

군수품 시험성적서 디지털화 방안에 대한 연구

윤성기, 김상덕, 정일호
국방기술품질원

e-mail:sgyoun@dtaq.re.kr, sdkim@dtaq.re.kr, ilhojung@dtaq.re.kr

A Study on Methods for Digitalizing Military Equipment Test Reports

Seong-Gi Youn, Sang-Deok Kim, Il-Ho Jung
Defence Agency for Technology and Quality

요약

4차산업혁명으로 일컬어지는 과학기술의 발달로 데이터 기반의 군수품 품질관리가 요구되고 있다. 이를 위해서는 전 순기 품질 데이터의 수집이 필수적이다. 그러나 품질정보가 방대하고 산재해 있으며, 비정형문서로 작성된 데이터는 관리와 분석에 어려움을 초래하고 있다. 본 논문은 시험성적서의 디지털화 필요성을 다루며, OCR(광학 문자 인식) 기술을 통한 비정형 문서로 작성된 군수품 시험성적서를 디지털화하는 방법을 검토하였다. 이를 통해 비정형으로 관리되고 있는 군수품 품질정보 데이터의 획득 및 분석에 효과적일 것으로 기대한다.

1. 서론

4차산업혁명으로 일컬어지는 과학기술의 발달로 데이터 기반 과학적 품질관리를 통한 군수품 품질수준 또한 발전이 요구되고 있다. 이를 위한 방법으로 방사청과 국방기술품질원에서는 개발·양산·운영유지 단계의 제반 품질정보 수집 및 빅데이터화 구축과 기품원 플랫폼을 활용, 소요군 등과 데이터를 공유하는 시스템의 구축, 점진적으로는 민·관·군 공동 활용 클라우드 기반 정보체계의 구축을 추진하고 있다.[1] 그러나 방대하고 산재된 데이터의 수집 및 디지털화는 여전히 큰 과제로 남아 있다. 특히, 개발·양산·운영유지 단계에서 생성되는 품질정보들은 비정형 문서형태로 작성되는 경우가 많아 이를 통합적으로 관리하고 분석하는데 많은 어려움이 따르고 있다. 이러한 문제를 해결하기 위해서 비정형 문서로부터의 데이터 수집기술은 빅데이터 구축과 효과적인 데이터분석의 필수조건이다.

본 논문에서는 비정형데이터인 군수품 시험성적서의 디지털화 필요성과 방안에 대해 논의하고, 이를 수행하기 위한 방안을 제안하고자 한다.

2. 군수품 시험성적서 디지털화

과거에 기술요소 확인 위주의 품질보증활동이 고도화 되었다면, 현재는 데이터에 기반을 둔 품질보증활동이 요구되고 있다.[2] 하지만 시험성적서 대부분이 비정형 문서 또는 수기

로 작성·관리되고 있어 데이터 관리에 어려움이 있다. 데이터를 지속적으로 관리 및 개선을 요하는 데이터 품질(Data quality)[3] 관점에서 보았을 때 우선적으로 관리를 위한 데이터 확보 단계에서부터 장애를 직면하게 되는 것이다. 특히 영세중소기업체에서는 인적, 시간적, 경제적 제한으로 수기로 시험성적서를 생성하는 경우가 많으며, 방산업체에서는 보안의 문제로 시험성적서 뿐만 아니라 외부로 반출되는 모든 문서에 대해 pdf, 그림파일 등 비정형 문서로 외부 반출하도록 제한하고 있어 이는 시험결과를 디지털화하여 활용하는데 어려움이 있다.

하지만 시험데이터를 디지털화하여 관리하면 다음과 같은 장점을 확보할 수 있어, 추진이 필수적이다.

첫째, 업무생산성을 향상할 수 있다. 디지털화된 시험성적서는 검토 과정에서 이상치 및 규격 불일치를 쉽게 식별할 수 있다. 특히 타산업군 대비 군수산업은 다품종 소량생산의 특성이 있어 인원당 담당하는 품목이 다양하여 시험성적서 작성 및 검토간 인적오류를 범할 가능성이 높아 위험도를 줄일 수 있다. 또한 데이터 조회 및 검토에 소요되는 시간을 대폭 줄일 수 있다.

둘째, 품질데이터 추적관리가 가능하다. 개발 및 양산단계에서부터 운영유지단계까지 품질 데이터를 추적 관리할 수 있어, 장기간 생산·운영유지 특성을 가진 군수품의 특성에 따라 4M 1E의 변동시 품질이 적절한 수준으로 관리되고 있는

지 수치화하여 추적관리가 가능하다.

