

OFFICE OPEN XML 개요



ECMA TC45

TOM NGO(NEXTPAGE), 편집자

1 소개

Office Open XML(OpenXML)은 워드 프로세싱 문서, 프레젠테이션 및 스프레드시트에 대해 제안된 오픈 표준으로서, 여러 플랫폼의 다양한 응용 프로그램으로 자유롭게 구현할 수 있습니다. OpenXML에 대한 게시는 해당 형식을 사용할 수 있는 응용 프로그램을 구현하려는 조직, 이러한 소프트웨어를 구입하려는 상업 기관과 정부 기관 및 OpenXML 형식을 가르치는 교사나 저자에게 도움이 됩니다. 따라서 모든 사용자는 문서에 대한 안정성, 보존성, 상호 운용성 및 지속적인 발전을 비롯한 XML 표준의 이점을 누릴 수 있게 됩니다.

OpenXML 표준화는 Apple, Barclays Capital, BP, 영국 국립 도서관, Essilor, Intel, Microsoft, NextPage, Novell, Statoil, Toshiba 및 미 의회 도서관의 대표로 구성된 Ecma의 TC45(Technical Committee 45)를 통해 Ecma International에서 수행되었습니다(1).

본 백서에서는 OpenXML에 대해 요약하며 그 내용은 다음과 같습니다.

- OpenXML의 목적과 OpenXML 사양의 구조를 이해합니다.
- OpenXML의 속성인 이전 버전과의 호환성, 보존성, 확장성, 사용자 지정 스키마, 하위 집합, 여러 플랫폼, 국가별 기능 및 내게 필요한 옵션을 다루는 방식을 확인합니다.
- OpenXML 파일에서 높은 수준의 구조를 수행하는 방법 및 보다 자세한 내용이 필요한 사양의 일부분을 신속하게 탐색하는 방법을 알아봅니다.

2 표준의 목적

OpenXML은 처음부터 Microsoft Corporation에서 정의한 이진 형식으로 인코딩된 기존의 워드 프로세서 문서, 프레젠테이션 및 스프레드시트 자료를 정확하게 표현할 수 있도록 하기 위해 설계되었습니다. 표준화 프로세스는 이러한 자료의 존속과 이들의 확장, 상세 문서의 제공 및 상호 운용성의 활성화를 나타내는 데 필요한 기능을 XML에 반영하는 것으로 구성되어 있습니다. 작성 당시 4억 명 이상의 사용자가 이진 형식으로 40억 개 이상의 문서를 생성하고 매년 수십 억 개의 문서가 작성되고 있는 것으로 추산됩니다.

이러한 파일의 원래 이진 형식은 공간을 중시하고 구문 분석 시간이 사용자 경험에 상당한 영향을 미치던 시기에 만들어졌습니다. 또한 Microsoft® Office® 응용 프로그램에서 사용되는 메모리 내 데이터 구조의 직접적인 직렬화에 기반을 두었습니다. 현대의 하드웨어, 네트워크 및 표준 인프라(특히 XML)는 여러 플랫폼에서 여러 공급업체별 구현을 지원하고 발전을 고려하는 새로운 디자인을 허용합니다.

이러한 기술적 진보에 따라 원래 문서 편집 프로그램의 단순한 세계에서는 계획되지 않았던 새로운 범위의 응용 프로그램을 포함할 만큼 시장이 다양해졌습니다. 이러한 새 응용 프로그램의 기능은 다음과 같습니다.

- 비즈니스 데이터에서 문서를 자동으로 생성
- 문서에서 비즈니스 데이터를 추출하여 비즈니스 응용 프로그램에 제공
- 편집 가능성을 유지하면서 문서의 작은 하위 집합에 작용하는 제한된 작업 수행
- 시각 장애인과 같이 특수한 요구 사항이 있는 사용자를 위한 내게 필요한 옵션 제공
- 모바일 장치를 포함한 다양한 하드웨어에서 실행

가장 난해한 문제는 장기 보존과 관련된 문제일 것입니다. 정보의 양은 기하급수적으로 증가하고 있습니다. 그러나 이러한 정보를 디지털 표현을 사용하여 인코딩해 왔으며 이를 만든 프로그램과는 서로 불가분의 관계에 있으므로 10년 혹은 20년 후에는 이러한 정보가 상당한 손실 없이 매우 읽기 어려워질 것이 자명합니다. 따라서 기존 문서든 새 문서든 이러한 문서에서 재정 및 지적 투자를 보존하는 일은 매우 긴급한 문제가 되었습니다.

이진 형식의 광범위한 채택, 기술적 진보, 다양한 응용 프로그램을 요구하는 시장 논리 및 장기 보존의 어려움 가중이라는 네 가지 압력이 대두되면서 OpenXML 형식을 정의하고 수억 개의 문서를 가능한 한 적은 손실을 통해 이 형식으로 마이그레이션할 필요성이 제기되었습니다. 또한 해당 OpenXML 형식을 표준화하고 오랜 기간 동안 유지 관리하면 모든 조직에서 사양의 지속적인 안정성을 안심하고 신뢰할 수 있으며 표준 절차에 따라 향후의 발전에 대한 견제와 균형 유지를 확신할 수 있는 환경을 만들 수 있습니다.

각종 문서 표준 및 사양에는 HTML, XHTML, PDF와 PDF의 하위 집합, ODF, DocBook, DITA, RTF 등이 있습니다. TIFF/IT, TIFF/EP, JPEG 2000 및 PNG와 같이 비트맵 이미지를 나타내는 수많은 표준과 마찬가지로 이들은 각기 다른 목적을 위해 생성되었습니다. OpenXML은 기존 문서 자료에 표시된 기능을 처리하는 표준에 대한 요구를 충족시킵니다. 지금까지 알려진 바로는 이진 형식의 모든 기능을 지원하는 유일한 XML 문서 형식은 OpenXML 뿐입니다.

3 표준의 구조

OpenXML은 워드 프로세싱, 프레젠테이션 및 스프레드시트 문서의 형식을 정의합니다. 각 문서 유형은 WordprocessingML, PresentationML 또는 SpreadsheetML과 같은 기본 표시 언어를 통해 지정됩니다. 포함 메커니즘을 통해 이 세 종류 중 한 가지 문서에 다른 기본 표시 언어로 된 내용과 지원하는 여러 표시 언어로 된 내용을 포함할 수 있습니다.

사양에는 규정 자료(OpenXML을 정의하는 자료) 및 정보 자료(독자의 이해를 돕지만 규정은 아닌 자료)가 들어 있으며, 다양한 독자의 요구를 충족시키기 위해 다음과 같이 여러 부로 나뉘어 구성되어 있습니다.

- 1부 – 기본 사항
165 페이지
 - 단어, 표기 규칙 및 약어를 정의합니다.
 - 세 가지 기본 표시 언어 및 지원하는 표시 언어에 대해 요약합니다.
 - 규칙의 조항을 설정하고 상호 운용성에 대한 지침을 제공합니다.
 - 각 문서 유형에 적용되는 오픈 패키징 규칙의 제약 조건을 설명합니다.

- 2부 – 오픈 패키징
규칙
125 페이지
 - OPC(오픈 패키징 규칙)를 정의합니다. 모든 OpenXML 파일은 파트라고 하는 바이트 스트림의 모음으로 구성되며 패키지라고 하는 컨테이너로 결합됩니다. 패키징 형식은 OPC에 의해 정의됩니다.
 - Zip 파일 형식을 사용하는 OPC에 대해 권장되는 물리적 구현을 설명합니다.
 - 전자 양식으로만 발행되는 부록에서 OPC의 XML 스키마를 XSD(XML 스키마 정의)(2)로 선언합니다. 부록에는 RELAX NG(ISO/IEC 19757-2)(3)를 사용한 스키마의 비규정적 표현이 포함됩니다.

