

국가 전자 공문서 포맷 선정에 관한 소고

A Study on the Selection of National e-Document Format

안승환, 유니닥스 (주), ash@unidocs.co.kr

남영준, 중앙대학교 문헌정보학과, namyj@cau.ac.kr

Seung-Hwan Ahn, Unidocs co, Ltd.

Young-Joon Nam, Dept. of Library and Information Science, Chung-ang University

본 연구에서는 전자문서 생성과 보관을 위한 전자자원 형식을 설정할 때 필요한 고려사항을 제안하였다. 이는 전자정부구현을 통한 국가정보와 지식의 공유환경 개발에 기본 요체이다. 조사결과, 국가 전자문서 형식은 활용을 위한 기준과 보존을 위한 기준으로 구분된다. 전자의 경우, 고려사항은 문서 변환과 문서 암호화, 데이터 객체 관리, 색인 생성, 워터마크, 국가 XML과 연동하는 표준 데이터 기술, 메타데이터 관리 등이 포함된다. 후자의 경우, 고려 사항은 문서 암호화와 디지털 저작물 관리 요소, 파일 사이즈(압축 효율) 등이 포함된다. 활용과 보존적 측면에서 공동으로 고려할 것은 행정자치부에서 권고하는 행정기관의 전자문서시스템 규격, 「행정기관간 전자문서유통 표준」, 「전자문서시스템과 행정정보시스템간 연계표준」이다.

I. 서론

전자정부라는 단어가 처음 사용되기 시작한 것은 1990년대 초 미국의 당시 Al Gore 부통령이 주도한 National Performance Review이다. 이때 정의된 전자정부의 의미는 정부를 고객위주의 효율적이고, 공공정보 및 공공서비스에 “보다 폭넓게, 보다 적시”에 접근하도록 하는 인터넷 서비스 기반의 정부체제였다. 이 과정에서 정보통신기술의 중요성이 강조되었고 이러한 정보통신기술을 활용하여 정부의 혁신을 지향한다는 점이 크게 부각되게 되었다. 한편, 한국에서의 전자정부는 정보화촉진 계획에서부터 논의가 본격적으로 이루어지기 시작했다. 국내에서 정의된 전자정부는 공통의

정보통신 기반 위에 하나로 연결된 각종 행정 서비스를 “언제, 어디서나, 어떻게 해서든” 제공할 수 있도록 하는 정부였다. 이러한 전자정부 구현을 위하여 정부에서는 G4C 사업을 비롯한 11대 과제를 선정하여 전자행정서비스를 수행하고 있다. 전자정부가 추구하는 목표는 고객 지향적 행정서비스 확립과 아울러 행정효율의 극대화를 통한 정부의 생산성 향상에 있다. 이를 위해 정부는 행정업무의 효율성과 국민에 대한 행정서비스를 제고하기 위해 각급 행정기관에서 구축·운영중인 행정정보시스템과 정부용 표준 전자문서시스템간에 원활한 연계가 가능하도록 필요한 기능 및 관련 표준들을 개발하여, 이 표준에 따라 전자정부의 요체인 전자문서시스템을 완성하였다(행정자치부 2002).

한편 각급 행정기관은 이러한 국가적 시책

에 따라 전자문서시스템을 도입하여, 법적 전자문서의 성립과 기관 간 문서유통도 많은 진전을 보이고 있다. 그러나 초기 전자문서시스템은 각 제조사 별로 각기 다른 시스템 구조와 전자문서 형식을 채택하고 있어 이를 위한 표준화가 필요하게 되었다. 시스템과 운용에 관한 표준화는 완성되었으나 전자문서와 저작권법 등과 같이 시스템 외적인 사항은 외부의 판단과 해석에 기준해야 하는 점 때문에 이에 대한 표준을 도출하기는 현실적으로 많은 문제점을 갖고 있다. 본 연구는 이 점에 착안하여 표준화된 전자문서 형식을 선택하는데 필요한 기준으로서의 고려요소를 제안하고자 한다.

2. 국내 전자 공문서의 관리

우리나라에서는 국가차원에서 인프라 구축과 함께 국가 주도 분야의 지식정보화를 주도하고 있다. 이는 부존 자원의 부족을 지식 자원의 구축을 통한 선진국 진입을 위한 장기적인 국가 투자적 성격을 띄고 있다. 이러한 추세에 따라 행정자치부의 주관아래 국가 지식정보화 확립을 위해 전자문서유통 및 관리체제로 모든 행정정보를 공유하도록 유도하고 있다.

