

이 동 윤 한국생산기술연구원 IT융합공정그룹 수석연구원

| e-mail : dylee@kitech.re.kr

윤 주 성 한국생산기술연구원 스마트제조기술그룹 선임연구원

| e-mail : jsyoon@kitech.re.kr

이 성 근 ㈜브릭 상무

| e-mail : sglee@brique.co.kr

이 글에서는 스마트 팩토리 구현을 위해 고려되어야 하는 다양한 속성들 중에서 ‘데이터’에 대해서 논하고자 한다. 특히, 제조업의 관점에서 데이터의 가치를 평가하는 방법과 도메인 지식의 중요성에 대해 설명하였고, 중소기업에서 데이터를 활용하기 위한 방안을 함께 고민해보았다.

### 데이터의 활용과 가치

전세계적으로 데이터의 양이 폭발적으로 늘어나고 있다는 것에 대해서는 독자들 모두 특별한 이견이 없을 것이다. 맥킨지의 빅데이터 관련한 분석보고서에 따르면 이러한 데이터 중 2%가 제조 및 엔지니어링 분야에서 생성되었다고 한다. 2%라는 숫자가 컸어 보이지 않지만, 최근 회자되고 있는 스마트 팩토리를 실현하기 위해 제조분야 데이터의 중요성이 점차 높아지는 것을 고려해보면 향후 제조분야에서 발생하는 데이터의 양과 활용도가 급격하게 높아질 거라는 것은 누구나 예상할 수 있을 것이다. 우리나라에서도 많은 회사들이 이미 ERP시스템을 이용하여 사업 및 고객과 관련한 데이터를 수집하고 있으며 생산 분야에 사용하고 있다. 상기 데이터 이외에도 CAQ(Computer Aided Quality) 및 MES(Manufacturing Execution System)의 사용이 급속하게 확대되어 가는 추세이다. 게다가, 최근 정부에서 시행하고 있는 스마트 팩토리 보급사업을 통해서 중소기업들의 제조데이터 수집 및 활용도가 높아지고 있다. 그러나 데이터의 수집과 활용은 여전히 괴리가 있는 것이 현장의 시각

이다. 수집은 물리적으로 한계가 없는 것으로 생각해 볼 수 있으나, 활용의 측면에서 볼 때 수집된 데이터로부터 의미 있는 데이터를 추출하여 제조 현장에 활용하는 것은 수준 높은 분석 작업이 필요하기 때문에 어려운 작업이며 실제로도 데이터를 활용하고 있는 기업들의 수는 많지 않다.

데이터를 활용한다는 것은 데이터분석법(data analytics)을 이용해서 정보를 추출하고 의미를 부여하는 것이라고 할 수 있다. 데이터분석법에 대해서는 많은 방법이 소개되어 있기 때문에 여기서 설명하진 않겠다. 데이터분석은 네 가지로 나누어서 생각해 볼 수 있다. 설명(Descriptive), 진단(Diagnostic), 예측(Predictive), 마지막으로 전망(Perspective)이다. 설명은 가장 기본적인 것으로서 다양한 수학적 방법을 이용하여 과거를 분석하고 시각화하는 것이다. 어제 ‘무슨 일’이 일어났는지? 혹은 불량률이 ‘얼마나’ 높은지? 와 같은 질문에 대해 일반적으로 평균, 군집화(clustering) 및 경향분석(trend analysis)을 통해 답을 찾는 것으로 이해하면 될 것이다. 진단에서는 이러한 기초적 분석기법을 바탕으로 과거에 일어난 일들의 원인을 찾아내는 것이다. 즉, 불량률이 ‘왜’ 높아졌

는지 등에 대한 답을 찾는 것으로서, 대부분의 경우 변수들 간의 상관관계를 찾을 수 있다고 알려져 있다. 설명과 진단은 과거에 발생한 사건들을 소급해서 분석하는 것인데 반하여, 예측은 다음 대통령이 누가 될 것인가? 혹은, 현재 가공 중인 부품의 품질 수준 등과 같이 미래에 발생할 사건들을 예측하는 것이다. 가장 어렵다고 알려진 전망에서는 미래에 발생할 수 있는 일을 막기 위한 조치를 시시각각의 변화에 따라 적절하게 알려준다.

