도록 지원하는 역할까지 수행할 수 있다.

빅데이터 예측 및 분석 알고리즘을 기반으로 물류업체들은 배송시간을 단축하고 고객과의 친밀성을 강화하며 역량 및 네트워크 단축을 최적화함으로써 프로세스의 효율과 서비스의 질을 현저히 향상시킬 수 있게 됐다.

신규 비즈니스 모델 창출에는 빅데이터를 활용하여 효율적인 모델링을 할 수 있다. 예를 들어 기상조건과 독감발생, 온라인 구매량 사이의 상관관계를 분석하여 고객의 행동을 예측하고, 그에 따라 새로운 비즈니스 모델에 대한 실마리를 발견할 수 있다는 것이다. 굿은 날씨가 온라인 구매량의 증가로 이어지고 이는 물류기업의 물량 증가로 연결된다. 이를 바탕으로 기업들은 프로세스를 최적화하여 개선된 고객 서비스를 제공할 수 있고, 새로운 비즈니스 창구를 개발해 시장에서 경쟁우위를 선점할 수 있다.<sup>4)</sup>

## 2. 빅데이터 정의

기존 빅데이터의 개념은 단순 데이터의 양이 많은 것을 의미했다면 최근의 빅데이터의 개념은 기존 데이터에 비해 너무 방대해 일반적으로 사용하는 방법이나 도구로 수집, 저장, 검색, 분석, 시각화 등을 하기 어려운 정형 또는 비정형 데이터의 집합을 의미한다.<sup>5)</sup>

4) 해양한국(2015), “해사산업계와 ICT 특집기획 물류분야”, 해양한국(5월), pp.70-71.

일찍부터 빅데이터에 대한 이슈를 적극적으로 제시한 매킨지 보고서에서는 빅데이터를 “크기가 지나치게 커서 기존의 데이터베이스 도구로 캡처, 저장, 분석이 불가능한 데이터 세트”라고 정의하였다(Abcuzeid etal 2009).

SERI(2010)는 “빅데이터를 기존 관리 및 분석 체계로는 감당할 수 없을 정도로 거대한 데이터의 집합”으로 정의하였는데, 대규모 데이터와 관계된 기술 및 도구(수집, 저장, 검색, 공유, 분석, 시각화 등)도 빅데이터의 범주에 포함된다고 제시하였으며, 빅데이터는 과거 특수 분야에 한정되었으나 ICT산업의 발달에 따라 전 분야로 확산될 것으로 예측하였다.

IT 전문 컨설팅그룹 가드너는 빅데이터에 대한 정의를 21세기의 원유로 다양한 종류의 데이터가 기업이 감당할 수 없을 정도로 빠르게 생성되는 현상으로 정의하여 데이터 활용에 초점으로 정의하였다.

Philip(2011)연구에서도 빅데이터에 대한 개념을 Volume, Velocity, Variety로 정의하고 있는데, 기존 3V에 대한 내용에서 확장된 개념으로 제시하고 있다. 빅데이터의 대용량(Volume) 측면에서 기존 연구에는 데이터의 양(Tera Byte, Peta Byte)에 초점이 맞추어져 있다면, 데이터의 종류 및 생성되는 시간에 따른 양, 지속적인 증가에 대한 부분이 반영되어야 한다고 제시하고 있다. 빅데이터의 속도(Velocity) 측면에서 기존 연구는 실시간으로 데이터를 획득하고 분석하는 것이 가능한 형태를 빅데이터로 정의하고 있다면, Philip(2011)의 연구에서는 이를 수행하기 위한 네트워크 측면도 반영해야 한다고 제시하고 있다 빅데이터의 다양성(Variety) 측면에서 기존 연구는 다양한 분야(텍스트, 이미지, 동영상 등)의 데이터 형태를 중심으로 정의하고 있다면, 데이터를 수집하기 위한 채널에 대한 부분도 감안해야 한다고 정의하고 있다.<sup>6)</sup>

빅데이터관련 정의는 학자에 따라 다양한 형태로 이루어지고 있으나 공통적으

5) 김성원(2014), “빅데이터가 기업의 경쟁력에 미치는 영향에 관한 연구”, 경기대 석사학위논문 p.6.

6) 김성원(2014), 전계서, p.17.

로 제시하고 있는 것을 살펴보면, 기존 기업에서 분석하고 있는 데이터에 비해 그 규모가 훨씬 큰 형태의 데이터 집합으로 다양한 형태로 데이터가 실시간으로 수집되며, 기존 방식으로는 분석하기 어려워 새로운 방식으로 분석해야하는 데이터로 개념을 정리 할 수 있다.

### 3. 빅데이터의 특징

빅데이터의 특징은 일반적으로 3V(Volume, Velocity, Variety)를 기본으로하고 1V(Value)나 1V(Veracity)가 추가되어 5V로 나타내는 학자도 있다. IBM은 빅데이터를 3V를 가진 새로운 타입의 데이터로 과거에는 답할 수 있었던 인사이트(Insight)를 얻어 기회로 정의하였으며, 5V로 많이 나타내고 있다.<sup>7)</sup>

#### 1) 규모 (Volume)

데이터의 발생량을 나타내는 것으로 페타바이트, 제타바이트 이상의 정보를 기준으로 물리적인 크기뿐만 아니라 데이터의 '속성'에도 연관되어 있으며 그것을 처리하는데 어려움이 있는지 혹은 없는지를 의미한다<sup>8)</sup>. 오늘날 다양한 문제 해결에 필요한 훨씬 복잡 미묘한 패턴을 잡아내기 위해 규모의 중요성을 매우 중요하게 인식되고 있다.

#### 2) 속도 (Velocity)

데이터를 처리하는 속도를 이용하여 배치 분석만을 의미하는 것이 아니라 필요에 따라서 수많은 사용자 요청을 실시간으로 처리한 후 처리 결과를 반환해주는

7) 신현신(2016), “빅데이터 기반의 재난정보관리 방안”, 서울시립대 석사학위논문, p.6.

8) 박두순(2014), “빅데이터 컴퓨팅 기술”, 한빛아카데미, pp.19~20.

기능에 대한 시간적 의미와 연관된다. 데이터의 생성과 처리가 진행되는 속도를 의미하며 의미는 세 가지로 구분된다고 볼 수 있다. 데이터가 발생 후 기업 내의 스토리지에 저장되기까지의 속도와 발생한 데이터의 불필요 부분과 무의미한 부분을 처리하여 가용하게 되는 수준까지의 속도, 그리고 정제된 데이터를 분석하고 의미를 추출하여 최종 목적을 달성하는 속도까지 말한다. 이러한 속도는 데이터의 접근성과 사용 가능성을 높이는 데 많은 영향을 준다.

### 3) 다양성 (Variety)

빅데이터 분석에는 일반적으로 정형데이터와 비정형데이터로 구분되는데 이중 통일된 구조로 정리하기 어려운 비정형 데이터 분석이 전체 중 90% 이상을 차지하며 사진, 동영상 등 기존의 구조화된 데이터가 아닌 다양한 형태의 데이터가 포함된다. 빅데이터를 수집하고 분석할 수 있게 되면서 훨씬 다양한 데이터의 활용이 가능해졌으며 특히, 다양한 비정형 데이터에서 생각해보지 못했던 결과를 얻는 일이 빈번해졌다.

### 4) 정확성(Veracity)

정확성은 데이터에 부여할 수 있는 신뢰 수준을 말하는데 보통 데이터가 신뢰에 따라 가치가 부여되기도 한다. 높은 데이터 품질을 유지하는 것은 빅데이터의 중요한 요구사항이자 어려운 과제이다. 하지만 최상의 데이터 정제(Data Cleaning) 기법을 사용해도 날씨나 경제, 고객의 미래 구매 결정 같은 일부 데이터의 본질적인 불확실성은 제거 할 수 없다. 소셜 네트워크 같은 인간 환경에서 생산되는 데이터는 신뢰하기 어렵고, 미래는 예측하기 어려우며 사람과 자연, 보이지 않는 시장의 힘 등이 빅데이터의 다양한 불확실성 형태로 나타난다.

## 5) 가치(Value)

가치는 빅데이터를 저장하려고 IT 인프라 구조 시스템을 구현하는 비용을 말하는데 최종적으로 데이터의 결과가 얼마가 가치를 발휘한다는 의미이다. 빅데이터의 규모가 상당하며 대부분은 비정형적인 텍스트와 이미지 등으로 구성되어 있다. 이 데이터들은 시간이 지남에 따라 빠르게 전파하면서 변하므로 그 전체를 파악하고 일정한 패턴을 발견하기가 쉽지 않아 가치의 중요성이 강조된다.

<표 2-2> 빅데이터의 대표적인 3가지(3V) 구성요소

구분	주요내용
규모의 증가 (Volume)	· 기술적인 발전과 IT의 일상화가 진행되면서 해마다 디지털 정보량이 기하급수적으로 폭증⇒제타바이트 시대로 진입
속도증가 (Velocity)	· 사물정보(센서, 모니터링), 스트리밍 정보 등 실시간성 정보 증가 · 실시간성으로 인한 데이터 생성, 이동(유통) 속도의 증가 · 대규모 데이터 처리 및 가치 있는 현재정보(실시간) 활용을 위해 데이터 처리 및 분석 속도가 중요
다양성 증가 (Variety)	· 로그기록, 소셜, 위치, 현실데이터 등 데이터의 종류가 증간 · 텍스트 이외의 멀티미디어 등 비정형화된 데이터 유형의 다양화

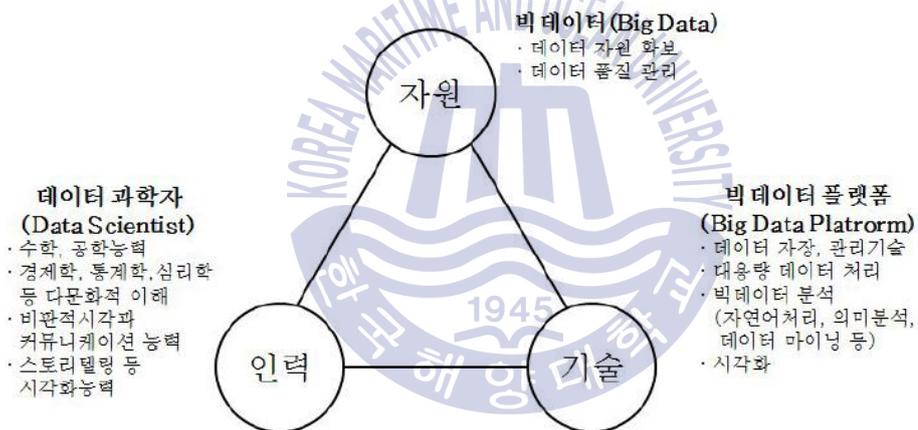
출처 : 정지선(2011), “신가치창출 엔진, 빅데이터의 새로운 가능성과 대응 전략”.

앞서 빅데이터의 3대 특징을 대용량(Volume), 다양성(Variety), 실시간성(Velocity)로 나타냈는데 이는 기존의 기술과 방법으로는 다루기 어렵다는 결론을 가지고 있다. 따라서 빅데이터라는 툴을 이용하여 새로운 발전의 기회를 제공하거나 문제점을 해결할 수 있는 단초로 제공하여 새로운 지적 통찰의 기회로 인식하

고 있다. 또한 빅데이터를 잘 활용하면 미래사회에서 새로운 기회를 창출하고, 위험을 해결하는 사회 발전의 역할 수행이 기대된다.

## 제2절 빅데이터의 구성요소

빅데이터 활용을 위한 요소에는 [그림 2-2] 과 같이 데이터 자원, 데이터를 가공하고 분석·처리하는 기술, 데이터의 의미를 통찰하는 인력 등 세 가지 분야의 전략 수립이 필수적이다.



<그림 2-2> 빅데이터 활용을 위한 3대 요소

출처 : 정지선(2012), “성공적인 빅데이터 활용을 위한 3대 요소”, 정보화진흥원

### 1. 자원

빅데이터를 위한 자원 확보, 즉 빅데이터 품질관리를 위한 자원 확보를 말하며 빅데이터를 관리·처리하는 측면과 함께 활용할 수 있는 기업의 내부, 외부 빅데이터 자원을 수집하는 전략이 필요하다. 빅데이터 자원은 조직의 내부 경계를 넘

어 외부에 걸쳐 존재하고 있다.

빅데이터 환경에서 자원을 확보하는 방안은 생성, 수입(검색), 공유와 연계, 협력, 참여, 오픈 등에 따라 단계적으로 확장 가능하며, 최근에는 웹, 소셜 미디어, 포털 검색 데이터 등을 분석하여 여론 확인, 기업의 마케팅 서비스 등에 적용하는 사례가 증가하고 있다. 또한 국내외적 공공 데이터의 개방을 확대하는 추세이며, 공공 오픈 데이터를 활용한 다양한 환경 분석이 가능하다 구체적인 방법은 진화에 따라 <표 2-3>과 같이 단계적으로 확보해야 한다.<sup>9)</sup>

<표 2-3> 데이터 자원 확보를 위한 단계적 방법

단계	내용과 과제	방법
1. : SILOS (Hoarding)	<ul style="list-style-type: none"> <li>· 조직의 독자적인 데이터 생성, 저장 중심의 단계</li> <li>· 외부 데이터는 인터넷을 통한 수집(검색) 기능</li> <li>· 데이터의 신뢰성과 품질제고 노력 필요</li> </ul>	생성, 저장, 수집(검색)
2. Exchanges (Sharing)	<ul style="list-style-type: none"> <li>· 기업의 데이터를 외부기관들과 상호 교환하는 단계 (1:1 또는 1: n의 공유, 연계)</li> </ul>	연계, 공유
3. Pools (Aggregating)	<ul style="list-style-type: none"> <li>· 특정한 활동이나 목적을 위해 모인 집합, 그룹, 클럽들이 상호협력과 공동의 장 형성(집단)</li> <li>· 표준화된 데이터 풀(Pool)의 연계를 통해 국경을 초월한 정보교환과 상호 이용 가능</li> <li>· 예시 : Global Data Synchronization Network</li> </ul>	참여, 협력
4. Commons (Co-creating)	<ul style="list-style-type: none"> <li>· 오픈 플랫폼을 통한 데이터 공유</li> <li>· 상호 협력과 참여를 통해 공동의 자원 창조</li> </ul>	오픈, 창조

출처: 정지선(2012), “빅데이터 활용을 위한 3대 요소”.

9) 정지선(2012), “빅데이터 활용을 위한 3대 요소”, 한국정보화진흥원.

## 2. 기술

스마트 시대에는 빅데이터의 등장으로 데이터 처리의 전체적인 과정이 업그레이드되고 정보화 시대에 차별적으로 발전하게 되었음 데이터 처리 과정에서 추론의 단계가 본격화 되었다는 것이 특징이다. 새롭게 추가된 추론의 영역은 IT 산업의 성장 동력으로 부상할 수 있으며 의료, 금융, 공공 등을 혁신 할 수 있는 신사업분야로 발돋움하고 있다 특히 상황인식 서비스는 미래전망, 사전대응, 자동화 서비스와 연계되어 차세대 서비스로 각광받고 있고, 개개인의 취향, 관심 있는 정보의 성격, 상태, 개인 의중에 맞는 맞춤형 개인화 서비스, 인공지능 서비스로 서비스의 패러다임이 변화하고 있다.

<표 2-4> 정보화 시대 VS 스마트 시대의 데이터 관련 이슈 변화

구분	정보화 시대(1세대)	스마트 시대(2세대)
저장	관계형/정형 데이터베이스, 데이터 웨어하우스	비관계형/비정형 데이터베이스
검색	검색엔진(text), 포털서비스	자연어/음성·영상/시멘틱 검색서비스
관리공유	KMS, Web 2.0	플랫폼, 소셜 네트워크, 집단지성
분석	경영정보/고객정보/자산정보 분석(ERP, CRM, 데이터마이닝 등)	빅데이터 분석(소셜 분석, 고급분석, BAO, 시각화)
추론	추론(전망) 전 단계 : 과거 데이터 및 현상 분석만 가능	추론에 기반한 서비스 상용화 : 상황인식서비스(미래전망, 사전대응, 자동화서비스), 개인화 서비스, 인공지능 서비스 등

출처 : ‘신가치창출 엔진, 빅데이터의 새로운 가능성과 대응전략’ 정지선(2011)

빅데이터의 처리 프로세스와 새로운 기술이 출현하고 있으며 데이터의 생성, 수집, 저장, 분석, 표현의 처리 과정을 거치며 각 프로세스 마다 세부 영역과 기술이 등장하고 있다. 빅데이터 플랫폼으로는 하둡(hadoop) 데이터 처리. 관련 기술에는

NoSQL, 분석기술에는 데이터마이닝, 기계학습, 자연어처리, 패턴인식, 소셜네트워크 분석 등이 있다. 또한, 분석한 데이터를 보여주는 시각화(Visualization) 기술 등이 있다.<sup>10)</sup>



<그림 2-3> 빅데이터 처리 프로세스

출처: 문혜정(2012), 빅데이터 기술구축과 사례를 중심으로 재구성

10) 김지숙(2012), “빅데이터 활용과 분석기법의 고찰”, 고려대 석사학위논문, p.12.

<표 2-5> 빅데이터 처리 프로세스별 기술 영역

흐름	영역	개요
소스	내부데이터	Database, File Management System
	외부데이터	File, Multimedia, Streaming
수집	크롤링(Crawling)	검색엔진의 로봇을 이용한 데이터 수집
	ETL(Extraction). Transformation. Loading	소스데이터의 추출, 전송, 변환, 적재
저장	NoSQL Databases	비정형데이터 관리
	Storage	빅데이터 저장
	Servers	초경량 서버
처리	MapReduce	데이터의 추출
	Processing	다중업무처리
분석	NLP(NeuroLingustic Programing)	자연어처리
	Machine Learning	기계학습을 통해 데이터의 패턴 발견
	Serialization	데이터 간의 순서화
표현	Visualization	데이터를 도표나 그래픽적으로 표현
	Acquisition	데이터의 획득 및 재해석

출처: Pete Warden(2011). Big data Giossary

### 3. 인력

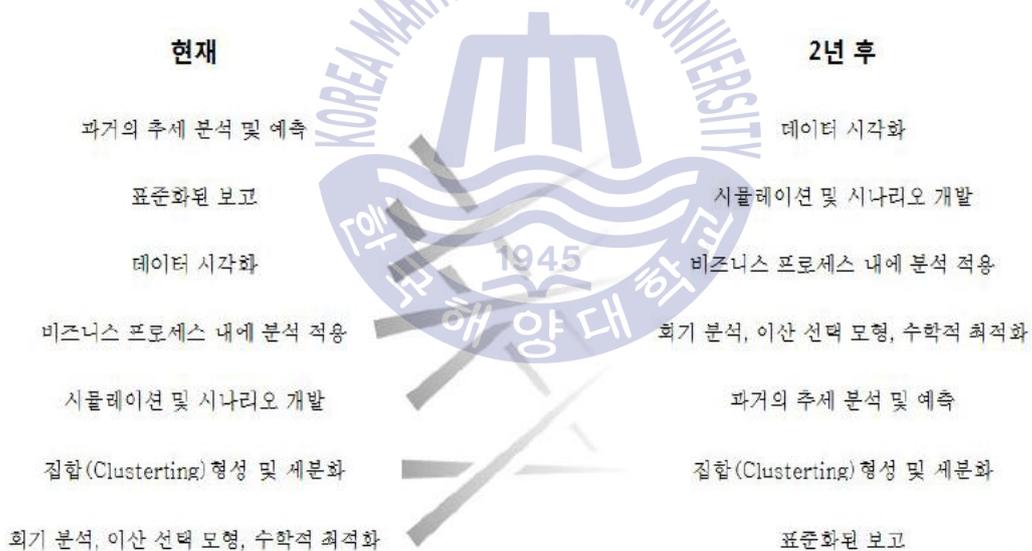
빅데이터 시대가 도래함에 따라 새로운 데이터 분석기술과 모델링 기법을 사용하여 통찰력 있는 결론을 발견해내는 일을 하는 새로운 직업이 주목받고 있다. 구조화되지 않은 대규모 데이터 속에서 정보를 찾아내는 데이터 과학자(Data Scientist)는 스마트 시대에 최고의 인재로 뽑히고 있으며 21세기 유망직업 중 하나로 부각되고 있다<sup>11)</sup>.

그러나 현재 데이터 과학자의 역량을 갖춘 인재는 매우 부족한 실정으로 McKinsey 2011에 따르면 미국에서는 2018년까지 14만 ~ 19만명의 전문가와 150만명 정도의 데이터 관리자와 분석인력이 부족할 전망을 예측하였고. EMC의 글로벌 조사결과, 향후 5년 동안 데이터 사이언티스트의 수요가 공급을 뛰어넘어 인재 부족이 심화될 것으로 예측하고 있다.<sup>12)</sup>

11) 정지선(2011), “신가치 창출 엔진 빅데이터의 새로운 가능성과 대응전략”, p.26.

데이터 사이언티스트는 기본적인 자질이 수학, 공학능력, 비판적 시각과 이를 발작성할 수 있는 글쓰기 능력, 대화능력, 호기심과 개인의 행동 소양<sup>13)</sup>이라고 말하며 대규모의 데이터를 분석한 결과를 생생하고 차별화되게 시각화하여 쉽게 전달하는 역량도 매우 중요하다고 하겠다.

‘MIT Slon’은 데이터 전문가들을 대상으로 조직에 가치를 제공하는 최고의 분석기법을 조사하여 주요 역량을 우선순위를 확인한 결과 <그림 2-4>와 같이 나타났다. 그리고 빅데이터의 도입을 통해 향후 2년 내 달성할 것으로 기대하는 바를 살펴보면, 과거의 추세 분석 및 예측이나 표준화된 보고서와 같은 기초적인 데이터 분석 결과는 그 중요도가 하락하는 반면, 데이터의 시각화 기술이나 시뮬레이션 및 시나리오 개발을 통한 비즈니스 프로세스 분석 적용이 점차 중요한 기능으로 부각될 것으로 전망하고 있다.



<그림 2-4> 데이터 중심 관리들은 어느 곳을 향하고 있는가?

출처 : MT Slon Management Review(2011), “Big Data” Analytics and the Path.

12) 정지선(2012), “빅데이터 활용을 위한 3대 요소”, 한국정보화진흥원.

13) Forbes. ‘Amazon’s John Rauser on “What is a Data Scientist?”(2011.10.7.) 이지영, 데이터 과학자가 되려면 블로터닷넷(2012.3.18.) 재인용.

