

HTML5기반의 PDF Viewer 제공을 위한 Middleware 구조 제안

정규병*, 이기범*, 박범**, 황성진***, 오상윤*

*아주대학교 정보컴퓨터공학과, **아주대학교 산업정보시스템공학부, *** (주)휴민텍
e-mail:{rbquddl16, rhlehwl, ppark}@ajou.ac.kr, seongjns@humintec.com, syoh@ajou.ac.kr

A Middleware Architecture for HTML5-Based PDF Viewer

Kyu-Byeong Jeong*, Gi-Beum Lee*, Peom Park**, Seong-Jin Hwang***, Sangyoon Oh*

*Department of Computer and Information Engineering, Ajou University,

**Department of Industrial and Information Systems Engineering, Ajou University,

***HuminTec Co.,Ltd.

요 약

Paperless 회의 및 행정에 대한 요구가 제기되면서 ISO 14001과 같은 표준들의 제안과 함께 이 요구에 대한 많은 연구와 구현이 제안되고 있다. 최근에는 다양한 플랫폼의 스마트 기기들의 보급이 확산됨에 따라 이 기기들의 플랫폼에 맞는 전자문서서식시스템의 요구도 높아지고 있다. 본 논문에서는 제시된 여러 연구들의 효율성 검토 결과를 기반으로 사용자 중심의 전자문서 공유 및 관리 기능을 제공하는 미들웨어의 구조를 제안한다. 제안하는 미들웨어는 Content Adaptation, 즉 File Format 및 Size Converting, 스마트기기내에서의 전자서식문서 보안 처리기능, 데이터 관리기능을 포함하는 PDF Viewer를 제공하며 최근 표준화 적용이 진행 중인 HTML5를 기반으로 하여 호환성과 적용성을 높였다. 제안된 구조와 구현방식에 대한 검증을 위해 구현 사례를 본 논문에서 제시하였다.

1. 서론

최근 국제표준화기구 기술위원회에서 제정한 ISO 14001 인증에 따른 ‘적극적인 사후관리를 위해 종이류와 토너류 절감대책 추진’이라는 정책 하에 종이 없는 회의문화를 정착하기 위한 대한민국 정부의 노력은 계속되어왔다. 종이에 쓰이는 행정비용만 줄여도 연간 오천만원 이상을 절감할 수 있다고 한다[1]. 실제로 회의에 쓰이는 종이는 회의를 위해 각 참석자들에게 회의 관련 자료로 배부되어지고, 회의가 끝나면 종이는 보관되거나 폐기된다. 이러한 회의의 반복은 자료의 보관 및 폐기 등에서 종이낭비문제를 야기한다. 따라서 종이 없는 회의문화를 정착하고 이런 문제를 해결하고자 모든 회의 관련 서류들을 전자문서로 관리하게 되었다[2]. 이로 인해 기존에 배부해야만 했던 회의관련 자료들은 모두 컴퓨터 파일로써 전자서식에 맞게 컴퓨터를 통해서 볼 수 있게 되었으며 또한 최근 전자문서 서식 시스템은 공인인증모듈과 문서위변조 방지 시스템을 통한 완전한 Paperless 전자문서 서식 시스템으로 가는 추세에 있다[3]. 하지만 스마트기기의 발전에 따라 PC뿐만이 아닌 다양한 플랫폼의 스마트기기에서도 구동이 가능한 전자문서 서식 시스템이 요구되고 있지만 다양한 스마트기기 내에서 PDF Viewer를 위해 각 플랫폼에 맞는 PDF Viewer 플러그인을 설치하면서 PDF File을 렌더링하여 보기에는 각 플랫폼에선 반드시 별도의 플러그인이 필요하다는 문제점이 있다. 이에 본 논문에서는 다양한 플랫폼의 스마트기기를 효과적으로 지원하기 위해 최근에 대두되고 있는 HTML5을 사용하여 PDF Viewer를 위한

미들웨어를 통해 Heterogeneous Device(이기종기기)의 다양한 플랫폼들 간의 협업을 지원하는 미들웨어 구조를 제안한다.