- 3부 – 입문서
466 페이지
 - 예제와 다이어그램을 통해 요소를 표시하고 컨텍스트를 제공함으로써 각 표시 언어의 기능을 소개합니다. 이 부분은 정보 자료(규정이 아님)에 해당합니다.
 - 비즈니스 데이터와의 통합을 지원하기 위해 패키지 내에 사용자 지정 XML 데이터를 저장하는 기능을 설명합니다.

- 4부 – 표시
언어 참조
5756 페이지
 - 모든 요소와 특성, 요소에 대한 부모/자식 관계의 계층 및 추가 의미를 적절하게 정의합니다. 이 부분은 요소 또는 특성에 대한 완전한 세부 정보가 필요한 경우에 참조로 사용됩니다.

- 사용자 지정 XML 데이터의 저장 기능을 정의합니다.
 - 전자 양식으로만 발행되는 부록에서 표시 언어의 XML 스키마를 XSD(2)로 선언합니다. 부록에서는 RELAX NG(ISO/IEC 19757-2)(3)를 사용하여 이 내용을 비규정적으로 설명합니다.
- 5부 – 표시
호환성 및
확장성
34 페이지
- OpenXML 문서 확장명의 기능을 설명합니다.
 - 상호 운용 가능한 다른 확장명을 사용하는 응용 프로그램으로 요소 및 특성을 지정합니다.
 - NVDL(ISO/IEC 19757-4)(4)을 사용하여 확장성 규칙을 설명합니다.

문서를 전체적으로 읽고 탐색하기 쉽게 하기 위해 전자 버전에는 많은 내부 활성 링크가 마련되어 있습니다. 특히 4부에는 부모 및 자식 요소에 전체적으로 링크가 연결되어 있습니다.

4 표준의 속성

이 섹션에서는 OpenXML의 일부 고급 속성에 대한 설명을 통해 OpenXML을 조사할 준비를 하도록 합니다. 각 하위 섹션에서는 이러한 속성 중 하나를 설명하고 OpenXML 내의 특정 기능을 참조합니다.

- “상호 운용성”에서는 OpenXML이 어떻게 독립적 형식, 기능 및 런타임 환경에 대해 독립적인지를 설명하여 개발자가 다양한 선택을 할 수 있도록 합니다.
- “국가별 기능”에서는 OpenXML이 모든 주요 언어 그룹을 지원하는 몇 가지 대표적인 방법에 대해 설명합니다.
- “개발자 채용 장벽 완화”, “간결성” 및 “모듈화”에서는 여러 단체에서 OpenXML을 구현하는 데 따른 실질적인 장애를 방지 또는 제거하는 특정 방식으로 곡선 배우기, 최소 기능 집합 및 성능을 나열합니다.
- “고성능 마이그레이션”에서는 OpenXML이 기존 문서와 새 문서에서 원래 작성자의 의도를 완벽하게 포함하면서 가장 중요한 목적인 정보 보존을 어떻게 충족시키는지 설명합니다.
- “비즈니스 데이터와 통합”에서는 OpenXML이 사용자 지정 스키마에서 비즈니스 정보를 어떻게 통합하여 생산성 응용 프로그램과 정보 시스템 간의 정보를 통합하고 다시 사용할 수 있도록 하는지를 설명합니다.
- “혁신을 위한 여지”에서는 더 나은 확장성 메커니즘을 정의하고 서로 다른 기능 집합을 포함하는 응용 프로그램 간의 상호 운용성을 제공함으로써 OpenXML의 미래에 대한 대비 방법을 설명합니다.

이 섹션과 문서의 나머지 부분에서는 OpenXML에 대해 주제별로 안내합니다. 사양에 대한 참조는 모두 §부:섹션.하위 섹션의 형식으로 되어 있습니다. 예를 들어 §1:2.5는 사양의 1부 섹션 2.5를 의미합니다. 문서 내의 다른 제목에 대한 참고 자료는 이름으로 표시합니다.

4.1 상호 운용성

개발자는 여러 플랫폼에서 OpenXML을 사용하고 생성하는 응용 프로그램을 작성할 수 있습니다.

우선 OpenXML의 상호 운용성은 Ecma TC45 기술 위원회(1) 구성원의 다양한 배경과 집단적 이해를 바탕으로 사양의 광범위한 기여, 수정 및 검토를 통해 이루어졌습니다. 해당 대표들은 다음과 같습니다.

- 여러 가지 운영 체제(Linux, MacOS 및 Windows)를 사용하고 다양한 목적에 따라 OpenXML을 사용하는 공급업체(Apple, Intel, Microsoft, NextPage, Novell 및 Toshiba)
- 미션 크리티컬한 트랜잭션 시스템을 비롯해 기존 콘텐츠에 막대한 투자를 한 회사(BP, Barclays Capital, Essilor, Statoil)
- 문서 보존에 가장 직접적인 관심을 가진 영국 국립 도서관 및 미국 의회 도서관

준비 기간 동안 위원회 구성원들은 정책, 명확성, 의미 및 환경 관련 가망 종속 요인을 고려한 수백 개의 문제를 제기하고 해결했습니다. 대표적인 문제점 및 기타 활동은 다음과 같습니다.

- 원래 이진 형식이 가진 메커니즘으로부터 플랫폼 독립성을 지원하는 기능
- 규칙의 조항
- 스키마의 콘텐츠
- RELAX NG(ISO/IEC 19757-2) 및 NVDL(ISO/IEC 19757-4)(4)을 사용하는 스키마와 확장성 메커니즘에 대한 대안
- 스키마를 자동으로 분석하고 시각화하기 위한 도구의 개발
- 국가별 기능
- 사양의 일부분을 구현해 본 결과 대다수의 경우에서 발견된 사양 전반에 걸친 설명의 완전성, 정확성 및 명확성

이 하위 섹션의 나머지 부분에서는 상호 운용성을 위해 원래 이진 형식에서 벗어난 OpenXML의 특정 영역에 대해 강조하고 있습니다.

상호 운용성의 주요 요구 사항 중 하나는 원본 콘텐츠의 특정 유형으로부터의 독립성입니다.

- OpenXML에는 이미지, 오디오 또는 비디오 유형에 대한 제한 없이 모두 포함될 수 있습니다. 예를 들어 GIF, PNG, TIFF, PICT, JPEG 또는 다른 어떤 유형의 이미지도 포함할 수 있습니다(§1:14.2.12).
- Java 또는 ActiveX 등 모든 유형의 포함된 컨트롤을 사용할 수 있습니다(§1:15.2.8).
- WordprocessingML 글꼴 사양에는 원래 글꼴을 사용할 수 없는 경우 대체 글꼴을 찾으려 도와줄 글꼴 메트릭 및 PANOSE 정보가 포함될 수 있습니다(§3:2.10.5).

또한 OpenXML은 문서를 작성한 응용 프로그램의 런타임 환경에 종속되지 않습니다.