데이터의 가치는 데이터가 만들어내는 가치라고 이해할 수 있으며 B2C와 같은 개인적인 영역에서는 많은 연구가 이루어져 왔다. 즉, 데이터분석법이 소비자의 요구를 확인하고 기존에 없던 시장을 만들어 이익을 만들어낼 수 있으며, 이에 따라 데이터의 가치를 평가하는 모델들이 소개되었다. 예를 들어, 인터넷 사용자 영역에서는 사용자 맞춤형 광고를 제공하기 위하여 데이터가 사용되며, 이런 데이터의 가치는 맞춤형 광고를 통해 구입한 물건을 통해 광고주에게 추가적으로 발생한 이익을 맞춤형 광고를 받은 사람들의 수로 나눈 값을 넘지는 못할 것이다. 데이터의 가치를 정량화하기 위하여 접근성, 정확성, 유연성, 품질, 타당성, 재사용성, 공유성, 시의적절성 등의 다양한 가치지배인자들이 소개되어왔으며, 이 글과 직접적인 관련이 없기 때문에 자세한 설명은 생략하였다. (Engelsman, W., "Information Assets and their Value," January 2017 & Garifova, L. F., "Infonomics and The Value of Information in The Digital Economy," Procedia Economics and Finance 23, pp. 738~743, 2015)

### 제조업에서 데이터의 가치

데이터의 가격(가치)은 데이터를 수집하고 재처리한

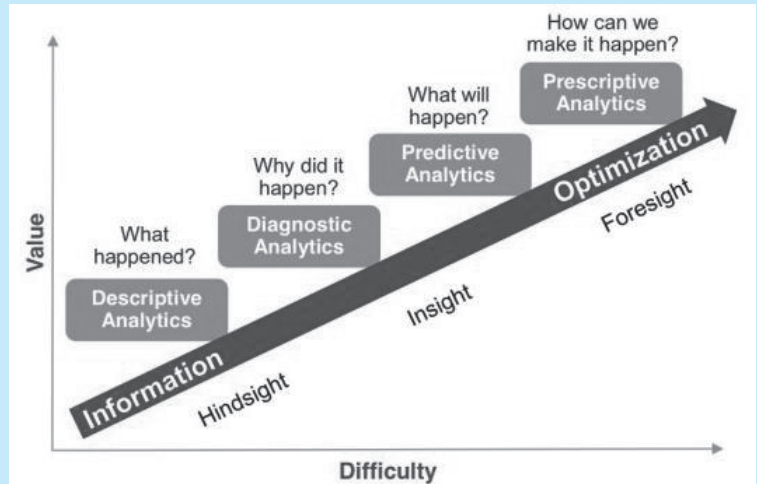
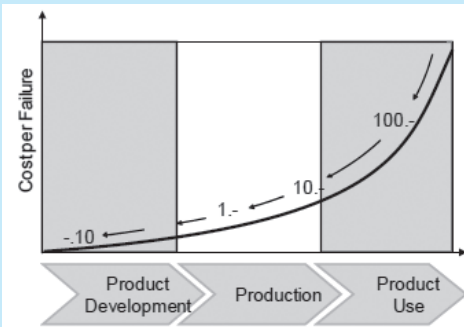


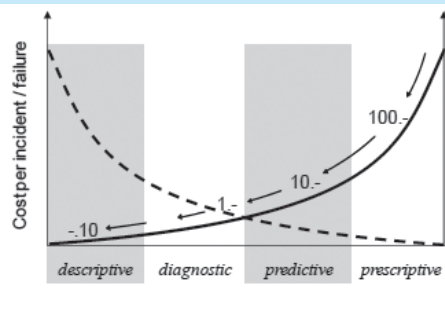
그림 1 데이터분석법의 난이도와 가치(Source: Gartner, March 2012)

