

사회경제발전의 중요자원인 대용량자료와 그 활용의 세계적추세

허 광 립

위대한 령도자 김정일동지께서는 다음과 같이 교시하시였다.

《우리는 21세기의 발전추세에 맞게 혁명투쟁과 건설사업을 벌려야 하며 새 세기에 맞먹는 사회주의강성대국을 건설해야 합니다. 그러자면 모든 문제를 새로운 관점에서 보고 새로운 높이에서 창조적으로, 우리 식으로 독특하게 풀어나가야 합니다.》(《김정일선집》 증보판 제 21권 73페이지)

오늘날 정보기술의 급속한 발전은 사람들의 생활방식을 끊임없이 새롭게 변모시키고있다.

정보기술이 빠른 속도로 발전하고 그것이 사람들의 일상생활에 적극 활용됨에 따라 사람들은 언제 어디서나 지능손전화와 판형컴퓨터, 휴대용컴퓨터와 같은 다양한 정보설비들을 리용하고있으며 이러한 정보설비들에 의하여 매 순간순간 방대한 규모의 자료가 끊임없이 생성되고있다.

사람들은 매일, 매 순간 지능손전화를 사용하면서 다른 사람들과 대화를 나누고 통보문을 주고받고있으며 사진도 찍고 촬영도 한다. 그리고 임의의 순간에 인터넷검색도 하면서 원하는 정보들을 언제 어디서나 찾아볼수 있다. 통계자료에 의하면 지구상의 사람들이 주고받는 Twitter통보문의 수는 하루에 1억건, 수자식사진수는 2억건에 달한다고 한다. 지리적국경을 초월하여 전지구가 하나로 연결된 인터넷상에도 매일같이 수많은 자료가 추가되고있다. 사람들이 수자식정보설비들을 사용하면서부터는 매일 상품구매나 전자결제, 인터넷검색 등 무엇인가를 할 때마다 그 흔적이 자료로 남는다. 거리와 건물에 설치되어있는 CCD카메라는 도로상태와 수많은 사람들과 차량들의 움직임 등을 실시간적으로 기록하며 각종 형태의 정보설비들과 수감부들도 실시간적으로 자료를 생성하고있다. 이와 함께 모든 형태의 국가관리나 기업활동에서도 정보설비가 리용되면서 매일, 매 순간 많은 자료가 축적되고있다.

이렇게 다양한 정보설비들의 리용이 급격히 늘어나면서 그로부터 실시간적으로 생성되는 자료의 규모는 세계적범위에서 2010년에는 0.9ZB, 2011년에는 1.7ZB, 2012년에는 2.8ZB로서 놀라운 속도로 증가하고있다. 지금까지 인류가 축적해온 모든 자료 가운데서 90%가 최근 2년(2011-2012년)동안에 생성된것이다.

앞으로 정보설비의 리용이 더욱 확대되면서 2020년경에 가서는 임의의 장소에서 인터넷과 무선연결이 가능한 말단기(손전화, 판형컴퓨터 등)의 리용개수가 세계적으로 7조개로 늘어나고 그로부터 생성되는 정보의 량은 지금의 50배이상 증가할것으로 예측하고있다.

이러한 현실을 반영하여 새롭게 출현한 개념이 바로 대용량자료이다. 대용량자료는 기존의 자료기저관리방법으로써는 처리하기 어려운 실시간적으로 끊임없이 생성되는 방대한 규모의 자료를 가리켜 이르는 말이다.

정보기술이 지금처럼 발전되지 못하였던 지난 시기에는 얻을수 없었고 또 그냥 흘려보내던 수많은 움직임들이 오늘날에는 발전된 정보기술과 설비들에 의하여 매 순간순간 모두 자료로 바뀌고있다.

이렇게 얻어지는 자료는 그 규모가 대단히 방대하고 지난 시기처럼 주어진 양식에 따라 형식화(또는 구조화)된 형태로뿐만아니라 문자나 음성, 사진이나 동화상 등과 같이 형식화되지 않은 다양한 형태로 주어진다.

자료에 의하면 2015년에 가서는 총체적인 자료가운데서 형식화되지 않은 자료가 차지하는 몫이 80~90%에 달할것으로 보고있다.

이러한 다양한 형태의 방대한 자료속에는 사람들의 기호와 취미는 무엇이며 매일 무엇을 얼마나 소비하는가, 무엇을 생각하고 무엇을 바라고있는가 하는것부터 시작하여 사람들이 살며 생활하는 지구상의 현재는 어떻게 변하고있으며 앞으로는 어떻게 변화발전 하겠는가에 이르기까지 많은 문제들에 해답을 줄수 있는 수많은 내용들이 담겨져있다.

오늘날 대용량자료는 빠른 속도로 변화발전하는 세계를 파악할수 있게 하는 매우 가치있는 중요한 핵심자원으로 되고있다.

