

데이터마이닝을 이용한 파이로프로세스 데이터 정량화 분석

김성용, 박희성, 최영, 김호동
 한국원자력연구원, 대전시 유성구 대덕대로 1045
 svkim@kaeri.rc.kr

1. 서론

사용후핵연료 건식처리공정인 Pyroprocess 기술을 종합적으로 평가하기 위한 파이로 일괄공정 연구시설인 PRIDE (Pyroprocess Integrated DEmonstration) 시설이 있다. 이를 구축하기 위한 일정관리 작업에는 SE (System Engineering)가 적용되는데 여러 단계의 세부공정 영역이 있고, 각 공정들은 다른 공정과 밀접한 관련이 있다. 사용자는 이 일정관리를 통해 공정사항을 파악하기에 관리 시스템 구축의 향상은 곧 공정관리 작업 능률 향상과 연결된다.

현재 파이로 공정관리와 관련하여 각 단계의 요구사항을 통합 관리하는 시스템 엔지니어링 환경(SPIRE)을 구축 하고 있다. SPIRE는 통합 데이터베이스 환경에서 사업관리자, 설계담당자, 시험담당자, 기타 이해관계자가 시스템 규격서, 시험 검증 계획서, 기타 필요 문서를 입력 및 조회할 수 있도록 함으로써 프로젝트를 관리한다. 스키마를 정의하고 해당 내용을 어떻게 사용자에게 보여줄 것인지에 대한 쿼리를 정의하는 내용은 해당 도구를 활용함에 있어 중요한 사안이다. 지금까지는 이러한 일정관리의 항목 배치를 사용자 정의를 통해 해왔으나, 각 담당자의 정의와 관리자의 정의 등을 취합하는 과정에서 많은 회의 및 차후 수정의 필요성이 있어 왔다.

본 연구에서는 이런 문제점을 최소화 하기 위해 데이터마이닝의 연관 분석 기법을 사용하여 접근하였다. 이를 통해 사용자는 항목들을 관리함에 있어서 데이터들이 가지고 있는 특징을 미리 파악함으로써 구축을 위한 근거를 확보함과 동시에 사람이 발견하지 못했던 패턴을 찾아냄으로써 해당 특성을 검토하여 활용할 기회를 제공받을 수 있게 된다.

2. 본론

2.1 통합 데이터베이스 환경의 특징

현재 진행 중에 있는 SPIRE는 각 공정단계의

요구사항을 관리함에 있어 해당 일정 상태는 물론 선행 항목들도 표현해 준다. 이를 통해 사용자가 특정 공정에서 문제되는 사항은 해당 틀에서 지원하는 통합 데이터베이스에 저장된 정보와 링크된 문서를 통해 파악할 수 있게 되며, 변경 이력을 통해 수정된 내역까지 공유하기 때문에 담당자들과 일일이 연락하지 않고 파악할 수 있는 편의를 제공해준다. Fig.1은 파이로 시스템 엔지니어링 도구인 SPIRE에서 파이로 공정에 대한 정보를 저장하는 부분이다.

The screenshot shows a web-based data entry form with the following fields and values:

- 이름: A장치(관형) 수경열계
- 관리 번호: 201107_01_01 (시달-번호)
- ACPF: Advanced spent nuclear fuel Conditioning Process Facility
- DFDF: DUFIC Fuel Demonstration Facility
- FRIDE: Pyroprocess IntegratedInActive DEmonstration facility
- 시작일: 01/07/11
- 종료일: 20/07/11
- 분야: 공정분야
- 진행현황: 미착수
- 제공일: 20/07/11
- 예산(원단위): 50,000,000
- 수달 내용: 입력 데이터 지면으로 일정 연기
- 가중치: 0.2
- 진행률(%): 0
- 공정률: (empty)
- 첨부자료명: (empty)
- 입력문서: HWP, PPT, EXCEL, PDF, WORD, PICTURE

Fig. 1. Data set.

2.2 연관규칙과 순차패턴 분석 적용의 필요성

파이로 데이터를 효과적으로 사용하기 위해선 가능한 자세하게 정보를 저장함과 동시에, 필요할 때 해당 정보만을 모아 사용자가 파악하는데 시간절약 및 이해도를 높이기 위해 맞춤 형식으로 제공하는 쿼리(Query)를 만드는 전략이 필요하다. Fig.2는 쿼리 실행 시 해당되는 정보들이 표시되는 것을 보여주고 있다.

지금까지는 개발자가 자체적으로 정의하고, 이를 실 사용하는 담당자들과 협의하는 방식을 반복하는 접근방식을 사용해왔다. 하지만 이는 시간

과 비용의 소모가 크고, 개개인의 견해를 바탕으로 하기에 후에 수정에 들어가는 비중도 커질 수 있다. 본 연구에서는 이러한 낭비를 줄이고 좀 더 효율적인 쿼리를 제공하기 위해 초기 쿼리 작성 시 데이터를 기반으로 나온 항목간의 연관성을 바탕으로[1]으로 할 것을 제안한다.

구분	구분명	구분코드	구분명	구분코드	구분명	구분코드	구분명	구분코드
1	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001
2	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001
3	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001
4	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001
5	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001	국립중앙도서관	010001

Fig. 2. Result of Query.

3. 결론

본 연구에서는 시스템 엔지니어링 도구를 도입함에 있어 쿼리 개발 및 유지보수 비용과 시간을 감소시키고 쿼리 구성에 대한 근거를 부여하기 위해 데이터마이닝 적용을 제안하였다. 실제 효율성을 정량화 하기 위해 추후 각 데이터의 정의에 대한 추가적인 연구가 필요하다.

4. 감사의 글

본 연구는 교육과학기술부의 원자력 연구개발사업의 일환으로 수행되었으며, 이에 감사드립니다.

5. 참고문헌

[1] Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Vipin Kumar, Pearson Education; ISBN 0-321-42052-7, pp 327-486.