- 이에 대한 전형적인 예는 디스플레이 표면의 일부분에 대한 이미지를 생성하는 외부 컨트롤 또는 응용 프로그램에서 발생합니다. 이러한 컨트롤 또는 응용 프로그램을 사용할 수 없거나 주어진 런타임 환경에서 실행할 수 없는 경우를 대비해 문서 파일에 이미지 표시를 포함시킬 수 있습니다. 이러한 메커니즘은 이전 이진 형식에도 존재합니다.
- OpenXML에서는 ACB(콘텐츠 차단 대안)(§3:2.18.4)라고 하는 보다 일반적인 메커니즘을 소개합니다. 이 메커니즘은 생성한 응용 프로그램에서 작성한 내용을 해석할 능력이 없는 응용 프로그램을 사용하는 다양한 상황에서 사용할 수 있으며 일반적으로 확장성 컨텍스트에서 사용됩니다. 이 메커니즘은 “혁신을 위한 여지” 하위 섹션에서 자세하게 설명합니다.
- OpenXML은 문서를 작성한 환경에서는 의미가 있지만 사용하는 환경에는 의미가 없는 매개 변수에 종속되지 않습니다. 예를 들어 CT_SYSCOLOR 매개 변수는 작성 환경에서의 색상표 색인입니다. 서로 다른 사용 환경에 대한 이식성을 지원하기 위해 PresentationML에서는 작성자가 문서 작성 시 사용한 시스템 색상을 캐시할 수 있습니다.

결국 가장 기본적으로 Office OpenXML은 XML(5) 및 XML 네임스페이스(6)와 같은 오픈 W3C 표준을 준수합니다. 이 사실만으로도 이러한 오픈 표준을 따르는 모든 플랫폼과 운영 체제에서 기본적인 수준의 상호 운용이 가능합니다.

4.2 국가별 기능

OpenXML은 아랍어, 중국어(세 가지 변이체), 히브리어, 힌디어, 일본어, 한국어, 러시아어, 터키어 등과 같은 다양한 언어에 필요한 국가별 기능을 지원합니다.

OpenXML도 XML이므로 기본적으로 유니코드를 지원합니다. 이와 더불어 OpenXML에는 수년에 걸쳐 구체화된 다양한 국가별 기능 집합이 있습니다. 국가별 기능의 대표적인 목록은 다음과 같습니다.

텍스트 방향: OpenXML은 LTR(왼쪽에서 오른쪽) 및 RTL(오른쪽에서 왼쪽) 언어를 지원합니다. 또한 아랍어, 페르시아어, 우르두어, 히브리어, 이디시어 등과 같이 원래는 오른쪽에서 왼쪽 방향으로 쓰지만 왼쪽에서 오른쪽 방향으로 된 텍스트가 포함된 세그먼트가 들어 있는 경우와 같은 양방향(“BiDi”) 언어도 지원합니다. WordprocessingML에서는 단락 수준(§4:2.3.1.6) 및 단락 내 런 수준(§4:2.3.2.28)의 텍스트 방향을 모두 제어할 수 있습니다. 마찬가지로 DrawingML 텍스트도 본문 수준(§4:5.1.5.1.1), 단락 수준(§4:5.1.5.2.2) 및 번호 매기기(§4.5.1.5.4) 내에서 방향을 제어할 수 있습니다.

텍스트 흐름: WordprocessingML에서는 텍스트 흐름의 방향을 섹션이나 표 수준(§4:2.3.1.41) 또는 단락 수준(§4:2.3.2.28)에서 제어할 수 있습니다. 섹션 및 표 수준에서 텍스트 흐름은 세로 및 가로 방향으로 제어할 수 있습니다. 따라서 OpenXML은 모든 잠재적 텍스트 레이아웃(예: 위에서 아래로 입력되는 세로 방향 텍스트가 왼쪽부터 차례로 나열되는 몽골어의 경우)을 지원할 수 있습니다. 이것은 목록, 표 및 기타 표현 요소의 레이아웃에 영향을 줍니다. DrawingML도 단락 및 런 수준에서 채워쓰기 설정을 이용하여 텍스트는 가로로, 숫자는 세로로 입력되도록 합니다(§4:5.1.5.2.3, §4:5.1.5.3.9). WordprocessingML(§4:2.3.1.16) 및 PresentationML(§4:4.3.1.15)에서는 한 행의 텍스트를 시작하고 끝낼 때 사용할 수 있는 문자를 지정하는 금칙 처리 설정을 사용하도록 문자 흐름을 지정할 수 있습니다.

숫자 표시: WordprocessingML의 필드 서식 지정(§4:2.16.4.3), WordprocessingML의 단락/목록 번호 매기기(§4:2.9) 및 DrawingML의 번호 매기기(§4:5.1.5.4, §4:5.1.12.61)의 경우 히라가나, 아랍어, 아브자드, 태국어, 기수 텍스트(예: “일백삼십이”), 중국어, 한국어(초성 또는 가나다), 히브리어, 힌디어, 일본어, 로마어, 베트남어 등 수십 개의 숫자 서식 중 하나를 사용하여 숫자의 서식을 지정할 수 있습니다. 또한 이 기능은 임의의 소수점 값(예: “1.00” 및 “1,00”)을 지원하고 구분 기호를 표시합니다. 국제화된 숫자 서식은 SpreadsheetML에서 특히 효과적이며, 셀 서식(§4:3.8.30)과 외부 데이터에 대한 참조(§4.3.13.12)에 이러한 기능이 모두 지원됩니다.

데이터 표시: WordprocessingML(§4:2.18.7) 및 SpreadsheetML(§4:3.18.5)에서 달력 날짜는 그레고리식(세 가지 변형), 히브리식, 회교식, 일본식(황제 시대), 한국식(단기), 사카식, 대만식 및 태국식을 사용하여 작성할 수 있습니다.

수식: SpreadsheetML의 수식 사양에는 BAHTTEXT(§4:3.17.7.22), JIS(§4:3.17.7.185), ASC(§4:3.17.7.11) 등의 여러 가지 국제화 관련 변환 함수가 제공됩니다.

언어 식별자: WordprocessingML(§4:2.3.2.18) 및 DrawingML(§4:5.1.5.3)에서 모든 단락과 런은 언어 식별자로 태그를 지정할 수 있으므로 응용 프로그램에서 적절한 언어 교정 도구와 기타 언어별 기능을 선택할 수 있습니다. 각 언어에 대한 식별자 외에도 OpenXML은 문자 집합, 글꼴 패밀리 및 PANOSE 값의 명명을 지원하여 로컬 지원이 제공되지 않는 경우 응용 프로그램이 문자의 적절한 하위 집합을 선택하도록 합니다.

4.3 개발자 채용 장벽 완화

숙련된 개발자라면 사양을 몇 시간만 읽고도 곧바로 간단한 OpenXML 준수 응용 프로그램을 작성할 수 있습니다.

사양에는 광범위한 기능 집합이 규정되어 있지만 OpenXML 준수 응용 프로그램에서 사양에 있는 모든 기능을 지원할 필요는 없습니다. 규칙 설명(§1:2)에서는 단지 규칙을 준수하는 소비자가 “원하는 문서 유형의 규칙에 맞는 문서를 거부하지 않고” 규칙을 준수하는 작성자는 “규칙에 맞는 문서를 생산할 수 있도록” 하고 있습니다(§1:2.5). 또한 요소 의미의 역할을 설명하는 상호 운용성 지침도 제공되어 있습니다(§1:2.6).

준수 응용 프로그램은 기능에 상당한 초점을 두었을 수 있습니다. 워드 프로세싱 문서 모음에서 저작권 표시만 업데이트하는 배치 프로세서나 슬라이드 프레젠테이션을 충분히 이해하여 사용자가 슬라이드를 탐색할 때 텍스트 콘텐츠를 오디오로 렌더링하는 텍스트 음성 변환 프로그램을 예로 들 수 있습니다. 파일 형식의 구조를 사용하여 OpenXML에 대한 최소한의 지식으로 이러한 프로그램을 작성할 수 있습니다. 특히 다음과 같은 경우가 이에 해당합니다.