지금 세계의 많은 기업들과 나라들에서는 대용량자료를 새로운 가치를 창조하고 국가를 발전시키기 위한 위력한 수단으로 적극 활용하고있다.

대용량자료의 리용가치와 그 활용이 가지는 중요성이 날로 높아짐에 따라 오늘날 세계의 많은 나라들에서는 발전하는 정보기술을 능동적으로 활용하여 필요한 대용량자료를 실시간적으로 생성하고 분석, 활용하고있다.

대용량자료활용의 세계적추세는 무엇보다먼저 대용량자료의 수집과 분석을 위한 기술개발과 인재육성사업을 적극 추진시키고있는것이다.

대용량자료는 그 규모의 방대성과 자료생성의 실시간성, 형태의 다양성을 그 특성으로 하는것으로 하여 지난 시기의 자료처리를 위한 체계와 방법으로는 대용량자료의 수집과 분석, 처리를 실현하기 어렵다.

대용량의 구조화 또는 비구조화된 자료를 저장, 분석하기 위한 계산체계인 Hadoop과 다양한 형태의 자료들을 통합적으로 리용할수 있게 하는 자료가상화기술, 저장, 분석기술인 BI(Business Intelligence)기술은 대용량자료의 수집, 분석, 처리를 위한 핵심기술로 되고있다.

대용량자료분석에서의 인공지능기술의 리용은 기존의 통계적분석방법의 제한성을 극복하고 실시간적으로 대량생성되는 대용량자료의 분석을 통하여 현실변화를 제때에 정확히 파악할수 있게 한다.

IBM, Google, Apple, Oracle, EMC, HP 등과 같은 정보기술기업들은 이 분야에서 주도권을 잡고 경쟁력을 높이기 위하여 대용량자료의 수집과 분석, 활용을 위한 다양한 기술의 개발, 대용량자료관련 정보봉사를 제공하는 등의 방법으로 시장을 확대해나가고있다.

방대한 량의 자료수집과 분석이 중요한 가치를 가지게 되면서 현시기 이 부문 전문가양성은 지식경제시대의 승패를 좌우하는 관건적문제로 제기되며 구조화되지 않은 대규모의 자료속에 숨겨진 보다 가치있는 정보, 지식을 찾아내는 능력을 가진 자료분석전문가는 현시대의 최고의 인재로 평가되고있다.

자료분석전문가가 되기 위해서는 그에 필요한 기본자질인 수학, 공학능력과 자료분석에 필수적인 가설을 세우거나 검증하는데 필요한 다양한 각도의 시각과 글쓰기능력, 다른 사람을 설득시킬수 있는 대화능력 등을 소유하여야 한다. 자료분석전문가들에게 요구되는 이러한 능력들을 한사람이 다 갖추기는 매우 어려운 일이다. 따라서 다양한 분야의 전문지식을 가진 사람들로 연구집단을 구성하고 그 성원들간의 원활한 의사소통, 의견교환 등

으로 제기되는 문제를 해결해나가야 한다. 연구집단에는 컴퓨터과학자, 공학자, 자연과학자, 기업경영전문가, 수학자, 인문학자, 정보관리전문가, 통계학과 경제학과 같은 사회과학전문가들과 기타 부문의 전문가들이 망라될수 있다.

대용량자료활용의 세계적추세는 다음으로 다양한 형태의 대용량의 자료원천들을 효과적으로 활용할수 있는 실시간적이며 종합적인 정보체제를 수립하고있는것이다.

정보기술이 발전하고 그에 따라 다양한 정보설비들의 리용과 각종 형태의 정보봉사들이 늘어남에 따라 리용가능한 자료원천들이 늘어나고있으며 기업관리와 국가관리에서는 요구되는 정보를 얻는데 필요한 다양한 자료원천을 적극 찾아내어 활용하는것이 중요한 문제로 나서고있다.

오늘날 정보기술관련 회사들뿐아니라 여러 부문의 일반기업들에서도 대용량자료를 기업관리의 중요한 자원으로 인식하고 리용가능한 기업내부와 외부의 자료를 충분히 수집, 분석, 활용함으로써 수익규모를 획기적으로 높여나가고있다.

소비품의 류통 및 소매를 위주로 하는 많은 기업들에서는 세계 각국의 수많은 봉사망들에서 기록되는 판매량이나 카드결제정보를 통한 소비자의 구매실적을 실시간적으로 집계하여 분석함으로써 불필요한 상품체화손실을 줄이고 소비자들이 원하는 상품과 수량을 제때에 보장함으로써 수익을 높이고있다.

실시간적으로 수집되는 판매 및 구매기록과 같은 기업내부정보와 함께 SNS(Social Network Service)자료나 인터넷검색기록 등을 통하여 얻어지는 외부정보를 분석하고자 자기 회사제품에 대한 사람들의 평가, 소비자의 기호와 취미, 관심 등을 파악하며 사람별로 그 특성에 맞는 정보봉사를 제공하는 방법으로 적극적인 판매활동을 벌리고있다.