한편, 국내·외 데이터 과학자에 대한 공급부족을 인식하고 유능한 인재육성 프로그램과 교육과정을 개설하고 있다. 충북대학교는 비즈니스 데이터 융합학과를 신설하여 종합적으로 데이터를 분석할 수 있는 인재를 육성하고 있으며, 서울대학교는 다양한 학문분야의 융합을 위해 데이터 과학과 지식창출 연구센터를 설립하였다. 이외에도 SAS는 자사의 분석 툴을 활용한 빅데이터 분석 및 적용기법을 활용하는 Analytics Conference를 개최하였고, 다음 소프트는 Opinion Mining Workshop을 개최하여 소셜 빅데이터의 분석과 활용방안을 공유하고 있다.<sup>14)</sup>

### 제3절 빅데이터 분석 인프라 및 분석기법

#### 1. 빅데이터 분석 인프라

빅데이터와 관련하여 인프라 기술들이 4차 산업혁명에 더불어 기존 인프라와 새로운 기술의 등장으로 인프라 기술들이 병합하고 있다. 테라바이트 또는 페타바이트 규모의 빅데이터 분석기법들은 인터넷에 대부분 활용되고 있고, 엄청난 규모의 데이터를 분석하기 위해서는 기존 기술과의 차별화된 구축이 필요하다. 현재 기술은 대용량 데이터 처리능력을 위한 분산처리 기술인 하둡, 의미 분석기술과 진보된 알고리즘 및 데이터 마이닝 기술을 보여주는 R, 비정형 데이터를 처리하기 위한 기술인 NoSQL 등 이 있는데 빅데이터를 분석하는 툴에 대해 자세히 알아보하고자 한다.

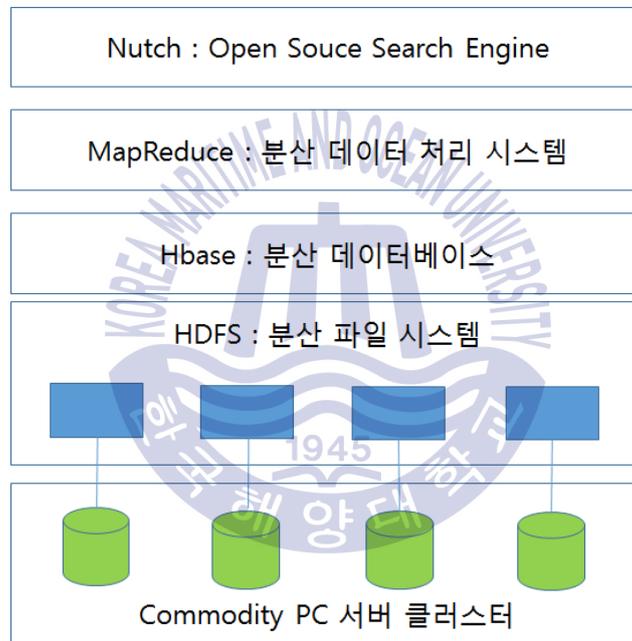
##### 1) 하둡(Hadoop)<sup>15)</sup>

하둡은 오픈 소스 분산처리기술을 지원하는 소프트웨어 프레임워크로 정형데이터 및 비정형데이터 분석에 가장 많이 사용되고 있는 대표적인 솔루션이다. 다수

14) 정지선(2012), “빅데이터 활용을 위한 3대 요소”, 한국정보화진흥원.

15) 김지숙(2012), “빅데이터 활용과 분석기법의 고찰”, 고려대석사학위논문, p.22.

의 서버를 묶어 분산 처리하는 플랫폼이며, 분산처리 하는 ‘Map’ 단계와 결과를 취합하는 ‘Reduce’ 단계로 이루어진 구글의 MapReduce 모델을 본떠 만든 것으로 Apache Group에 의해 제작되고 배포되었다. HDFS(Hadoop Distributed File System), HBase는 구글의 파일 시스템인 GFS(Google File System)와 빅데이터에 영향을 받으며, 비용 효율적인 X86 서버로 가상화된 대형 스토리지(HDFS)를 구성하고, HDFS에 저장된 빅데이터를 손쉽게 분산 처리할 수 있는 Java 기반의 MapReduce 프레임워크를 제공한다.



<그림 2-5> 하둡의 구조와 분산처리기술

하둡 자체는 파일 시스템과 분산 처리 플랫폼이지만 하둡을 중심으로 다양한 예코 시스템이 구축되어 있다. 분산 데이터베이스인 HBase, 검색엔진 너치(Nutch), 대규모 데이터셋 탐색용 흐름 언어와 실행 환경 및 관계형 대수 쿼리 언어 인터페이스인 피그(Pig), HDFS에 저장된 데이터를 관리하는 데이터웨어하우징 솔루션인

하이브(Hive) 등이 포함되어 있다. 하둡은 아파치 소프트웨어 라이선스에 따라 전체 프레임워크가 오픈 소스이기 때문에 기본 소프트웨어에 대한 라이선스 비용이 없고, 또한 고가의 하드웨어나 고성능 프로세서가 필요하지 않고 저가의 상용 서버를 사용할 수 있어 비용 절감 효과가 높은 것으로 평가되고 있다<sup>16)</sup>.

## 2) R

오픈 소스 프로젝트 RDMS 통계 계산 및 시각화를 위한 언어 및 개발환경을 제공하며 R언어와 개발환경을 통해 기본적인 통계 기법부터 모델링, 최신 데이터마ining 기법까지 구현 및 개선이 가능하다. 이렇게 구현한 결과는 그래프 등으로 시각화할 수 있으며 Java나 C, Python 등의 다른 프로그래밍 언어와 연결도 용이하다. Mac OS, 리눅스/유닉스, 윈도우 등의 대부분의 컴퓨팅 환경을 지원하는 것도 장점이다.

또 R의 경우에는 GNU 프로젝트로 만들어져서 무료로 사용될 수 있고, 다양하게 활용할 수 있는 패키지들이 빠르게 추가되고 있어 최근 들어 더욱 주목받고 있다. 위의 장점들로 인하여 R은 통계분석 분야에서 인지도를 높여왔으며, 하둡 환경 상에서 분산처리를 지원하는 라이브러리 덕분에 구글, 페이스북, 아마존 등의 데이터 분석이 필요한 기업에서 대용량 데이터 통계분석 및 데이터 마이닝을 위해 널리 사용되고 있다.

## 3) NoSQL

NoSQL은 Not-Only SQL의 줄임말로써 기존 관계형 데이터베이스에서 관계형 특징(스키마, 테이블간 조인 연산 등)을 포함하지 않고 다른 특징을 가지는 데이터베이스를 일컫는 말이다. 즉 전통적인 관계형 데이터베이스와 다르게 설계된 비관

16) 김지숙(2012), “빅데이터 활용과 분석기법의 고찰”, 고려대 석사학위논문, p.23.

계형 데이터베이스를 의미한다. 대표적인 NoSQL 솔루션으로는 Cassandra, Hbase, MongoDB 등이 존재한다. NoSQL은 테이블 스키마(Table Schema)가 고정되지 않고, 테이블간 조인(Join)연산을 지원하지 않으며, 수평적 확장(Horizontal Scalability)이 용이하다는 특징을 가진다. 이러한 NoSQL은 데이터의 양이 증가하고 저장할 데이터가 지속적으로 변화하며 데이터의 요구가 일관적이지 않은 웹시장의 요구를 수용하고 기존의 관계형 데이터베이스의 수평적 확장성의 한계를 해결하고자 대두되었다. 관계형 데이터베이스의 경우 일관성(Consistency : 모든 노드는 같은 시간에 같은 데이터를 보여줘야 한다.)과 유효성(Availability : 일부 노드가 다운되어도 다른 노드에 영향을 주지 말아야한다)에 중점을 두고 있는 반면, NoSQL 기술은 분산가능성(Partition Tolerance : 네트워크 전송 중 일부 데이터를 손실하더라도 시스템은 정상 동작을 해야 한다)에 중점을 두고 일관성과 유효성은 보장하지 않는다. 이것은 일관성, 유효성 분산가능성 중 2가지만 보장이 가능하다는 분산 데이터베이스 시스템 분야의 CAP 이론에 따른 것이다. 따라서 유연한 데이터 처리를 위해서는 NoSQL 기술이 적합하지만, 안정성이 중요한 시스템에서는 오랫동안 검증된 관계형 데이터베이스를 채택할 필요가 있다.

## 2. 빅데이터 분석 기법

빅데이터 처리의 특징을 만족하게 하기 위해 다양한 스토리지, 컴퓨터 기술 및 분석 기법들이 개발되었다. 디지털 시대에는 쌓여 있는 데이터 속에서 유용하고 가치 있는 정보를 찾기 위한 노력이 끊임없이 진행 중이며 실생활 속에서 축적되는 다양한 유형의 데이터가 증가할수록 데이터의 활용가치는 무한히 상승하고 있다.

대부분의 분석기법들은 통계학과 전산학, 특히 기계학습/데이터 마이닝 분야에서 이미 사용되던 기법들이며, 이 분석기법들의 알고리즘을 대규모 데이터 처리에 맞도록 개선하여 빅데이터 처리에 적용시키고 있다. 빅데이터 분석기법들에는 다

음과 같은 기법들이 있다.<sup>17)</sup>

<표 2-6> 빅데이터 활용에 따른 분석 방법의 변화

분류		현재 (AS-IS)	유합지식기반(TO-BE)
활용 변화	데이터개방	웹 기반 인터페이스	원본/분석/가시화 3계층
	이슈접근	후 집계/원인 파악하는 사후대책반	실시간 이슈 탐지를 통한 선 대응적 기획반
	활용형태	부처별 수직적 활용	범부처/민간 수평적 분석
	주체	업무 운영담당자	분석, 기획 담당자
분석 변화	분석대상	정형화된 DB 데이터	정형데이터+논, 질문 등의 비정형데이터
	규모	기가~테라바이트급	페타~제타 바이트급
	분석범위	단일 저장소	다중 저장소
	적용시간	일괄처리	인타임 처리
	데이터	저장 후 분석	흐림(on-the-fly) 분석

출처 : 국가정보화전략위원회(2011), 빅데이터를 활용한 스마트 정부 구현(안).

#### 1) 데이터 마이닝(Data Mining)

데이터 마이닝은 대용량의 데이터, 데이터베이스 등에서 감춰진 지식, 기대하지 못했던 경향, 새로운 규칙 등의 유용한 정보를 발견하는 과정 데이터 마이닝을 통해 정보의 연관성(순차 패턴, 유사성 등)을 파악함으로써 가치 있는 정보를 만들어 의사결정에 적용하는 기술이다. 다시 말해 기업이 보유하고 있는 일일 거래자료, 고객자료, 상품자료, 마케팅 활동의 피드백 자료와 기타 외부자료를 포함하여 사용 가능한 데이터를 기반으로 숨겨진 지식, 기대하지 못했던 패턴, 새로운 법칙과 관계를 발견하고 이를 실제 경영의 의사결정 등을 위한 정보로 활용하자 하는 것이

17) 이철웅(2013), “글로벌 온라인 비즈니스를 위한 빅데이터 활용에 대한 연구”, 한성대석사학위논문, pp.15~16.

다. 또한 단순한 질의어로 추출할 수 없는 형태의 ‘유용한’ 정보를 찾아내고 이를 바탕으로 데이터에 대한 통찰(Insight)을 얻는 것을 의미하기도 한다<sup>18)</sup>.

<표 2-7> 텍스트 마이닝 과정

과정	주요내용
Text 문서	· DB Contents. 텍스트 기반의 문서(웹문서, 오피스문서, 메일 등)
Text 전처리	· 문서 내의 표현되어 있는 단어·구절에 해당하는 내용을 언어 분석·처리 과정으로 가공하여 데이터로 표현
의미정보변환	· 전처리된 데이터 중 의미 있는 정보를 선별하여 저장(불용어 처리, 대소문자처리, Stemming 처리)
의미정보추출	· 복잡한 의미정보의 표현을 단순화하고 도메인에 적합한 정보를 문서의 의미데이터로 저장
패턴 및 경향분석	· Feature 정보를 기반으로 문서를 자동 군집하거나 자동 분류하는 등의 정보 재생산
정보표현 및 평가	· 새롭게 생성된 정보를 사용자에게 시각화하여 효과적으로 표현, 평가과정을 통해 텍스트 마이닝의 처리과정 중 문제가 되는 부분을 수정 및 보완하여 품질과 성능을 높이는데 활용

출처 : 하연 ‘빅데이터와 DBMS의 시장전망’(2015)

## 2) 텍스트마이닝(Text Mining)

텍스트 마이닝은 비/반정형 텍스트 데이터에서 자연어처리(Natural Language Processing)기술을 기반으로 유용한 정보를 추출, 가공하는 것을 목적으로 하는 기술이다. 텍스트 마이닝 기술을 통해 방대한 텍스트 문치에서 의미 있는 정보를 추출해내고, 다른 정보와의 연계성을 파악하며, 텍스트가 가진 카테고리를 찾아내거나 단순한 정보검색 그 이상의 결과를 얻어낼 수 있다. 컴퓨터가 인간이 사용하는 언어(자연어)를 분석하고 그 안에 숨겨진 정보를 발굴해 내기 위해 대용량 언어자원과 통계적, 규칙적 알고리즘이 사용되고 있다<sup>19)</sup>.

18) 넥스트기술연구소 데이터마이닝팀(2001), “웹마이닝 Microsoftware 특집”, Vo.2, pp.242-243.

텍스트 마이닝 애플리케이션의 성공 사례가 늘어남에 따라 정형 및 비정형화된 데이터를 동시에 분석하는 텍스트 마이닝 기술은 전 세계 모든 곳에서 건전한 조직의 필수 요소로 자리 잡게 될 것으로 전망되고 있다. 주요 응용분야로는 문서 분류(Document Classification), 문서 군집(Document Clustering), 정보 추출(Information Extraction), 문서 요약(Document Summarization) 등이 있다<sup>20)</sup>.

### 3) 오피니언 마이닝(Opinion Mining)

최근 새로운 여론 분석기술로 주목받고 있는 오피니언 마이닝은 소셜미디어와 웹사이트 등에 나타난 여론과 의견을 분석하여 유용한 정보로 재가공하는 기술이라고 하며 평판분석(Sentiment Analysis)라고도 한다. 오피니언 마이닝 혹은 감정 분석을 활용하면 네티즌이 그들에 대해 이야기하는 댓글이나 포스팅 등을 긍정, 부정, 중립으로 분류하여 더 객관적이고 정확하게 평판을 파악할 수 있다. 오류를 피하기 위하여 사용되는 오피니언 마이닝 연구의 3단계는 다음과 같다. 첫째 ‘주관성 분석’으로 주어진 텍스트가 주관적인지 객관적인지를 결정하는 것으로, 주어진 텍스트에 나타난 저자의 태도를 판단하는 단계이다. 둘째 ‘극성분석’으로 텍스트가 주관적인 의견을 갖고 있는 경우 긍정인지 부정인지를 분류하는 단계이다. 셋째 ‘극성의 정보 분석’으로 주관적인 텍스트에 대하여 긍정적인 정도와 부정적인 정도를 측정하는 단계이다. 오피니언 마이닝은 특정 서비스 및 상품에 대한 시장규모 예측, 소비자 반응, 입소문 분석 등에 활용되고 있으며 공공분야의 경우 민원의 원인이나 문제점 등을 파악하는 것이 용이해 서비스를 개선할 수 있다. 또한 기업은 특정제품에 대한 고객의 반응을 빠르게 파악하고 선호도를 역으로 추론하는데 효

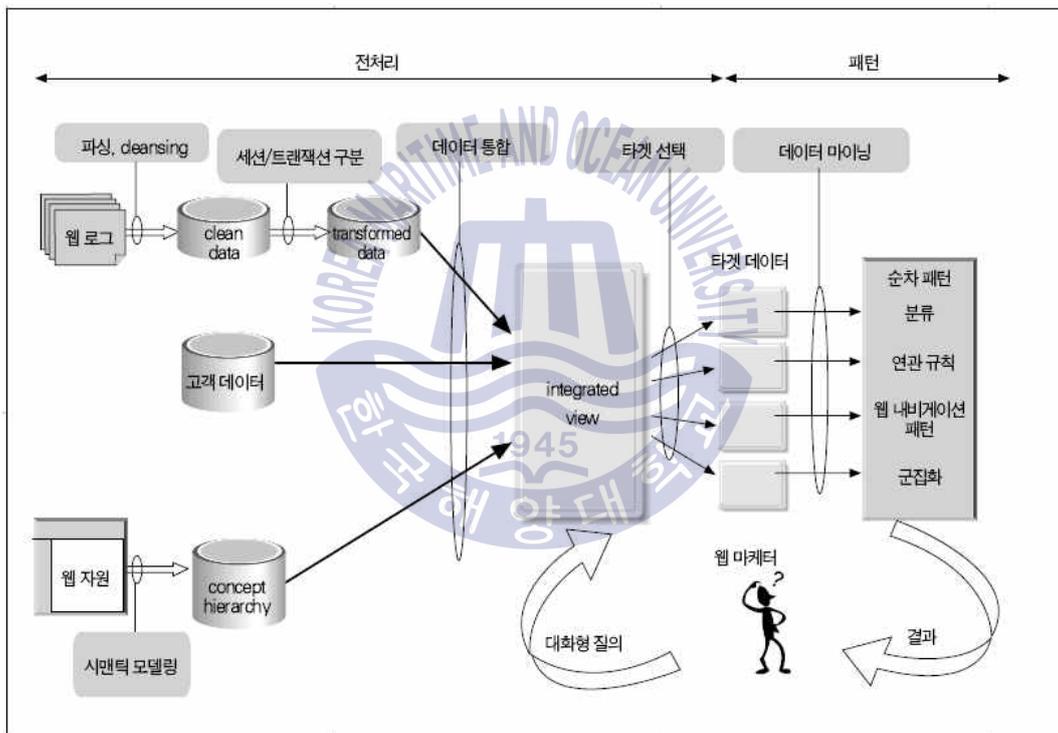
19) 조성우(2011), “빅데이터 시대의 기술”, KT 종합기술원 중앙연구소.

20) 이철웅(2013), “글로벌 온라인 비즈니스를 위한 빅데이터 활용에 대한 연구”, 한성대 석사 학위논문, pp.15~16.

과적으로 활용할 수 있다. 정확한 오피니언 마이닝을 위해서는 전문가에 의한 선 호도를 나타내는 표현·단어 자원의 축적이 필요하다<sup>21)</sup>.

#### 4) 웹마이닝(Web Mining)

인터넷상에서 수집된 정보를 마이닝 방법으로 분석하는 통합하는 기법 웹 마이 닝은 콘텐츠 마이닝(웹 검색, 수집 데이터) 구조마이닝(웹 사이트 구조), 활용 마이 닝(사용자 이용형태) 등으로 세분될 수 있다<sup>22)</sup>.



<그림 2-6> 웹 사용 마이닝 시스템

출처 : 넷스투기술연구소 데이터마이닝팀, 2001. 5.

21) 조성우(2011), “빅데이터 시대의 기술”, KT 종합기술원 중앙연구소.

22) 정지선(2012), “성공적인 빅데이터 활용을 위한 3대 요소”, 한국정보화진흥원.

웹마이닝은 웹에서 발생하는 모든 데이터를 분석 대상으로 삼는다. 이러한 데이터로는 서버 접속 로그 데이터(server access log data), 사용자 등록 정보(user registration data 또는 profile), 사용자 세션(session), 또는 트랜잭션(transaction), ERP 데이터(enterprise resource planning data)가 있다. 웹 마이닝은 데이터 마이닝의 한 분야이기도 하지만, 기존의 데이터 마이닝 알고리즘, 웹 데이터의 전처리(preprocessing)를 위한 데이터 웨어하우징 기술, 그외에 웹 환경 관련 기술이 연관된 데이터 마이닝을 포함하는 개념으로도 이해할 수 있다. 일반적으로 웹 마이닝은 대상이 되는 웹 데이터(구조, 내용, 사용)에 따라 웹 구조 마이닝(web structure mining), 웹 내용 마이닝(webcontent mining), 웹 사용 마이닝(web usage mining) 분야로 나눌 수 있다. 웹 구조 마이닝(web structure mining)은 웹 사이트와 웹 페이지의 구조적요약 정보를 얻는 것을 목표로 한다. 웹 사이트의 구조적 정보란, 웹 페이지사이의 하이퍼링크(hyperlink)를 통한 그래프(graph) 구조를 뜻한다. 웹 내용마이닝(web content mining)은 실제 웹 사이트를 구성하고 있는 페이지로부터 의미 있는 내용을 추출하는 기법이다. 이는 일종의 정보 추출(information retrieval)이라고도 할 수 있고, 텍스트 마이닝(text mining) 기술과도 밀접한 관련이 있다. 다시 말하면 온라인에 있는 방대한 웹 데이터(텍스트, 그림, 사운드 등)에서 유용한 정보를 자동으로 찾는 기술이다. 웹 사용 마이닝(web usage mining)은 웹 사용자의 사용 패턴을 분석하는 것이다. 이를 통해 웹사용자의 행동을 접속 통계 정보 이상으로 이해할 수 있고, 또한 웹 페이지의 이용 패턴을 알 수 있게 된다. 결국 이 정보는 사용자에게 더욱 친숙하게 페이지를 재구성하거나, 웹 서버 로드 밸런싱, 사용자별 맞춤형 웹 페이지 구성 등에 이용된다.<sup>23)</sup>

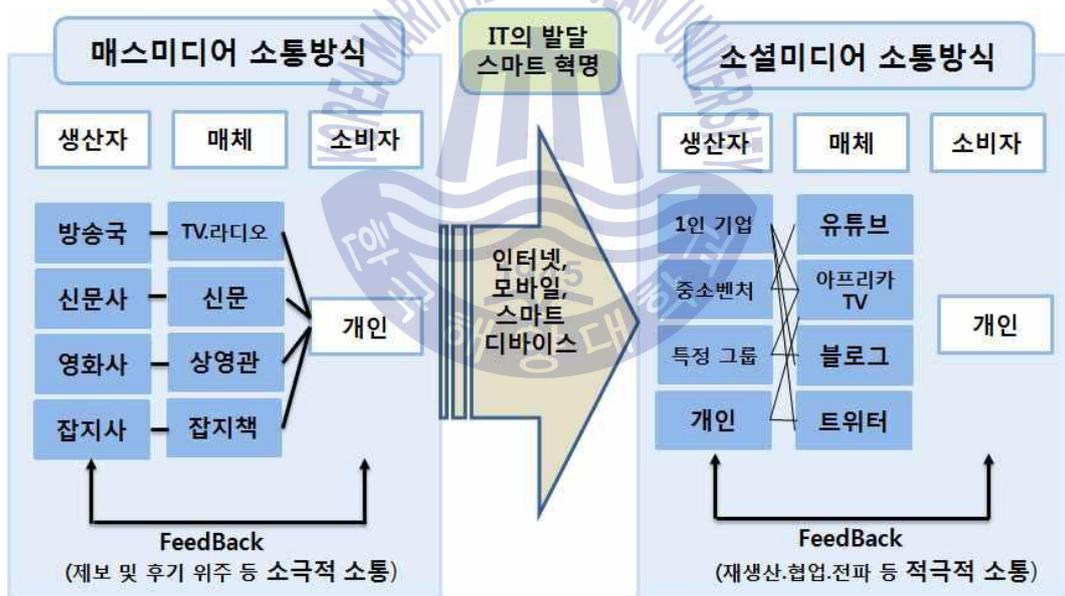
---

23) 이철웅(2013), “글로벌 온라인 비즈니스를 위한 빅데이터 활용에 대한 연구”, 한성대학교 석사학위논문. pp.20~22.