후 분석하는 비용과 그로 인해 시장에서 추가적으로 발생하는 이윤에 의해서 결정될 것이다. 제조업에서는 대부분의 데이터분석이 공정최적화를 목표로 해왔다. 다시 말해, 손실을 최소화하거나 효율과 양품율을 극대화함으로써 가치를 만들어내는 것이다. 데이터를 통하여 공정에 대한 이해도를 높이고 안정적인 공정변수 영역이나 최적화된 운영지표를 설정할 수 있다. 장비 및 공정 데이터가 가치를 만들어 내는 단계에서 중요할 것이라는 점에는 대부분 공감할 수 있을 것이다.

제조업의 관점에서 바라보는 데이터의 가치를 고려할 때 '반응속도'가 중요한 요인이 될 수 있다. 즉, 진단단계에서는 왜 특정한 공정이 제대로 동작하지 않았고 그것을 어떻게 개선할 수 있는지를 알려주는 데 며칠 혹은 몇 주의 시간이 소요될 수 있다. 그 동안 손실은 계속해서 누적될 것이고, 이때 사용되는 데이터의 가치는 낮을 것이다. 이에 반하여, 예측단계에서는 실시간으로 잘못된 공정을 검출하기 때문에 잘못 가공된 부품을 생산라인에서 사전에 제거하거나 일정 범위를 넘어서는 공정을 중단시킴으로써 손실비용을 최소화할 수 있다. 즉, 예측단계에서 사용되는 데이터의 가치는 진단단계에서 사용되는 데이터보다 가치가 높다고 할 수 있겠다. 예방단계에서는 불량발



**그림 2** 손실비용 10배 증가의 법칙(Plutz, M. et al., "Method for the Monetary Quantification of Failure Risks in Production," Safety, Reliability and Risk Analysis: Beyond the Horizon, CRC Press, London, ISBN: 978-1-138-00123-7, pp. 492-499)



**그림 3** 제조데이터의 가치

생을 사전에 막기 위한 조치를 취하게 함으로써 데이터의 가치를 극대화하는 것이 가능할 것이다. 다시 말해, 제조업에서 데이터의 가치는 그 데이터가 각 데이터분석법의 어느 단계에 적용할 수 있는지, 즉 데이터의 반응 속도에 의해서 결정된다고 할 수 있다.

품질관리분야에 언급되는 실패비용을 동일하게 적용할 수는 없지만 비유해서 생각해볼 수 있을 것이다. 기술적 위험을 관리함에 있어서, 실패로 인한 손해는 그림 2에서와 같이 제품의 개발, 제조, 사용으로 갈수록 10배씩 증가하는 것으로 생각해볼 수 있다. 진단에서 예측, 예방이라는 단계를 상기 품질관리분야에 적용해 보면 이해가 쉬울 것이다. 독자들의 이해를 돕기 위해 절삭가공을 예를 들어 설명을 해보겠다. 터빈블레이드는 상당히 고가의 제품으로 중간크기 제품의 가격을 개당 1,500만 원으로 일단 생각해보겠다. 가공 후 측정단계에서 불량 발생을 확인한 경우라면 최대 1,500만 원을 손실비용으로 생각해 볼 수 있으며, 수리가 가능한 경우라면 가격의 10%인 150만 원의 손실을 가정해 볼 수 있다. 이 경우 추가적인 장비 및 작업자의 투입과 함께 엔지니어링 비용이 손실비용에 포함된다. 반면, 가공 중에 불량을 예지하고 미리 조치를 취한다면, - 이 경우 절삭속도 및 이송속도를 늦추는

방법을 생각해 볼 수 있다.  
- 손실비용이 추가적으로 발생하는 가공시간만을 고려하면 되므로 150만 원보다는 훨씬 줄어들 것으로 예상해 볼 수 있다. 각각의 단계에서는 서로 다른 데이터분석법이 사용되기 때문에, 그 데이터분석에 사용

된 데이터의 가치는 그림 3과 같이 서로 달라진다. 정리해서 말하면, 가공 중에 발생하는 불량을 사전에 인지하여 조치를 취해서 불량발생을 회피해 손실비용을 크게 줄인다면, 이때(전망단계) 사용되는 데이터의 가치는 높다고 할 수 있다. 반면, 가공 후 수리를 결정할 때(예측 및 진단단계)에서 사용되는 데이터의 가치는 상대적으로 낮다고 할 수 있으며, 폐기를 결정할 때(설명단계)에서 사용되는 데이터의 가치는 큰 손실비용을 고려할 때 가장 낮다고 할 수 있다.