- 파일 형식은 특히 XML 및 ZIP과 같이 완성된 도구가 있는 잘 만들어진 표준을 따릅니다.
- 파일 형식은 표준 메커니즘에 XML 및 ZIP을 결합하여 파일 내에서 관계를 표현하는 오픈 패키징 규칙을 사용합니다. 이러한 이유로 OpenXML의 기본 또는 지원되는 표시 언어에 대한 태그 의미도 모른 채 파일 콘텐츠를 탐색하는 경우가 종종 있습니다.
- 다른 구조에 대한 방해없이 XML 트리의 깊은 곳에 있는 요소에 액세스하여 이를 수정할 수 있습니다.

파일 형식 전반에 걸쳐 있는 약간의 세부 사항 중 이진 형식에 나타나지 않는 일부 내용은 캐시된 값을 제공하여 최소 기능이 있는 응용 프로그램을 지원합니다. 예를 들면 다음과 같습니다.

- 페이지 수 매기기를 구현하지 않고도 시각 장애인용 판독기와 같은 응용 프로그램에서 최근 계산된 페이지 나누기(§4:2.3.3.13)를 사용하여 페이지 탐색 기능을 제공할 수 있습니다.
- 수식을 구현하거나 외부 데이터 원본과 통합하지 않고도 스프레드시트 프로그램은 캐시된 계산(§3:3.2.9) 및 캐시된 외부 데이터(§4:3.14 및 §4:3.10.1.76)에서 작업을 수행할 수 있습니다.

최소 준수 문서는 매우 단순합니다. “최소 WordprocessingML 문서”를 참조하십시오.

4.4 간결성

OpenXML 파일 형식은 고성능 응용 프로그램의 작성을 지원합니다. 이 하위 섹션에서는 간결한 파일을 생성함으로써 처리 및 구문 분석을 신속하게 수행하는 몇 가지 설계 요점을 설명합니다. 다음 하위 섹션에서는 문서의 작은 하위 집합만 구문 분석하거나 수정하여 모듈식 파일 구조가 응용 프로그램을 통해 많은 작업을 수행하는 방식을 보여 줍니다.

OpenXML 파일은 패키징 및 압축을 위해 보통 ZIP 보관 파일로 저장되며 오픈 패키징 규칙의 구현 권장 사항을 따릅니다. 놀랍게도 OpenXML 파일은 이진 파일보다 평균 25% 더 작고, 많게는 75% 더 작은 경우도 있습니다. 예를 들어 이 백서의 이진 형식은 85% 더 큼니다.

간결성의 두 번째 간단한 근원은 XML의 식별자 길이로, 이에 대해서는 특히 자세한 설명이 필요합니다. 자주 사용되는 태그 이름은 짧습니다. 구현 시 짧은 네임스페이스 접두사를 사용하도록 권장되기도 합니다. 예를 들어 WordprocessingML 네임스페이스의 기본 접두사는 “w”입니다.

파일 형식 전체에서의 반복 방지를 통해 길이가 더욱 짧아집니다. 큰 개체의 중복 저장 요소를 제거하는 경우의 예는 다음과 같습니다.

- SpreadsheetML에서는 반복된 문자열이 통합 문서의 문자열 테이블에 저장되고 인덱스에서 참조합니다(§3:3.3).
- SpreadsheetML에서는 아래로 채워져 있거나 여러 셀에 걸쳐 있는 수식이 왼쪽 맨 위의 셀에 단일 “마스터” 수식으로 저장되어 있으며, 채우기 범위 내에 있는 다른 셀은 인덱스를 그룹화하여 참조합니다(§3:3.2.9.2).
- DrawingML에서는 도형 이름(§4:5.1.12.56), 텍스트 기하학(§4:5.1.12.76) 및 기본 설정(§3:5.8, §3:5.9 및 §4:5.1.12 전반의 여러 곳)을 명시적인 표현이 아닌 이름 또는 숫자로 표시합니다. 이 경우 이름과 숫자의 의미는 해당 파일이 아니라 사양에 존재합니다. 여기서, 선택된 표현은 표준 절차 동안의 명시적인 교섭 결정의 결과입니다. 표현은 간결하며 올바른 추상화 수준에서 편집할 수 있습니다. 예를 들어 하나의 특성을 변경하여 사각형을 타원형으로 바꿀 수 있습니다(§4:5.1.11.18).

다른 종류의 예에서는 계층이 상속 의미를 알려 주는 데 사용됩니다. 다행스러운 부차적 결과로 파일 크기가 감소하여 성능이 향상됩니다.

- WordprocessingML에서는 스타일이 계층적입니다(§3:2.8.9).
- DrawingML에서는 도형이 계층적으로 그룹화됩니다(§4:5.1.2.1.20).
- PresentationML에서는 기본 계층이 슬라이드 마스터, 슬라이드 레이아웃 및 슬라이드와 관련됩니다(§3:4.2).

OpenXML의 이면 역시 효율적인 구현이 가능하도록 설계되었습니다. 예를 들어 SpreadsheetML에서 셀 테이블에는 비어 있지 않은 셀만 저장하고 병합된 셀을 단위로 표시할 수 있습니다. 이 기술의 효율적인 사용은 희소 스프레드시트에 있어 중요한 역할을 합니다.

4.5 모듈화

응용 프로그램은 문서의 작은 하위 집합을 구문 분석하거나 수정하여 많은 작업을 수행할 수 있습니다.

OpenXML 형식의 세 가지 기능이 함께 작동하여 이러한 모듈화를 제공합니다.

- 문서는 일체식 구조가 아니라 여러 파트로 구성되어 있습니다.
- 각 파트 간의 관계는 자체적으로 파트에 저장됩니다.
- OpenXML 문서를 지원하기 위해 일반적으로 사용하는 ZIP 보관 파일 형식은 각 파트에 대한 임의 액세스를 지원합니다.

예를 들면 다음과 같습니다.

- 응용 프로그램은 슬라이드 콘텐츠를 전혀 구문 분석하지 않고 이미지, 레이아웃 등의 리소스와 함께 슬라이드를 프레젠테이션 간에 정확하게 이동할 수 있습니다(§3:13.3.8). 이 예에서는 명시적 관계라고 하는 데이터를 사용하여 슬라이드와 해당 리소스를 찾습니다. 명시적 관계는 오픈 패키징 규칙에 의해 정의되며 PresentationML 태그 의미를 전혀 몰라도 구문 분석될 수 있습니다(§1:9.2, §2:8.3).
- 응용 프로그램은 해당 콘텐츠를 전혀 구문 분석하지 않고도 WordprocessingML 문서에서 모든 주석을 제거할 수 있습니다(§1:11.3.2). 이 예에서는 암시적 관계라고 하는 데이터를 사용하여 주석을 찾습니다. 암시적 관계는 OpenXML과 관련이 있으므로 관련 표시 언어에 대한 약간의 지식이 필요합니다(§1:9.2).

4.6 충실한 마이그레이션

OpenXML은 Microsoft Office 97-2003 이진 형식의 모든 기능을 지원하도록 설계되었습니다.

이러한 목적을 수행하는 데 따른 어려움과 이를 수행하는 OpenXML의 필연적인 독자성을 과장하여 말하기는 어렵습니다. PDF와 같은 일부 형식은 완성된 문서의 시각적 복사본을 최종 사용자에게 전달하고자 설계되었습니다. 이와 반대로 OpenXML은 원래 작성자와 동일한 추상화 수준을 사용하여 추후에 편집하거나 처리할 수 있도록 하기 위한 것입니다. 예를 들어 벡터 그래픽을 비트맵으로 줄이는 것은 이러한 취지에 맞지 않으며 스타일 계층을 독자적 스타일로 축소하는 것도 마찬가지입니다. 게다가 문서에는 원래 작성자가 유지하고자 의도한 계산적 의미 즉, 오류 코드 또는 동적인 동작을 만드는 애니메이션 규칙을 비롯하여 중간 계산 결과에 영향을 받는 수식 논리 등이 포함될 수 있습니다.