2.1 전자 공문서의 정의

전 세계 모든 국가에서는 자국의 역사를 기록하고 이를 바탕으로 국가의 미래를 설계하기 위해 모든 국가 문서를 보관하는 기관이나 단체를 운영하고 있다. 이러한 기관에서 운용하는 대상은 대부분 종이기반의 문서였다.

전통적으로 이러한 국가 공문서는 다음 네 가지의 성립 요건을 갖고 있다. 첫째, 사람이 인지 할 수 있는 수단으로 표시되어야 하고 둘째, 일정한 사항 또는 관념을 표시하여야 한다. 셋째, 기호, 문자, 그림 등 가시적인 물체

의 형태를 포함해야 하고 넷째, 특정인의 의사 표시가 내재되어야 한다.

이러한 전통적 문서는 사람이 눈으로 확인할 수 있는 가시성과 내용을 판독할 수 있는 판독성, 그리고 비교적 오랫동안 보존할 수 있어야 하는 영속성을 갖고 있다. 지금까지 문서의 개념은 비단 종이문서에만 한정된 것은 아니지만 그 사용의 편리함과 사후 참조의 용이함을 비롯한 많은 장점들로 인하여 현재까지도 종이문서가 가장 널리 이용되고 있다.

한편, 전자문서의 법률적 정의는 전자거래기본법, 전자서명법 등 유관 법률, 시행령, 시행규칙을 포함하여 116곳에 정의되어 있으며, 법률마다 다소 상이하게 정의하고 있다. 그 가운데 전자문서에 대해 “정보처리 시스템에 의한 전자적 정보”, “전자적 형태로 송·수신 또는 저장된 정보”의 개념을 가장 폭 넓게 사용하고 있다. 이와 같은 전자문서가 갖는 문서성을 살펴보면 다음과 같다. 첫째, 법적인 증거 능력이 있어야 한다. 컴퓨터로 생성한 정보 또는 보관하고 있는 정보에 대하여 법적인 증거로 채택될 수 있어야 한다. 최근에 와서는 과거와 달리 정보기술의 기술적 뒷받침으로 인해 전자문서 관계법에 있어 이러한 전자문서가 증거능력으로 인정되도록 개정되고 있다. 둘째, 오류로부터 신뢰성을 확보해야 한다. 이 역시 정보기술의 발전으로 많은 신뢰성을 회복하고 있다. 셋째, 일정한 형식을 갖추어야 한다.

전자문서가 이러한 요소를 갖추고 일반 공문서와 동일한 효력을 갖기 위해서는 법률에서 제기하는 공인된 서명 등과 같은 인증(authenticity)과정이 필요하다. 현재 금융권에서는 서명 역할을 하고 있는 공인인증서를 널리 이용하고 있으며 정부에서도 각종 전자공문서 내에 인증서를 적용하는 기술을 2005년부터 시행하여 전자문서 활용을 위한 초석을 마련하였다.

2.2 전자공문서의 유형 및 관리

현재 전자정부 구현을 위해 사용되어지는 문서의 유형은 다양하다. 기관 간 문서유통을 위한 영역에는 일반적으로 XML형식의 문서가 활용되고 있으며, 일반 공문서의 본문영역은 한글문서를, 그리고 첨부영역에는 마이크로소프트사의 파워포인트문서, 엑셀문서, 워드문서와 기타 훈민정음 문서 등의 다양한 형태의 문서를 활용하고 있다. 최근 전자문서의 호환성 문제를 해결할 것으로 주목받았던 XML언어를 이용한 저장방식은 표준화의 벽에 부딪히고 전환비용이 높은 단점을 안고 있음으로 인해 그 이용 및 활용 범위가 폭넓게 전파되지 못하고 있는 실정이다. 또한, 현재 정부에서는 일부 기관을 제외하고 행정정보 공개에 있어 다양한 문서형식들을 그대로 서비스하고 있다. 때문에 국민들은 각기 다른 형식의 문서들을 열람하기 위해서 개별 뷰어프로그램을 설치하여 열람해야 하는 불편을 겪고 있다. 이것이 바로 고객 지향적 행정 서비스 확립과 아울러 행정 효율의 극대화를 목표로 하는 전자정부의 현실이다. 반면 국내 IT 산업의 발전과 일반 국민들의 높은 IT 산업의 관심도에 의해 국민들의 정보통신 이용률과 정보통신 이해도는 갈수록 높아져 가고 있다. 이로 인해 다양한 형태로 작성되어진 전자문서들을 단일한 형태로 보관, 관리, 유통함으로써 보다 효율적인 업무처리와 행정서비스를 할 수 있는 방안들이 요구되어지고 있는 추세이다. 이것은 오래전부터 논의된 사항으로 이미 생성되어진 종이문서를 비롯한 전자문서들을 통합된 형태로 관리하며, 이를 활용하고자 하는 것이었다. 이를 위해 정부에서는 오래 전부터 종이문서의 전자 문서화를 추진하였으나, 이는 보관의 의미에 초점이 맞춰진 것으로 종이문서를 단순한 형태의 이미지 형태로의 문서와 이를 식별할 수 있는 서지정보만을 내포하고 있었다. 그러나 전자정부의 활성화로 인해 이러한 제