5) 소셜분석, 소셜 마이닝(Social Mining)

소셜 네트워크의 분석은 수학의 그래피론에 뿌리를 두고 있으며 소셜 네트워크 연결구조 및 연결 강도 등을 바탕으로 사용자의 명성 및 영향력을 측정하여 소셜 네트워크상에서 입소문의 중심이나 허브 역할을 하는 사용자를 찾는 데 주로 활용된다. 소셜 미디어에 올라오는 글과 사용자를 분석해 소비자의 흐름이나 패턴 등을 분석하고, 판매나 홍보에 적용 마케팅 분야뿐만 아니라 사회의 흐름과 트렌드, 여론변화 추이를 읽어내는 소셜 미디어 시대의 새로운 마이닝 기법이다.

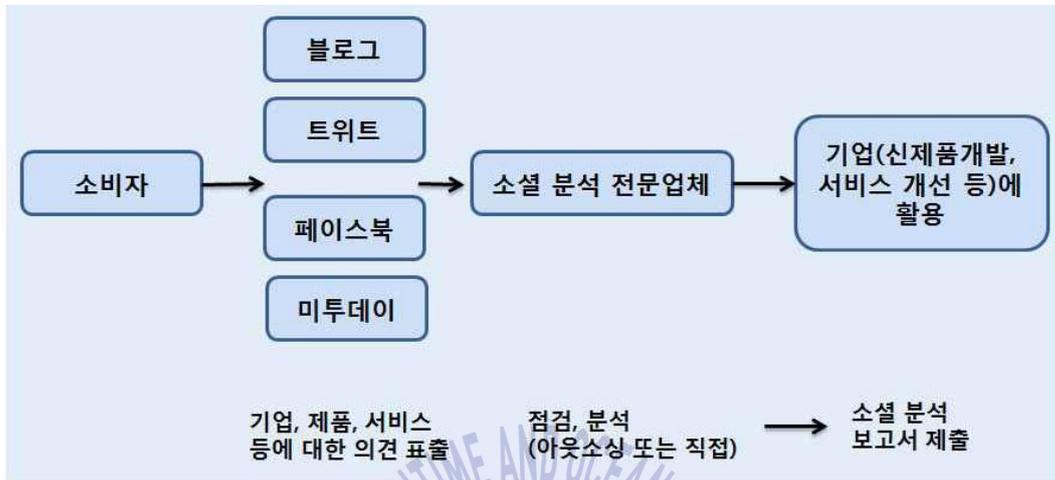
IT의 발달과 스마트 혁명의 본격화로 더욱 활성화된 소셜 미디어는 <그림 2-7>과 같이 정부와 국민, 기업과 소비자, 개인과 개인의 소통방식에 혁신적 변화를 가져왔다.



<그림 2-7> IT의 발달과 소통방식의 변화

출처 : 한국정보화진흥원.

소셜 네트워크 분석은 텍스트 마이닝 기법에 의해 주로 이루어져 있다. 소셜 프로세스과정은 <그림 2-8>과 같다.



<그림 2-8> 소셜 프로세스

소셜 네트워크 분석 기술은 이전부터 있었으나 20세기 말 네트워크 사이언스의 출현으로 네트워크 분석 기술이 급격히 발전하게 되면서 그 결과 기존 기술로는 불가능했던 대규모 소셜네트워크 데이터 분석을 빠르고 정확하게 할 수 있게 된 것이다<sup>24)</sup>.

SNS는 개인을 노드(Node) 개인의 사회적 관계를 링크(Link)로 간주하면 소셜 네트워크를 구할 수 있고 이렇게 형성된 소셜 네트워크에서 다음의 4단계를 통해 정보를 추출 및 분석할 수 있다. 첫째 소셜 네트워크의 위상학적 구조(Network Topology Structure) 분석으로 네트워크 전반적 특성을 파악한다. 둘째 네트워크 구조의 시간에 따른 진화를 분석한다. 셋째 네트워크상의 각 노드(사용자)가 생산 확산시키는 콘텐츠(포스트, 댓글, 리트윗, 동영상, 링크 등) 흐름을 분석한다. 넷째 종합하여 각 개인 또는 그룹의 소셜 네트워크 내 영향력, 관심사, 성향 및 행동 패턴을 분석 추출한다. 소셜 네트워크 분석의 활용 효과는 이미 각 기업이 빅데이터

24) 하연(2012), “빅데이터와 DBNS의 시장전망”, 하연편집부.

에 주목하면서 SNS 데이터 분석기술을 통해 방대한 비정형 데이터들을 분석하고 이를 비즈니스에 활용하고 있다. 하지만 그 부작용 최소화를 위한 자체 모니터링, 위험 완화 프로그램 개발 등 관련 기업·시장의 사회적 책임이 강조되어야 하며, 프라이버시 보호 등 부작용 대응을 위한 기술개발 및 산업육성이 지원되어야 한다

#### 6) 현실 마이닝(Reality Mining)

사람들의 행동패턴을 예측하기 위해 사회적 행동과 관련된 정보를 기기(휴대폰, GPS 등)를 통해 얻고 분석하는 기술이다. 휴대폰 등 모바일 기기들을 통해 현실에서 발생하는 정보를 기반으로 인간관계와 행동 양태 등을 추론한다.

#### 7) 군집분석(Cluster Analysis)

군집분석은 각 객체(대상)의 유사성을 측정하여 높은 대상 집단을 분류하고, 군집에 속한 객체들의 유사성과 서로 다른 군집에 속한 객체간의 상이성을 규명하는 방법을 말하는데 유사 특성의 군을 발굴하는데 사용된다.

군집분석을 위하여 가장 흔히 사용하는 자료는 그림과 같이 간격척도 혹은 비율척도로 측정된 거리 값(Distance measures)의 경우에 따라 서열척도로 측정된 값들로 군집분석이 가능하다. 계층적 군집분석은 문제점은 없지만 군집의 수를 사전에 지정해 주어야 한다. 현실적으로 계층적 방법에 의해 군집화를 한 다음 그 결과로부터 가장 적절한 수의 군집 수를 결정하여 다시 비 계층적 방법에 의해 분석하면서 그 수를 지정하는 방법을 사용한다. 아울러 계층적 군집 분석에서 나타나는 예외 값들을 이때 제거하는 것이 바람직하다.<sup>25)</sup>

25) 김정숙(2012), “빅데이터 활용과 관련기술 고찰”, 「한국콘텐츠학회지」, 제10권 제1호, pp.9-116.

## 제4절 빅데이터를 이용한 부산항 정보의 구축과 활용 사례

### 1. 항만물류정보 역할 및 필요성

#### 1) 물류정보관리시스템의 개념

과거 물류관리에 필요한 정보시스템은 물류활동에 필요한 각 공정별(생산, 판매, 조달, 반품 등)로 원하는 정보를 얻거나 그 정보를 미리 받아 물류 단계의 효율성을 높이고자 노력하였다.

물류관리의 목적은 최소의 비용으로 고객에 대한 서비스를 극대화시키는 것으로써 최소의 비용을 들여 적절한 품질과 수량의 물품을 적절한 시기와 정보에 공급하는 목적이다. 또한 고객의 욕구에 경제적으로 부응하도록 수송과 보관뿐만 아니라 이에 수반되는 하역, 포장, 정보처리 등의 물류 활동을 총괄적으로 계획하고 통제하는 것이 물류관리이다.<sup>26)</sup> 이러한 물류관리를 위해서는 정보의 뒷받침이 불가결한 요소이다. 소비자가 원하는 것을 원하는 때에, 일관성 있게 어떻게 제공할 것인가에 대한 것이 기업의 물류시스템이며 이를 가능토록 지원하는 것이 물류정보시스템이다.<sup>27)</sup>

물류정보시스템은 계획과 통제를 위해 필요한 정보를 효과적으로 제공함으로써 물류활동을 지원하고 있으며, 이를 두 가지로 구분할 수 있는데 그 하나는 원료조달에서 시작하여 생산품이 최종소비자에게까지 전달되는 물리적 과정이고 다른 하나는 이러한 활동을 조정하고 통제하기 위해 필요한 물류정보체계이다. 따라서 컴퓨터의 발전과 더불어 신속·정확한 자료처리와 함께 가능해진 효과적인 물류정보체계는 효율적인 물류시스템 구축과 운영에 필요한 요소이다.<sup>28)</sup>

26) 허동욱(1998), “물류정보시스템과 물류성가에 관한 실증적 연구”, 원광대 박사학위논문.

27) 이철영(1998), 「항만물류시스템」, 효성출판사, p.459.

28) 양익모(1996), “물류기점과 성과에 관한 연구”, 건국대 박사학위논문.

즉, 물류관리에 관계된 정보를 활용하여 물류 제 기능의 원활화와 효율화를 도모하여 물류관리과정에서 발생하는 정보의 전달과 처리를 정확하고 신속하게 함으로써 수송, 배송, 재고관리, 주문처리 등 물류의 모든 기능 영역들을 지원하는 것이다.

## 2) 물류정보관리시스템의 역할

물류정보시스템은 기업에 있어서 단순한 원가절감, 생산성 향상 차원을 벗어나서 경쟁우위 확보, 시장 확인 및 방어, 기업혁신, 생존 전략을 펼칠 수 있는 분야로서 리엔지니어링의 최적 대상이며 가장 큰 성과를 기대할 수 있는 분야이다. 물류정보시스템을 효율적으로 활용함으로써 기업은 무재고 시스템의 토대를 마련하고 물류정보의 경영 전략적 활용가치를 창출하며 적량 적기 납품 실현은 물론, 나아가서는 생산과 영업부서간의 불신까지도 해소할 수 있다.<sup>29)</sup>

Stock, Lambert(1987)는 물류정보시스템의 역할을 구체적으로 나타내면 다음과 같이 말하고 있다.<sup>30)</sup>

### ① 주문정보를 정확하게 파악하고 전달하는 역할

주문은 여러 가지 형태로 기업에 도착하는데 이러한 주문을 정확히 파악하는 것이 물류정보시스템의 첫 번째 기능이다.

### ② 물건의 움직임을 정확히 파악하고, 전달하는 기능

물건의 움직임을 정확히 파악하여 전달하는 것은 물류정보시스템의 기본적인 기능이다. 왜냐하면 주문한 물건이 어디에 있는지 정확히 알지 못하면, 창고업무, 배송업무 등 제반 업무가 원활히 이루어질 수 없기 때문이다.

29) 김수옥 물류정보시스템의 기능과 역할, 서울대학교 경영대학 전자상거래 지원센터 2005. pp.8~9

30) Stock, J. R & D. M. Lambert. Strategic Logistics Management. 2nd. Homewood. Illinois. Irwin. 1987. pp.10~22.

### ③ 고객에게 정보를 제공하는 기능

고객에게 납기정보, 주문정보, 화물추적정보, 시장상황정보 등을 제공함으로써 고객에게서 신뢰를 얻을 수 있으며, 고객중심의 물류업무가 가능해진다. 즉, 물류정보시스템을 통해서 고객만족수준을 높이는 것이 가능해진다.

### ④ 여러 계획과 실적을 잘 통제하는 기능

물류 각 기능의 지표를 설정하여 관리함으로써 실적과 대비하여 개선해 나갈 수 있다. 또 물류정보시스템의 역할은 첫째, 물류활동의 구체적인 실현을 위해 수송, 보관, 하역, 유통가공 및 포장 등의 모든 활동에 필요한 정보를 신속하게 전달하는 것이며, 둘째는 모든 물류 활동을 종합한 물류관리의 토털시스템을 구성하고 전체로서 그 효율화를 도모를 그 역할로 제시하고 있다.<sup>31)</sup>

결국 물류정보시스템은 여러 계획과 실적을 잘 통제하는 역할을 수행하고 물류 전반에 걸쳐서 적정한 재고수준의 유지, 수요와 공급을 조정, 리드타임의 감소, 수송효율의 향상, 하역작업 효율의 향상, 사무관리 자동화, 출하 및 배송의 정확성을 향상, 판매 기능에 대한 적절한 지원과 총 물류비용을 절감시켜 주는 역할을 하는 것이다.

### 3) 물류정보관리시스템의 필요성

기업들은 시장 개방과 국내외 기업 간 경쟁 심화로 매우 어려운 환경에 직면해 있다. 가격 및 서비스 경쟁의 심화, 수요의 고도화 및 개성화, 정보와의 진전 등에 의해 기업의 생산과 판매, 물류전략에 많은 변화가 일어나게 되었다.<sup>32)</sup> 이러한 시점에서 원가절감의 최후영역 이라고 불리는 물류에 대해 기업은 물론 학계에서도

31) 박준성(1998), “물류관리 정보시스템의 역할과 요건”, 물류신문(6/17).

32) 윤소영(2007), “물류정보시스템에 있어서 RFID시스템의 도입방안에 관한 연구”, 부경대 석사학위논문 국제통상물류학과 pp. 9

관심을 가지고 우선적으로 연구되어야 할 과제로 인식되고 있다. 이에 따라 물류에 관한 많은 연구가 학계나 산업계 등에서 진행되고 있다. 그러나 물류는 그 성격상 경영분야에서 시스템 개념이 잘 적용되어야 할 분야임에도 불구하고 이에 대한 인식의 부족으로 기업의 물류생산성 향상에 장애요인이 되고 있다.<sup>33)</sup>

물류정보시스템은 물류활동을 촉발시키는 역할을 하며 물류시스템의 유지에 필수적이다. 즉, 상거래를 구체적으로 실현하기 위해 수송, 보관, 포장 등의 물류기능을 현실적으로 실행할 수 있도록 그 내용을 정확하고 신속하게 전달하는 기능과, 이러한 여러 기능을 통합된 시스템으로 구성하여 전체적으로 효율을 극대화 하는 기능을 한다. 따라서 앞으로의 물류조직의 성공여부는 이러한 물류정보시스템의 효율성을 어떻게 높일 수 있느냐에 달려 있다고 해도 과언이 아니며, 물류정보의 효율성을 높이기 위해서는 결국 다양한 정보기술들을 적절히 평가, 수용하고 이로부터 얻어지는 정보를 인수, 처리, 가공할 수 있는 체계적인 시스템을 개발해야 한다. 이와 동시에 이를 적시에 기업 내부는 물론 물류 기업 외부까지 전달할 수 있는 네트워크를 구축해야 한다.<sup>34)</sup>

그러므로 물류 관리를 보다 효과적으로 수행하기 위해서는 개별적인 물류기능들을 단일 시스템으로 통합하여, 전체 기업 내의 특정 세부 기능만을 최적화하거나 전체 시스템 내 가 기능의 목표가 상충되는 것을 피하고, 조정과 통합을 통해기업 전체의 이익을 극대화시켜야 할 것이다. 이는 고객 서비스 수준의 향상과 물류비용의 절감, 그리고 시장 환경의 변화에 대한 민첩한 대응 등과 같은 이점을 기대할 수 있을 것이다.

---

33) 최성배(2006), “기업 물류특성이 물류정보시스템활용 및 물류성과에 미치는 영향에 관한 연구”, 한양대 석사학위논문.

34) Novac. R. A., L M Rinehart, M.V. Wells(1992), “Rethinking Concept: Foundating in Logistics Management”, *Journal of Business Logistics*, Vol.13 No.2. pp.223-237.

## 2. 항만물류정보관리 문제점과 사례

일반 기업은 물류비용의 증가로 인해 기업 내 생산성 향상을 위해 많은 비용을 투자하여도 기업 내외의 물류문제를 해결하지 않고서는 경쟁력 제고와 고객 만족, 경영 효율화 등의 효과를 보는 것이 어려운 상황이다.

국내 물류업계의 경우, 최근 전자상거래 확산 등의 영향으로 인터넷을 기반으로 하는 물류정보화의 추진이 점차 확대되고 있으나 다음과 같은 문제점 들이 있다. 첫째, 물류 정보화의 부족이다. 일부 물류 대형업체를 제외하고는 인터넷 홈페이지를 단순 홍보수단의 기능으로 활용하는 정도에 그치는 경우가 많아 물류업계 전반적으로 물류정보화의 수준은 저조한 상태이다.

둘째, 물류정보시스템간 연계성 부족이다. 국내 물류정보 시스템은 부문별 물류 정보화가 추진되고는 있으나, 아직 제공 가능한 서비스가 부족하여 관련 정보망간 연계가 미흡하여 이용자의 사용 불편을 초래하고 있다. 또한 물류관련 데이터 베이스간 연계검색체계가 별도로 구축되고 통합 데이터베이스 미비 등의 이유로 물류관련 정보의 일괄 서비스가 거의 불가능하다. 특히, 항만 이용자간 One-STOP 서비스가 되지 않아 불편을 야기하고 있다. 시스템간 연계성 부족은 물류정보의 단절 현상을 발생시키고 이로 인해 다빈도 소규모 배송에 따른 물류비 증가, 재고를 줄일 수 있는 적기 배송 미흡 등의 비효율이 초래되고 있다.

셋째, 물류정보화의 기반요소 빈약이다. 상품의 수배송, 하역, 보관, 재고관리 등에 관한 정보를 처리하는 개별 기업 차원의 물류 정보화는 대기업을 중심으로 지속적으로 확대되고 있다.<sup>35)</sup> 특히, 대형 제조, 유통업체의 경우 유통업무와 관련하여 상품 보관, 재고관리를 위한 바코드, POS시스템, EDI시스템 등의 정보화 요소 확충에 매우 적극적인 것으로 파악되고 있다. 그러나 위치추적정보(GPS)시스템,

35) 윤소형(2007), “물류정보시스템에 있어서 RFID시스템의 도입방안에 관한 연구”, 부경대석사 학위논문, pp.12~13.

무선 통신기술 등의 정보화 기술과 수배송, 하역 등의 물류업무와의 통합은 아직 미흡한 실정이다.<sup>36)</sup>

이처럼 물류정보관리는 항만관리주체 이외에는 전체적으로 모두 구축하기가 어려운 실정이며 만약 각 주체별(선사, 운영사, 포워딩 등) 다른 물류정보가 있어 이를 네트워크로 연결하여야만 가능한 일이다.

항만물류정보관리 사례(VAN EDI)로는 현재 국내 항만물류와 관련된 정보망은 정보부문에서는 해양수산부의 항만물류정보시스템(Port-MIS), 건설교통부의 수출입물류정보시스템, 관세청의 통관망 등이 있다.

이중 항만 신고와 관련된 Port-MIS는 선박입출항관련업무, 수출입화물 반출입에 관한 업무, 항만시설물에 관한 업무 및 의사결정지원시스템 등 크게 4개 업무로 구성되어 있다.<sup>37)</sup>

Port-MIS는 VAN을 통한 EDI를 이용하여 서로간의 정보를 주고 받는다. VAN(Value added Network : 부가가치통신망)이란 전형적인 순수한 통신업자들로부터 통신설비를 빌려서 대규모의 통합적인 정보의 제공과 통신서비스를 제공하기 위해 대형컴퓨터와 결합시켜 구축된 네트워크를 말한다. 즉, 제3자(데이터 통신처리업자 또는 회사)를 매개로 하여 자료를 교환하는 통신망이다.

현재 Port-MIS는 물동량을 이용한 EDI업무를 수행하고 있어 VAN EDI인 물류망 전용선과 인터넷을 이용하여 데이터를 전송하고 있다. 하지만 인터넷을 통한다고 하여도 데이터는 물류망 서버로 전송되어 이는 다시 물류 전용선을 통하여 해양수산청으로 EDI 문서로 전달되기 때문에 결국 사용자들은 폐쇄망과 동일한 수준의 통신료를 부담하게 된다. 이런 과정을 거쳐 해양수산청으로 전달된 문서는 특별한 변환기를 통해서 다른 데이터 형태로 해양수산청 데이터 베이스 서버로 전

36) 진향찬(2004), "RFID 물류정보시스템 구축방안", 명지대학교 석사학위논문, pp.10-11.

37) 최형림, 박남규, 김현수(2000), "항만정보시스템용 EDI 소프트웨어 개발", 한국경영정보학회, 제2권 2호, p.343.

달리며 이를 받은 데이터 베이스 서버는 문서가 도착하는 즉시 이를 해석하여 저장하게 되어있다. 이 모든 전송은 물류망을 통한 전용선을 이용하고 있으며 이는 폐쇄형 EDI 전송 시스템의 형태를 띠고 있다.

### 3. 최신 ICT 기술을 이용한 스마트 물류정보관리 사례<sup>38)</sup>

그렇다면 과연 스마트 물류란 무엇일까, 차세대 물류 IT 기술연구 사업단에서는 물류 IT를 ‘컨테이너, 팔레트 또는 상품상자에 전자 태그를 부착하여 언제 어디서나 물류의 흐름을 실시간으로 파악할 수 있는 기술로서 유비쿼터스 컴퓨팅(Ubiquitous Computing) 기술과 항만물류 기술이 접목된 기술’로 정의하고 있다.

KSA 한국 표준협회에서는 스마트물류를 물류활동의 모든 단계에서 RTI(Returnable Transport Item:회수기능 운송용 아이템)가 가능한 ULD(Unit Load Device : 단위탑재용기)를 기본으로 한 I-ULS(Intelligent Unified Logistics System) 체계로, 스마트물류는 IT와 물류기술을 복합적으로 융합한 형태로써 수송, 포장, 보관/하역, 정보/보안의 물류활동간 Traceability, Environment Friendly, Green Logistics, Interface-Oriented, Human Centered, Asset Management 등 6가지 요소의 특징을 갖는다고 정의하였다.

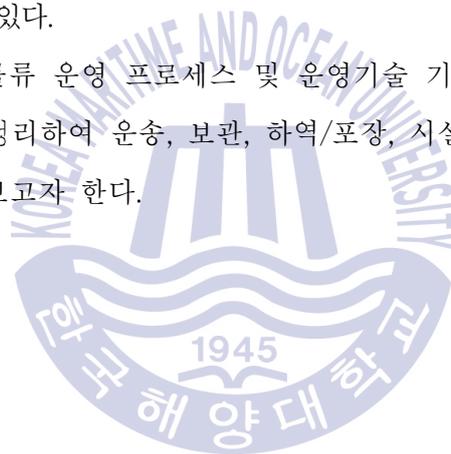
스마트 물류(Smart Logistics)는 운송, 보관, 하역, 포장,시설, 장비 및 물류시스템 등 물류의 전 분야에 걸쳐 IT 기술, 센서, 정보통신 및 제어기술을 접목함으로써 물류운영의 효율화와 물류비용의 절감을 목표로 하고 있는 물류로 정의하고자 한다. 즉, 기술의 결합으로 상호보완적인 통합시스템을 의미하는 것이다. 스마트 물류의 궁극적인 목표는 물류기업 에서는 요구되는 물류비용 절감 및 물류의 효율적 운영을 달성하기 위하여 인프라 중심 물류에서 나아가 첨단 물류 IT 기반의 종합적인 통합 물류 서비스를 제공하는 것이다. 고도화 되고 선진화된 종합 물류

38) 정성용, 정기연(2012), “스마트 물류 IT 기술 및 사례”, 전자공학회지, 제39권 제5호.