### 데이터분석보다 중요한 현장지식(Domain knowledge)

이번 절에서는 스마트 팩토리를 구축하면서 간과하기 쉬운 점들에 대해서 논의하고자 한다. RAMI 4.0(그림 4)의 구조에서 보듯이 스마트 팩토리의 구성 흐름은 설비에서 발생하는 데이터에서 시작해 정보의 추출과 분석, 그리고 비즈니스적 판단에까지 이르게 된다. 즉, 스마트 팩토리에서의 정보화는 설비와 기기들을 연결하여 데이터를 수집, 저장하고 이 데이터를 목적에 맞게끔 분석하여 의미 있는 결과를 도출하고, 도출한 결과를 비즈니스 환경과 결부하여 최종적인 비즈니스 인사이트를 얻는 과정이다. 결국 얻고자 하는 비즈니스 인사이트에 따라 획득하는 데이터의 종류, IT 인프라 시스템, 데이터 분석 방법에도 차이가 있을 수 있다. 특히 대규모 스마트 팩토리에서 발생하는 데이터는 주로

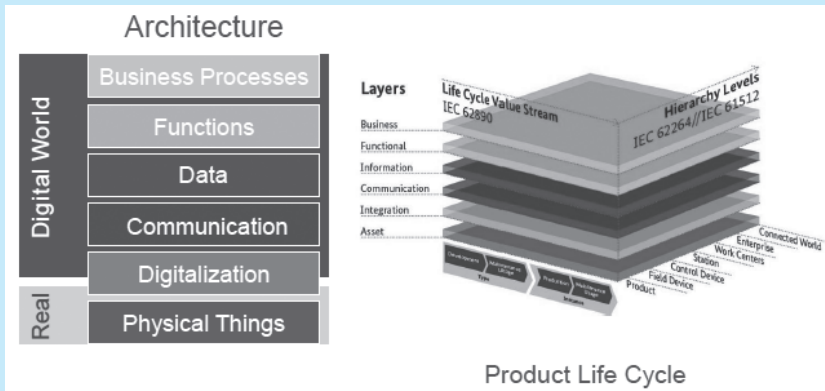


그림 4 Reference Architectural Model Industrie 4.0(Source: Platform Industry 4.0, April 2016)

시계열 데이터이기 때문에 데이터 수집, 저장에 관한 IT 인프라 시스템의 구축비용은 데이터 양에 따라 기하급수적으로 증가할 수 있다. 예를 들어 미세 공정 제어가 필요하고 수익 규모도 거대한 반도체 제조에서는 페타바이트(Peta Byte) 단위의 데이터를 고가의 관계형 데이터베이스에 저장하더라도 투자 대비 원하는 기대효과를 얻을 수 있다. 하지만, 단순한 설비 가동률만 필요한 수십 대의 생산 설비를 갖춘 중소 제조업체에서 초기 투자에서부터 각종 시계열 데이터를 수집하는 시스템을 도입할 경우 투자 대비 효과는 매우 낮을 수 있다. 이러한 예는 빅데이터시스템을 도입하고 데이터는 확보하였지만 의미 있는 결과를 얻지 못하고 시스템 도입 실패 사례를 남기는 것과 일맥상통하다고 볼 수 있다. 이 부분에서 강조하고 싶은 것은 스마트 팩토리 구축 시 IT 및 데이터 분석에 관한 전문 지식뿐만 아니라 도메인 엔지니어의 역할이 매우 중요하다는 것이다. 즉, 도메인 엔지니어가 연고자 하는 비즈니스 인사이트를 명확히 정의하고 이를 위한 시스템 구성 및 데이터 분석 방법론에 관해 IT 및 데이터 사이언스 엔지니어들과 논의하여 최적안을 도출하여야 한다.