HTML5는 모바일 등의 플랫폼에 종속되지 않고 다양한 플랫폼에서 호환되는 차세대 웹 표준 규격으로 별도의 플러그인 설치 없이 플래시, 미디어 플레이어 등 다양한 웹 애플리케이션 기능을 PC나 모바일 기기 환경에서 모두 적용할 수 있다[4]. 따라서 본 논문에서는 궁극적으로 이기종기기에 대한 협업의 토대를 마련하기 위해 HTML5 기반의 PDF Viewer를 위한 미들웨어 구조를 제안하며, 여러 명의 사용자가 각기 다른 작업환경(이기종기기)에서 통합된 하나의 PDF Viewer를 사용할 수 있도록 트랜스코딩이 가능한 호환성을 제공함으로써 전자서식과 관련된 업무의 활동성을 증대 시킨다.

본 논문의 구성은 다음과 같다. 2장에서는 트랜스코딩, Content Adaption, PDF.js 등 PDF Viewer를 위한 미들웨어 구조 제안에 활용한 관련연구들에 대해 소개하고 3, 4장에서는 PDF Viewer를 위한 미들웨어의 구조에 대해 기술한다. 그리고 5장에서는 토의 및 향후 연구를 제시한다.

2. 관련연구

2.1. 트랜스코딩

트랜스코딩이란 여러 가지 다양한 장비에서 애플리케이션 디스플레이와 웹 내용을 표시하는 기술. 변환 부호화 서버 및 서비스를 통해 통신 규약, 애플리케이션, 화면 크기, 사용 언어에 관계없이 다양한 플랫폼에 알맞은 양식으로 변환한다[5]. 스마트 폰이나 개인 휴대 정보 단말기(PDA)

같은 무선 장비에서는 사용 전력, 메모리, 해상도, 화면 크기 등의 제약으로 일반적인 웹 내용을 표시하는 데 어려움이 있어 단순한 내용이나 이미지가 없는 주식 시세 같은 데이터만이 가능하다. 변환되지 않은 내용은 각 장비마다 데이터를 업데이트 할 때 마다 다시 작성해야 하는데, 그때마다 항상 오류 가능성이 존재하기 때문에 모든 장비를 동시에 업데이트할 수 없다. 따라서 이를 해결하기 위해 관련 서버나 관련 서비스를 활용한다.

2.2. Contents Adaption

특정 서비스를 목적으로 제작된 콘텐츠들이 또 다른 서비스의 용도로도 사용될 수 있게 하는 것이다. 예를 들어, 개인용 컴퓨터(PC) 기반으로 작성된 웹 콘텐츠 및 서비스를 휴대폰이나 개인 휴대 정보 단말기(PDA)와 같은 다른 서비스 환경의 단말기에서도 이용할 수 있도록 하는 것이다 [6]. 콘텐츠 적응화 방법으로는 저작 시에 하이퍼 텍스트 생성 언어(HTML), 무선 마크업 언어(WML), cHTML(Compact HTML), 확장성 생성 언어(XML) 등 다중의 플랫폼별 콘텐츠를 제작해 두어 클라이언트 요청이 들어올 때 알맞은 버전을 선택하는 방법과 적응화를 위해 필요한 콘텐츠 모델링 정보를 XML 형태로 제작하여 두는 방법, 콘텐츠 모델링 정보 없이 문서 분석과 적응화 알고리즘으로 적응화하는 방법들이 있다.

2.3. 렌더링 라이브러리 : PDF.js

PDF.js는 모질라 재단에서 PDF 파일의 안전 및 웹 표준 준수, 웹 브라우저 렌더링을 위한 HTML5로 PDF 파일을 변환하기 위한 자바 스크립트 라이브러리이다[7]. 일반적으로 PDF 문서는 브라우저에 설치된 플러그인을 통해 보거나 Adobe의 PDF Viewer를 통해 보게 되는데 PDF.js를 이용하면 별도의 플러그인 없이 PDF File을 볼 수 있고 HTML5로 렌더링해서 보여주는 방식을 통하여 구글 크롬 웹브라우저에서 웹클라이언트의 반환 메시지를 이용하여 불법 인증 및 정보를 유출하는 공격인 코드 인젝션공격을 방지하기 위해 PDF 렌더링에 대한 샌드박스를 만든 것과 달리 이런 코드 인젝션공격에 대해 안정성을 보장할 수 있어 효율적이다.