셋째, 통계적 품질관리를 안정적으로 수행할 수 있다. 통계적 기법을 활용한 모니터링은 불량품 출하나 리콜과 같은 문제를 방지할 수 있으며, 관리도를 통해 품질 데이터의 변동이 허용 범위 내에 있는지를 확인할 수 있어 공정간 활용되고 있다. 하지만 이를 위해서는 통계학적 전문 지식과 데이터 분석 기술이 요구되며, 잘못된 기법 적용을 방지하기 위한 사용자 교육이 필요하다. 최근 운영유지단계의 유지부품들은 영세업체들의 계약으로 불량과 납품지체가 빈번히 발생하고 있어, 군의 무기체계 가용도를 떨어뜨릴 뿐만 아니라 계약업체에서도 매출 미달성 및 계약해지에 따른 불이익을 받고 있다. 따라서 데이터 축적 및 분석환경 구축을 통해 운영유지단계에 참여하는 대부분의 영세업체에 도움이 클 것으로 예상되며, 이를 통해 군의 안정적인 부품확보기반도 다질 수 있다.

2.1 OCR을 활용한 시험성적서 디지털화 방안

비정형문서의 데이터를 디지털화 하기위해 일반적으로 사용되는 OCR(Optical Character Recognition, 광학 문자 인식)은 이미지나 스캔된 문서에서 문자 데이터를 추출하는 기술이다.[4] 이 과정에서 성적서 내의 데이터는 디지털화되어 전산 시스템에 저장될 수 있으며, 이는 이후 데이터 분석 및 품질 관리에 유용하게 활용되며 일반적인 OCR을 활용한 시험성적서의 디지털화 방안은 다음의 절차를 따른다.

2.1.1 OCR 데이터 추출

OCR 기술을 사용하여 이미지나 스캔된 문서로부터 텍스트 데이터가 자동으로 추출한다. 이 과정에서 수작업으로 입력해야 할 데이터를 자동화하여 추출함으로써 효율성을 높일 수 있으며, 특히 대량의 성적서를 빠르게 처리하는 데 유용하다.

2.1.2. 데이터 전처리 및 검증

추출된 텍스트 데이터는 원본 문서에서 바로 사용하기 어려운 오류나 불필요한 요소를 포함할 수 있다. 따라서 전처리 단계에서는 잘못된 값을 수정하는 알고리즘을 사용하여 데이터를 정제하는 과정이 필요하다. 이후 전처리된 데이터가 오류나 불일치 여부를 검증이 필요하다.

2.1.3. 데이터 저장

검증된 데이터는 이후 조회, 분석, 통계 처리를 위해 데이터베이스에 저장된다. 이로써 데이터는 체계적으로 관리되고, 필요할 때 신속하게 접근하여 사용할 수 있다.

2.1.4. 데이터 분석 및 결과 판단

저장된 데이터를 분석하여, 규격기준에 따라 합격 또는 불합격을 판단하는 과정이다. 이 단계에서는 필요시 통계적 분석 알고리즘이 사용되며, 이를 통해 품질관리의 효율성이 크게 향상될 수 있다. 관리도 등을 사용하여 데이터를 모니터링

하고 품질변동이나 리스크를 사전에 예방할 수 있다.

2.2 OCR을 활용한 시험성적서 디지털화의 제한사항

하지만, OCR 기술이 절대적인 해결방법은 아니며, 다음과 같은 제한이있어 군수품 시험성적서의 디지털화에 적용전 사전검토가 필요하다.

2.2.1 이미지 품질에 대한 의존성

OCR 기술은 이미지의 해상도와 품질에 크게 의존한다. 성적서 스캔 이미지의 품질이 낮거나, 이미지가 흐릿한 경우, 정확한 인식이 어려울 수 있다. 따라서 고해상도 스캔 등 전처리가 필요하다.

2.2.2. 인식 오류 발생

시험성적서의 양식이 일정하지 않거나, 손글씨가 포함된 경우, OCR의 인식률이 떨어질 수 있다. 손상된 문서, 글씨체가 불분명한 경우, OCR은 올바르게 인식하지 못해 데이터의 정확도가 떨어질 수 있다.

2.2.3. 다양한 양식 처리의 어려움

시험성적서의 레이아웃이 다양하거나 비표형식일 경우, OCR이 문서 구조를 제대로 인식하지 못하는 경우가 있다. 이러한 비정형 데이터를 처리하기 위해서는 추가적인 전처리 작업이나 양식별 맞춤형 OCR 설정이 필요할 수 있다.

2.2.4. 사용자 교육 필요성

새로운 전산화 시스템의 도입에 따라 사용자들이 이를 효과적으로 활용할 수 있도록 교육이 필요하다. 사용자가 시스템을 제대로 이해하지 못하면 오히려 효율성이 떨어질 수 있으며, 새로운 시스템에 대한 저항이 발생할 가능성도 있다.

2.3 군수품 시험성적서 디지털화 방안

현재의 OCR 기술의 적용 및 제한사항을 극복하여 군수품에 적용하기위해 다음과 같은 방안을 제안하고자 한다.

