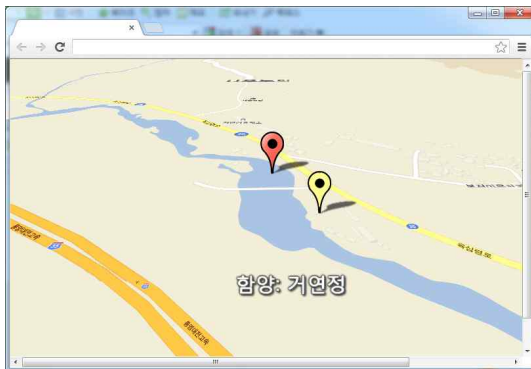


시맨틱 웹이란?

‘시맨틱 웹’의 이상은 월드와이드웹 상에서 ‘지식의 조각’이라고 할 수 있는 개체들(individuals)이 서로서로 어떤 의미로 관계를 맺는지 명시적으로 드러내고, 그 바탕 위에서 개체와 개체 사이의 관계를 추론하여 새로운 사실을 발견하거나 종합적인 지식을 얻으려는 것이다.¹⁾ 이 이상적인 목표를 이루기 위한 실천적 노력의 첫걸음은 한 편의 월드와이드웹 문서를 작성할 때 그것이 다른 무엇과 어떠한 관계를 갖는지 표시해 주는 것이다. 서로에 대해 ‘관계가 있음’을 알리는 데이터의 집합을 ‘링크드 데이터’(Linked Data)라고 하며, 그 링크드 데이터가 인터넷상에서 공개되어, 누구라도 그 연결을 확장해 갈 수 있는 것을 ‘링크드 오픈 데이터’(Linked Open Data, LOD)라고 한다.²⁾

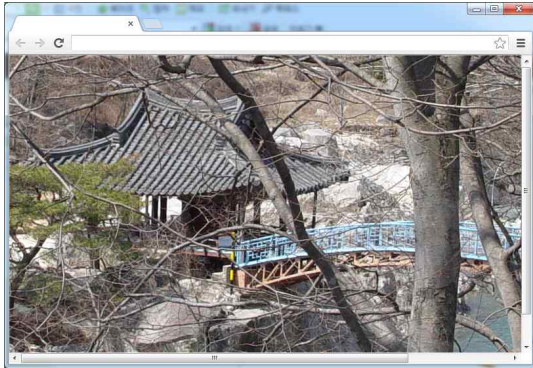
시맨틱 웹을 통해 무엇을 얻을 수 있는지 간단한 예시를 통해 알아보기로 하자.

[예시] 내용적으로 유관한 월드와이드웹 자원



[1-1]
함양 거연정 위치
Google Maps

- 1) 시맨틱 웹에 대해 팀 버너스리는 “현재의 웹에 명확한 의미의 정보를 부가하여 사람과 컴퓨터가 협업할 수 있도록 하는 것”이라고 하였다. (Berners-Lee et al. The Semantic Web, Scientific American (May 2001))
- 2) 시맨틱 웹(Semantic Web)과 LOD(Linked Open Data): 팀 버너스리는 시맨틱 웹에 다가가기 위한 실천 전략으로 ‘Linked Data’와 ‘Linked Open Data’의 개념을 소개하였다.
☞ Tim Berners-Lee, “Linked Data”, W3C Design Issues (2006, 2009), <https://www.w3.org/DesignIssues/LinkedData.html>
‘시맨틱 웹’이 데이터의 연결망으로서의 미래의 웹에 대한 포괄적인 구상이라고 한다면, LOD는 HTTP URI(Uniform Resource Identifier)와 RDF(Resource Describing Framework) 등의 데이터 기술 형식에 의존하여, 일정한 영역에서 데이터의 연결을 구현하는 구체적인 방법의 제안이라고 할 수 있다.