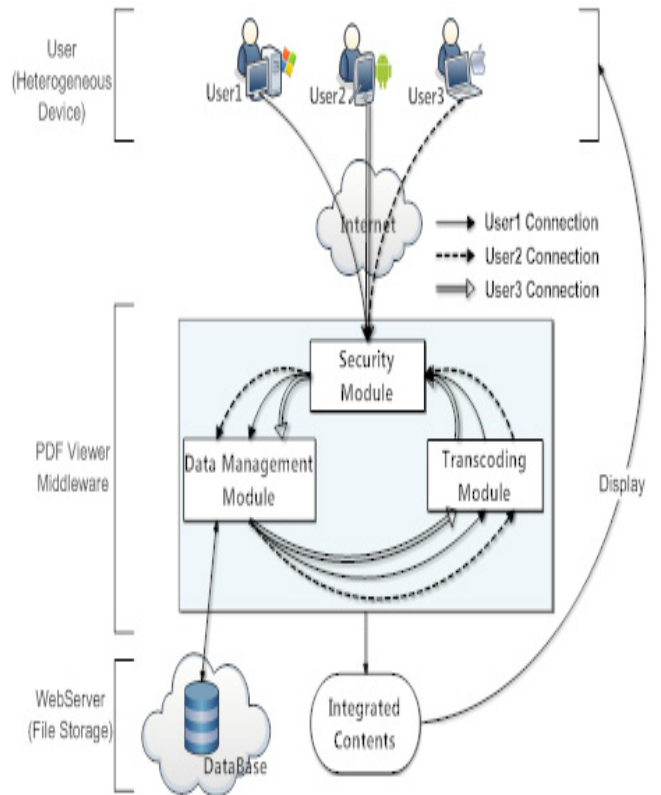
3. PDF Viewer를 위한 미들웨어 제안

3.1. Overview

서론에서 제시한 HTML5기반의 PDF Viewer를 위한 미들웨어 모델은 그림 1에 나타내었다. 미들웨어 구조는 크게 사용자의 접속, 파일변환, 보안모듈 등 전반적인 PDF Viewer의 역할을 총괄하는 PDF Viewer 미들웨어와 다수 사용자들의 각기 다른 시스템 및 PDF File이 저장되어 있는 웹서버의 데이터베이스로 구성된다.

구조를 살펴보면 사용자들은 서버에 접속할 때, 공통적으로 플랫폼에 독립적인 HTML5기반 웹 브라우저를 이용하여 원하는 데이터를 가져오기 위해 미들웨어를 사용한다. 이를 바탕으로 미들웨어에서는 서로 다른 플랫폼을 사용하는 다수의 사용자 기기에서 호환되기 위하여 Content Adaptation을 적용한다. 보안모듈에서는 각 기기의 플랫폼에

접근하여 기기의 콘텐츠 프로파일과 보안모듈을 가져오게 되며, 이 정보를 기반으로 데이터 관리모듈과 트랜스코딩 모듈이 협업하여 동일 콘텐츠의 기기종 display가 가능하도록 지원한다.



<그림 1> PDF Viewer를 위한 미들웨어 제안 구조

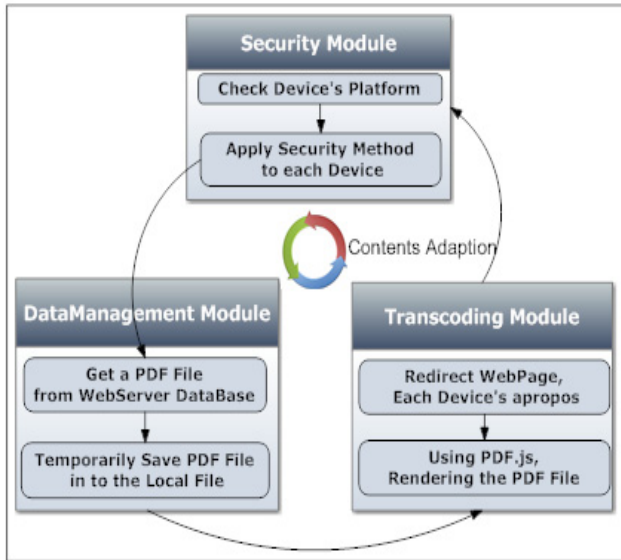
즉, 여러 명의 사용자가 각기 별개의 작업환경(이종기기)에서 통합된 하나의 시스템을 동시에 수행할 수 있도록 해주는 미들웨어를 제안한다. 또한 웹 서버의 데이터베이스와 연결하여 데이터베이스로부터 사용자가 원하는 데이터를 찾을 때, 사용자의 요청 내용에 따라 각 사용자들의 플랫폼에 맞도록 웹브라우저 내 데이터를 나타내게 하여, 각 플랫폼에 맞는 전자서식시스템의 공인인증모듈과 문서위변조방지모듈을 사용하도록 하였다. 더 구체적인 내용은 다음 3.2절에서 PDF Viewer를 위한 미들웨어의 구조와 함께 기술하겠다.