서비스 제공을 위하여 실시간 대상제품의 가시성을 확보하고 이와 관련하여 수집된 빅 데이터(Big Data)를 기반으로 분석된 정보를 택배, 제3자 물류, 국제물류 서비스를 영역에 제공하는 최적화된 종합 물류 서비스를 목표로 하는 것이다.

최적화된 종합 물류 서비스를 제공하기 위하여 각각의 운영 프로세스 상에 물류 IT가 활용되고 있으며 이는 기존의 공급망 분야에서 활용되어 온 ERP(Enterprise Resource Planning)에서부터 WMS(Warehouse Management Systems), TMS(Transportation Management systems), RFID, 센서 및 RTLS(Real Time Location System)에 이르는 무선통신 인프라까지 물류의 기능을 첨단화하고 지능화기 위한 기술을 포함하고 있다. Supply Chain 분야에서 IT 기술은 경쟁력을 향상시키기 위한 주요 역할을 담당해 오고 있다.

<표 2-8>과 같이, 물류 운영 프로세스 및 운영기술 기반으로 물류산업에서 활용되는 물류기술을 재정리하여 운송, 보관, 하역/포장, 시설/장비, S/W로 구분하여 기술개발에 대해 알아보고자 한다.



<표 2-8> 물류 프로세스별 주요 기술

구분	주요기술
운송	<ul style="list-style-type: none"> <li>· 저탄소, 친환경 운송 수단 개발</li> <li>· 운송 수단간 연계 강화</li> <li>· 첨단 무인 자동화 운송 시스템 개발</li> <li>· 글로벌 물류정보망 실시간 통합 운영기술 개발</li> <li>· 운송차량 이력관리(위치추적 및 상태 모니터링)</li> <li>· 차량 및 교통 통합관리 시스템(ITS 기술 연계)</li> </ul>
보관	<ul style="list-style-type: none"> <li>· 물류센터의 가능 자동화, 지능화, 안전성 강화 시스템 개발</li> <li>· 재고/자산관리 시스템 개발</li> <li>· 물품 입출고 관리 시스템 개발</li> <li>· 자동 운송/지능형 터미널</li> </ul>
하역/포장	<ul style="list-style-type: none"> <li>· 첨단 컨테이너 개발</li> <li>· 고효율/첨단 하역장비(크레인, 이송차량 등) 개발</li> <li>· 웹/인터넷 기반 하역/포장 정보 시스템</li> <li>· 물류 하역/보관 시스템 간 상호 정보 교환 서비스 제공</li> <li>· 친환경 포장 용기 개발</li> </ul>
시설/장비	<ul style="list-style-type: none"> <li>· 첨단 친환경 장비 및 부품 개발</li> <li>· 통합적 인터모달 시스템 개발</li> <li>· 원격관리/보수, 모니터링 시스템 개발</li> <li>· 시설 장비간 상호 운용 제어 정보 시스템 개발</li> <li>· 무선 데이터 취합, 가공 시스템 개발</li> <li>· 화물 운송 인프라(H/W, S/W) 구축</li> <li>· 에너지 절감형 운송 시설 및 장비 개발</li> </ul>
S/W	<ul style="list-style-type: none"> <li>· 글로벌 물류정보망 실시간 통합 운영 기술 개발</li> <li>· 첨단 물류 운영 시스템 개발</li> <li>· 보안기술 개발</li> <li>· 물류용 단말기 운영 기술 개발</li> <li>· 수송경로 계획/관리 시스템</li> </ul>

출처: 정성용, 정기연(2012), “스마트물류 IT기술사례”, 「전자공학회지」, 제39권 제5호.

## 1) 주요사례

### (1) 네덜란드 ECT(Europe Container Terminal)

ECT는 네덜란드의 로테르담 항만에 개발된 세계 최초의 자동화 컨테이너 터미널로 장치장 배치방법을 수평에서 수직2)으로 변경 설계한 신개념 터미널이다. 현

재는 수직배치의 컨테이너터미널이 다수 운영되고 있으나 당시는 세계에서 최초로 배치형태를 바꾼 획기적인 물류시스템이었다. 동 터미널은 컨테이너 운반차량(AGV: automated guided vehicle)과 야드하역장비(ASC: automated stacking crane)를 최초로 자동화한 무인 하역시스템을 개발하였다. 현재 세계에서 가장 큰 4곳의 자동화 컨테이너터미널을 운영하고 있으며 전체면적은 265ha, 안벽길이 3.6km, 수심은 DL(-) 16.65M에 이르는 초대형 항만이다.

ECT중 DMU를 제외한 3곳, Euromax는 완전무인자동화 운영되고 있다. 즉 안벽장비와 이송장비의 연계작업, 컨테이너 이송, 이송장비와 야드장비의 연계작업, 야드 적재, 인출작업 등이 모두 무인으로 수행되고 있다. 그러나 자동화 장비와 유인장비가 연계 작업이 이루어지는 곳은 원격조정으로 작업을 수행한다. 예를 들어 외부에서 운전자가 탑승한 차량에 대해 작업할 경우 중앙통제센터에서 원격조정으로 컨테이너를 싣고 내린다. 또한 안전을 위해 해측(자동화 구간) 과 육측(원격구간)의 분리운영을 수행한다.

## (2) 독일 CTA(Container Terminal Altenwerder)

CTA는 1997년 항만생산성 증가, 서비스 향상, 물류비 절감을 위한 목적으로 자동화터미널 개발에 착수(운영사 HHLA)하였다. 5년의 개발과 시험운영 끝에 2002년 세계에서 2번째의 완전 무인자동화 컨테이너터미널 운영을 시작했다. 전체면적은 1.0km<sup>2</sup>, 안벽길이 1.4km(4berth), 수심은 DL(-) 16.7M인 초대형 터미널이다.

네덜란드 ECT를 벤치마킹하여 터미널을 개발하였으나 ECT와는 다소 다른 하역시스템을 운영하고 있다. 기본적인 운영은 완전무인자동으로 운영되고 있으나 안벽장비, 야드장비의 컨테이너 이동, 적재방식이 ECT와 다른 DHST(Dual Hoist Second Trolley) C/C, 5단10열의 ARMGC(블록당 2대 운용)를 운영하고 있다.

CTA 또한 완전무인자동화로 운영되고 있다. 안벽장비와 이송장비의 연계작업,

컨테이너 이송, 이송장비와 야드장비의 연계작업, 야드 적재, 인출작업 등이 모두 무인으로 수행되고 있다. 더불어 자동화 장비가 작업하는 구간인 해측과 유인트럭이 진입하는 육측구간을 엄격히 분리하여 운영한다.

<표 2-9> CTA 장비 운영 현황

구분	세부내용	비고
선박작업(해측)	선박-C/C	유인화
선박작업(육측)	C/C-AGV	무인자동화
이송작업	AGV	무인자동화
야드작업(본선)	AGV-ARMGC	무인자동화
야드작업(반출입)	ARMGC-RT	원격조정

### (3) 중국 ZPMC의 친환경 자동화터미널

중국 ZPMC(Shanghai Zhenhua Heavy Industry Co.,Ltd.)는 세계 최대의 항만하역장비 생산업체로서 2007년 세계에서 최초로 탠덤리프트(Tandem Lift)3 방식의 완전무인 자동화 컨테이너터미널을 개발하였다. 이는 세계 최대의 하역장비 중공업 업체인 ZPMC에서 자체비용으로 개발하여 현재 시험운영중에 있다. 개발의 목적은 전기를 이용하는 ①친환경적 터미널, ②고생산성의 컨테이너터미널을 개발 ③ 미래의 하역장비 신규시장 창출 ④ 세계 항만표준모델 선도를 목적으로 하고 있다.

터미널 운영은 수직배치 형태의 야드 시스템을 가지고 있으며 안벽과 야드 사이에 컨테이너 이동을 기존의 야드 트럭이나 AGV가 아닌 별도의 레일시스템을 이용하도록 개발한 것이 특징이다. 안벽장비, 야드 장비, 이송장비 모두 Tandem 컨테이너 취급 운영한다. 모든 장비는 전기식으로 운영되며 수평/수직 플랫폼롤리와 플랫폼카(Flat Car)를 사용한다. 또한 동 시스템은 100% 전기식으로 운영되어 친환경

경 터미널이기도하다.

현재 개발후 ZPMC 부지내 1선석 규모의 실물을 제작하여 테스트 진행 중에 있으며 1단계로 중국 와이가차오 컨테이너 터미널 배후지에 공컨테이너 취급을 위한 실규모 장비를 건설, 운영중에 있다. 2단계로 2011년까지 중국 카오웨이디안 터미널(허베이 성)에 2선석 규모를 건설할 예정에 있다.

동 시스템은 안벽과 야드 구간의 컨테이너 이동을 레일시스템을 이용하기 때문에 사고나 비상상황시 다른 시스템에 비해 유연성이 떨어지며 초기 투자비용이 높은 것이 단점이다.

<표 2-10> ZPMC 장비 운영 현황

구분	세부내용	비고
선박작업(해측)	선박-C/C	유인화
선박작업(육측)	C/C-수평FT(Flat Trolley)	무인자동화
이송작업	LT(Lifting Trolley)-수직FC(Flat Trolley)	무인자동화
야드작업(본선)	수직FT-RMGC	무인자동화
야드작업(리마살링)	RMGC	무인자동화

#### (4) 미국 스피드포트(Speed Port)

스피드포트는 미국의 ACTA Maritime Development Corporation이 제안한 신개념 컨테이너 터미널로, 대량의 안벽크레인과 이송장비를 투입하여 선박재항시간을 획기적으로 줄일 수 있도록 설계된 시스템이다.

터미널은 도크형태를 가지고 있으며 선박의 폭방향에 걸쳐 상부빔이 설치되어 있다. 도크 안에 설치되어 있는 상부빔(overhead beam)에는 독립 이송장치인 스파이더(spider)가 있으며 이를 이용해 컨테이너 양적하 작업과 이동작업을 수행하는

형태를 가지고 있다.

따라서 동시스템은 빠른 하역속도, 높은 안전성, 보안성, 저렴한 비용 등이 장점으로 평가되고 있으나 현재 개념개발이 완료된 상황이다. 운영개념은 모든 시스템이 완전무인자동으로 운영되며 기존 터미널과 달리 하역시스템 측면에서 안벽과 야드 영역의 구분 없이 통합된 운영을 기본으로 설계된 개념이다.

현재 운영되는 터미널은 없으며, 개발 시 고가의 토목 건설비 소요, 유연성 저하, 야드 장치 및 이송부문에서의 병목현상 발생 등의 단점이 있으나, 기존보다 토지효율성 면에서는 우수한 것으로 보인다.



<그림 2-9> 스피드 포트

#### (5) 고층 자동야드적재시스템

고층 자동야드적재시스템은 기존의 자동화 물류창고에서 적용하던 개념으로 이를 항만에서 응용하여 기존 1~6단적으로 사용하여 항만의 야드부지를 20층 이상으로 적재함으로서 획기적인 부지절감을 이룰 수 있는 시스템이다. 동 시스템은 AS/RS(automated storage and retrieval system)와 유사한 구조형태로 컨테이너를 자동으로 적재·불출할 수 있고, 좁은 부지에 구축이 가능하며, 적재 층수를 늘려

용량을 증가시킬 수 있다.

고층 적재시스템은 컨테이너 위에 컨테이너를 적재하는 방식이 아닌 랙 내부의 셀별 컨테이너 적재장소가 있는 방식이다. 따라서 컨테이너 재취급 작업 없이 적재 구조 내에서 무작위로 컨테이너에 접근이 가능해(무작위 접근[random access]) 생산성이 높다. AS/RS는 이처럼 생산성이 높고 적재 능력이 탁월해 부지가 매우 제한적이거나 이용료가 비싼 곳에서 각광 받고 있다.

동 시스템의 운영은 모두 무인자동으로 운영되고 있으며 기본적으로 플랫폼, 리프터, 셔틀 등의 장비가 필요한 개념이다. 기존 자동물류창고의 AS/RS와 차별되는 기술은 40톤 이상의 컨테이너를 이동, 승하강 시킬 수 있는 장비기술 개발이 매우 어렵다는 것이 단점이다. 국내에서는 유일하게 (주)이지인더스에서 동기술에 대한 특허를 보유 중에 있다. 동 시스템은 여러 종류가 개발 중에 있으며 HSS(High Stack System), Computainer System, Krupp Fast handling system 등이 있어 연구개발되어 있다.

#### (6) 다목적 물류정보 시스템(MPS)

RFID 기술을 활용한 다목적 물류정보 시스템(MPS, Multi Process System)은 입고, 출고, 재고조사 등 물류 운영 작업에 해당 상품과 수량을 자동으로 작업자에게 알려주어 업무의 정확성과 효율성을 향상시키도록 지원하는 시스템으로, MPS의 단말기(MPI, Multi Purpose Indicator)를 창고 내에 랙(Rack)에 부착하여 시스템을 통하여 작업 지시를 내리면 MPI는 자동으로 상품과 수량을 표시한다.

기존에 물류센터에서 사용한 DPS(Digital Picking System)이나 DAS(Digital Assort System)은 유선시스템으로 선반에 고정된 형태로만 운영이 가능하였으나, MPS는 유선케이블 방식을 사용하지 않고 데이터 송수신이나 전원공급을 100% 무선 환경(Optional)에서 할 수 있고, 자석을 이용한 탈부착 방식을 되어 설치나

유지보수가 훨씬 용이하여 설치 및 운영 환경에 제약을 받지 않는다.

MPS의 초기 도입 비용도 45% 가량 낮아 물류업체들의 비용 부담도 크게 감소할 전망이다. 또한 물류 현장에서 일일이 종이를 들고 다니며 손으로 체크하여 작업을 할 필요가 없어 기존 대비 평균 40%의 작업시간을 단축시킬 수 있으며, 작업 현장의 종이를 없앴으로써 500대의 MPI 단말기가 도입된 물류센터의 경우 연간 4만장의 종이를 절약할 수 있어 녹색물류 실현에도 뒷받침할 수 있을 것으로 보인다.



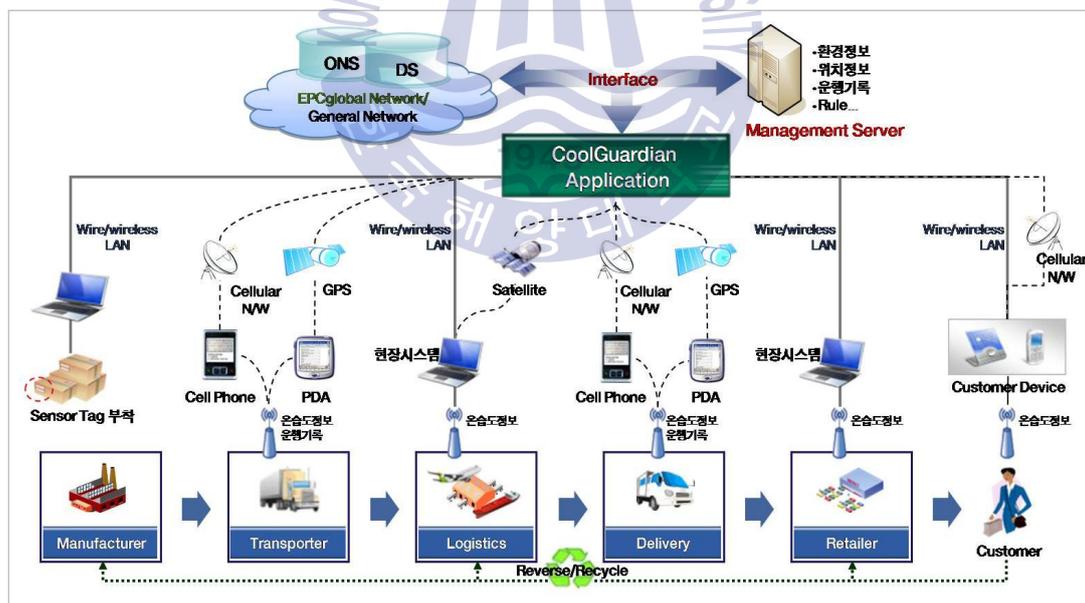
<그림 2-10> 다목적 물류정보시스템

(7) 온습도 관제시스템(CoolGuardian)

CoolGuardian은 차량 적재함 온습도 관제, ECO Driving 관리를 위한 운행기록 관제, Rule에 의한 실시간 온습도 제어, 차량 적재함 보안관리 기능을 수행하는 통합장비로 물류센터 내에서는 Cell단위 온습도 관리용으로 활용 가능하다.

CoolGuardian은 유선기반 센서를 무선으로 전환하여 Location별 온습도를 관리하는 온습도 모니터링, ODB를 활용한 무선 운행 기록 관제 및 ECO 기능을 수행하는 운행기록관제(Taco), Rule 기반의 통합 온습도 제어 시스템으로 실시간 온습도 제어 및 도착지 이외의 지점에서 차량 Lock 해제시 이와 관련된 Event를 처리하는 차량 보안 관제(PEA) 시스템으로 구성되어 있다. SCM 프로세스 상에서 온습도에 민감한 정온관리 식품 및 어패럴을 대상으로 활용되고 있다.

CoolGuardian의 적용을 통하여 기존의 타코미터 대비 운영비용의 절감을 나타냈으며, 저온물류센터의 Location별 온도 관리가 가능하여, 온도 민감도를 기준으로 Location별 적재 상품을 보관하는 것이 가능해 졌다. 이를 통하여 상품의 품질을 보장하여 보다 안정적인 서비스를 제공할 수 있게 되었다. 또한 상온 물류의 경우에도 실시간 온습도관리가 가능하여 화장품이나 어패럴을 대상으로 체계적인 온도관리를 해서, 온도관리의 체계화를 통한 통합관리가 가능하도록 하였다.



<그림 2-11> CoolGuardian 시스템 구성

#### (8) 고품질 택배 시스템(HPDS)

택배서비스는 다품종 소량의 물량을 운송하여 공급자로부터 소비자에게 제품을 전달하는 서비스를 제공한다. 최근에는 귀중품이나 부패·파손되기 쉬운 품목의 취급, 빠른 배송, 심야 또는 지정시간배달 등 소비자들의 수요 및 니즈가 다양해지고 있는 추세이며, 택배서비스 이용이 증가에 따른 택배화물 분실·훼손·지연 등 택배 민원이 급증하고 있어 서비스수준 향상을 위한 새로운 대책이 필요한 실정이다.

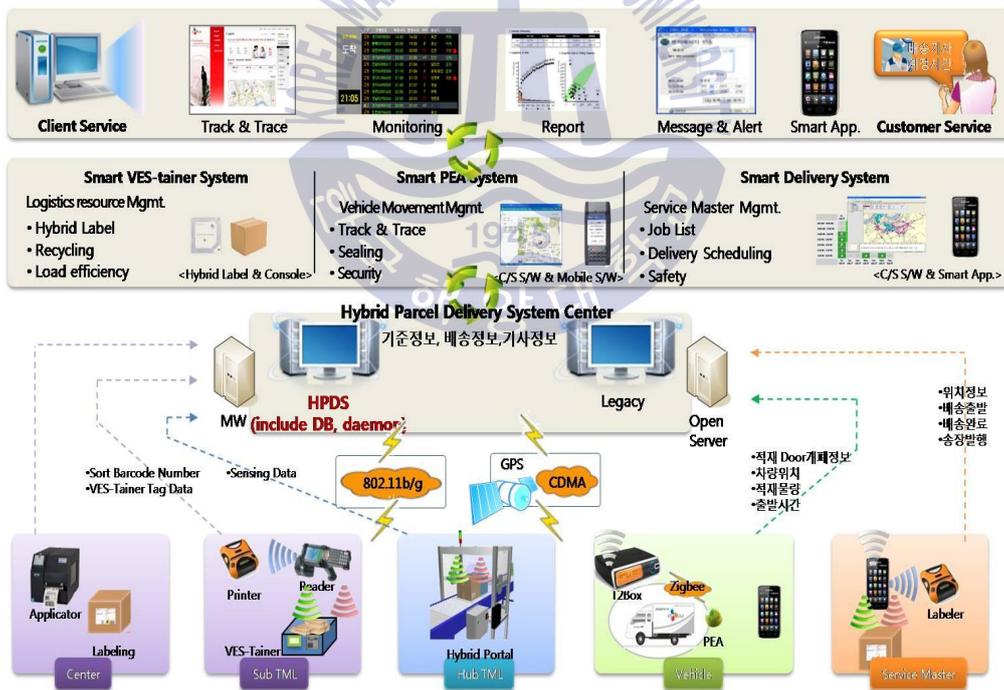
고품질 택배 시스템(HPDS, Hybrid Parcel Delivery System)은 택배 송장에 Passive RFID Tag를 삽입하여 출고지부터 최종 도착지까지 상품관리를 Barcorde와 RFID Tag를 병행하여 사용할 수 있도록 개발되었다. 소비자 권익보호를 위해 위탁한 화물을 안전하게 인도하지 못하여 화물이 멸실된 경우 정확한 책임소재를 판단하기 위한 이력관리 정보와 RFID/USN 기술 적용을 통해 대상 화물의 추적에 대한 정보를 제공한다.

국내 택배산업 전반에 RFID 적용을 목표로 핵심 시스템을 개발하여 물류전반의 Visibility를 제공하고, Smart Client Service 및 Smart Customer Service) 현장 적용을 추진하고자 현재 사업진행 중으로 시스템을 개발하고 있다. 국내 홈쇼핑 등에서 취급하는 고가품에 단품박스 단위로 RFID/USN Tag를 부착하여 1다음 년도(2011)에는 서울권 물량을 대상으로 적용하였고, 현재 2다음 년도(2012)에는 수도권 물량을 대상으로 적용 범위를 확대 적용을 목표로 하고 있다. 대상 품목은 분실률이 비교적 높은 고가품과 화장품 등 진품확인이 필요한 제품에 초점을 맞추어 적용 하였으며, 이를 통해 공급자 측면에서는 택배 배송업무 효율성 향상, 택배비용 절감, 업무 처리시간 단축의 효과를, 수요자(최종 소비자) 측면에서는 진품 정보 및 이력정보 등의 제공을 통한 고객만족도를 향상시키고자 한다.

최근 물류/유통 산업은 운영 시스템 분야를 중심으로 급속한 IT 기술 적용을 운영 개선이 이루어지고 있다. 택배 서비스 분야에도 RFID/USN 등 u-IT 신기술을

활용한 업무프로세스 효율화를 통해 고객에게 신뢰성 및 안전성 측면의 향상된 서비스를 제공하고자 한다. 이는 궁극적으로 택배산업을 고부가가치 산업으로 발전시키는데 결과를 보일 수 있을 것으로 기대된다.

첨단 물류 IT를 활용하여 Smart SCM을 성공적으로 구축하여 적용하기 위해서는 단순히 가장 최신의 Auto-ID 시스템이나 센서 시스템을 적용하는 것이 아니라, 적용 가능한 기술의 구성 요소 및 특징을 살펴보고, 적용시 예측 가능한 타당성 및 경제성 검증을 통하여 가장 적합한 기술이 적용될 수 있도록 해야 한다. 부분적 효용성 및 ROI 가 아닌 전체 시스템 관점에서 시스템 구성 요소간의 상호용성과 현장 환조건을 충분히 검토 한 후, 기술 적용을 위한 표준화 연구, 운영 전략 수립, 비즈니스 모델 개발 및 관련 법·제도 정비 등이 복합적으로 추진되어야 할 것이다.