이러한 사양에 대한 참조는 OpenXML 기능의 예로 이진 형식의 미묘한 측면을 나타냅니다.

- SpreadsheetML 설명에는 광범위한 서식 사양이 포함되어 있습니다(§4:3.17.7).
- WordprocessingML 사양은 직접적인 형식을 구성하는 단락, 문자, 번호 매기기 및 표 속성의 규칙을 문서화합니다(§3:2.8, 특히 §3:2.8.10).
- PresentationML 사양은 애니메이션 기능을 문서화합니다(§3:4.4).

OpenXML을 사용하면 논리에 맞지 않는 모든 세부 사항을 맞춰 보지 않고도 여러 가지 구현을 통해 적용할 수 있습니다. 이는 레이아웃, 효과 렌더링, 수식 평가 등의 숫자 계산과 관련된 경우에 특히 중요합니다. 실용성보다 일관성을 더 요구할 경우 개발자가 준수해야 하는 불필요하게 높은 장벽만 만들어질 뿐입니다. 이러한 관점에서 위원회가 만든 샘플 결정에 대한 설명이 아래에 강조되어 있습니다.

- OpenXML은 개발자에게 결과의 모든 세부 내용을 맞추도록 강요하지 않으면서 표면 모양과 같은 결과(§5.1.12.50)를 정의합니다.
- OpenXML은 페이지 여백(§4:2.6.11), 글꼴(§4:2.8) 및 양쪽 맞춤(§4:2.3.1.13)과 같은 매개 변수를 정의합니다. 개발자는 이러한 매개 변수를 고려하는 범위에서 다른 흐름 알고리즘을 구현할 수 있습니다.
- SpreadsheetML 수식 사양(§4:3.17.7)에서는 부동 소수점 계산에서의 변형 제거를 시도하지 않습니다. 그 이유는 일반적으로 변형을 제거하려고 하면 준수 응용 프로그램이 원시 하드웨어에 의존하는 대신 느린 에뮬레이션을 구현해야 하기 때문입니다. 대신 수치 계산에 대한 전체 자릿수의 최소 비트 수를 지정합니다(§4:3.17.5).
- SpreadsheetML 수식 사양에서는 미래의 혁신이 가능하도록 특정 조건 결정을 정의된 구현 그대로 둡니다. 예를 들어 SpreadsheetML 수식 사양에서는 NORMINV(§4:3.17.7.227)와 같은 계산의 반복 횟수를 제한하지 않습니다. NORMINV 함수는 반복 검색을 수행하여 정규 분포를 역으로 수행합니다.

VML(§3:6)과 같은 여러 기존 기능에는 이전 버전과의 호환성이 기본적으로 포함되어 있습니다. 새 문서를 작성할 때는 DrawingML(§3:5)과 같이 이미 OpenXML에 있는 최신 표준을 사용하도록 권장됩니다.

4.7 비즈니스 데이터와의 통합

OpenXML을 사용하면 조직에서 OpenXML 문서 내의 사용자 지정 스키마를 사용하여 비즈니스 프로세스를 관리하는 정보 시스템과 생산성 응용 프로그램을 통합할 수 있습니다. 조직이 이러한 방식을 취하는 목적은 비즈니스 응용 프로그램에서 읽거나 쓸 수 없는 문서 내에 불확실하게 묻혀 있는 비즈니스 정보를 재이용하고 사업 정보의 처리를 자동화하기 위함입니다.

응용 프로그램에 포함된 기능은 다음과 같습니다.

- *검색*: 최종 사용자는 이익률이 20%가 넘는 회사의 스프레드시트 모음을 검색할 수 있습니다.
- *메타데이터 태깅*: 회사는 규제 관점에서 승인된 프레젠테이션에 태그를 사용할 수 있습니다.
- *문서 어셈블리*: 제안 그룹은 기본 데이터의 준비를 자동화함으로써 제안 생성을 간소화할 수 있습니다.
- *데이터 재이용*: 영업부 임원은 주어진 데이터 범위, 고객 목록, 거래 규모 및 수정된 규정과 내용 내에서 모든 영업 계약에 대한 보고서를 생성할 수 있습니다.
- *업무 라인 응용 프로그램*: 전문화된 단계의 전문가는 익숙한 생산 환경에서 배송 상품을 준비할 수 있지만 작업 제품을 자동으로 비즈니스 시스템으로 이동했습니다.

이러한 목적을 수행하려면 문서의 클래스에 포함할 수 있는 데이터 유형과 구조를 정의하고 각 문서의 흐름 내에서 저절로 정보가 발생할 수 있는 곳마다 정보가 표시되도록 허용해야 합니다. 간단한 예로 이력서를 살펴봅시다. 이름, 전화 번호, 주소, 경력 목표 및 자격 사항이라고 하는 필드가 포함된 데이터 구조를 정의합니다. 그런 다음 작성자가 문서에 이러한 필드를 배치할 때마다 해당 필드가 나타나도록 배열합니다. 금융 그룹 또는 의학 센터와 같이 서로 다른 비즈니스 설정에 따라 구조와 데이터 필드가 달라집니다.

OpenXML을 사용하면 표준화된 방식에 따라 이러한 과정을 수행할 수 있습니다.

첫째, 비즈니스 데이터의 구조가 사용자 지정 XML 스키마를 사용하여 처음 표시됩니다. 이를 통해 조직은 비즈니스 관점에서 의미 있는 데이터를 태그를 사용하여 표시할 수 있습니다. 조직은 자신의 고유한 스키마를 만들거나 재무 보고를 위한 XBRL(7), 건강 정보를 위한 HL7(8) 등과 같은 산업 표준 스키마를 사용할 수 있습니다. 스키마는 회사 내의 공공 부문에서 생성되며 출생 증명서에서 보험 정보에 이르기까지 목적에 따라 산업 표준으로 생성됩니다. 스키마가 XSD 양식(2)으로 표시되는 한 모든 사용자 지정 스키마를 사용할 수 있습니다.

둘째, 사용자 지정 데이터는 사용자 지정 XML 파트(§3.7.3)의 OpenXML 문서에 포함되며 사용자 지정 XML 데이터 속성 파트(§4.7.5)를 사용하여 설명될 수 있습니다. 프레젠테이션에서 이러한 사용자 지정 데이터를 구분함으로써 문서, 양식, 슬라이드 및 스프레드시트를 비롯한 광범위한 컨텍스트 내에서 최종 사용자는 프레젠테이션과 처리가 가능한 반면 OpenXML에서는 완전한 데이터 통합이 가능합니다. 따라서 상호 운용성은 좀 더 기본적이고 의미적으로 정확한 수준에서 수행될 수 있습니다.

4.8 혁신을 위한 여지

OpenXML은 Office 이진 형식이 정의되었을 때뿐 아니라 OpenXML이 정의되었을 때에도 고려하지 못한 새 응용 프로그램을 작성하도록 개발자를 장려합니다.

먼저 서로 다른 기능 집합을 포함하는 응용 프로그램 간의 상호 운용성을 가능하게 하는 확장성 메커니즘에 대해 알아봅니다. OpenXML에서 문서화되지 않은 새로운 기능을 포함하는 *상위* 수준 응용 프로그램과 이러한 기능을 인식하지 못하는 *하위* 수준 응용 프로그램을 살펴봅니다. 확장성의 세 가지 기본 목적은 다음과 같습니다.