3.2. PDF Viewer를 위한 미들웨어

3.1절에서 제시한 PDF Viewer를 위한 미들웨어 구조는 그림 2에 나타내었다. 미들웨어 구조는 크게 보안모듈, 데이터관리 모듈, 트랜스코딩 모듈로 구성된다.

각기 다른 이 모듈들은 Contents Adaption을 통해 사용자들이 사용하는 다양한 플랫폼에 따라 적합한 각 모듈들의 기능을 찾아 사용자가 원하는 데이터 형식을 제공하게 된다. 따라서 사용자가 PDF File에 대한 데이터를 요청하면 미들웨어에서 내 보안모듈에서는 각 사용자의 플랫폼에 맞는 보안기능을 제시하게 된다. 그리고 데이터관리 모듈에서 사용자가 요청한 데이터를 웹서버 데이터베이스로부터 가져와 Local File로 저장 한 뒤, 보안모듈에서 제시한 데이

터에 대한 설명과 서명된 데이터에 대한 검증이 이루어진 보안기능을 이용하여 인증 및 접근제어에 대한 처리를 행한다. 이 후 트랜스코딩 모듈에서 보안모듈에서 확인한 사용자의 플랫폼에 맞게 웹 페이지를 포워딩하며, PDF.js를 이용, PDF 파일의 안전 및 웹 표준 준수 웹 브라우저 렌더링을 위한 HTML5로 PDF 파일을 변환한다.



<그림 2> PDF Viewer를 위한 미들웨어 구조

이로 인해 네트워크상에서 받은 데이터를 각 스마트기기 내부에서 자체적으로 트랜스코딩을 처리하는 기능을 네트워크상의 미들웨어에 통합하면서, 기존 방법과 비교하였을 때 사용자인증, 접근제어 및 데이터 관리, 표기 등의 처리를 연쇄적으로 행할 수 있어, 시스템의 복잡한 구조를 완화함과 동시에 미들웨어에서 각 기기들의 플랫폼에 맞는 전자문서서식시스템을 트랜스코딩하여 기기종기기간의 호환성 및 적용성을 높이게 되었다.

4. PDF Viewer 미들웨어 구현

4.1. 개발 환경

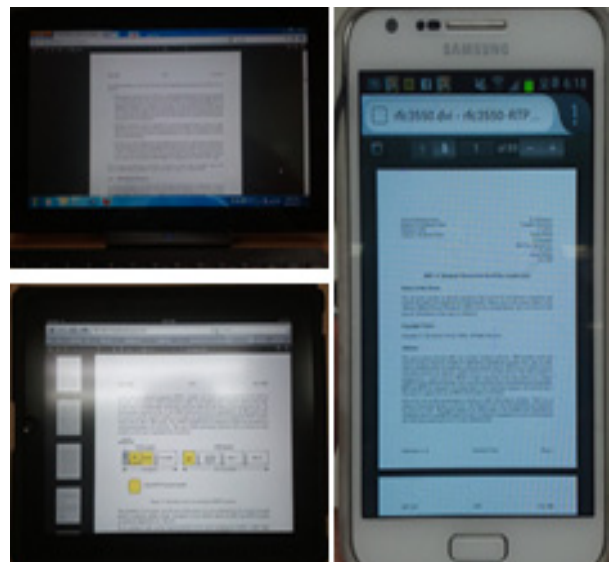
표 1 에는 PDF Viewer를 위한 미들웨어를 구현한 환경에 대해 서술하였다. 본 미들웨어 시스템은 .NET(IIS 7.0)을 이용한 웹서비스를 구축하였다. 웹서버는 JSP를 기반으로 구성하였고, HTML5 환경에서 Javascript기반의 PDF Viewer를 하기 위해 pdf.js를 사용하였다.

