

한국고용정보원 고용정보통합분석시스템 구축사례

Independent DW Appliance System Insights and Reports

Focus on: IBM 'Netezza 1000' and DW Efficiency

데일리그리드 리서치 센터

소프트웨어그룹

2012. 07.

Sponsored By :



“대용량 연계분석으로, 대국민 서비스 날개 달다”

복합적인 주제 다차원 분석 실현...고용정책 결정에 핵심정보 제공

DW 어플라이언스로 IBM 네티자 적용...데이터 조회속도 10배 향상



분석기법을 이용해 적시에 필요한 고용 통계 정보를 사용자들에게 제공한다. 정책 결정자는 의사결정에 DW 정보를 적극 활용할 수 있고, 연구자나 현업 사용자는 시스템에서 필요한 노동시장 관련 정보를 직접 추출할 수 있다.

모두 11개 전산망을 보유하고 있는 고용정보원. 이 중 정보원을 대표하는 시스템은 고용보험전산망, 워크넷, 직업훈련전산망, 외국인고용관리전산망 등 4개다. 이 4개의 시스템에 차례로 도입된 DW시스템은, 2000년대 초반에 준비를 거쳐 2000년대 중반부터 본격적으로 도입됐다. 뿌리 데이터에 해당하는 고용보험시스템을 시작으로 워크넷, 직업훈련, 외국인 등으로 도입을 확대했으며, 3년 전 도입을 완료했다. 이후 운영 노하우를 꾸준히 쌓아가고 있다.

데이터 품질 개선해 고품격 서비스 제공

고용노동시장의 이슈에 대응하고 고용정책용 현황정보를 적기에 제공하기 위해 한국고용정보원은 지난 2009년 고용정보통합분석시스템(DW)을 구축했다. 데이터 품질을 높여 대국민 서비스를 개선하는 것도 DW구축의 배경이 됐다.

고용정보통합분석시스템은 이름 그대로 고용 관련 정보를 통합해 보관하고 있는 거대한 정보 창고다. 이 시스템은 각종 통계 및 다양한

DW가 취급하는 데이터 양은 실로 방대하다. 운영계서 만드는 데이터 양이 많기 때문이다. 1998년 처음 서비스를 시작한 구직구인정보를 서비스하는 공공 취업포털 워크넷(work.go.kr)의 하루 평균 이용자수는 2011년 10월 기준으로 35만 명에 달한다. 또한 워크넷을 이용하는 유효구직인원은 665,780명이고, 유효구인인원은 305,266명(75,278개 기업)이다.

고용보험전산망은 고용보험관련 행정업무(피보험자, 실업급여, 모성보호, 고용안정 및 직업능력개발사업 등)를 처리하는 전산시스템으로, 지난해 기준으로 하루 평균 외부 이용자수는 4만 7,000명. 내부망 이용자수(고용센터, 근로복지공단 등 고용보험시스템 이용자)도 4,896명에 달한다. 직업훈련전산망과 외국인고용관리전산망에서 쏟아내는 데이터의 양도 두 시스템에 못지 않게 많다.

4개 시스템에 담긴 방대한 양의 빅데이터 중 핵심 데이터인 주요 고객정보와 중요 업무데이터만 DW에 모았다. 데이터의 양은 매우 방대하다고 한다.

운영계에선 원 자료나 목록 등 마이크로 데이터를 DW에 제공하는 역할을 하지만, DW에선 사용자들에게 매크로 데이터를 공급한다. 주로 집계성 데이터인데, 특정 항목에 관련된 정보를 한데 모아서 보여준다. 예를 들면 “서울시

00구엔 00나라에서 입국한 외국인이 00명 살고 있다”와 같은 데이터 유형이다.

이런 집계성 데이터를 다양한 기관이 공유하고 있다. 먼저 고용정보원의 내부 사용자가 필요한 업무에 이용한다. 또한 고용 관련 업무를 취급하는 고용노동부는 물론 근로복지공단이나 한국산업인력공단 등이 이 데이터를 자주 쓴다.