2.3.1. 관리품목 선택

시험성적서 전산화를 효과적으로 구현하기 위해서는 관리품목을 선정하는 것이 중요하다. 관리대상의 기준은 재고번호, QAR번호, 도면번호, 부품번호 등으로 분류할 수 있다. 제조 공정품 기준으로 관리함이 합리적이므로 최하위 기준인 부품번호를 기준으로 품목을 선택함이 데이터 품질관점에서는 최선의 조건이다. 하지만 제조업체에서는 공정이 안정될 경우 관리항목이 많은 것은 손실이므로 안정될 시 해당 데이터가 개발업체 내부공정 품질관리단계에서 폐기되어 운영유지 단계의 신규업체까지 전달되지 않을 수 있다. 따라서 품질정보 데이터 관리 및 데이터 축적을 통한 빅데이터 구성의 관점에서는 부품번호 기준이 타당하나, 데이터의 지속적인 관리를 위해서는 운영유지단계까지 지속적으로 관리되는 최소한 재고번호의 기준으로 관리되어야 한다.

2.3.2. 관리 데이터 설정

관리 데이터는 해당 품목의 주요 치수, 성능 데이터(전기적 성능 등)와 같은 핵심 데이터를 포함하여야 한다. 또한 주요 공정 데이터는 통계적 기법을 적용하여 손쉽게 품질변동을 모니터링하고 리스크를 사전에 관리할 수 있으므로, 필요한 관리 데이터를 식별하여야 한다. 하지만 협력업체 또는 특수 공기업에서 생성되는 원자재, 보호피막처리, 공인기관 성적서 등의 데이터는 성적서 양식을 표준화 할 수 없어 OCR을 활용한 데이터 관리에서는 제외할 수 밖에 없는 한계가 존재한다. 다만, 수량이 많은 품목, 관리가 필요한 품질 데이터, 재 활용도가 높은 품질 데이터는 표준화 양식을 제조전 양식 작성 및 사전배포를 통해 관리 데이터에 포함할 수 있다.

2.3.3. 표준화 성적서 양식 개발

OCR로 얻어낸 데이터의 신뢰성을 확보하기 위해서는 표준화된 성적서 양식을 개발하는 것이 필수적이다. 이를 통해 데이터 전처리 단계를 줄일 수 있다. 뿐만 아니라, 후속 양산 단계에서는 품질관리 데이터의 사전확인으로 관리 항목을 사전에 검토가능하여 품질문제를 예방할 수 있다.

2.3.4. 데이터 전처리 및 분석 SW 구현

수집된 품질 정보의 신뢰성을 높이기 위해서는 효과적인 데이터 전처리 기술의 적용이 필요하며, 품목별 적절한 데이터 분석 알고리즘을 구현 하여야한다. 데이터 분석 알고리즘을 통해 품목의 합격/불합격 여부를 판단뿐만 아니라, 통계 분석결과를 제공하여 품질관리 수준을 높일 수 있다.

2.3.5. 데이터 관리

마지막으로, 적절한 데이터의 관리가 필요하다. OCR 데이터를 활용하기 위해서는 ‘성적서 양식 데이터’, ‘데이터 전처리 및 분석 SW’, ‘군수품 성능 및 분석 데이터’ 3종의 데이터의 관리가 필요하다. ‘성적서 양식 데이터’와 ‘데이터 전처리 및 분석SW’는 지속적인 업데이트 및 배포를 위해 국방규격 책임기관으로서 방위사업청에서 운용하고 있는 국방표준종합정보시스템(KDSIS)에 최소단위인 부품기준으로 관리하고, 품질분석 데이터인 ‘군수품 성능 및 분석 데이터’는 국방기술 품질원에서 운용하고 있는 품질정보서비스(IQIS)에 저장하는 것이 효과적일 것으로 판단된다.

3. 결론

본 연구에서는 군수품 시험성적서 디지털화의 필요성과 이를 효과적으로 구현하기 위한 방안을 제시하였다. 기존의 비정형 문서기반 시험성적서 관리 방식은 데이터의 관리가 어려워 OCR 기술을 활용한 데이터의 디지털화 방안을 제시하였다.

본 연구에서 제안된 방안들이 실제로 적용될 경우, 군수품 품질관리 수준이 한층 강화될 것으로 기대된다. OCR 등 정보

기술의 발달로 군수품 품질 데이터의 획득은 지속적으로 개선될 수 있을 것으로 예상되나, 품목별 관리 데이터의 설정 또는 범위에 대한 연구는 지속적으로 구체화가 필요할 것으로 판단된다.

참고문헌

- [1] 방위사업청 외, “24~28 군수품 품질관리 기본계획”, pp.29-31, 9월, 2023년
- [2] 전수연 외, “빅데이터 기반 군수품 품질정보 활용방안에 대한 연구”, pp.315-325, 한국산학기술학회논문지, 제17권 제6호, 2016년
- [3] 한국정보화진흥원, “공공데이터 품질관리 매뉴얼”, pp.13, 2018년.
- [4] 박동원 외, “Certificate of Conformance(COC)에서 정보추출을 위한 OCR기반 알고리즘 개발”, pp.97-98, 한국컴퓨터정보학회 하계학술대회 논문집 제32권 제2호, 7월, 2024년