구분	환경
서버	.NET Web Server(IIS 7.0)
웹 언어	JSP, HTML5, Javascript
개발 환경	Eclipse Indigo 3.7.2
테스트환경	노트북(window 7) / 갤럭시탭 10.1 안드로이드폰(진저브레드 2.3.4) 아이패드(iOS 4.2)

<표 1> 구현 환경

4.2. 프로토타입 구현 결과

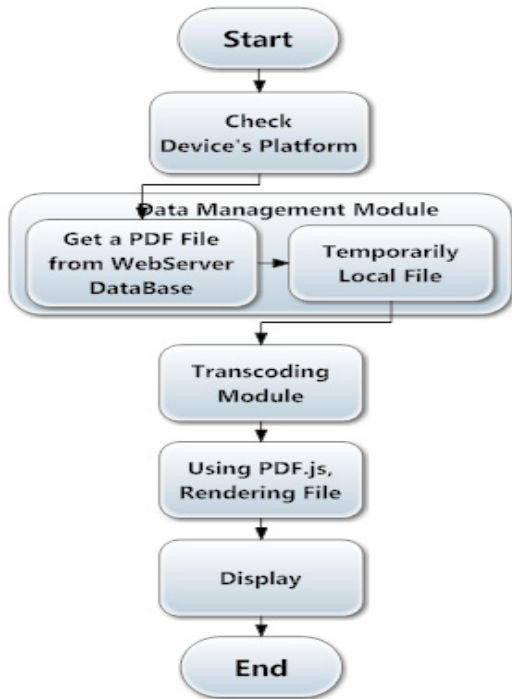
본 논문에서 제안한 PDF Viewer를 위한 미들웨어구조와 달리 프로토타입 구현 결과에서는 보안 모듈 내 각각의 기기에 따른 보안기능을 적용시키는 점이 구현이 되지 않았다. 이는 현재 본 미들웨어 서비스에서 필요한 HTML5기반 PDF.js 기능을 Firefox, Safari, Opera 브라우저에서만 지원하고 있고 Internet Explorer 및 Chrome 브라우저에서는 PDF.js기능을 제공하고 있지 않아, 각각의 브라우저에 맞는 보안기능을 적용시키지 못하기 때문이다. 또한 각 스마트기기 내부에서 Task들을 처리하는 방법이 아닌 미들웨어 레벨에서 다른 모듈들과 함께 보안 모듈을 실행시킴으로써 하나의 통합된 미들웨어에서 모든 Task를 처리하는 것이 효과적이기 때문에 프로토타입 모델에서는 궁극적인 모델에 비해 구조적인 형태에서 다른 모습을 갖게 되었다. 하지만 현재 Internet Explorer 및 Chrome에서도 HTML5 표준이 진행 중인 관계로 PDF.js 기능도 추후 지원함에 따라 각각의 브라우저에 따른 보안기능을 하나의 통합된 미들웨어에서 처리할 수 있게 되어 본 프로토타입의 한계점을 극복해나갈 수 있다.



<그림 3> 프로토타입 구현 결과 화면

그림 3은 PDF Viewer 미들웨어 중 웹 서버 데이터베이스에서 관련 PDF File Data를 받는 데이터관리 모듈과 각 사용자의 스마트기기 플랫폼에 따라 트랜스코딩 해주는 모듈의 프로토타입 모델을 구현한 사진이다. 각각 노트북, 안드로이드 폰, 아이패드를 이용해서 PDF File을 보는 모습을 보여주는데 왼쪽 위에서부터 노트북에서 Firefox 브라우저를 이용한 스크린, 아이패드에서 Safari 브라우저를 이용한 스크린, 오른쪽은 안드로이드 폰에서 Firefox 브라우저를 이용한 스크린을 나타낸다.

앞에서 제시한 HTML5기반의 PDF Viewer를 위한 미들웨어 프로토타입 모델의 동작하는 순서를 사용자관점에서 나타내기 위하여 그림 4에 플로우차트를 이용하여 나타내었다.



<그림 4> 프로토타입 플로우차트

처음에 PDF Viewer를 위한 미들웨어는 보안모듈 내부에서 'mobile_chk' flag값을 이용하여 사용자가 어떤 플랫폼의 기기에서 PDF File 데이터를 요청하였는지 또는 웹 브라우저에 PC로 접속했는지 모바일 기기로 접속했는지 여부를 확인한다. 그리고 데이터관리 모듈을 통해 웹서버 데이터베이스로부터 사용자로부터 요구받은 PDF File을 가져와 임시파일에 저장시켜 놓은 뒤, 이 전 단계에서 확인했던 사용자 기기의 플랫폼에 따라 트랜스코딩 모듈을 통해 각 사용자의 플랫폼에 따라 데이터를 트랜스코딩하며, 만일 사용자가 모바일 기기로 접속하였을 때 PC에서 보는 웹 페이지와 다르게 따로 개발 및 구성해 놓은 모바일 웹 페이지로 포워딩하게 구현하였다. 그리고 웹 브라우저 렌더링을 통해 PDF File을 변환하기 위해 HTML5 기반인 PDF.js를 이용하였다.