- *시각적 충실도*: 높은 수준의 응용 프로그램이 표시하는 것을 나타내기 위한 하위 수준의 응용 프로그램을 위한 기능입니다. 원래 이 기능에서는 동일한 데이터에 대한 여러 가지 내용을 파일에 저장해야 합니다.
- *편집 가능성*: 하나 이상의 내용을 편집할 수 있는 기능입니다.
- *개인 정보 보호*: 다른 내용을 편집한 후 삭제되거나 수정되었다고 믿고 있는 이전 버전의 정보가 예기치 않게 남지 않도록 하는 기능입니다. 응용 프로그램은 내용을 제거하거나 동기화하여 이 기능을 수행할 수 있습니다.

OpenXML 기능 집합의 확장에 대한 개발자의 바람에는 두 가지 주요 옵션이 있습니다.

- *다른 콘텐츠 차단*: 다른 콘텐츠 차단(§3:2.18.4 and §5:9.2)은 동일한 콘텐츠에 대해 여러 가지 표현을 저장하며 각각의 선택에 대해 서로를 차단하는 것입니다. 하위 수준 응용 프로그램에서는 한 가지 선택을 읽으면 읽기 기능이 차단됩니다. 편집 시에도 작성하는 선택의 수만큼 쓰기 기능이 차단됩니다.
- *확장명 목록*: 확장명 목록(§3:2.6)에는 임의의 사용자 지정 XML이 시각적 내용 없이 저장됩니다.

개발자에게는 확장성 메커니즘 이상으로 혁신시킬 여지가 있습니다.

- *다른 상호 작용 패러다임*: OpenXML에서 문서 구문을 더 많이 지정할수록 응용 프로그램 기능은 더 적어집니다. 규칙 설명에서 설명한 대로 OpenXML은 의미에 중점을 둡니다(§1:2.2, §1.2.3). 결국 준수 응용 프로그램은 최종 사용자와 다양한 수단을 통해 마음대로 통신할 수 있거나 지정된 의미를 고려하는 한 최종 사용자와 전혀 통신하지 못합니다.
- *Novel 컴퓨팅 환경*: 규칙 설명에서는 낮은 수준의 기능이 있는 응용 프로그램을 허용하여 이러한 응용 프로그램이 소형 장치와 OpenXML(§1:2.6)의 하위 집합만 구현하는 응용 프로그램을 실행하도록 할 수 있습니다. 추가적인 특성 메커니즘에서는 작성된 응용 프로그램에 자신의 기능 제한(§3:8.1)을 허용합니다.

이전 하위 섹션에서 설명한 것처럼 가장 중요한 혁신 기회 중 일부는 직접적인 사용자 상호 작용을 위한 문서 렌더링이 아닌 XML 웹 서비스(9)를 통하는 경우와 같이 XML 메시지 형식을 사용한 컴퓨터 간 처리와 관련이 있습니다. 이런 응용 프로그램에는 OpenXML 문서 내에 포함된 데이터의 운영 외에 사용자가 볼 수 있는 기능은 없지만 완전한 구문 형식의, 의미가 통합된 상호 운용성 지침(§1:2.6)인 문서 규칙(§1:2.4)과 응용 프로그램 규칙(§1:2.5)을 따릅니다.

사용자 지정된 XML 처리에 대한 모든 가망 이용 사례를 열거할 수는 없지만 XML 중심 서비스의 경우에는 가능합니다. XML 중심 서비스는 사용자 지정 데이터의 자동 추출 및 삽입을 위해서나, XML 디지털 서명(10), XML 암호화(11) 등의 사용자 지정 보안 서비스나, XML 형식 간에 서로 변환하는 임의적인 XSLT 변환(12)을 위해 OpenXML 문서를 처리합니다. OpenXML에는 이러한 처리 도중 금지 사항이나 제한 사항이 없습니다.

5 OFFICE OPEN XML 문서의 구조

이 백서의 기본 목적은 독자가 모든 OpenXML 파일의 고급 구조에 대해 이해할 수 있도록 하는 것입니다. 이를 위해 OPC(오픈 패키징 규칙)에 대한 보통 수준의 세부 사항과 개별 표시 언어에 대한 개요가 여기에 설명되어 있습니다.

5.1 오픈 패키징 규칙

OPC(오픈 패키징 규칙)는 문서를 완전히 표시하기 위해 여러 유형의 콘텐츠(예: XML, 이미지 및 메타데이터)를 ZIP 보관 파일과 같은 컨테이너에 저장하는 방식을 제공합니다. 또한 제약 및 관계를 나타내기 위한 논리적 모델을 설명합니다.

OPC의 권장 구현 방법은 ZIP 보관 파일 형식을 사용하는 것입니다. ZIP 뷰어를 사용하여 OpenXML 파일의 구조를 조사할 수 있습니다. 이 설명을 읽으면서 작은 OpenXML 파일의 콘텐츠를 이런 식으로 조사하면 도움이 됩니다. Windows 운영 체제에서는 파일 이름에 “.zip” 확장명을 추가하고 해당 파일을 두 번 클릭하기만 하면 됩니다.

논리적으로 OpenXML 문서는 OPC *패키지*(§5:8)입니다. 패키지는 *파트*(§5:8.1)의 평면적 집합입니다. 각 파트에는 대/소문자를 구분하지 않는 *파트 이름*이 있으며 “/pres/slides/slide 1 .xml”(§5:8.1.1)과 같이 세그먼트 이름을 슬래시로 구분하는 시퀀스로 구성됩니다. 각 파트에는 또한 *콘텐츠 유형*(§5:8.1.2)이 있습니다. 물리적으로 ZIP 보관 파일은 하나의 패키지로써, 보관 파일에 있는 각 ZIP 항목이 하나의 파트에 해당하고 ZIP 보관 파일 내의 경로 이름은 파트 이름과 바로 일치합니다.

ZIP 구현 시 “/[Content_Types].xml”을 통해 소비자는 패키지에서 각 파트의 콘텐츠 유형(§2:9.2.6)을 확인할 수 있습니다. 미디어 유형에 대한 구문 및 정의는 RFC 2616의 3.7 섹션(13)을 따릅니다.

패키지와 파트는 외부 리소스뿐만 아니라 패키지 내에 다른 파트에 대한 *명시적 관계*(§1:9.2)를 포함할 수 있습니다. 모든 명시적 관계에는 파트의 콘텐츠가 파트를 참조할 수 있는 관계 ID와 응용 프로그램이 이를 처리할 방법을 결정할 수 있는 유형이 있습니다. URI를 사용하여 관계 유형이 명명되며 파트가 조정되지 않더라도 충돌 없이 새 유형을 안전하게 생성할 수 있습니다.

지정된 원본 패키지나 파트에 대한 명시적 관계의 집합은 *관계 파트*에 저장됩니다. 패키지의 관계 파트 전체를 “/_rels/.rels”라고 하며, “/a/b/c.xml”이라는 파트를 위한 관계 파트는 “/a/b/_rels/c.xml.rels”라고 합니다. 관계 파트(ZIP 구현에서의 콘텐츠 유형 파트)는 패키지에서 유일하게 특별히 명명된 파트입니다. 패키지를 열려면 응용 프로그램에서 패키지-관계 파트를 구문 분석하여 적절한 유형의 관계를 따라야 합니다.

OpenXML 문서의 다른 모든 파트에는 OpenXML, 사용자 지정 XML 또는 멀티미디어 개체와 같은 임의 유형의 콘텐츠가 포함되어 있습니다. 사용자 지정 XML을 보유하는 파트의 기능은 비즈니스 데이터와 메타데이터를 포함하기 위한 특히 강력한 메커니즘입니다.