<그림 2-12> 고품질 택배시스템 구성도

## 제3장 빅데이터를 이용한 부산항신항 ITT 이동패턴 분석

### 제1절 타부두 TS 현황분석

#### 1. 부산항 환적화물 현황

부산항은 동북아 환적중심항만으로 물동량이 점차 증대하고 있으나 중국 주요항만의 급성장으로 '16년 물동량은 세계 6위항만으로 추락하였다. 세계 중심항만은 환적화물 유치를 위해 국가와 선사가 공동으로 마케팅활동을 펼치고 있으며, 환적중심항만이 되고자 인센티브 정책을 등 글로벌 항만간의 경쟁이 점차 심화되고 있다.

부산항 전체 컨테이너 화물은 연평균 5% 정도 증가되고 있으나 수출입물동량은 2%내외의 증가세로 주춤한 반면 환적화물은 연평균 약 9%을 증가세로 부산항 물동량 증가의 견인차 역할을 하고 있다.

<표 3-1> 부산항 물동량 처리현황

(단위 : 천TEU)

구분	2011	2012	2013	2014	2015	2016
총계	16,184	17,046	17,686	18,683	19,468	19,432
수출입	8,708	8,808	8,933	9,253	9,363	9,608
환적	7,352	8,147	8,748	9,429	10,105	9,823
연안	124	90	4	-	-	-

출처 : 부산항만공사 PORT-MIS Data 기준.

## 2. 환적화물의 변화와 전망

최근 환적화물의 비중이 전체 물동량에 약 51%로 절반 이상의 수준을 차지하고 있어 부산항 물동량 증가에 가장 중요한 역할을 차지하고 있다. 특히 한진해운 사태 이후로 중국 화물은 부산항 환적물량의 최대 실적국으로 대부분이 북중국 환적화물 증가분에 기인하고 있으며 동시에 이들 항만에 의존도가 높다. 부산항의 환적화물 중 중국의 비중은 2009년부터 꾸준히 증가하였으나 이후에는 글로벌 경기 침체와 중국의 성장 둔화로 부산항의 환적화물의 증가세 둔화나 감소를 줄 것으로 예상된다.

일본 서안 지역은 환동해 경제권으로 부상하고 있으며 컨테이너 화물량이 증가하고 있다. 특히 중국, 러시아 경제 활성화로 인해 일본의 교역 확대에서 일본 서안 항만들은 부산항을 통해 중국의 닝보, 상하이, 청도와 러시아의 블라디보스토크, 보스치니로 환적화물이 이동 중인 것이고 일본 서안 지역은 부산항과 지역적으로 인접하고 무역, 물류 분야에서 강력한 경쟁 및 협력관계를 유지하고 있어, 일본 서안 지역의 항만 집중 육성 전략이 활성화될 경우 부산항의 환적물량에 직접적으로 영향을 줄 가능성이 있다고 부산항만공사는 전하고 있다.

## 3. 타부두 TS 현황

부산항 전체 물동량 중 환적화물이 차지하는 비중이 매년 증가 추세를 보임에 따라 타부두 TS 비중도 점차 늘어나고 있다. 특히, 글로벌 선사의 얼라이언스 재편으로 인한 환적물동량 증가와 신항 터미널이 쪼개져 있어 타부두 TS 증가는 당분간 불가피한 전망이다.

<표 3-2>와 같이 '16년 부산항 환적화물 물동량 9,608천TEU 중 자부두 TS 전년 대비 증가세 보다 타부두 TS 증가세가 무려 4.7배의 증가세를 보이고 있고, 자부두 TS 환적비중은 약 5.3% 증가세를 보이고 있어 타부두 TS가 증가추세를 짐작할 수 있다.

<표 3-2> 신항 ITT(타부두T/S) 발생 현황

(단위 : 천TEU)

구분	2014년	2015년	2016년	증감율 (15년/16년)
환적화물 계(A)	6,814	7,570	9,608	26.9%
자부두 환적	6,041	6,769	8,073	19.2%
타부두 환적(B)	773	801	1,535	91.6%
타부두 환적 비중(B/A)	11.3%	10.6%	15.9%	

출처: 부산항만고사 Port-MIS Data 기준

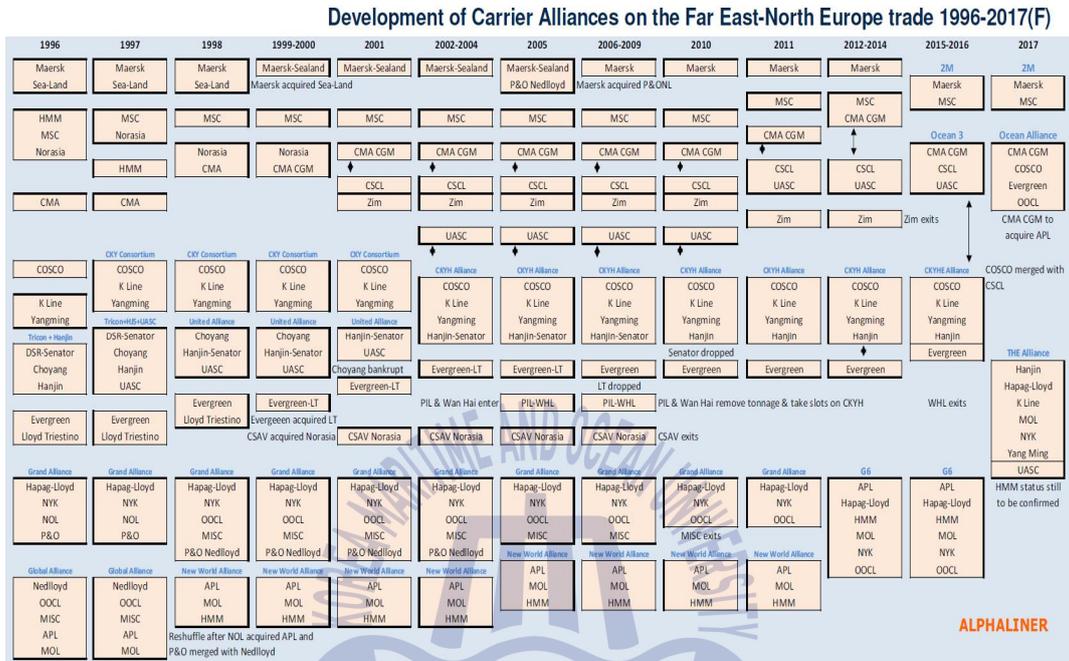
타부두 TS 물동량 중 신항 터미널의 출발지와 도착지를 기준으로 물동량 분석을 하면 <표 3-3>과 같다. 이 중 신항 1부두(PNIT) 부두가 타부두 TS물동량이 현저히 높은 것으로 확인할 수 있고, 신항 3부두(HJNC)가 가장 낮은 것으로 확인되었다. 이는 2016 얼라이언스 재편(G6, CKYHE)에 따른 하역계약부두 분산으로 타부두 TS 물동량이 증가한 것으로 판단된다.

<표 3-3> 2016년 터미널별 타부두 TS 물동량(From/To)

(단위 : 천TEU)

(From \ To)	PNIT (신항1부두)	PNC (신항2부두)	HJNC (신항3부두)	HPNT (신항4부두)	BNCT (신항5부두)	합계
PNIT (신항1부두)	0	151,383	45,269	189,576	23,423	409,651
PNC (신항2부두)	254,786	0	35,254	114,042	47,473	254,786
HJNC (신항3부두)	75,574	37,163	0	50,538	45,883	75,574
HPNT (신항4부두)	159,348	74,126	57,509	0	29,718	320,701
BNCT (신항5부두)	23,574	57,196	37,739	17,295	0	135,804
합계	513,282	319,868	175,771	371,451	146,497	1,526,869
평균	102,656	63,973	35,154	74,290	29,299	453,897

<그림 3-1>에서 보면 얼라이언스가 1개의 단일부두로 하역계약이 된 것이 아니라 2개 부두로 쪼개져서 얼라이언스 물동량의 이동이 많은 것으로 파악되고 있다.



<그림 3-1> 연도별 얼라이언스 재편 현황

출처 : 알파라이너(2017.1월).

## 제2절 부산항신항 ITT 이동패턴 분석

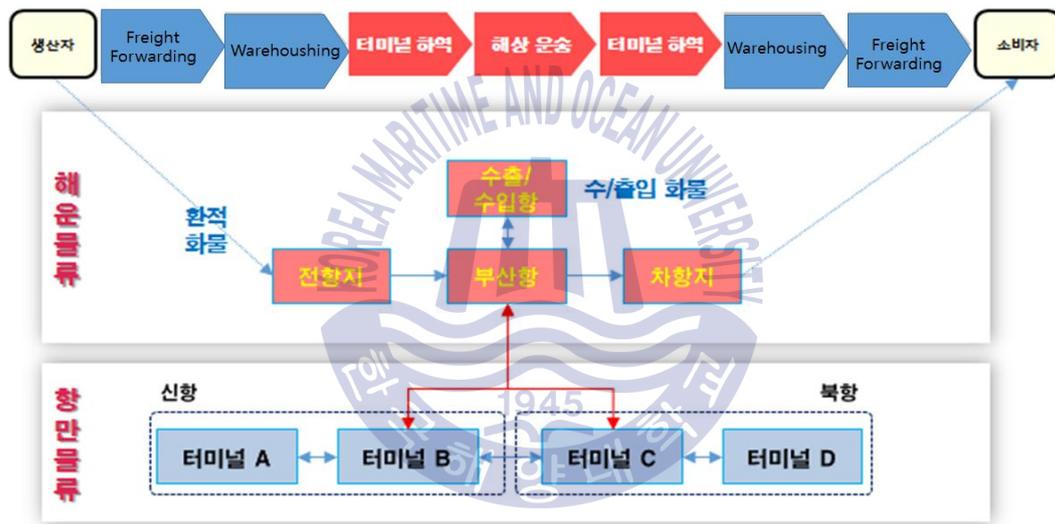
### 1. 분석 방법 및 타부두 ITT 패턴분석

신항 터미널 내 혼잡도 분석은 우선적으로 신항 터미널을 기준으로 '16. 1. 1. ~ '16. 1. 31까지 하역장비(Container Crane, Transfer Crane)간 이동 및 양적하 시간, 게이트 반출입 시간 등을 감안한 실제 데이터를 기준으로 분석하였고 분석 틀은

Qlik Sense라는 제품을 이용하였다.

Qlik Sense는 미국 제품이며 전 세계 가장 많이 사용되는 분석가용 툴로서 셀프 서비스 및 시각화 및 유도분석에 대한 직관적인 솔루션 제공하는 제품이다.

구체적인 분석의 방법으로는 선박의 양적하 기준인 수입화물(Inbound)와 수출화물(Outbound)화물을 터미널내 CY 반입 게이트 반출입 시간 및 종류별로 선박에서 터미널내 양하시간, CY에 반입된 시간, 컨종류별 턴어라운드 타임 등을 정리하여 분석시스템으로 나타내었다. 아래 [그림 3-2] 은 부산항을 기점으로 선박 입항과 출항에 따른 터미널간 ITT 등 종합적인 개념을 나타낸 것이다.

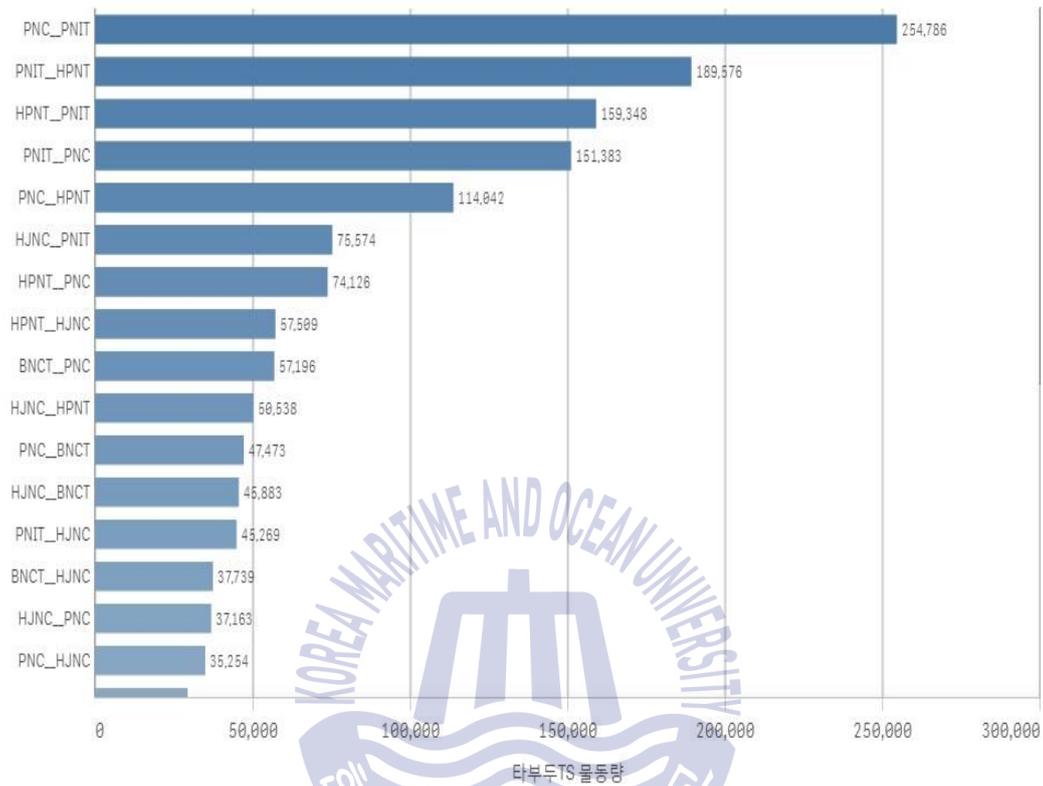


<그림 3-2> 선박입출항에 따른 타부두 TS 및 ITT 개념도

타부두 ITT 패턴분석의 방법은 Port-MIS와 CODECO(컨테이너반출입결과 정보)의 정보를 Raw데이터로 하여 2016년을 기준으로 빅데이터 툴을 사용하여 분석하였다.

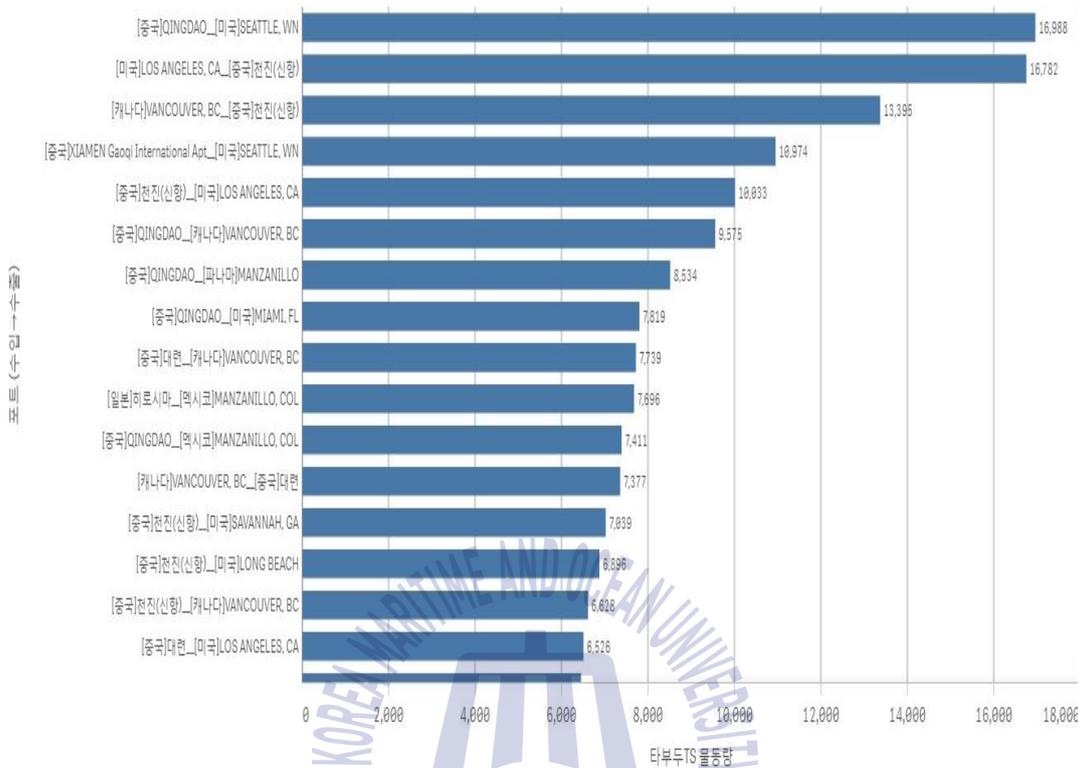
<그림 3-3>에서 보면 '16년 기준으로 부산항 신항 터미널간 가장 많이 발생한 터미널은 PNC에서 PNIT로 가는 타부두 TS 물동량이고 다음은 PNIT에서 HPNT

로 가는 화물, HPNT에서 PNIT로 가는 화물의 순이다.



<그림 3-3> 터미널별 신항 ITT 발생도(From/To)

부산항을 기점으로 타부두 TS가 가장 많이 발생하는 전항지그 뒤로 타부두 TS의 가장 많은 물동량을 차지하는 전항지와 차항지 항만은 <그림 3-4>와 같다. 이 중에서 타부두 TS가 가장 많이 발생하는 항만은 중국(칭다오)에서 미국(시애틀)로 가는 물동량이며, 미국(LA항)에서 중국(천진항)으로 가는 물동량을 나타냈다.



<그림 3-4> 전항지/차항지별 신항 ITT 발생도(From/To)

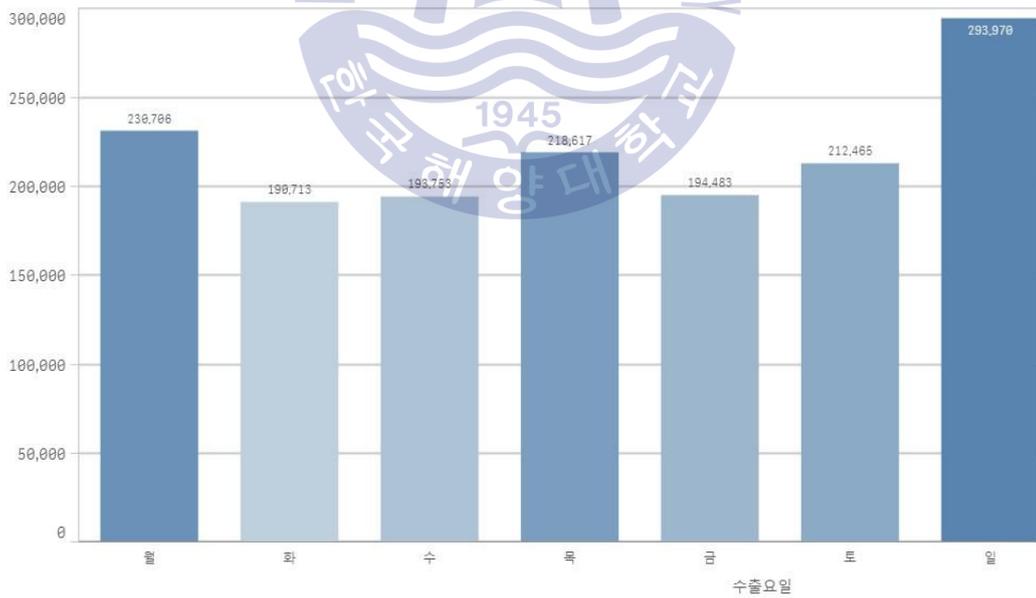
부산항에서 타부두 TS가 가장 많이 물리는 요일은 수입 환적기준으로 화요일과 토요일, 수출 환적기준으로 일요일과 월요일을 들 수 있으면 타부두 TS 총물동량 기준(수입환적+수출환적)으로는 가장 혼잡한 요일은 일요일과 월요일을 들 수 있겠다.

수입요일별 타부두 TS 물동량



<그림 3-5> 수입환적 요일별 처리 물동량

수출요일별 타부두 TS 물동량

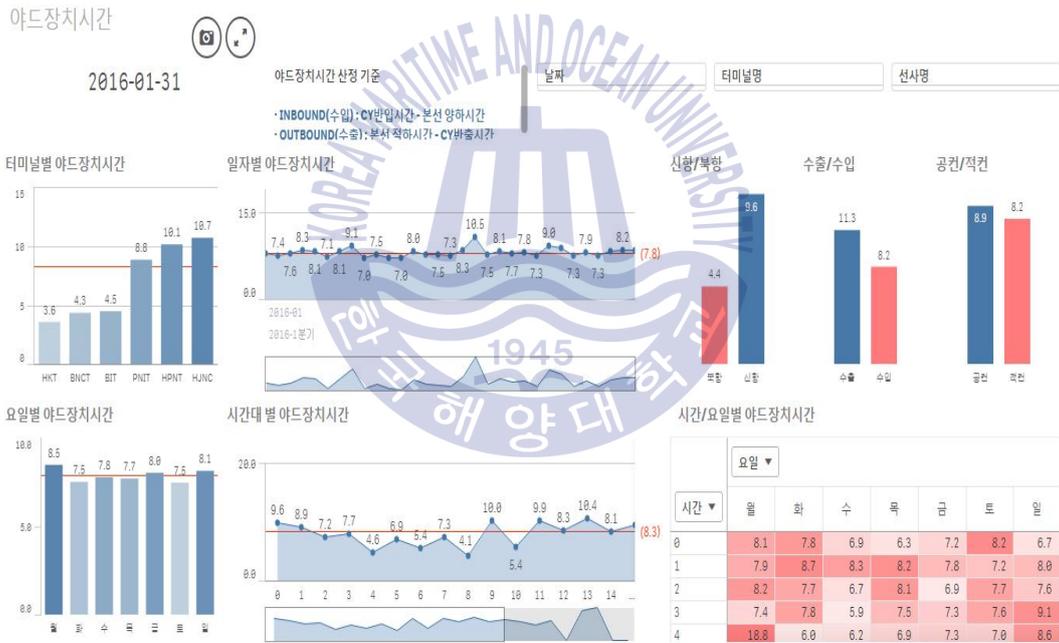


<그림 3-6> 수출환적 요일별 처리 물동량

## 2. 터미널내 혼잡도 분석

### 1) 야드장치 시간 분석

신항 터미널내 혼잡도는 '16년 1월간 실제 데이터를 기준으로 빅데이터로 분석하였는데 주로 야드내 운영현황 즉 야드 장치시간을 기준으로 분석하며 세부적으로 날짜, 시간, 요일, 컨테이너의 종류별로 장치시간으로 산정하였다. 수입화물(Inbound)와 수출화물(Outbound)화물을 기준으로 산정하였는데 수입화물(Inbound) 경우 CY 반입시간-본선 양하시간을 빼는 방식과 수출화물(Outbound)은 본선 적 시간-CY 반출시간을 빼는 방식으로 산정하였다.