5. 토의 및 향후 연구

본 논문에서 제안하고 있는 HTML5기반의 PDF Viewer를 위한 미들웨어 구조는 Contents Adaption을 활용, 각기 다른 사용자들이 사용하는 플랫폼에 따른 각 모듈들의 기능을 찾아 사용자가 원하는 데이터에 도달할 수 있게 한다. 또한 HTML5 기반의 PDF.js를 활용하여, 사용자의 다양한 플랫폼에서 제공되는 브라우저를 통해 이기종간의 Collaboration이 가능하다. 이에 따라 각 사용자의 스마트기기에서는 미들웨어에서 제공하는 하나의 통합된 전자서식문서로 종이 없는 회의, 즉 Paperless 전자문서 서식환경을 조성할 수 있으며, 기존 전자문서 서식환경에서는 PC를 이용한 전자문서조회로 국한되어 있지만 스마트기기도 이용 가능함으로써 전자문자 서식수단이 확장되었다. 또한 스마트기기 내부적으로 전자서식문서를 트랜스코딩하는 것을 네트워크상의 미들웨어와 통합함으로써 스마트

기기로 데이터 접근에 문제시 되었던 보안문제를 Contents Adaption을 통해 각 플랫폼에 맞는 보안모듈을 사용하게 함으로써 하나의 통합인증서버에서 모든 Task를 처리할 수 있어 효과적으로 되었다.

지속가능성	전자문서 효과	사 례
경제적성과	종이서류 비용 절감 (연 매출 1~3%)	전자무역망을 통한 전자문서 및 수출입절차 개선, 비용절감
환경적성과	지구온난화 방지, 탄소절감	종이 1톤당 30년생 나무 17그루, 물 25만 리터 절감
사회적성과	기업의 대내외 투명성제고	종이문서의 전자화 사업을 통한 일자리 및 고용창출

<표 2> 전자문서를 통한 지속가능경영 성과

따라서 본 제안 모델을 이용함으로써 표 2와 같이 종이 없는 회의환경을 조성하여 경제적, 환경적, 사회적 측면에서의 비용을 절감할 뿐만 아니라 회의문서의 출력 필요성을 감소시켜 Paperless를 통한 환경 친화적이고 효율적인 시스템 모델을 구축할 수 있으리라 판단된다[8].

6. Acknowledgement

본 연구는 지식경제부 및 정보통신산업진흥원의 서울 어코드 활성화 지원 사업(아주대학교 소프트웨어융합학과)의 산학프로젝트 결과로 수행되었으며, 'N-스크린과 멀티플랫폼을 지원하는 효율적인 양방향 회의시스템'에 대한 연구결과(한국정보과학회, 하계컴퓨터통신 워크샵, pp. 49-51, 2012)를 일부 소개하였으며, (주)휴민텍 산학프로젝트 'E-IRB 회의시스템'의 일환으로 수행되었음.

참고문헌

- [1] 산업자원부 미래생활산업본부 디지털 혁신팀, 종이 없는(Paperless)시대로 한걸음 성큼, 한국개발연구원, 2007
- [2] 양영모, Paperless 환경을 위한 전자문서 포맷 PDF, 한국기록관리협회 제10회 기록관리 워크샵, 2009.9, 109-120
- [3] 지식경제부 성장동력실 정보통신산업정책관 정보통신활용과, 전자문서로 녹색경제를 선도하자, 한국개발연구원, 2010
- [4] HTML5 working draft, <http://www.w3.org/TR/2012/WD-html5-20120329/>
- [5] wikipedia, <http://en.wikipedia.org/wiki/transcoding>
- [6] R.Bagrodia 외 11명, iMash: Interactive Mobile Application Session Handoff, UCLA Computer Science, Los Angeles, CA 90095-1596
- [7] pdf.js, <https://github.com/mozilla/pdf.js>
- [8] 안세기, Paperless를 통한 저탄소 녹색성장 “지구사랑 프로젝트” 공공 및 민간부문 Paperless 적용사례, KNET 공인전자문서센터