한적 형태의 전자문서는 그 의미가 상실되어 가고 있으며, 또한 정보통신기술의 발전으로 인해 전자문서의 생성, 보관, 관리, 유통 등의 폭넓은 요구조건 등을 수용할 수 있는 대안들이 대두되고 있다.

2.3 전자 공문서의 포맷

이러한 대안 중에는 단순 그래픽 형태의 BMP와 TIF 형태의 파일을 소장하고 이를 관리하는 방안이 특별한 뷰어나 압축기술없이 운용되고 있다. 또한 압축알고리즘을 사용하여 기존의 그래픽 파일이 갖는 파일사이즈를 줄여 저장공간과 전송시간을 절약하는 상업적 소프트웨어를 이용하여 특정 뷰어와 관리기술이 필요한 형태로 운용되고 있다. 즉, 전자의 경우는 스캔된 문서를 기본적인 프레임과 메타언어를 부여하여 관리하는 방법이다. 후자의 경우는 스캔 과정에서 저장 및 활용을 위한 전송까지 여러 가지 기능을 부여하여 관리하는 방법이다. 후자의 경우는 여러 가지 방법이 있으나 국내의 주요 관공서에서 활용하는 것은 PDF(portable document file) 알고리즘에 기반하는 방법과 DjV(데자뷰) 알고리즘에 기반한 방법, 기존의 문서작성 소프트웨어에서 제공하는 기능을 이용하는 방법으로 구분할 수 있다. PDF 문서 형식은 미국의 Adobe사의 원천기술을 이용한 것으로서 아크로벳 리더라는 공개뷰어를 통해 문서를 열람하는 방식을 취하며, 국내외 많은 대학도서관의 문서보관 포맷으로 사용하고 있다. DjVu(데자뷰)는 미국 AT&T 연구소에서 개발되어진 것으로, 미 항공우주국의 자료와 국내 일부 국가 기관에서 사용하고 있으며, 별도의 뷰어를 통해 전자문서를 열람하는 방식을 취하고 있다.

이러한 상업용 전자 포맷 알고리즘은 기존 그래픽문서가 갖고 있는 정보양의 손실없이 파일 형태를 압축하는 방법으로 워드프로세서 파일이나 TIF파일과 같은 원본데이터를 강제

로 압축하는 방식을 취하고 있다. 따라서 압축 과정에서 정보량의 손실이 발생할 수 있으며, 압축된 파일은 원본 파일로 변환이 불가능하다는 단점 때문에 일부 대학에서는 TIF와 같은 원본 형태의 그래픽파일을 보관·운용하고 있다.

2.4 전자문서 포맷 설정에 따른 고려요소

일반 기업이나 개인이 아닌 국가가 국가차원의 사업으로 국가 생산 정보와 지식을 파일화하는 작업은 규모의 방대성과 사업실패에 대한 후유증이 크기 때문에 사업 수행전 정책 기획과 결정은 반드시 이루어져야 한다.

국가 전자 공문서를 표준 포맷으로 결정하는 것도 국가지식정보화 사업에 가장 근간이 되는 핵심 요체이다. 따라서 이에 대한 결정은 단순한 실험이나 문헌적 고찰에 의존할 수 없으며 면밀한 검토가 이루어져야 한다. 이 과정에서는 객관성이 확보되어야 하며, 객관화를 위한 조건으로써 포맷결정에 기준이 되는 고려요소가 먼저 개발되어야 한다.