<그림 3-7> 신항 터미널별 야드장치시간 빅데이터 분석 화면

신항 터미널별 CY내 평균 장치시간은 8.3시간을 나타나고 있으며 HJNC, HPNT, PNIT가 평균 장치시간을 웃도는 것으로 나타났다. 요일별로는 평균 장치

기간이 7.8시간이며 월요일, 금요일, 일요일이 평균 장치기간보다 높은 것으로 나타났다. 특히 시간과 요일을 합산하면 <그림 3-8>과 같이 붉은색인 진한 정도에 따라 CY 장치율이 높고, 장치기간이 장기간인 것을 확인할 수 있다 이는 향후, 터미널의 생산성과 컨테이너의 회전율과 연계될 수 있을 것이다.

시간/요일별 야드장치시간

시간	요일						
	월	화	수	목	금	토	일
0	8.1	7.8	6.9	6.3	7.2	8.2	6.7
1	7.9	8.7	8.3	8.2	7.8	7.2	8.0
2	8.2	7.7	6.7	8.1	6.9	7.7	7.6
3	7.4	7.8	5.9	7.5	7.3	7.6	9.1
4	18.8	6.0	6.2	6.9	7.3	7.0	8.6
5	7.6	6.9	6.9	7.4	7.9	8.4	6.7
6	14.4	6.9	7.6	7.5	9.5	8.4	6.9
7	9.8	7.6	8.9	7.9	8.0	7.5	9.3
8	8.9	8.2	7.2	8.8	7.6	8.0	7.5
9	7.6	7.4	6.9	7.6	7.9	7.6	8.0
10	7.1	7.6	7.5	7.3	6.9	7.5	7.0
11	10.1	7.2	8.7	7.9	7.2	8.2	7.6
12	7.6	7.1	8.8	7.1	7.6	6.6	8.6
13	8.1	6.9	7.7	8.2	7.9	7.9	9.3
14	7.9	6.3	7.4	7.2	8.6	7.1	8.4
15	7.1	8.2	7.4	7.5	8.2	7.1	8.5
16	7.3	6.7	7.1	8.2	7.7	6.7	8.3
17	8.2	8.9	7.4	10.4	11.7	7.5	9.2
18	9.5	11.1	9.9	8.1	9.0	8.0	9.0
19	7.4	7.8	8.5	9.0	9.2	7.3	9.6

<그림 3-8> 신항 터미널별 야드장치시간(요일/시간별) 혼잡도 분석 화면

공컨테이너 v03



<그림 3-9> 터미널별 공'컨' 장치기간 및 적재율 분석 화면

특히, 공컨테이너의 경우 장치기간을 좀더 세분화 하여 분석하면 공컨테이너의 회전율과 각 터미널별 계약된 선사간의 공컨테이너 장치현황을 한눈에 알 수 있어 터미널 생산성과 효율성에 효과가 있을 것으로 기대된다.

부두내 공컨테이너는 ON-DOCK<sup>39)</sup>계약시 Free-Time(자유장치)로 하는 터미널이 대다수라 향후 공컨테이너 수급관리 및 정책수립에 유용한 자료로 활용될 것으로 예상된다.

39) On-dock CY의 약어로 컨테이너부두내 선사소유 컨테이너를 장치할 수 있도록 하는 계약형태

## 2) 터미널별 턴어라운드(Turn-around) 타임 분석

터미널내 턴어라운드 타임에 대한 개념은 보통 게이트를 통과해서 컨테이너 부두 야적장내 컨테이너를 상하차 후 게이트를 통과하는 시간을 말하는데 이 개념은 화물의 종류 즉 수입화물, 수출화물이 다르게 적용된다. 이로 인해 각 터미널운영사는 턴어라운드 타임에 대한 기준 설정이 모호하여 터미널 이용 고객에게 혼란을 가중시키고 있다. 예를 들어 A터미널의 경우 터미널 기준 수입화물<sup>40)</sup>(Inbound)과 수출화물(Outbound)<sup>41)</sup>의 경우 모든 턴어라운드 타임을 게이트 기준으로 설정하여 화물기사의 터미널내 휴식시간 등을 포함하여 명확한 턴어라운드 타임을 설정하기 매우 어렵다.

빅데이터 분석에서 터미널의 턴어라운드 타임은 화물의 종류(수입화물, 수출화물) 마다 다르게 설정하여 실제적인 턴어라운드 타임을 설정하고자 한다.

수입화물의 경우(Inbound) 선박에서 부두 하역 후 CY(Container Yard)에 이송 적재된 컨테이너에 대하여 컨테이너 외부반출을 위한 CY내 TC(Transfer Crane)가 컨테이너를 움직인 시간(①) 과 외부 차량이 게이트를 통해 터미널에 통과한 시간(②)을 산정하였다.

수출화물(Outbound)는 이와 반대의 경우로 외부차량이 우선적으로 게이트를 통과한 시각(③)과 CY내 TC가 컨테이너를 움직인 시간(④)의 차이로 선정하였다. 이것이 실제로 외부차량이 컨테이너를 움직이기 위한 최소의 시간으로 설정하였다.

---

40) 관세법 기준으로 외국에서 국내로 수입되는 화물

41) 관세법 기준으로 국내에서 외국으로 수출되는 화물

## GATE 반출입 시간



<그림 3-10> 터미널별 턴어라운드 타임 분석 화면

<그림 3-10>에 따르면 터미널별 평균 턴어라운드 타임은 19.4분으로 PNIT와 HPNT는 다소 평균보다 높은 것으로 나타났다. 특히 시간대별로는 새벽이나 밤시간대가 낮으며 날짜별로는 10~20일간이 턴어라운드 타임이 매우 적은 것으로 나타났다. 이는 제조사의 월말기준 물동량이 폭증하는 영향이 항만에서도 적용된다고 할 수 있다, 또한 요일별은 월, 금, 토이 높아 선박과 화물이 집중하는 것이 알 수 있었다. 이는 부산항이 타 항만에 비해 항만시설 사용료가 저렴한 효과가 데이터로서 입증되는 결과이다.

시간/요일별 GATE 반출입 시간

시간 ▼	요일 ▼						
	일	월	수	목	금	토	일
0	34.1	42.2	35.9	39.1	32.5	42.2	38.0
1	41.2	32.0	33.2	39.7	35.2	41.9	39.0
2	35.4	39.2	35.9	29.2	32.5	38.9	36.3
3	27.7	30.8	32.8	38.3	30.7	33.2	35.7
4	30.3	40.0	27.6	38.2	24.7	41.8	31.3
5	28.0	33.6	35.7	43.7	31.6	43.8	27.6
6	34.4	38.1	34.0	48.1	31.1	40.6	24.1
7	34.7	32.7	37.9	38.3	35.5	41.6	35.5
8	33.9	37.6	40.4	42.5	32.4	37.3	37.2
9	30.2	39.3	47.6	35.8	35.1	34.3	36.0
10	32.6	44.3	38.3	35.8	42.4	31.5	40.7
11	34.5	35.8	31.6	37.7	38.7	36.9	40.2
12	35.8	38.0	34.5	37.8	42.0	35.8	44.8
13	34.6	41.5	32.1	38.0	40.2	36.0	36.3
14	34.1	40.2	37.5	41.0	37.7	38.6	43.4
15	34.1	34.7	34.1	38.1	36.1	36.3	40.0
16	34.0	39.1	31.1	35.3	39.0	36.6	37.8
17	35.9	37.7	33.4	42.4	44.4	40.2	42.0
18	34.3	32.7	34.2	38.4	36.9	35.9	39.1
19	39.9	35.1	33.9	36.1	44.5	34.4	37.4
20	48.3	36.5	30.9	33.7	39.3	34.7	35.8

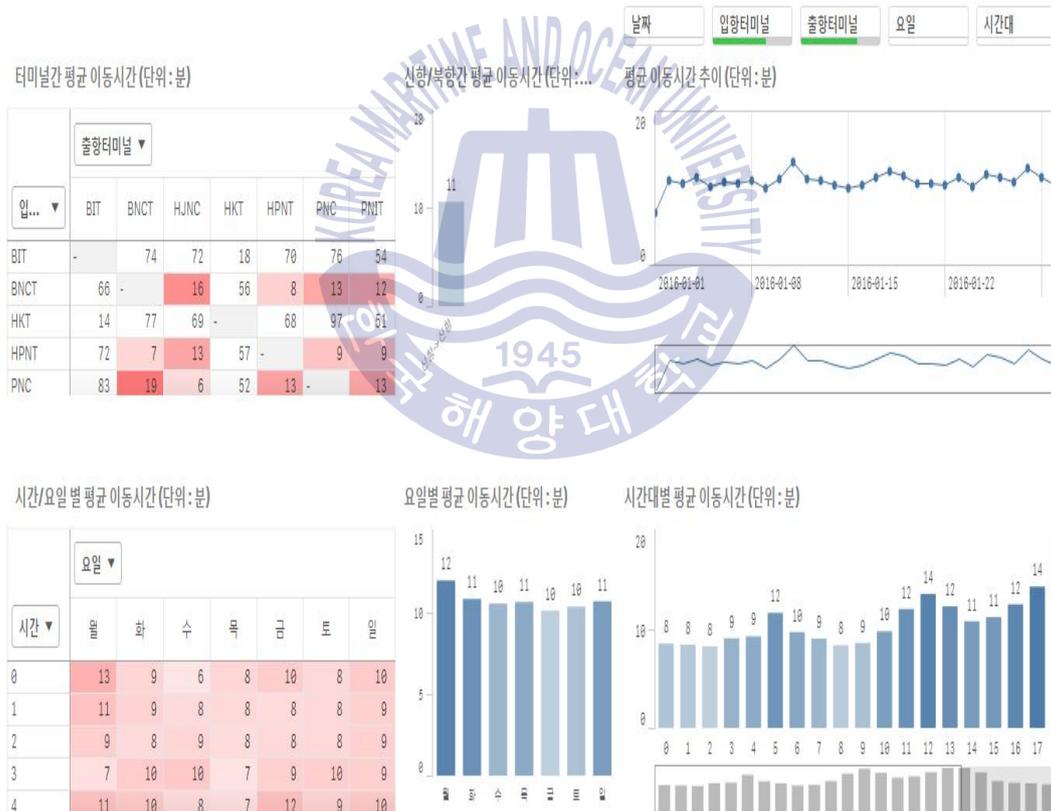
<그림 3-11> 터미널별 턴어라운드 타임분석(요일/시간별)

특히,<그림 3-11>과 같이 요일 및 시간대별 턴어라운드 타임이 오래 걸리는 정도를 나타내어 향후 외부트럭운전자들이 동 시간대별 상하차 정도 및 대기시간을 줄일 수 있는 용도로 활용될 것이다.

### 3) 터미널간 이동 시간 분석

신항 터미널별 이동시간을 '16년 1월달 기준으로 분석하였다. PNC와 BNCT간의 평균이동시간 19분으로 상대적으로 높고, 다음으로 HJNC와 BNCT간의 순으로 나타났다. 가장 짧은 시간대는 PNC와 PNIT가 가장 이동시간이 낮았다. 이는 터미널간 이동시간은 터미널간의 거리와 상관이 있으며, 오후 시간대가 차량이 혼잡하여 이동 시간이 다소 지체되는 것으로 파악되었다. 요일별은 선박이 많이 입항하는 월요일, 토요일이 상대적으로 높았다.

#### 터미널간 이동시간 정보



<그림 3-12> 터미널별 턴어라운드 타임분석(요일/시간별)

4) 복화율을 높이기 위한 요일 및 시간별 업체별 분석(신항터미널 위주)

'16년 1월 1일 선박 입항일 기준으로 1일간 타부두 TS를 ITT를 한다는 가정하에 <표 3-4>로 나타냈다.

<표 3-4> 1일 기준 복화율

(단위 : 총건수/복화율건수)

From \ TO	PNIT	PNC	HJNC	HPNT	BNCT	복화율 평균 (%)
PNIT		217(82)	16(16)	15(15)	32(32)	51
PNC	229(39)		15(15)	33(33)	21(21)	36
HJNC	34(34)	10(10)		2(2)	14(14)	100
HPNT	113(21)	175(35)	107(38)		17(17)	27
BNCT	3(3)	29(29)	31(31)	9(9)		100

출처: 부산항만공사 내부자료

<표 3-4>는 1일간 타부두 TS 총 물동량 중 운송사가 다른 것을 복화율을 한다는 가정하에 복화율 평균을 산정하였다. 평균 100건의 미만인 터미널은 모두 복화율을 통하여 통행량을 50%이상 줄일 수 있고, 그로인해 생산성과 효율적인 측면을 나타낼 수 있다. 향후 ITT 플랫폼을 통한 운송오더가 1곳으로 모일시 복화율은 더욱 상승하게 될 것으로 기대되고 있다.

## 제4장 부산항신항 ITT 운영방안

### 제1절 부산항신항 ITT 현황 및 문제점

#### 1. 부산항신항 터미널 운영현황

2017년 현재 5개 부두(민자 2개, 임대 3개) 6.85km의 선석이 운영 중이며, 2020년 이후 4개 부두에 4.65km의 선석이 단계적으로 개장될 예정 최대 9개 운영사가 상호 경쟁할 수 있는 구조로 향후 북항 운영상 경영난 재현이 우려된다.

<표 4-1> 부산항신항 운영사 현황

구분	1부두	2부두(민자)	3부두	4부두	5부두(민자)
운영사	부산신항 국제터미널(주) (PNIT)	부산신항만(주) (PNC)	한진해운 신항만(주) (HJNC)	PSA 현대부산 신항만(주) (PSA HPNT)	(주)비엔씨티 (BNCT)
운영개시	2010.3.	2006.1.	2009.2.	2010.2.	2012.1.
부두길이	1,200m	2,000m	1,100m	1,150m	1,400m
전면수심	16m	16~17m	18m	16~17m	17m
하역능력	2,090천TEU	3,680천TEU	2,310천TEU	1,940천TEU	2,440천TEU
부지면적	840천㎡	1,210천㎡	688천㎡	553천㎡	735천㎡
지분관계	PSA(싱가포르) 60% (주)한진(한국) 40%	DPW(아랍에미리트) 66.03% 기타 33.97%	(주)한진(한국) 50%(+1주) 펠리샤(한국) 50%(-1주)	PSA(싱가포르) 40%+1주 현대상선(한국) 10% 와스카(한국) 50%(-1주)	MKIF(호주) 30% 타미라(프랑스) 외1 18.5% BPA(한국) 9% 기타 42.5%

출처 : 부산항만공사.

<표 4-2> 부산항 신항 추가 부두 건설 현황

구분	2-4단계(민자)	2-5단계(확장구간 포함)	2-6단계	3단계
운영개시	2021년(예정)	2020년(예정)	미정	미정
부두길이	1,050m	1,050m	700m	1,850m
전면수심	17~18m	20m	20m	18m 이상

출처 : 부산항만공사.

한편, 신항의 문제점을 살펴보면, 첫째 운영사 과다문제이다. 선박 대형화로 선석 길이 부족, 장치장 혼잡 등 선사 측면에서 항만이용에 대한 매력도가 떨어지고 있으며, 신규 부두 개장으로 운영사 간 효율경쟁 및 반감 증대될 수 있다. 둘째 하역료 불균형 수출입과 환적(자부두, 타부두) 화물 간 효율 불균형으로 신항 경쟁력 저하가 우려된다. 셋째 5개 부두로 쪼개진 부두 특성상 신항 내 환적물동량 효율성 저하된다.

세부적인 내용은 현재 운영 중인 신항 5개 컨테이너 터미널 총 길이는 6,950m로 부두 당 평균 길이는 1,400m(약 3선석)에 못 미치는 실정으로 최근 선박 대형화 추세와 더불어 여러 척의 선박 동시 접안 시 선석 길이 부족, 장치장 혼잡 사태 등 생산성에 한계 발생한다.

<표 4-3> 주요 시기별 선박 대형화 추세

1968년	1972년	1980년	1997년
Encounter Bay (1,530TEU급)	Hamburg Express (2,950TEU급)	Neptune Carnet (4,100TEU급)	Susan Maersk (8,000TEU급 이상)
2005년	2012년	2014/2015년	2018년
Gjertrud Maersk (10,000TEU급 이상)	Marco Polo (15,000TEU급 이상)	CSCL Globe/MSOscar (19,000TEU급 이상)	미상 (22,000TEU급)

타부두 환적 발생 증가되면 이는 5개 부두로 쪼개진 부산항 신항 운영체제로 인해 환적화물의 부두 간 이동과 비용발생이 불가피 한 실정으로 부산항 환적경쟁력 확보에 최대 걸림돌이며, 2020년 이후 운영개시 예정인 2-4, 2-5, 2-6단계, 3단계 부두 간 환적문제도 잠재되어 있다.

특히, 동서 간 환적은 신항 북'컨' 3개 부두 및 남'컨' 2개 부두 간 화물 이동 단절로 인해 비용 발생되고 남북 간 환적(신항 북'컨'과 남'컨' 부두 간 이동 단절로 인해 비효율이 발생하고 있다. 장기적으로 서'컨'과 북'컨' 간 화물 이동에 대한 대안 부재가 문제점으로 거론되고 있다.

## 2. 부산항신항 ITT 운영 현황 및 문제점

현재 부산항신항 ITT는 터미널간 운영사의 이해관계로 ITT운송을 위한 터미널 간 울타리는 철거된 상태로 이송에 대한 인프라적 문제는 없지만, ITT에 대한 운송비용 주체 및 터미널내 안전사고 등을 감안하여 최적은 운영 방안을 모색하기 위해 항만관리자(부산항만공사)와 개선방안을 논의 중에 있다고 한다.

부산항신항 ITT의 문제점을 정리하면 첫째 ITT 운송차량을 당초 YT가 아닌 외부 RT(Road Tractor)로 변경이 필요하다는 것이다. 이유는 신항 북'컨'부두와 남'컨'부두 끝단까지 최장 약 7km 거리를 두고 있어 YT 안정성 등을 감안할 때 운송차량이 변경이 불가피하다는 것이다. 둘째 선사가 부담하는 ITT 운송요금에 대한 지속적인 덤핑에 대한 것이다. 이는 터미널 운영사가 하역요금외에 ITT 운송요금을 선사에게 청구하면 선사들은 자부두 TS 이송을 핑계로 타부두 TS 이송에 대한 요금을 덤핑을 요구할 것이다라는 터미널 운영사의 주장이다. 이중에서 ITT 운송요금에 대한 선사의 덤핑은 ITT 실행에 대한 심각한 걸림돌이 되고 있다. 당초 ITT 시행은 타부두 셔틀비용 절감으로 선사, 운송사, 터미널운영사의 WIN-WIN 효과로 환적화물 유치에 반드시 필요한 것으로 시행되었으나, 그 외의

선사의 운송요금 덤핑은 ITT 시행을 더욱 어렵게 한다고 볼 수 있다. ITT플랫폼 설치에 따른 운송비용의 절감 효과에 대하여 자세히 살펴보고자 한다.

타부두 TS를 처리시 하역료와 셔틀 운송비를 포함하여 20ft 115천원, 40ft 168천 원 소요되나 부산항만공사에서 신항 내 ITT로 인한 추가 운송비용 보전을 위해 인센티브 제도 시행(20ft 5천원, 40ft 6.5천원, 연 43억원 '16년도)하였다.

<표 4-4> 부산항 신항 화물 유형별 처리 비용

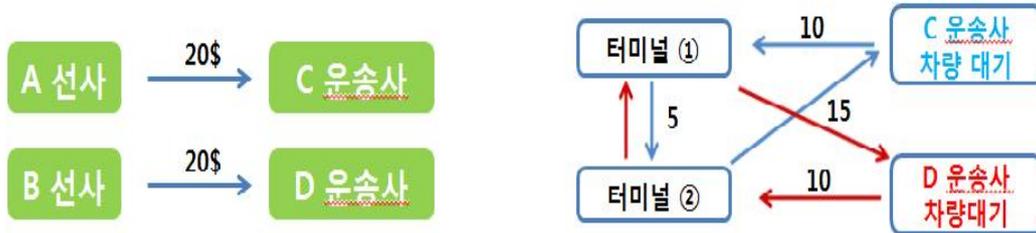
(단위 : 천원)

구 분	계		하역료		셔틀비	
	20ft	40ft	20ft	40ft	20ft	40ft
수출입	50	75	50	75	-	-
자부두 T/S(A)	80	120	80 (50×2회×0.8)	120 (75×2회×0.8)	-	-
타부두 T/S(B)	115	168	95 (50×2회×0.95)	143 (75×2회×0.95)	20	25
비용 차(B-A)	35	48	15	23	20	25

출처 : 부산항만공사 자체 자료(폴컨데이터기준 시장 평균치 추정)

현 ITT 운영체제 문제점으로는 경쟁항만의 환적화물 하역료 및 신항 내 자부두 TS 대비 신항 ITT 처리 비용이 높아 선사의 물류비 부담 증가되고 있다. 20ft T/S 하역료 기준 Ningbo항 45천원(부산항 대비 39%), Shanghai항 54천원(부산항 대비 50%)으로 부산항 신항 115천원 대비 저렴한 수준이며, 신항 내 ITT 시 20ft 35천원, 40ft, 48천원의 추가비용이 발생하여 자부두TS 대비 약 40%가 비싼 구조이다.

한편, 복화율을 높임으로 단순 이동 구간 감소 및 공차율은 줄임으로써 이익이 증대되는 것을 알 수 있는데, 이는 <그림 4-1>과 같다. 즉, ITT Platform 도입으로 선사 부담 감소, 운송사 매출 생산성 증대됨을 알 수 있다.



- 2 box 셔틀 운송을 위해서 2대의 트럭 투입. C 트럭 운행거리 30. D 트럭 운행거리 30. 총 매출액 40\$
- 즉, 두 운송사이 단위거리당 매출액 = 총 매출액 (40\$) / 총운행거리 (60) = 0.67\$
- 두 선사 부담 셔틀비용 총 40\$



- 2 box 셔틀 운송을 위해서 1대의 트럭 투입. 트럭 총 운행거리 최대 40. 총 매출액 34\$
- 즉, 단위거리당 매출액 = 총 매출액 (34\$) / 총운행거리 (40) = 0.85\$
- 두 선사 부담 총 셔틀비용 36\$

<그림 4-1> 현행과 ITT 플랫폼 도입시 비용 비교

또한, 터미널의 하역 플래닝시 본선 작업을 최우선으로 진행되고 있어, ITT 운송 차량의 터미널 내 장시간 대기 빈번하게 발생한다 평균적으로 ITT 입항 터미널 게이트 출발하여 출항 터미널 상하차 완료까지 1 move당 평균 약 1시간 20분 소요로 운송사의 ITT 운송 기피를 대체적으로 기피하는 현상이다.

신항 ITT 효율화 추진방안으로 제도적으로 크게 문제될 것이 없으나 체계적으로 ITT 운영을 진행하기 위해서는 ITT플랫폼을 구축하여 시스템을 구축하는 것이 우선적으로 필요하다고 항만주체인 부산항만공사를 밝히고 있다.