5.2 WORDPROCESSINGML

WordprocessingML 문서는 텍스트 영역의 모음(§3:2.1)으로 구성되어 있습니다. 각 텍스트 영역은 주 문서(§3:2.2), 용어집 문서(§3:2.13), 하위 문서(§3:2.18.2), 머리글(§3:2.11.1), 바닥글(§3:2.11.2), 주석(§3:2.14.5), 프레임, 텍스트 상자(§3:2.18.1), 각주(§3:2.12.1) 또는 미주 (§3:2.12.2) 중 하나입니다.

유일한 필수 텍스트 영역은 주 문서입니다. 주 문서는 패키지 관계의 대상이며 그 유형은 다음과 같습니다.

<http://schemas.openxmlformats.org/officeDocument/2006/relationships/officeDocument>

XML 트리의 루트에서 리프까지 일반적인 경로는 다음과 같은 XML 요소(§3:2.2)로 구성됩니다.

- document — 주 문서의 루트 요소(§3:2.3).
- body — 본문(§3:2.7.1). 여러 단락을 포함할 수 있으며 sectPr 요소에 지정된 섹션 속성도 포함할 수 있습니다.
- p — 단락(§3:2.4.1). 하나 이상의 런을 포함할 수 있습니다. pPr 요소에 지정된 단락 속성을 포함할 수도 있으며 rPr 요소(§3:2.4.4)에 지정된 기본 런 속성(문자 속성이라고도 함)이 차례대로 포함될 수 있습니다.
- r — 런(§3:2.4.2). 주로 텍스트 범위인 런 콘텐츠의 여러 유형을 포함할 수 있으며 런 속성(rPr)을 포함할 수도 있습니다. 런은 OpenXML 내의 기본 개념입니다. 런은 동일한 속성을 가진 인접한 텍스트 조각으로 런에는 추가적인 텍스트가 포함되지 않습니다. 예를 들어 문장에 “this is **three runs**”라는 단어가 포함된 경우 적어도 “this is”, “**three**” 및 “runs”라는 세 개의 런으로 표현될 수 있습니다. 이러한 면에서 OpenXML은 속성이 임의로 중첩될 수 있는 HTML과 같은 형식과 상당한 차이가 있습니다.
- t — 텍스트 범위(§3:2.4.3.1). 서식 지정, 줄 바꿈, 표, 그래픽 또는 기타 텍스트가 아닌 내용을 포함하지 않는 임의 분량의 텍스트를 포함합니다. 텍스트에 대한 서식 지정은 런 속성 및 단락 속성에서 상속됩니다. 이 요소는 종종 xml:space="preserve" 특성을 사용합니다.

이 하위 섹션에서는 단락과 런 속성을 지정하여 텍스트의 서식을 직접 지정했습니다. 직접 서식을 지정하는 것은 응용 프로그램 순서의 맨 마지막에 해당하며, 여기에는 문서 기본 설정(§3:2.8.10)은 물론 문자, 단락, 번호 매기기 및 표 스타일이 포함됩니다. 이러한 스타일은 스스로 상속 계층으로(§3:2.8.9) 구성됩니다.

아래에 있는 “최소 WordprocessingML 문서”의 하위 섹션에 WordprocessingML 문서가 모두 나열되어 있습니다.

5.3 PRESENTATIONML

PresentationML 문서는 프레젠테이션 파트에 의해 설명됩니다. 프레젠테이션 파트는 패키지 관계의 대상으로 해당 유형은 다음과 같습니다.

<http://schemas.openxmlformats.org/officeDocument/2006/relationships/officeDocument>
프레젠테이션은 다음과 같은 기본 구조(§3:4.2)를 참조하며 기본 계층에서 맨 위에서 아래쪽으로 나열합니다.

- 슬라이드 마스터, 노트 마스터 및 유인물 마스터(§3:4.2.2)는 모두 프레젠테이션의 속성을 상속합니다.
- 슬라이드 레이아웃(§3:4.2.5)은 슬라이드 마스터에서 프레젠테이션 속성을 상속합니다.
- 슬라이드(§3:4.2.3) 및 슬라이드 노트 페이지(§3:4.2.4)는 슬라이드 레이아웃 및 노트 마스터에서 각각 프레젠테이션 속성을 상속합니다.

각 마스터, 레이아웃 및 슬라이드는 자신의 파트에 저장되어 있습니다. 각 파트의 이름은 프레젠테이션 파트의 관계 파트에서 지정됩니다. 프레젠테이션 외에 6개의 파트가 각각 기본적으로 동일한 방식으로 구성됩니다. XML 트리의 루트에서 리프까지 일반적인 경로는 다음과 같은 XML 요소(§3:2.2)로 구성됩니다.

- `sld`, `sldLayout`, `sldMaster`, `notes`, `notesMaster` 또는 `handoutMaster` — 루트 요소.
- `csld` — 슬라이드(§4:4.4.1.15). 다음 두 개의 클머리 기호에서 설명한 DrawingML 요소와 아래에서 설명한 다른 구조적 요소를 포함할 수 있습니다.
- `spTree` — 도형 트리(§4:4.4.1.42). `grpSpPr` 요소(§4:4.4.1.20)의 그룹 도형 속성 및 `nvGrpSpPr` 요소(§4:4.4.1.28)의 비시각적인 그룹 도형을 포함할 수 있습니다. 이 노드와 이 노드의 하위 항목은 모두 DrawingML 요소입니다. PresentationML에서 이 요소들의 중요한 역할 때문에 일부 DrawingML 요소를 여기에 나열하였습니다.
- `sp` — 도형(§4:4.4.1.40). `spPr` 요소(§4:4.4.1.41)의 도형 속성 및 `nvSpPr` 요소(§4:4.4.1.31)의 비시각적인 도형 속성을 포함할 수 있습니다.

DrawingML 도형 콘텐츠를 비롯하여 다음 표에 요약되어 있는 것처럼 루트 요소의 위치에 따라 cs1d에는 다른 구조적 요소가 포함될 수 있습니다.

	슬라이드	슬라이드 레이아웃	슬라이드 마스터	유인물 마스터	노트 마스터	노트 페이지
일반 데이터	X	X	X	X	X	X
전환	X	X	X			
시간	X	X	X			
머리글/바닥글		X	X	X	X	
일치하는 이름		X				
레이아웃 유형		X				
보존		X	X			
레이아웃 목록			X			
텍스트 스타일			X			

기본 계층에서 더 낮은 개체에 의해 지정된 속성(슬라이드 마스터, 슬라이드 레이아웃, 슬라이드)은 계층에서 더 높은 개체에 의해 지정된 해당 속성을 무시합니다. 예를 들어 전환이 슬라이드에 지정되지 않은 경우 이 전환은 슬라이드 레이아웃에서 제거되며, 슬라이드 레이아웃에 지정되지 않은 경우 슬라이드 마스터에서 제거됩니다.

5.4 SPREADSHEETML

SpreadsheetML 문서는 통합 문서 파트의 최상위 수준에서 설명됩니다. 통합 문서 파트는 패키지 관계의 대상으로 해당 유형은 다음과 같습니다.

<http://schemas.openxmlformats.org/officeDocument/2006/relationships/officeDocument>

통합 문서 파트는 통합 문서와 파일 버전, 작성한 응용 프로그램 및 수정할 암호와 같은 이 문서의 구조에 대한 정보를 저장합니다. 논리적으로 통합 문서는 하나 이상의 시트(§3:3.2)를 포함하지만 물리적으로 각 시트는 각각의 파트에 저장되고 통합 문서 파트로부터 일반적인 방식으로 참조됩니다. 각 시트는 워크시트, 차트 시트 또는 대화 상자 시트일 수 있습니다. 여기서는 가장 일반적인 유형인 워크시트에 대해서만 설명하겠습니다. 워크시트 개체에서 XML 트리의 루트에서 리프까지 일반적인 경로는 다음과 같은 XML 요소로 구성됩니다.