2.4.1 활용적 요소

국가 전자 공문서는 크게 활용적 요소와 보존적 요소로 구분할 수 있다.

활용적 요소는 국가에서 공표하거나 생산한 (행정)정보를 모든 국민이 공감함으로써 국가나 개인에게 이익이 될 수 있도록 하는 것이 최우선적으로 고려되어야 한다.

활용적 요소에서 고려할 사항을 정리하면 다음과 같다.

- ① 문서 변환 및 가독의 용이성 : 기존의 인간들에게 친숙한 형태로써 종이기반의 형식과 동일한지를 고려한다.
- ② 문서 암호화의 편이성 및 신뢰도 : 국가공문서는 그 자체만으로 제3자에게 신뢰를 갖는다. 이는 공문서의 형식과 신뢰에 기인한 것으로써 전자공문서도 이러한 것을 보완할

수 있는 방법으로 문서에 대한 암호화와 신뢰도를 동시에 확보해야 한다.

③ 데이터 객체 관리

④ 색인 생성의 안정성과 용이성

⑤ 워터마크 : 이 기능은 문서의 변환을 허용하되 원본 문서의 출처를 알려주는 기능으로 저작권과는 약간의 적용상 차이점을 갖고 있다.

⑥ 국가 XML 동하는 표준 데이터 기술 : 이를 위해 전자문서 포맷에서는 XML 정보를 수용할 수 있는 기술을 내재하여야 한다. XML의 설계 유연성을 수용하는 전자 포맷은 모든 문서를 통합하여 인간의 가독성을 극대화시킬 수 있기 때문이다. 따라서 XML 데이터를 손상없이 수용하여 문서의 구조적 저장과 체계적 관리, 최적의 문서유통을 지원할 수 있다. 이를 위해서는 표준 포맷에서는 W3C의 표준(XML 표준)을 준수해야 하며, 이를 바탕으로 표준화된 XML 메타데이터 체계는 많은 이기종 시스템 또는 어플리케이션간의 메타데이터 교환기능을 제공해야 한다. 이러한 XML기반의 메타데이터를 관리함으로써 어플리케이션이나 시스템은 이 XML 메타데이터를 읽어서 대상 자원의 속성을 판단하고 자동 처리 및 자동 분류 등의 작업을 진행할 수 있기 때문이다.

⑦ 메타데이터 관리 : 기존의 MARC이나 Dublin Core 등과 같은 메타데이터를 수용할 수 있어야 한다.

2.4.2 보존적 요소

보존적 요소의 주된 관점은 전자공문서의 보존적 성격을 부여하여 사적 가치를 갖도록 하는 것이다. 이는 이용자의 서비스적 특성보다 전자 공문서에 대한 국가 사료로써 보존적 특성을 고려하여야 하는 것을 의미한다.

보존적 요소에서 고려할 사항을 정리하면 다음과 같다.

- ① 문서 암호화
- ② 디지털 저작물 관리요소
- ③ 파일 사이즈(압축 효율)

이상과 같은 사항과 함께 고려되어야 하는 것은 보존된 자료의 배포과정에서 저작권의 보호와 원본의 진본성이라는 측면을 고려한 디지털 저작물 관리(DRM: Digital Record Management)요소가 반드시 고려되어야 한다.

2.4.3 공통 요소

이상과 같은 활용과 보존적 측면이외에 전자 공문서 구축에 따른 국가 표준의 수용여부는 국가 공통 전자 포맷의 선정에 기본적으로 고려되어야 한다. 행정자치부에서는 행정기관의 전자문서 교환에 따른 표준을 지시하여 원칙적으로 모든 공공 기관에서 생산되는 전자 공문서는 「행정기관의 전자문서시스템 규격」을 비롯하여 「행정기관간 전자문서유통 표준」, 「전자문서시스템과 행정정보시스템간 연계표준」을 준수하고 있다(행정자치부 2002).

이와는 별도로 2004년 4월 정보통신부가 정부혁신지방분권 위원회(전자정부 전문위원회)와 공동으로 관계 기관 및 전문가의 의견수렴을 거쳐, 정보 시스템 간 상호 운용성 확보 및 정보보호 강화 등을 위해 정보 시스템 구축·운영 시 준수해야 할 최소한의 기술 기준으로 제시한 “정보시스템 기술가이드”에서는 PDF 문서를 전자문서 교환의 표준 포맷으로 발표하였다. 즉, 행정자치부의 지침은 XML에 기반한 공문서 구축을 권장하고 있으며, 전자정부 전문위원회에서는 전자 문서 포맷형식으로 PDF 파일을 권장하고 있다.