부산항만공사는 ITT 시행을 중단기로 시행을 계획하고 있으며, 우선 1단계로

다목적부두를 ITT 전용도로로 활용하고 PNIT ↔ HPNT 간 내부 운송 개시하는 것으로 하며, 중장기인 2단계로는 신항 전체 터미널 간 내부운송 및 ITT 플랫폼을 통한 공동배차제도 도입으로 ITT의 효율성을 극대화 하겠다는 의견이다. 현재로서는 1단계 사업이 진행중이며 2단계 플랫폼 설계 중인 알려지고 있다.



<그림 4-2> 신항 ITT 효율화 개념도

세부적인 ITT 운행 방법은 터미널 내부도로를 탄력적으로 활용한다는 개념으로 터미널 간 울타리에 내부 운송을 위한 게이트를 설치하고 북'건'과 남'건'을 연결하는 ITT 전용도로 운영한다. 인접 터미널은 내부도로를 통해 YT(Yard Tractor)로 운송하며 운영사가 운송 주체가 되고, 비인접 터미널은 임항도로를 통해 RT(Road Tractor)로 운송하며 운송사가 운송 주체가 된다.

향후, 신항 ITT 정보를 집적화하여 공동배차·운송 할 수 있는 통합정보시스템 기반의 공동배차제도 운영할 계획이다.

주요 문제점으로는 ITT 공동배차시스템을 통한 선사, 운영사, 운송사의 ITT 정보를 집적화하여 환적화물 상세정보 제공하여 복화율을 높이고, 터미널의 혼잡도 등을 감안하여 차량대기 시간 및 차량을 분산하여 터미널의 생산성과 효율성을 향상시키는 방안이 필요하다.

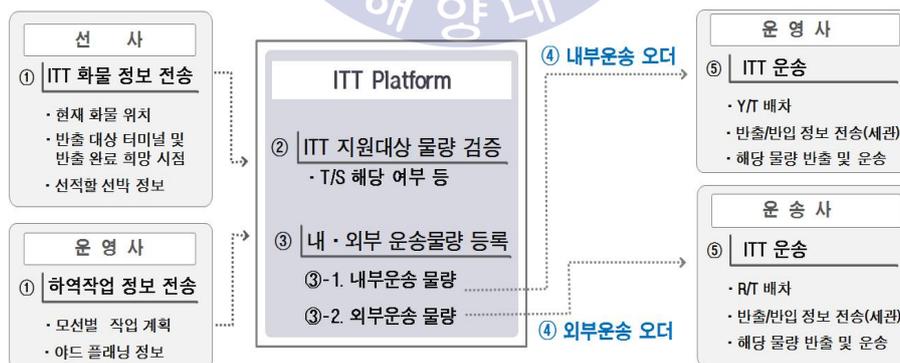
## 제2절 ITT 플랫폼 설치를 위한 기본 방향 및 운영방안

### 1. ITT플랫폼 설치를 위한 기본방향

부산항신항 내 인접 터미널간은 터미널 운영사가 주체가 되어 YT로 운송하며 비인접터미널간 타부두 TS는 외부운송을 위한 ITT 플랫폼을 통한 공동배차제도 도입을 부산항만공사는 적극 운영할 계획이라고 한다.

ITT 플랫폼의 주요목적으로는 기존 운송사 플랫폼화를 통한 ITT 운송 물량 공유로 차량 대기시간 최소화 등 운송 효율 제고하는 것이라 할 수 있다. 시행을 위한 기본방향은 부산항만공사 주체로 공동배차시스템을 개발·운영하여 운송사 간 ITT 정보 공유 및 비인접터미널 간 외부운송 ITT 물량 공동 운송하는 것이다. 특히 ITT 플랫폼 및 공동배차제도는 우선적으로 항만관리자가 시행하고 향후 안정화시기때는 민간에게 위탁하는 방식이 기본방향이라 할 수 있다.

세부적인 ITT 플랫폼을 통한 공동배차제도 운영방식은 <그림 4-3>과 같다. 선사를 통하여 ITT 화물 정보를 플랫폼으로 전송받고, 운영사의 하역작업 정보를 플랫폼에 담아 ITT 운송정보 결합, 복화율등을 감안하여 내부운송 및 외부운송으로 오더로 분류하여 운송사에 전달하게 된다.



<그림 4-3> 공동배차 방식 개념도

## 2. ITT 플랫폼 운영 및 구축 방향

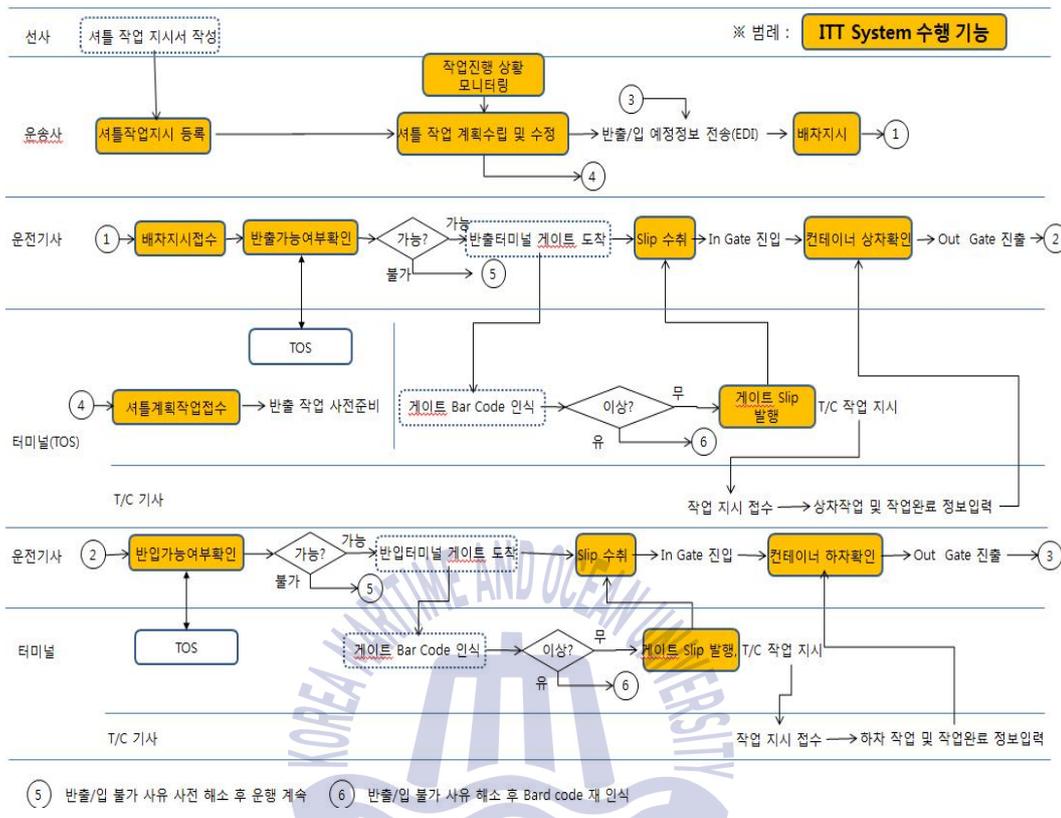
부산항신항 ITT 플랫폼은 공동배차방식의 일환으로 항만관리 주체인 부산항만공사가 직영으로 ITT플랫폼을 설립·운영하고, 기존 운송사가 Alliance 체계를 구축한 후 BPA로부터 위탁받아 ITT 운송하는 방식이 필요하다. 각 주체별 역할 분담은 아래와 같다.

<표 4-5> ITT 플랫폼 운영 주체별 역할

구 분	부산항만공사(BPA)	운송사
주요역할	<ul style="list-style-type: none"> <li>• ITT 운송비용 부담</li> <li>• ITT 지원대상 물량 검증, 운송오더 등 Platform 운영</li> </ul>	<ul style="list-style-type: none"> <li>• ITT 외부운송(차량수배, 배차, 운송)</li> <li>• 사고발생 시 책임 주체</li> </ul>

즉 선사, 운영사, 운송사가 가진 ITT 정보를 집적화하여 내·외부운송 ITT 공동배차를 위한 「ITT Platform」 설립·운영이 필요하다.

이제 ITT 플랫폼에 대한 구축 방향은 선사, 터미널운영사, 운송사 시스템의 연계되어 운영이 필요하며 선사는 다음 내용을 포함한 “셔틀 작업 지시서”를 ITT Platform에 송부하고 ITT Platform은 접수된 셔틀 작업 지시서 내용을 근거로 셔틀 작업 계획 수립한 후 셔틀 작업 계획에 따라 배차 지시 및 운송완료 후 비용 청구 및 지급의 절차가 필요하다 구체적인 프로세스는 아래 ITT 플랫폼 기능 체계도 <그림 4-4>와 같다. 동 플랫폼의 기능 체계도는 업무프로세별로 순서도를 통하여 나타내었다.



<그림 4-4> ITT Platform 기능 체계도

### 제3절 빅데이터를 이용한 ITT 운영 시스템 구축방안

앞서 분석한 빅데이터를 이용한 결과를 ITT 플랫폼에 설계단계부터 활용하면 복화울 뿐만아니라 터미널의 혼잡도 까지 분석이 가능하여 향후 터미널 생산성과 효율성을 증가시킬 수 있을 것이다.

ITT운영 시스템 구현방안은 우선적으로 기존 운영사가 사용하는 개별 프로그램을 기준으로 복화울과 터미널기준 내외부 운송을 구분할 수 있는 프로그램 개발이 필요하다. 특히 선사 및 터미널 운영사로부터 화물정보, 내외부 오더 정보 등을 주

고 받기 위해서는 EDI(Electronic Data Interchange) 방식으로 추진하는 효율적이다.

또한 시스템 운영시 가장 ITT 운송을 많이 하는 선사와 계약된 운송사를 대상으로 우선적으로 플랫폼 적용을 하여 대고객 서비스 일환으로 컨테이너 운송 기사에게 터미널별 시간대별로 혼잡도, 작업완료 추정시간, 그리고 20분내에 작업완료할 확률 등 실시간 현황을 서비스 할 수 있으면 RT 운전기사들이 불필요한 시간 낭비를 줄이 수 있게되어 효율성이 높다고 판단된다. 특히 빅데이터 분석을 통해 본선작업 정도는 작업을 완료할 수 있는 시간과 크게 영향력이 있는 것으로 나타나서 향후 빅데이터 분석이 좀 더 필요한 것으로 사료된다.

<그림 4-5>는 최종적으로 운송기사 및 운송사가 보는 예상되는 화면을 구현해 보았다.

터미널 선택 : **PNIT** 날짜 선택 : **2016-5-4** **수요일**

시간	혼잡도	본선 작업 정도	작업완료 추정시간	20분 이내 작업완료할 확률
07시~08시	■		9분	98%
08시~09시	■■	■	13분	82%
09시~10시	■■■	■■	15분	71%
10시~11시	■■■	■■■	12분	84%
11시~12시	■■■	■■■	14분	78%
12시~13시	■■■		12분	86%
13시~14시	■■■■		17분	66%
14시~15시	■■■■■	■	21분	56%
15시~16시	■■■■■	■	21분	56%
16시~17시	■■■■■	■	21분	56%
17시~18시	■■■■■		27분	43%
18시~19시	■■■■■	■■■	20분	57%
19시~20시	■■■■■	■■■■	20분	55%
20시~21시	■■■■■	■■■■■	19분	60%
21시~22시	■■■■■	■■■■■	19분	60%
22시~23시	■■■■■	■■■■■	20분	58%
23시~24시	■■■	■■■■■	13분	78%

<그림 4-5> 터미널내 혼잡도 최종 UI 화면 예시

## 제5장 결론

### 제1절 연구결과의 요약 및 시사점

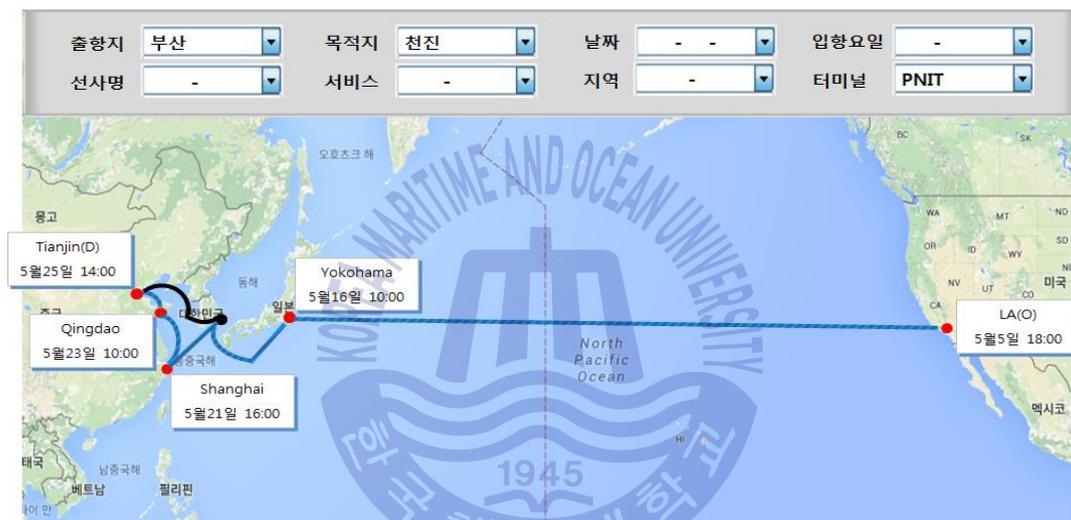
부산항이 글로벌 허브항만으로 선점하기 위해서는 경쟁항만에 비해 불필요한 고정비용을 절감하는 효과정책이 우선적으로 필요하다. 고정비용 중 선사들이 부산항의 경쟁력을 높이는 데 가장 많이 요구하는 사항이 타 부두 TS 이송 비용 절감을 요구하고 있다. 이를 위해 중 가장 시급한 대안을 요구하는 사항은 바로 ITT 시행에 따른 터미널 간 운송비용 발생이라고 할 수 있다. 부산항에서는 신항내 터미널 간 환적물량 운송비뿐만 아니라 북항-신항 간 환적 물량 운송으로 추가적인 환적비용이 발생하고 있어 향후 환적화물 유치에 제약조건으로 발생하고 있다.

부산 신항 내 터미널 간 환적 물량 운송이 발생하게 되면 터미널 운영사는 추가적으로 발생하는 운송비를 선사들에게 떠넘기게 되고 이로 인해 선사들에게는 직접적인 추가환적비용이 발생하게 된다. 선사동맹이 확대될수록 물동량은 터미널 전반에 분산될 수밖에 없기에 ITT 관련 비용은 향후 환적물동량 유치를 조속히 해결하지 않을 수 없는 문제라 할 수 있다.

신항내 ITT를 실행하기 위해 플랫폼 구축하고, 운영에 필요한 데이터를 우선적으로 알고리즘을 만들어 빅데이터라는 최신 ICT 기술을 활용하여 분석해 보았다. 빅데이터 분석을 통해 타부두 TS 복화율을 높이기 위해 플랫폼을 구축을 위해 타 운송사의 운송시간과 선사를 맞추어 복화율을 높이는 시뮬레이션과 각 터미널 별 혼잡도 및 턴어라운드 타임을 분석하여 요일별, 시간대별 한가한 시간을 알려 줌으로써 터미널의 운송기사에게 효율성고 편리성을 줄 수 있다. 특히, 턴어라운드 타임 등 항만이용고객의 편리성을 위해 빅데이터가 많은 효과를 나타내는 것으로

판단된다.

향후 부산항신항 ITT 뿐만 아니라 부산항 물동량 패턴분석, 위험물, 공‘컨’ 이상감지 등 빅데이터를 이용하면 효과적이고 의미있는 데이터를 분석할 수 있을 것 같다. <그림 5-1>은 빅데이터 분석 외에 환적화물을 선대서비스를 화면으로 나타낸 것도 가능할 것으로 판단된다 이는 최종 화물과 선박의 부킹서비스도 동시에 나타낼 수 가 있어 고객에 대한 서비스가 한층 좋아 질것으로 판단하고 있다.



<그림 5-1> 출발지 및 목적지 입력에 따른 선대서비스(예시)

## 제2절 연구의 한계점과 과제

본 연구는 부산항 신항 ITT 플랫폼 구성·운영하기 위하여 우선적으로 빅데이터 틀을 이용하여 복화율과 터미널 혼잡도 등을 반영하도록 최적으로 데이터 분석을 제시하였다. 그러나 빅데이터를 이용한 부산항 신항 ITT 플랫폼에는 아직 적용하지 못한 것이 동 연구에 한계점이다. 또한 빅데이터를 혼잡도를 분석하는 것

이 현장에서 실제 얼마나 혼잡한가? 라는 측면에서 향후 연구에서는 실증검증이 필요하다.

그러나 빅데이터를 통해 효율적으로 ITT 플랫폼에 적용하기 위해서는 빅데이터로 분석된 데이터에 의한 신항 ITT 플랫폼의 데이터 알고리즘에 반영한다면 보다 효과적인 부산항신항 ITT 시행 도움을 줄 수 있을 것이다.



## 참고 문헌

### <국내 문헌>

- 이창훈(2014), “RHive를 활용한 빅데이터 분석 연구”, 가천대학교 석사학위논문.
- 한경록(2015), “광주 주요분야의 빅데이터 구축 및 가공·활용방안 연구”. 광주발전연구원.
- 김현곤외(2009), “주요국의 미래전망 및 ICT 활용전략 연구”, 한국정보화진흥원.
- 데이코산업연구소(2013), “빅데이터 관련시장 실태와 전망”, 데이코.
- 김진영(2015), “물관리에서 빅데이터 활용 가능성 탐구”, 물정책 학회지.
- 김동환(2014), “국토도시 정책을 위한 공간 빅데이터 분석과 활용”, 부동산포커스.
- 박원준(2012), “빅데이터 활용에 대한 기대와 우려”, 방송통신전파저널.
- 권기동외(2015), “빅데이터 플랫폼 구축을 활용한 운전습관 및 운전형태 분석”. 한국교통대학교.
- 유영준(2015), “관광마케팅 전략수립을 위한 빅데이터 활용 방안”. 울산발전연구원.
- 오석문외(2014), “부산신항 타부두 환적의 효율적인 처리방안 연구”. 한국해양과학기술진흥원.
- 송영미(2015), “클라우드 및 빅데이터 기반 침수도로 탐색 시스템 개발”, 부경대학교.
- 박남규(2014), “부산신항 ITT 플랫폼 설치 운영에 관한 공동 연구”, 동명대산학협력단·BPA.
- 김성원(2014), “빅데이터가 기업의 경쟁력에 미치는 영향에 관한 연구”, 경기대학교 석사학위 논문.
- 신현자(2016), “빅데이터 기반의 재난정보관리 방안”, 서울시립대학교 석사학위논문.
- 가회관(2015), “빅데이터 도입의도에 미치는 영향요인에 관한 연구”, 중앙대학교 박사학위논문.
- 김연수(2016), “빅데이터 분석 기법을 이용한 기후변화 복원탄력성 지표 개발”, 인하대학교 박사학위논문.
- 김일환(2014), “빅데이터 분석기법을 활용한 국가주요시설물의 안전등급 예측 모델에 관

한 연구, 고려대학교 석사학위논문.

김지 (2012), “빅데이터 활용과 분석기법 고찰”, 고려대학교 석사학위논문.

<외국 문헌>

Marchment Hill Consulting(2010), “Smart Water Metering Cost Benefit Study”.

Betty.M.(2013), Big data, smart cities and city planning, *Dialogues in Human Geography* 3(3)

D.H. Lee. J. G. Jin(2012), “Terminal and Yard Allocation Problem for a Container Trasnshipment Hub with Multiple Terminal”, *Transportation Resarch Part E*, Vol. 48.

Gartner (2011), *Big Data Analytics*, Gartner Group

Gartner (2011), “Pattern-Based Strategy : Getting Value form Big Data”, Gartner Group

Gartner (2011), “How to plan Participate and Prosper in the Data”, Economy, Gartner Group

McKinsey (2011), *Big Data : The Next Frontier for Innovation, Competition, and Productivity*, McKinsey & Company

Philip Russom (2011), “Big Data Analytic Report”, TDWI Research

Steve LaValle etc. (2011), “Big Data, Analytics and the Path From Insights to Value”, *MIT Slon Management Review*

Vertica (2009), “Managing Big Data with Hadoop & Vertica, Vertica Systems”.

## 부 록

### 1. 주요업체(기업별) 활용사례<sup>42)</sup>

업체	활용사례
구글	<p>감기'와 관련된 검색어 분석을 통한 독감 예보 시스템 제공</p> <ul style="list-style-type: none"> <li>미국 질병통제 예방센터의 데이터와 비교 결과, 검색 빈도 및 실제 독감증세를 보인 환자 숫자 사이에 매우 밀접한 상관관계가 있는 것을 확인하고, 구글 홈페이지에서 독감, 인플루엔자 등 독감과 관련된 검색어 쿼리의 빈도를 조사, '구글독감 동향 (Google Flu Trends)'이라는 독감 확산 조기 경보체계 마련</li> <li>미국 보건 당국보다 한발 앞서 시간 및 지역별 독감 유행 정보를 제공 실시간으로 내부생산 및 재고를 관리하여 패스트 패션(fast fashion) 지향</li> </ul>
자라	<ul style="list-style-type: none"> <li>MIT 대학과 함께 전 세계 매장의 판매와 재고 데이터를 분석하여 최대매출을 달성할 수 있는 재고 최적 분배 시스템 개발하여 자사 기획 상품의 판매현황을 실시간으로 분석</li> <li>시장에서 인기 있는 제품에 대한 분석을 통해 생산 시스템에 직접적으로 연결함으로써 빠르게 변화하는 트렌드에 맞추어 의류 제공할 수 있는 시스템 마련</li> </ul>
코카콜라	<p>SNS, 인터넷 게시판 등의 데이터를 실시간으로 분석하여 내부 자산으로 활용</p> <ul style="list-style-type: none"> <li>다양한 소셜 미디어에서 얻어진 데이터를 분석하여 제품 판매에 연관된 의사결정에 반영하여 코카콜라에 비우호적인 정보가 증가하는 국가나 지역을 대상으로 홍보를 강화하는 등 실시간 대응 가능</li> </ul>
월마트	<p>각 지점의 모바일과 소셜 쇼핑의 특징을 이용한 '@월마트랩(Walmartlabs)' 운영</p> <ul style="list-style-type: none"> <li>소셜미디어 회사인 코스믹스(Kosmix) 인수를 통해 소셜 네트워크와 콘텐츠를 관리함으로써 유통과 전자상거래(e-commerce)간의 시너지 발휘</li> </ul>

42) 고준철, 이해욱, 정지윤, 강경식(2012), "빅데이터의 새로운 고객 가치와 비즈니스 창출을 위한 대응 전략", 「대한안전경영과학회지」, 제14권 제4호, pp.229~238.