DW 정보는 철저한 보안 속에 관리되고 있다. ID와 PW를 받은 사용자만 시스템에 접근할 수 있고, 사용자들은 개인정보보호 규칙에 따른 통제를 받는다.

도입 전 상황은 어땠을까? 정보를 추출하고 이용하는데 이런 저런 불편을 겪었지만, 특히 연계분석이 어려웠다고 한다. 이전엔 하나의 이슈를 파악하기 위한 정보를 개별 운영계 시스템에 별도로 요청했는데, 최종 결과값이 나오기까지 시간이 너무 지체돼 제때에 정보를 활용할 수 없었다.

고용정보통합분석시스템(DW)의 주요 서비스 내용

- **연계분석:** 고용보험, 취업알선(워크넷), 직업훈련 등의 다양한 고용정보를 연계 분석한 자료 조회
- **고용동향:** 워크넷, 외국인, 자활, 사회적일자리, 구인업체개척사업, 경제활동인구조사 등 관련 통계 자료 조회 및 분석
- **고용보험:** 사업장, 피보험자, 모성보호, 지원금, 실업급여 관련 통계 자료 조회 및 분석
- **직업훈련:** 실업자, 재직자, 고용촉진, 지도점검, 훈련교사 관련 통계 자료 조회 및 분석
- **부서현안자료:** 고용노동부 각 팀 별 해당 통계 자료 제공

DW 구축으로 이런 문제점은 말끔히 해결됐다. 또한 사용자가 원하는 모양으로 통계현황표를 자유롭게 변형할 수 있는 것도 장점이다.

솔루션 선택도 잘 됐다. 고성능 DW 어플라이언스인 "IBM 네티자1000"을 도입해 데이터 조회속도를 크게 올린 것. 오류데이터 검출이나 수정 작업이 용이해져 데이터 품질도 높였다.



한국고용정보원은 DW시스템 구축 후 신뢰성 있는 정보를 제공해 정부 고용정책의 실효성을 높였다.

다양한 연계분석으로 고용정책 실효성 높여

연계분석은 DW를 구축해 얻은 가장 큰 성과다. 먼저 구직자에 대한 연계분석 사례를 살펴보자. 시스템 이용자가, 구직자에 대한 현재 상황을 파악하려면 구직자에 대한 고용 변경 추적 및 현재 상태를 파악해야 한다. 이를 위해 사용자는 워크넷 서비스에 접근해 구직 등록한 사람을 파악한 후, 그 구직자의 취업여부 확인을 위해 고용보험시스템에 접근해 정말 취직했는지 확인한다. 고용정보시스템에선 개인특성,

사업장특성 등 구직자에 대한 세부사항을 조사할 수 있다. 이처럼 취직 현황을 파악할 수 있고, 취직한 후에도 얼마나 고용을 유지하고 있는지 바로 확인할 수 있다.

맞춤분석 적용해 사용자 편의성 향상

직업훈련시스템에 접근해선 취업자가 어떤 재직자훈련을 받았는지, 혹시 실직했다면 어떤 실업자훈련을 받았는지 알 수 있다. 또한 직업훈련을 받은 사람과 받지 않은 사람을 비교해, 해당사업장이 고용을 유지한 추이를 파악할 수 있다. 더불어, 개인특성(성별/연령/직종 등)이나 사업장특성(규모/산업 등)을 특성 별로 조합해 파악할 수 있기에, 고용정책 결정에 유용한 자료로 활용할 수 있다.

워크넷에서 취업한 구직자의 고용현황을 자세히 볼 때도 연계분석이 필요하다. 피보험자(상용직)의 근속년수를 분석하려면 워크넷 취업자 정보와 고용보험 사업장 및 피보험자 이력 자료를 연계분석 한 후 워크넷 취업자의 현재 고용유지 현황을 파악할 수 있다. 이 연계작업을 위해 사용자는 워크넷과 고용정보 데이터를 함께 이용하며, 이 때 적용되는 지표는 워크넷 취업자 고용보험 사업장 취업률, 워크넷 취업자 고용보험 사업장 취업률, 워크넷 취업자 월 기준시점별 고용유지 현황, 워크넷 취업자의 전 직장에서 근속년수, 규모별 취득자 현황, 피보험자 근속년수 구간별 분포 등이다.