- `worksheet` — 워크시트의 루트 요소(§3:3.2).
- `sheet Data` — 워크시트에 비어있지 않은 모든 셀을 나타내는 셀 테이블(§3:3.2.4).
- `row` — 셀 테이블에 있는 한 행의 셀(§3:2.8).
- `c` — 하나의 셀(§3:3.2.9). `r` 특성은 A1-스타일 좌표를 사용하여 셀의 위치를 나타냅니다. 셀에는 스타일 식별자(`s` 특성) 및 데이터 유형(`t` 특성)이 있습니다.
- `v` 및 `f` — 셀의 값(§3:3.2.9.1) 및 선택적 수식(§3:3.2.9.2). 셀에 수식이 있는 경우 값은 가장 최근의 계산 결과입니다.

문자열과 수식이 공유 테이블(§3:3.3 및 §3:3.2.9.2.1)에 저장되어 있어 저장의 중복을 피하고 로드와 저장 속도가 빨라집니다.

5.5 표시 언어 지원

여러 가지 지원 표시 언어를 사용하여 OpenXML 문서의 콘텐츠를 설명할 수도 있습니다.

- DrawingML(§3:5) — 도형과 그래픽으로 렌더링한 기타 개체를 문서에 표시하는 데 사용합니다.
- VML(§3:6) — 이전 버전과의 호환성을 위해 벡터 그래픽의 형식이 포함되었으며 결국 DrawingML로 교체됩니다.
- 공유 ML: 수학(§3:7.1), 메타데이터(§3:7.2), 사용자 지정 XML(§3:7.3) 및 참고 문헌(§3:7.4)이 있습니다.

5.6 최소 WORDPROCESSINGML 문서

이 하위 섹션에는 세 파트로 구성된 최소 WordprocessingML 문서가 포함되어 있습니다.

콘텐츠-유형 파트인 “[Content_Types].xml”은 두 가지 다른 필수 파트의 콘텐츠 유형을 설명합니다.

```
<Types xmlns="http://schemas.openxmlformats.org/package/2006/content-types">
  <Default Extension="rels"
    ContentType="application/vnd.openxmlformats-package.relationships+xml"/>
  <Default Extension="xml"
    ContentType="application/vnd.openxmlformats-officedocument.wordprocessingml.document.main+xml"/>
</Types>
```

패키지-관계 파트인 “/_rels/.rels”는 패키지과 주 문서 간의 관계를 설명합니다.

```
<Relationships xmlns="http://schemas.openxmlformats.org/package/2006/relationships">
  <Relationship Id="rId1"
    Type="http://schemas.openxmlformats.org/officeDocument/2006/relationships/officeDocument"
    Target="document.xml"/>
</Relationships>
```

이 경우 문서 파트는 “/document.xml”이며 문서 콘텐츠를 포함합니다.

```
<w:document xmlns:w="http://schemas.openxmlformats.org/wordprocessingml/2006/main">
  <w:body>
    <w:p>
      <w:r>
        <w:t>Hello, world.</w:t>
      </w:r>
    </w:p>
  </w:body>
</w:document>
```

사양에는 최소 문서와 WordprocessingML(§1:11.2), PresentationML(§1:13.2) 및 SpreadsheetML(§1:12.2)에 대한 자세한 내용이 설명되어 있습니다.

6 요약

OpenXML은 다양한 배경과 조직적 이해를 가진 여러 산업 및 공공 기관 대표들의 상당한 노력의 산물입니다. 기존 문서 자료에서 사용된 전체 기능 집합뿐 아니라 전 세계 모든 주요 언어 그룹의 기본적인 요구인 국가별 옵션까지 다루고 있습니다. Ecma TC45(1)의 표준화 작업과 공개적인 논의를 통한 기여의 결과 OpenXML은 높은 수준의 상호 운용성 및 플랫폼 독립이 가능합니다. 그리고 OpenXML 문서는 광범위한 참고 자료를 통해 완성도를 높이고 비규정적 설명을 통해 사용이 용이해졌습니다. OpenXML에는 문서를 올바르게 처리할 수 있도록 보조 기술 제품에 대한 충분한 정보가 포함되어 있습니다. OpenXML 구현은 아주 작은 규모로 수행하거나 기능에만 중점을 두거나 혹은 전체 기능 집합을 망라할 수도 있습니다. 이 형식에 포함된 확장성 메커니즘은 혁신의 여지를 보장합니다.

해당 형식 사양을 표준화하고 오랜 기간 동안 유지 관리하면 여러 부문에서 사양의 안전성을 신뢰할 수 있으며 오픈 표준 절차에 따라 향후 발전에 대한 견제와 균형이 유지되리라고 확신합니다. 기존의 이진 형식으로 작성된 수십억 개의 문서와 매년 지속적으로 작성되는 수십억 개의 문서를 보존할 수 있는 오픈 문서 형식 표준에 대한 절실한 요구는 하드웨어, 네트워킹 및 표준 기반 소프트웨어 인프라 기술의 발전으로 충족되었습니다. 또한 주요 핵심 비즈니스 시스템에 대한 기존의 대규모 투자를 비롯하여 시장 요구의 폭발적인 다양화도 인해 오픈 문서 형식 표준이 필수 요소로 대두되었습니다.

7 인용

1. **Ecma international.** TC45 - Office Open XML 형식. *Ecma International*. [온라인]
<http://www.ecma-international.org/memento/TC45.htm>.
2. **W3C.** XML 스키마. *World Wide Web 컨소시엄*. [온라인]
<http://www.w3.org/XML/Schema>.
3. **ISO.** ISO/IEC 19757-2:2003. *국제 표준화 기구*. [온라인]
<http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=37605&ICS1=35&ICS2=240&ICS3=30>.
4. **ISO.** ISO/IEC 19757-4:2006. *국제 표준화 기구*. [온라인]
<http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=38615&ICS1=35&ICS2=240&ICS3=30>.
5. **W3C.** Extensible Markup Language(XML) 1.0(Fourth Edition). *World Wide Web 컨소시엄*. [온라인] 2006.
<http://www.w3.org/TR/2006/REC-xml-20060816/>.
6. **W3C.** Namespaces in XML 1.0(Second Edition). *World Wide Web 컨소시엄*. [온라인] 2006.
<http://www.w3.org/TR/2006/REC-xml-names-20060816/>.

7. **XBRL International.** XBRL 사양. *Extensible Business Reporting Language*. [온라인]
<http://www.xbrl.org/Specifications/>.
8. **Health Level Seven.** HL7 ANSI-승인 표준. *Health Level Seven*. [온라인]
http://www.hl7.org/about/directories.cfm?framepage=/documentcenter/public/faq/ansi_approved.htm.
9. **W3C.** W3C 웹 서비스 아키텍처. *World Wide Web 컨소시엄*. [온라인] 2002.
<http://www.w3.org/2002/ws/>.
10. **W3C.** W3C XML 서명. *World Wide Web 컨소시엄*. [온라인]
<http://www.w3.org/Signature/>.
11. **W3C.** W3C XML 암호화. *World Wide Web 컨소시엄*. [온라인] 2001.
<http://www.w3.org/Encryption/2001/>.
12. **W3C.** XSL 및 XSLT. *World Wide Web 컨소시엄*. [온라인]
<http://www.w3.org/Style/XSL/>.
13. **W3C.** Hypertext Transfer Protocol - HTTP/1.1. *World Wide Web 컨소시엄*. [온라인]
<http://www.w3.org/Protocols/rfc2616/rfc2616.html>.