스마트 팩토리에서 효율적인 데이터 수집은 이러한 노력의 첫 걸음이다. 빅데이터의 주요 특징은 3V(Volume, Variety, Velocity)로 정의되는데 스마트 팩토리에서 발생하는 데이터를 이것으로 표현해보면

다음과 같을 것이다. 스마트 팩토리의 생산 설비에는 많은 센서가 존재하는데 어떤 센서 데이터를 수집할 것인가의 문제는 Variety로 표현할 수 있고, 센서 데이터에서 수집하는 데이터의 분해능은 Velocity로 표현할 수 있을 것이다. 그리고 이 두 변수의 조합에 따라 최종적으로 수집하는 데이터의 양, Volume은 결정될 것이다. 반도체나 디스플레이 생산시스템에서는 1초 정도의 분해능으로 센서

데이터를 대부분 수집하고 최근에는 공정의 미세 변동을 감지하기 위하여 0.1초 단위의 데이터를 수집하기도 한다. 절삭 가공 분야는 반도체 공정만큼 미세 공정 제어는 없지만 가공 공정의 특성 상 공정을 모니터링 하는 센서의 분해능은 높다. 예를 들어 채터를 감지하기 위한 진동센서는 10kHz 이상의 데이터를 수집하는데 이를 끊임없이 수집한다면 저장해야 할 데이터의 양은 매우 커질 것이고 시스템은 비효율적으로 운영될 것이다. 그렇다고 시간 간격을 두고 데이터를 수집하게 되면 감지 자체를 하지 못하는 경우가 발생하게 된다. 따라서 이러한 데이터는 크기가 축소된 형태의 유의미한 데이터로 변환하고 이를 저장하고 분석하는 것이 올바른 방향일 것이다. 그리고 변환된 데이터가 원 데이터의 특성을 설명할 능력이 된다면 원 데이터는 저장하지 않고 변환 데이터만 저장하는 것도 좋은 방법이 될 수 있다. 예를 들어 대량의 데이터를 빠르게 연산할 수 있는 기기와 데이터 처리 기술을 통해 IT 인프라 규모를 축소시킬 수 있다. 또한, 제철, 정유 등 특정한 산업 분야에 있어서는 데이터의 분해능보다 유의미한 정보를 획득할 수 있는 센서의 확보나 배치가 더욱 중요할 수 있다. 즉, 도메인을 중심으로 각 분야 엔지니어링 활동을 통해 투자비용을 감소시키고 설명력이 높은 데이터를 수집함으로써 최종적인 비즈니스 영역에서의 스마트 팩토리 활용도를 높일 수 있다.



## 중소기업에서의 데이터 활용성 제고 방안

스마트 팩토리는 이미 모든 제조 기업들에게 선택이 아닌 필수가 되었다. 앞 장에서 설명한 데이터의 가치 및 도메인 중요성을 다시 한 번 정리하면서 특히 중소기업의 입장에서 고려할 내용들을 살펴보고자 하겠다. 제조 현장에서 발생하는 다양한 데이터들은 다양한 방식으로 가공 및 처리되어 제조기업의 의사결정을 지원하는 중요한 정보로 활용된다. 그러나 실제로 제조기업 현업에서는 인프라 부족으로 이러한 데이터를 육안으로 확인하거나 일부만 수기로 작성하는 수준으로 관리하여 데이터 수집 자체가 어려운 경우가 대다수이고, 데이터를 수집하여 모아 두고 있더라도 분석 기법이나 활용 방안이 마련되어 있지 않아 어디에 어떻게 활용할지 모르는 경우도 많다. 이러한 데이터 활용의 한계는 규모가 영세하고 현장이 열악한 중소기업으로 갈수록 더욱 심각한 실정이다. 특히 중소 제조기업에서는 데이터 활용을 위해 ICT 인프라를 구축하고 전문 인력을 투입하는 것이 부가적인 투자를 필요로 하는 것으로 인식되어 큰 비용부담으로 작용한다. 이러한 데이터 수집, 데이터 활용, ICT 인프라에 대한 문제 및 한계점을 보완하여 중소기업의 데이터 활용성을 제고하는 방안을 좀더 구체적으로 살펴보자.