지난 시기에는 기업이 새로운 제품을 개발하여 판매하면 1년이나 2년후에야 제품에 대한 소비자의 평가를 받아 제품을 다시 수정할수 있었지만 이제는 SNS자료 등을 적극 활용하여 제품의 판매와 함께 즉시적으로 분석하여 제품개선에 반영할수 있게 되었다.

이와 함께 많은 제조기업들에서는 생산공정의 정보화수준이 높아짐에 따라 그로부터 실시간적으로 생성되는 각종 정보들과 기상정보, 지리정보, 위성정보, 각종 형태의 수감부정보 등을 종합적으로 수집, 분석하여 최량적인 생산조건을 마련하여 원가를 낮추고 수익성을 높여나가고있다.

어느 한 나라의 보건부문에서는 오랜 기간 축적된 림상치료경험과 약물작용에 대한 대용량의 자료를 누구나 리용할수 있게 하는 의료정보봉사를 제공함으로써 사람들이 어떤 약에 관심을 가지는가, 어느 지역의 사람들이 어떤 병을 앓고있는가, 해당 지역이나 나라에서의 질병들의 분포정도, 해당 질병의 전파가능성 등 보건상태와 관련한 필요한 많은 문제들을 파악하고있으며 이 자료들은 질병치료와 앞으로의 보건정책을 세우는데 적극 활용되고있다.

이외에도 각종 수감부로부터 실시간적으로 전송되는 자료를 활용한 지능형 교통안내체제의 운영과 축적된 기상자료와 위성, 선박, 항공기, 부표, 수감부 등에서 매일 35억건 이상 수집되는 자료를 분석한 실시간적인 기상정보의 제공, 운전시 룰전기재의 흔들림을 감시할수 있는 수감부로부터 얻는 자료를 통한 도로상태확인파 합리적인 보수유지봉사 등을 비롯하여 사회안전과 보건, 체육, 교통, 날씨, 물자원과 전력자원의 효과적리용 등을 위한 여러가지 목적에 대용량자료가 적극 활용됨으로써 사람들의 일상생활에서의 편리성이 보장되고있다.

대용량자료활용의 세계적추세는 다음으로 대용량자료의 실시간적인 수집과 분석을 통하여 끊임없이 변화발전하는 현실을 옹계 파악하고 그에 과학적으로, 전략적으로 대응할수 있는 기업관리와 국가관리체계를 수립하고있는것이다.

현실에서 끊임없이 생성되는 대용량의 자료를 분석하여 앞으로 제기될수 있는 문제들을 찾고 그에 전략적으로 대응하는것은 기업이나 국가에 있어서 지속가능한 발전을 위한 필수적인 문제로 된다.

세계 여러 나라들에서는 주변환경변화에 대한 대용량의 각종 자료들을 실시간적으로 수집하고 첨단분석리론 및 방법을 리용하여 기업이나 국가발전에 큰 영향을 미칠수 있는 미래의 발생가능한 문제들을 분석하며 그에 따르는 중장기대응전략을 마련하기 위한 사업에 대용량자료관련 기술과 방법을 적극 활용하고있다.

유럽동맹은 기술, 경제, 사회체계의 복잡성 및 호상의존성이 강화됨에 따라 새로운 안목을 가지고 세계의 변화발전에 대처하기 위하여 “FutureI CT” 와 “iknow” 라는 전략계획을 추진하고있다.

이 계획은 미래연구를 위해 다양한 분야의 사람들을 적극 참가시키고 서로의 방대한 자료와 지식을 공유함으로써 앞으로 발생할 다양한 문제들을 발견하고 다가올 위험에 효과적으로 대응하며 새로운 기회를 리용할수 있는 합리적이고 전략적인 방안을 제시하는 것을 목적으로 한다.

이를 위하여 지구신경망체계를 구축하고 이로부터 수집되는 방대한 수감정보를 리용하여 세계적인 사회, 경제, 환경 및 기술분야에 대한 대량의 자료를 실시간적으로 수집하며 그 자료와 전지구모의체계를 리용하여 현실세계를 모형화하기 위한 사업을 추진하고있다.

이와 함께 다른 나라들에서도 현재 나라의 안전과 발전을 위하여 필요한 대용량자료의 수집과 분석, 활용사업을 정부적인 사업으로 적극 추진시켜나가고있다.

이와 같이 오늘날 세계의 많은 나라와 기업들은 지속가능하며 경쟁력있는 사회경제발전을 이룩하기 위하여 대용량자료를 적극 활용하고있으며 현실은 이 문제가 얼마나 절박하고 중요한 문제인가를 잘 보여주고있다.

우리는 현시기 나라의 경제를 비약적으로 발전시키는데서 대용량자료의 활용이 노는 역할에 대하여 옹계 인식하고 나라의 정보기술을 적극 발전시키며 그것을 경제관리에 적극 활용해나감으로써 이 땅우에 사회주의강성국가를 반드시 일떠세워야 할것이다.