사용자가 통계현황표를 원하는 모양으로 자유롭게 변형하는 맞춤분석이 가능해진 것도 시스템 구축의 성과다. GUI 환경으로 구성된 화면

에서 사용자는 필요한 업무 데이터를 직접 조회할 수 있고, 데이터 항목도 자유롭게 조합할 수 있다. 마치 엑셀로 만든 장표처럼 원하는 모양을 만든 후, 출력해서 분석용으로 쓰거나 보고용으로 제출한다. 데이터 출력 과정도 간편해 산출버튼을 누르면 결과값이 바로 나온다.

네티자 도입으로 데이터 조회속도 개선

데이터 조회속도를 크게 높인 것은, 새로운 장비 도입의 효과였다. DW 장비와 소프트웨어의 사용연한이 지나 새 시스템 도입이 필요했던 고용정보원은 2011년 "IBM 네티자 1000"을 도입했다. 이전 시스템보다 데이터 조회 속도가 10배 이상 빨라졌다. 2011년 7월에 네티자 시스템을 정식 오픈했다.

DW 구축 후 데이터 품질이 매우 좋아졌다. 오류데이터 검출과 수정작업이 쉬워졌기 때문이

다. DW는 운영계 데이터를 단순히 카피하지 않는다. 운영계 데이터를 쌓아 놓고 마감작업을 진행하는데, 통상적으로 월 기준으로 마감한다. 이 때 나오는 오류 데이터를 운영계 팀에 전달해 수정한다. 이런 과정을 거쳐 신뢰성 있는 데이터를 확보할 수 있고, 이를 통해 사용자들은 고품질의 데이터를 제공받을 수 있다.

고용정보원은 DW시스템 구축으로 대국민 서비스를 크게 개선했다. 변화가 심한 노동시장에 속히 대응할 수 있는 정책 구상과 실행이 가능해졌기 때문에, 국민의 삶의 질 향상에 기여하고 있다. 고용정보원은 앞으로 사용자 요구에 부응하는 통계 분석 지원을 위한 DM 기능을 더욱 강화할 예정이다. 또한 사용자의 통계 정보 활용 편리성을 높이기 위해 사용자 인터페이스를 개선하며, 고용정보에 대한 연계 분석을 강화할 방침이다.



"고용정책 신뢰성 높였다"

한국고용정보원 데이터관리팀 이상훈 과장

- DW시스템을 구축한 배경은? "고용정책을 위한 현황정보를 적기에 제공하기 위해서다. 데이터의 품질을 높여 대국민 서비스를 개선하는 것도 이유가 됐다."

- 솔루션 선정과정은? "6개월 이상 시간을 써가며 우수한 DW 솔루션을 선택했다. 이를 통해 성능이나 기능면에서 만족스러운 솔루션을 선택했다."

- 가장 큰 성과는? "연계분석이다. 한가지 지표에 대한 다차원 분석이 가능해 고용정책의 실효성을 크게 높였다. 이를 통해 신뢰성 있는 고용정책 결정과 실행이 가능해졌다."

- 향후 계획은? "통계 분석 지원을 위한 DM 기능을 더욱 강화하고, 사용자 인터페이스를 개선할 것이다. 고용정보에 대한 연계 분석을 강화할 예정이다."

COPYRIGHT NOTICE

- 본 보고서의 판권은 데일리그리드가 소유하고 있으며, 데일리그리드의 허가 없이 무단 배포 및 복사를 금합니다.
- 본 보고서의 내용은 전재(轉載)할 수 없으며, 인용할 경우 본지와 상의 후 그 출처를 반드시 명시해야 합니다.

DailyGrid

Phone : 02-749-3205

E-mail : grid@dailygrid.net

Web : www.dailygrid.net

This publication is copyrighted by *The DailyGrid Research Center*. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the *The DailyGrid Research Center*. Should you have any questions, please contact us at (02) 749-3205 or grid@dailygrid.net. This Case study was developed with the assistance and funding of *IBM Netezza*. Visit www.dailygrid.net to learn more about IT information and market trend.