한편 문헌정보학적 관점으로 이러한 국가 전자공문서 형식 선정에 고려해야 할 요소는 크게 메타데이터 연관기술과 색인기술이다.

① 메타데이터 연관기술

메타데이터 연관기술은 기존의 도서관에서

사용하는 목록기술 방법을 수용하면서 전자공문서 형식에 맞는 새로운 메타데이터를 활용할 수 있도록 파일 형식이 유연해야 함을 의미한다. 예를 들면, GILS(Government Information Locator Service)와 같은 메타데이터를 적용할 수 있도록 헤더부분 관리가 용이해야 한다. 또한 기존의 인쇄형 저작물의 메타정보와 통합될 수 있는 관리요소정보를 삽입할 수 있어야 한다.

② 색인기술

전자 공문서 운용의 장점에 속하는 것은 기존의 인쇄형 자료의 색인작업은 수작업에 의존하는 것에 비해 전자 문서의 색인작업은 자동 색인방법을 적용할 수 있는 점이다. 또한 전자 공문서는 페이지 단위의 접근 기술과 검색어에 대한 하이라이트 처리가 가능하기 때문에 필요자료에 대한 접근을 실시간으로 할 수 있는 것이다. 예를 들면, 일반적으로 사용자가 키워드를 입력하여 제공받은 검색결과로 116건의 문서 리스트가 나열되었다고 가정하면 사용자는 116건의 문서를 열람하여야 하고 1건의 문서를 약200면으로 기준으로 한다면 전체문서를 열람하기 위해서는 23,200면을 열람하여야 한다. 그러나 검색어에 대한 하이라이트 처리 기술을 적용한다면 약 200면을 기준으로 1건의 문서에 사용자가 입력한 키워드가 일치한 페이지는 평균 7면 정도만 해당되는데 이를 전체 검색결과에 적용할 때 사용자는 단지 812면(7면*116건)만 열람하더라도 원하는 정보를 찾을 수 있을 것이다.

3. 결론

지난 1980년대 이후로 전자정부 구현을 위해 종이 문서에 대한 광파일링 도입, 전자적으로 생성되는 각종 자료들을 쉽게 검색할 수 있는 전자문서 관리시스템 및 이를 기반으로 한 전자결재와 지식관리시스템 등 효율적 문

서관리를 위한 다양한 노력들이 지속적으로 시도되어 왔다.

본 연구는 이 점에 착안하여 국가 전자 공문서 포맷 설정에 고려해야할 요소를 조사하였다. 조사 결과는 다음과 같다.

포맷설정을 위한 고려요소는 다음 세가지를 구분할 수 있었다.

- 활용적 요소를 정리하면 다음과 같다.

- 문서 변환과 문서 암호화
- 데이터 객체 관리
- 색인 생성 기능
- 워터마크의 활용성
- 국가 XML과 연동하는 표준 데이터 기술
- 메타데이터 관리

활용과 보존적 측면에서 공동으로 고려할 것은 행정자치부에서 권고하는 「행정기관의 전자문서시스템 규격」, 「행정기관간 전자문서유통 표준」, 「전자문서시스템과 행정정보시스템간 연계표준」

- 보존적 요소를 정리하면 다음과 같다.

- 문서 암호화
- 디지털 저작물 관리요소
- 파일 사이즈(압축 효율)

- 공통적 요소를 정리하면 다음과 같다.

- 국가 표준 권고안의 수용
- 메타데이터 연관기술
- 색인 기술

참고문헌

행정자치부. 2002. 「행정기관의 전자문서시스템 규격」, 동부처

행정자치부. 2002. 전자문서유통 표준」, 동부처

행정자치부. 2002. 「전자문서시스템과 행정정보시스템간 연계표준」, 동부처

오동근. 2004 「메타데이터의 이해」, 대구 : 태일사

이구민. 2003. PDF-XML. 한빛소프트

박진덕. 2002. Adobe Magazine. April.

손원성 외. 2001. “전자책 관련기술과 발전동향.” 한국멀티미디어학회지. 제5권1호. : 25-40