먼저, 시간과 비용을 고려하여 효과적인 데이터 수집을 위해서는 활용처를 기준으로 데이터 수집 포인트를 잡는 것이 중요하다. 어디에 어떻게 사용할 것인지 고려하지 않고 수집하는 데이터는 의미 없이 쌓이다가 비용적인 부담을 가중시키는 골칫거리가 되기 일쑤이다. 또한 데이터의 활용처와 현업의 상황에 맞게 자동적인 수집 방식과 수동 방식을 적절하게 조합시키는 것도 필요하다. 생산장비의 컨트롤러나 PLC 점접 인터페이스, 다양한 센서모듈 연결, 별도의 측정 장치 설치, 키오스크나 터치스크린을 통한 사용자 입력 등 여러 가지 방식들 중에서 적합한 것을 선정하는

과정이 매우 중요하며, 다양한 방식의 데이터 수집 인프라와 표준화된 인터페이스를 통한 통신 네트워크가 구성되어야 한다. 그리고 고도화된 데이터 활용을 위해서는 품질불량이나 설비이상과 같이 제조공정의 컨텍스트를 파악할 수 있는 정보들이 확보되어야 한다.

다음으로 데이터 활용 관점에서는 해결하고자 하는 문제나 활용하고자 하는 업무에 따라 적합한 엔지니어링 모델과 데이터 분석 기법을 적용하는 것이 중요하다. 특히 데이터 분석 기법은 관련 전문지식이나 지원 도구 활용이 필수적이기 때문에 중소기업 현업에서 대응하기 어렵고 제조분야에 적용된 성공사례가 충분하지 않아 제조문제에 특화된 데이터 분석 기법이 개발 및 검증되어야 한다. 품질불량 판정, 설비고장 진단, 공구수명 예측 등 특정 제조문제에 좋은 성능을 보이는 데이터 분석 기법을 중소기업 현업에서 활용 가능한 수준의 소프트웨어 지원 도구 형태로 개발하거나 서비스 형태로 지원하는 것이 필요하다.

마지막으로, ICT 인프라 문제는 국내 중소기업의 현실을 고려해야 하는 매우 중요한 이슈이다. 데이터 활용을 위해서는 소프트웨어 시스템 및 관련 하드웨어 장비·장치의 도입이 뒤따르기 마련이며, 이는 도입 비용과 운영 인력의 문제와 직결된다. 하나의 대안으로 클라우드 서비스를 고려할 수 있으며, 클라우드 서비스는 사용 기반의 과금으로 초기 인프라 투자 비용이 절감될 수 있고 시스템 운영 및 유지보수 측면에서 장점을 보이기 때문에 중소 제조기업을 위한 데이터 활용 서비스로 하나의 모델이 될 수 있다. 클라우드 서비스에 대한 데이터 보안이나 시스템 접근성 문제로 서버만 관리해주고 사용기업이 원하거나 여건이 갖추어졌을 때 서버를 그대로 기업 내로 이전할 수 있는 코로케이션 서비스를 활용할 수도 있다. 실시간 데이터의 네트워크 부하 문제와 데이터 보안에 대한 염려가 극복된다면 클라우드 서비스 방식이 보다 효과적인 것은 분명하다.

[주] 본 기고문 중 제조데이터의 가치와 관련한 내용은 2017년 캐나다 몬트리올에서 개최된 국제 학술대회 VMPT 2017에서 발표된 "Determining the value of data - a manufacturing perspective(R.H. Schmitt, et. al.)"에서 발췌하여 요약하였다.