	<ul style="list-style-type: none"> <li>• Social Genome은 소셜 미디어를 통해 대규모 데이터를 수집하여 리얼타임으로 해석된 추출된 정보를 이용하여 상품판매를 촉진하는 기법으로 시시각각 변화하는 소비자의 패턴을 분석하여 적재적소에 필요한 물품을 빠르게 제공함으로써 불필요한 재고 낭비 방지하고 고객이 원하는 물품을 충분히 공급할 수 있기 때문에 점포당 고객 만족도 향상으로 이어져 기업 발전에 선순환적인 역할</li> </ul>
넵플릭스	<p>고객 데이터 분석 활용한 차별화를 통해 콘텐츠미디어 기업으로 성장</p> <ul style="list-style-type: none"> <li>• 고객이 과거에 대여한 영화목록과 시청한 영화에 부여한 평점 등의 데이터를 분석하는 시스템 개발하여 10만 개의 영화정보, 1,600만 명 고객의 시청이력 정보에 대한 분석추진</li> <li>• 각 고객별 웹사이트 내 실시간 행동패턴을 분류하여 개인별 맞춤형 페이지를 구축하고 최적화된 영화 콘텐츠를 추천하여 하루 평균 50억 개 추천</li> </ul>
포스코	<p>원료가격을 효율적으로 구매함으로써 가격 경쟁력을 향상</p> <ul style="list-style-type: none"> <li>• 철광석 가격에 영향을 미치는 남미, 호주 광산의 상황과 런던 금속거래소(LME)를 통해 수집한 광물 가격 데이터를 실시간으로 분석</li> <li>• 고객사의 수요 데이터와 전 세계 철광산 및 현물 거래소의 가격 데이터를 조합한 후 비교하여 철광석 구매의 최적 타이밍과 가격대 결정</li> <li>• 자원 투기 세력에 의한 원자재 가격 급락의 위험으로부터 보호</li> </ul>
GS EPS	<p>전력시장 분석을 통해 합리적인 전력 시장 분석 및 예측</p> <ul style="list-style-type: none"> <li>• 전력에 영향을 미치는 다양한 변수를 고려한 합리적인 전력 시장 예측</li> <li>• 송전 제약, 연료 제약, 국내탄 발전, 열 공급 발전 제약과 같은 다양한 조건적 요소를 반영한 시뮬레이션 결과 제공</li> <li>• 전력 수요 및 SMP(Standard Market Price) 예측, 사업 계획 수립 및 통합 DB 구축</li> </ul>
포스코	<p>소셜 네트워크에서의 여론동향 분석을 통해 기업에 필요한 콘텐츠 제공</p> <ul style="list-style-type: none"> <li>• 기업의 SNS 계정을 통합 관리하고 지원하는 기능을 통해 한국어트윗의 약65%정도 분석</li> <li>• 기업의 평판을 실시간으로 모니터링 하여 기업의 대응전략 마련을 지원</li> </ul>
다음	<p>소셜 네트워크의 이슈와 관심 키워드의 실시간 모니터링으로 상황에 맞는 대응전략 및 마케팅 전략 수립에 지원</p>
SAS	<p>기분이나 정서(Mood)를 분석하여 사회변화 예측</p> <ul style="list-style-type: none"> <li>• 최근2년 동안 미국과 아일랜드에서 인터넷채팅, 블로그, 페이스북, 트위터 등 소셜미디어 데이터의 기분이나 정서를 분석</li> </ul>

	<ul style="list-style-type: none"> <li>• 미국에서 ‘우울하다’, ‘열 받는다’와 같은 채팅이 늘어나면 4개월 뒤 실업률이 폭등함을 확인</li> <li>• 아일랜드에서는 실업률 증가 5개월 전 ‘불안하다’는 분위기가 퍼져나갔고, 2개월 전에는 ‘확신한다’는 채팅이 크게 감소</li> </ul>
EMC	<ul style="list-style-type: none"> <li>• 데이터 저장부터 관리, 분석까지 빅데이터와 관한 모든 것을 제공하기 위해 그린플럼, 아이실론 등 빅데이터 솔루션 업체 및 데이터 관련 다수업체 인수</li> <li>• 빅데이터 스토리지 솔루션(아이실론, 아트모스), 콘텐츠 관리 솔루션(다큐멘텀) 등 제공</li> <li>• ‘EMC 애널리틱스 랩’을 운영하여 데이터 사이언티스트(Data Scientist) 육성</li> </ul>
IBM	<ul style="list-style-type: none"> <li>• 지난 5년간 140억 달러 이상을 투자하여 비즈니스 분석 관련업체 인수 <ul style="list-style-type: none"> <li>- 분석용데이터 저장관리업체(네티자), 데이터통합 업체(에센셜), 분석솔루션 업체(코그너스)</li> </ul> </li> <li>• 빅데이터 솔루션 : InfoSphere BigInsight(Hadoop), InfoSphere Streams</li> <li>• 지속가능한 지구를 만들기 위해 지구 데이터(기온, 토양상태, 교통 흐름 등)를 분석하는 ‘스마트 플래닛(smart planet)’ 프로젝트 전개</li> </ul>
오라클	<ul style="list-style-type: none"> <li>• 세계적인 DB 업체, ‘하이페리온社’를 인수해서 분석기술 확보</li> <li>• 오라클 빅데이터 어플라이언스 제품 출시</li> </ul>
SAP	<ul style="list-style-type: none"> <li>• 업무용 어플리케이션 업체에서 최근 DB 전문업체로 변신</li> <li>• 메모리 기반 DB 어플라이언스(HANA) 제시</li> <li>• BI 소프트웨어, 플랫폼을 제공하는 ‘비즈니스 오브젝트社’인수</li> </ul>
HP	<ul style="list-style-type: none"> <li>• BI 솔루션 업체 ‘버티카’, 기업용 검색엔진 업체 ‘오토노미’인수</li> <li>• 버티카와 오토노미를 결합하여 빅데이터 분석 시장에 진입</li> <li>• ‘인스턴트-온 엔터프라이즈(Instant-On Enterprise)’ 솔루션으로 기업경영 의사결정, 경영정보 분석 등 경영지원 전략 수립 서비스 제공</li> </ul>
MS	<ul style="list-style-type: none"> <li>• 윈도 애저(Windows Azure)와 윈도 서버 플랫폼용 아파치 하둡 개발 계획</li> <li>• 하둡(Hadoop)9 기술 전문업체 ‘호튼웍스’와 협력</li> </ul>

## 2. 각국의 빅데이터 활용 사례<sup>43)</sup>

국가	활용사례
미국	<p>[국토보안]</p> <p>9.11 이후 미국은 국토안보부를 중심으로 테러, 범죄 방지를 위한 범정부적 빅데이터 수집, 분석 및 예측체계를 도입</p> <ul style="list-style-type: none"> <li>• 부시행정부의 국토안보부 장관인 Michael Chertoff는 국토보안을 위한 빅데이터 추진 현황언급</li> <li>• 국내외 금융 시스템인 개인, 기관의 금융거래 감시로 자금세탁 및 테러 자금조달 색출 강화</li> </ul> <p>[치안]</p> <p>FBI, 유전자 색인 시스템 활용한 단시간 범인 검거 체계 마련(CODIS)</p> <ul style="list-style-type: none"> <li>• 유전자 정보은행 CODIS(Combined DNA Index System) 구축</li> <li>• DNA포렌직, 클라우드DNA 분석 등 빅DNA데이터의 활용을 통해 2007년 45,000건의 범인 DNA Hit rate 달성</li> <li>• CODIS에는 미제 사건 용의자 및 실종자에 대한 DNA 정보 1만 3,000건을 포함한 12만 명의 범죄자 DNA 정보 저장</li> </ul> <p>(FBI는 매년 2,200만 명의 DNA 샘플을 추가하여 범죄 수사에 활용)</p> <ul style="list-style-type: none"> <li>• 1시간 안에 범인 DNA 분석을 위한 주정부 데이터 연계 및 빅데이터 실시간 분석 솔루션 확보 빅데이터 분석을 활용하여 탈세 및 사기범죄 예방 시스템 구축</li> <li>• 소셜 네트워크 분석에 기반을 둔 범죄 네트워크 분석 기능을 통해 데이터에서 이상 징후를 찾아내고, 예측 모델링을 통해 과거 행동 정보를 분석해 사기 패턴과 유사한 행동을 파악</li> <li>• 통합형 탈세 및 정부사기 방지 시스템을 통해 연간 3,450억 달러에 달하는세금 누락 및 불필요한 세금 환급 절감 효과</li> </ul> <p>[의료]</p> <p>오바마 health.20 - 필박스 프로젝트(Pillbox)</p> <ul style="list-style-type: none"> <li>• 국립보건원(National Library of Medicine)의 사이트로 약 검색을 서비스</li> <li>• PillBox를 통해 수집된 빅데이터를 통해 후천성면역결핍증(HIV)등 관리대상 주요 질병의 분포, 연도별 증가 등에 대한 통계치 확보 가능</li> </ul>

43) 양혜영(2012),“빅데이터를 활용한 기술기획 방법론”, 한국과학기술기획평가원 보고서, 제14권.

	<p>유전자 데이터 공유를 통한 질병치료체계 마련</p> <ul style="list-style-type: none"> <li>• 국립보건원과 75개 기업 및 기관들이 파트너십을 통해 공동으로 진행한 1000 유전체 프로젝트의 일환으로 200TB의 유전자 정보 확보</li> <li>• 1000 유전체 프로젝트는 전 세계에의 2,662명의 유전자 정보를 저장하고, 질병 연구를 위해 1% 이상의 빈도를 나타내는 유전적 다양성을 분석</li> <li>• 유전자 정보를 공유함으로써 새로운 질병에 대한 빠른 진단 서비스 제공 가능 퇴역군인의 전자의료기록 분석을 통한 맞춤형 의료 서비스 지원</li> <li>• 빅데이터 분석을 위해 2년간 25개의 데이터 웨어하우스를 배치하여 2,200만 퇴역 군인들에게 의료 서비스 지원</li> <li>• 페타바이트 수준의 임상적, 유전적 데이터 분석으로 보다 효과적인 의료 서비스 지원에 기여</li> </ul>
영국	<p>[정보공개]</p> <p>공공·민간·개인 등이 보유한 정보를 개방하여 상호 공유하게 하는 서비스</p> <ul style="list-style-type: none"> <li>• 2010년 영국에서 시작되었으며 공공기관, 민간기업, 개인 등이 보유하고 있는 전기, 가전, 휴대폰, 가로등의 센서로부터 제공된 정보를 저장하여 분석·제공</li> <li>• 개방된 소스로서 재난 안전 관리 시스템의 상호 연계를 지원</li> <li>• 공유데이터를 기반으로 웹 프로그램, 스마트폰 앱 개발 등에 응용 및 활용</li> <li>• 일반적인 수준의 데이터 이용은 대중에게 무료로 공개</li> <li>• 앱 개발 또는 분석된 데이터의 독점적 이용을 원하는 고객들에게는 유료로 제공 정부사이트(data.gov.uk)를 통해 공공부문의 정보 공유 및 활용을 위한 데이터 원스톱 서비스 제공</li> <li>• 시민의 자발적인 참여와 창의성을 기반으로 정책수립에 기여</li> <li>• 시민들이 공공정보에 대한 검색 및 재사용을 쉽게 할 수 있도록 정부가 보유한 다양한 데이터를 제공</li> </ul>
호주	<p>[정보공개]</p> <p>호주 정보관리청은 정부 2.0을 통한 정보 개방</p> <ul style="list-style-type: none"> <li>• 방대한 양의 정보를 검색하고 분석 및 재사용할 수 있도록 자동화된 틀을 활용하여 시간과 자원을 절감</li> <li>• AGIMO 산하 정부 2.0 전략/서비스 팀에서는 정부 데이터에 대한</li> </ul>

	리포지터리 및 검색 툴을 서비스하는 data.gov.au 웹사이트 운영
싱가포르	<p>[국가위협관리]</p> <p>싱가포르 정부는 빈번히 발생하는 테러 및 전염병으로 인한 불확실한 미래 대비를 위하여 2004년부터 빅데이터 기반 위험 관리 계획을 추진</p> <ul style="list-style-type: none"> <li>• RAHS(Risk Assessment &amp; Horizon Scanning) 시스템을 통해 질병, 금융위기 등 모든 국가적 위험을 수집 및 분석하여 위험을 선제적으로 관리</li> <li>• 수집된 위험 정보는 시뮬레이션, 시나리오 기법 등을 통해 분석되어 사전에 위험을 예측하고 대응 방안을 모색함</li> </ul> <p>[복지]</p> <p>주민위원회 센터 네트워크를 기반으로 맞춤형 복지사회 구현</p> <ul style="list-style-type: none"> <li>• 싱가포르 PA(People's Association)는 약1,800개의 주민위원회센터를 네트워크로 연결</li> <li>• PA는 서비스 만족도 제고를 위한 장기적 분석 및 시스템 구축</li> <li>• PA는 맞춤형 오피를 통한 개인별 서비스 활용 데이터를 약 25% 이상 확보하게 되고, 다양한 맞춤형 서비스 캠페인을 통해 2배 이상 가동률 향상</li> </ul>
네델란드	<p>데이터를 분석하여 소의 사육밀도를 높여 건강한 사육과 우유 생산량 증대에 기여</p> <ul style="list-style-type: none"> <li>• 소에 센서를 부착하여 소에 관한 정보를 실시간으로 확보</li> <li>• 기후 변화와 소의 상관관계 분석을 통해 외부 환경에 맞춘 농장 운영 방안 마련</li> <li>• 소의 움직임과 기후조건 등 1년 치의 소에 관한 약 200MB의 정보를 이용하여 소 사육에 적용</li> </ul>
덴마크	<p>풍력 발전에 필요한 터빈 관리 및 배치의 효율적 사용</p> <ul style="list-style-type: none"> <li>• IBM의 분석 솔루션과 '파이어스톰'슈퍼컴퓨터를 활용한 데이터 분석으로 터빈(동력장치)의 에너지 효율성 증대</li> <li>• 풍력터빈 및 풍력단지 설계를 위한 기상 및 지형데이터 분석시스템 구축을 통해 각 지형적 위치당 2.8 페타바이트에 이르는 공공 및 상업용 기상데이터를 분석하는데 수주가 걸리던 풍력예측정보 모델링 시간을 1시간 이내로 단축</li> </ul>
프랑스	<p>스마트폰의 마이크로부터 얻은 소음 정보를 종합해 소음지도(Noise Tube) 제작</p> <ul style="list-style-type: none"> <li>• 스마트폰 애플리케이션을 통해 각 지역별 실시간 소음정도 분석</li> <li>• 파리에 위치한 소니 컴퓨터과학 연구소는 스마트폰의 마이크 기능을</li> </ul>

	<p>사용하여 소음을 측정하는 응용 프로그램 개발</p> <ul style="list-style-type: none"> <li>스마트폰의 마이크와 GPS 정보를 종합하여 데이터 분석함으로써 실시간 데이터를 종합하여 새로운 정보 공유 가능</li> </ul>
캐나다	<p>미숙아 모니터링을 통한 감염 예방 및 예측(온타리오 공과대병원)</p> <ul style="list-style-type: none"> <li>미숙아는 병원균에 매우 민감하기 때문에 미숙아의 심장 박동 및 호흡 데이터를 이용하여 감염을 조기 판단(미숙아 모니터링 장비에서 생성되는 환자당 일9,000만 건 이상의 생리학 데이터스트림을 실시간으로 분석)</li> <li>병원(환자치료), 대학(데이터 분석), 기업(기술과 소프트웨어 지원)의 협력체계 구축을 통해 환자의 데이터를 공유하고, 개인정보보호를 위해 데이터 수집 윤리 적용</li> </ul>
케냐	<p>우샤히디, 집단지성으로 이루어진 재난관리 오픈소스 플랫폼</p> <ul style="list-style-type: none"> <li>우샤히디(Ushahidi)는 스와힐리어로 '증언'이라는 뜻으로 현재 세계 각국에서 발생하는 폭력, 테러, 자연재해 등을 실시간으로 정보 공유하는 시스템</li> <li>이메일, 트위터, 휴대폰 등으로 다양하게 취합된 재난현장정보를 웹 지도상에서 보여 주는 국민 참여형 플랫폼</li> <li>아이티 지진, 러시아 산불, 칠레 지진, 영국 지하철 파업 등 다양한 재난현장에서 활용되었으며, 신종플루, 쓰나미, 일본 지진발생 지역 등 전 세계의 사건·사고 정보 공유 가능</li> <li>문자, 이메일, 트위터 등 다양한 채널에서 수집된 정보를 실시간으로 시각화 하고, 인터넷 매핑 솔루션을 통해 상호 매핑을 가능하게 하는 오픈소스 플랫폼 구축</li> <li>우샤히디는 다양한 소스의 정보를 지도에 표시하는 오픈소스 플랫폼으로 방대한 실시간 정보를 단시간 안에 처리하기 위한 스마트 필터링 오픈소스 툴(Swift River), 단시간 안에 사업을 시작할 수 있도록 호스팅된 우샤히디 플랫폼(Crowdmap)등 3가지 서비스를 제공</li> <li>재난·재해, 사건·사고 등이 발생하면 관련 정보를 실시간으로 중계하는 웹사이트를 만들어 구조를 청하는 피해자들이 보내는 문자정보로 GPS 좌표를 설계</li> <li>IT기반 재난안전 플랫폼의 제공을 통해 국민들의 자발적인 참여에 의한 재난 상황 전파와 연계협력으로 안전을 관리하는 국민참여형 안전관리 구현</li> </ul>

독일	<p>독일연방 노동기구에서는 고용관련 빅데이터를 분석</p> <ul style="list-style-type: none"> <li>• 실업자 이력, 고용 중재, 구직기간, 고용주 등에 대한 방대한 정보를 분류, 분석하여 고용 중재에 활용(이와 같은 분석결과를 맞춤형 고용에 적용하여 3년간 백억유로의 비용절감효과 발생)</li> </ul>
일본	<p>센서데이터를 활용한 지능형 교통안내 시스템</p> <ul style="list-style-type: none"> <li>• 일본 전역 지정도시의 택시 약 11,000 대와 데이터 제공에 동의한 사용자로부터 실시간으로 교통정보 수집</li> <li>• GPS 데이터에서 자동차의 주행스피드를 계산하여 도로 교통정보를 예측한 후 사용자의 스마트폰으로 송신</li> <li>• 다양한 사용자에 의해 취득된 정보를 바탕으로한 실시간 교통 정보를 공유함으로써 최적의 교통 안내 서비스 가능</li> </ul>
한국	<p>고객 목소리 분석 시스템을 통한 서비스 혁신(한국도로공사)</p> <ul style="list-style-type: none"> <li>• 도로공사의 고객의 목소리(VOC, Voice of Customer) 분석을 통한 개선 활동</li> <li>• 콜 상담서비스, 민원관리 시스템, 채팅 상담 시스템을 고도화된 언어처리 기법으로 분석하여 고객만족활동에 도움이 될 수 있는 지표와 이슈 도출</li> <li>• 분석 주제를 기반으로 키워드와 토픽을 추출하여 유형별, 시간별 분석 실시</li> <li>• 연관도와 추이 분석, 토픽간의 연관성을 빠르고 쉽게 확인 가능</li> <li>• 정보통신 기술을 접목한 물 관리 시스템(한국수자원공사)</li> <li>• 수도관 수천 개를 하나의 네트워크 개념으로 통합하고 전체 네트워크가 효율적으로 운영되도록 각각의 수도관을 관리</li> <li>• 하나의 관제 컴퓨터가 시규모의 지역 수도관 정보 통합·관리</li> <li>• 수도관 중간에 유량, 수질, 유수율(물 공급량과 수도요금의 비율) 등을 관측하는 센서가 설치되어 있어 수도관의 정보를 실시간으로 관제 컴퓨터에 전송</li> <li>• 수도관 파손으로 누수가 발생하면 수도관에 설치된 센서에서 누수량감지, 누수위치 파악</li> <li>• 관제 컴퓨터가 사고 지점과 연결된 수도관을 파악하여 원격 자동감시 장치에 처리명령을 내려 수압 조절</li> <li>• 상수도관에 설치된 센서를 통해 실시간으로 물 관리가 가능하여 원격자동 감시장치를 통해 최적화된 물 관리 시스템 구현</li> </ul>

<p>찾아가는 서비스를 통한 맞춤형 고객관계 관리(근로복지공단)</p> <ul style="list-style-type: none"> <li>• 전문적·단계적인 산재보상 서비스를 제공하기 위한‘찾아가는 서비스’도입 후 고객 만족도 향상</li> <li>• 전문직으로 구성된 찾아가는 서비스팀을 통해 환자 중심의 서비스 제공 가능</li> <li>• 찾아가는 서비스 도입 후 산재환자에 대한 초기 접근이 빨라져 업무상 재해를 결정하는데 소요되던 기간이 19.3일에서 10.3일로 감소</li> <li>• 임금근로 일자리 통계로 일자리 현황 파악 지원(통계청)</li> <li>• 청장년층의 취업난과 실업 비정규직 고용불안 등 일자리에 대한 관심이 증가로 이에 맞춘 일자리 지표 마련</li> <li>• 고용보험, 국민연금, 건강보험 자료와 사업장 정보 파악을 위한 산재보험 자료를 활용하여 일자리 통계 지표 마련</li> <li>• 신규 일자리와 줄어든 일자리가 즉각 파악되어 일자리의 안정성과 변동성을 쉽게 측정 할 수 있는 지표 마련</li> </ul> <p>지자체 공무원들의 복지 행정 처리를 지원하는 정보 시스템(보건복지부)</p> <ul style="list-style-type: none"> <li>• 각종 사회복지 급여 및 서비스 지원 대상자의 자격 및 이력에 관한 정보를 통합관리</li> <li>• 복지대상자의 신청·접수·조사, 지원 여부결정, 급여지급, 사후관리 등 지자체의 복지업무 처리 지원</li> <li>• 복지급여 지급과정에서 지급내역의 임의수정을 통한 부정 소지 차단 및 실명 확인 후 입금으로 재정의 투명성 제고</li> <li>• 1회 방문 신청 시 개인 및 가족의 복지 요구를 종합적으로 파악하여 가능한 모든 서비스를 안내, 상담할 수 있는 시스템으로 발전</li> </ul>
--

## 感謝의 글

입학한지 엇그제 같은데 어느덧 졸업논문을 쓰게 되니 감회가 새롭습니다. 지난 2년의 시간 동안 저에게 도움을 주신 분들이 무척이나 많습니다. 많은 미흡하지만 논문을 마무리 하면서 감사의 인사를 드립니다.

특히, 밤낮으로 바쁘면서 세심하고 꼼꼼한 손길로 지도해주신 저의 지도교수님이신 신영란 교수님께 감사의 인사를 드리고 싶습니다. 또한 바쁘신 와중에서도 따뜻한 격려와 조언을 아끼지 않으신 신재영 교수님, 신용준 교수님께 머리숙여 감사드립니다.

직장 생활을 하면서 저의 논문에 자료 등에 협조해 주신 박민주 대리님, 홍정우 책임님, 백영주 과장님께 이 자리를 빌어 감사의 말씀을 전해 드립니다.

무엇보다도 대학원 다닌다고 가족에게 많이 소홀했는데 지원을 아끼지 않은 우리가족 백진규, 백주희 사랑하는 김성하 여사님에게도 감사의 인사를 드리며 늘 사랑합니다.

마지막으로 하늘에 계시고 늘 저를 응원해 주신 부모님 감사드립니다.

2017년 8월

백 용 주