[1-2]
함양 거연정 사진
문화재청 문화재 DB



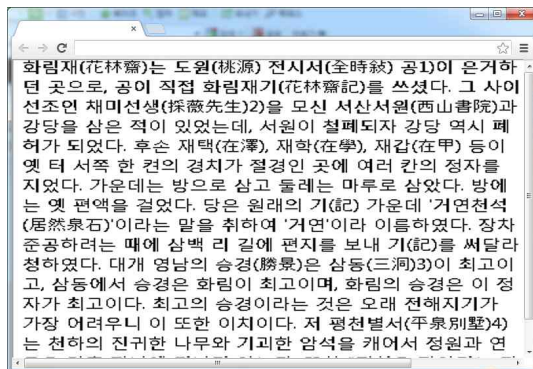
[1-3]
함양 거연정 기사
한국학중앙연구원 한국민족문화대백과사전



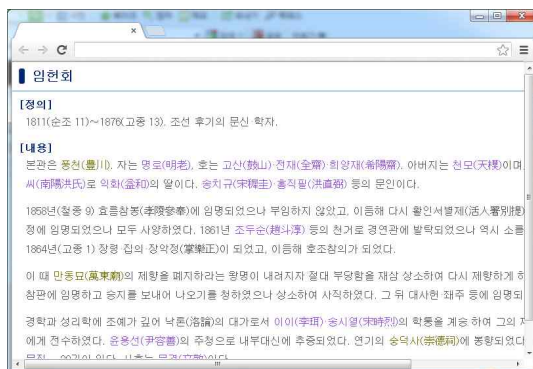
[1-4]
거연정기 편액 사진
개인 웹 페이지



[1-5]
거연정기 한문 원문
고전번역원 문집총간 DB



[1-6]
거연정기 번역문
개인 웹 페이지



[1-7]
임현회(거연정기 저자) 기사
한국학중앙연구원 역대인물 DB

위에 예시로 든 7개의 디지털 자료는 모두 경상남도 함양군 남강천에 있는 유서 깊은 누정 ‘거연정’과 관련이 있는 것이다. 그것들은 내용면에서 관련성이 있지만 각각 다른 곳에 별개의 문서로 존재한다. 이러한 자료를 고립된 자료원으로 두지 않고, 의미의 연결 고리를 좇아 문맥을 이루게 하려면 어떠한 노력이 더해져야 할까?

현재의 월드와이드웹 환경에서는 그 문서들에 공통으로 포함되어 있는 어휘-이름테면 ‘거연정’-를 이용해서 서로 관계가 있을 수 있는 후보 문서를 모두 탐색한 후에, 이용자의 눈으로 그 결과 하나하나를 확인하면서 필요한 것을 취하는 방법이 가능할 뿐이다.

시맨틱 웹은 데이터가 생산될 때 유관한 자료의 의미적 연관 관계를 약속된 방식으로 명시함으로써 보다 지능적인 데이터 연계가 이루어지게 할 것을 제안한다. 위에 예시한 월드와이드웹 상의 데이터에 대해 그 관계성을 정의하는 데이터를 다음과 같이 만들었다고 가정해 보자.

- 개체의 고유한 이름(IRI: Internationalized Resource Identifier) 정의³⁾

3) IRI(Internationalized Resource Identifier): 인터넷상의 특정 자원을 유일하게 식별할 수 있게 하는 이름. 종래의 URI(Uniform Resource Identifier)는 ASCII 문자(영문 알파벳과 기호)를 사용했던 데 반해 IRI는 한글, 한자를 포함하는 세계문자부호계(Unicode/ISO 10646)의 모든 문자를 사용할 수 있

N1: 월드와이드웹 상에서 '함양 거연정'을 유일하게 지목하는 식별자
N2: 월드와이드웹 상에서 '거연정기'를 유일하게 지목하는 식별자
N3: 월드와이드웹 상에서 '임헌회'를 유일하게 지목하는 식별자
D1: [1-1] 문서의 월드와이드웹 주소
D2: [1-2] 문서의 월드와이드웹 주소
D3: [1-3] 문서의 월드와이드웹 주소
D4: [1-4] 문서의 월드와이드웹 주소
D5: [1-5] 문서의 월드와이드웹 주소
D6: [1-6] 문서의 월드와이드웹 주소
D7: [1-7] 문서의 월드와이드웹 주소

- 개체간의 관계성 정의

N1의 누정기는 N2이다.
N2의 저자는 N3이다.
D1은 N1의 지도이다.
D2는 N1의 사진이다.
D3는 N1에 대한 설명이다.
D4는 N2의 사진이다.
D5는 N2의 한문 원문 텍스트이다.
D6는 N2의 번역문이다.
D7은 N3에 대한 설명이다.

이러한 방식으로 관계성을 기술한 정보가 [예시 3]에서 보인 개별 문서와 함께 인터넷상에 존재할 때, 우리가 현재의 월드와이드웹에서는 얻을 수 없는 어떤 것을 더 얻게 될 것인지는 쉽게 이해할 수 있을 것이다. 그 이전까지는 웹 상에서 고립된 형태로 존재하던 지식의 조각이 다른 조각들과 함께 모여 커다란 의미체를 이루게 된다.

도록 하였다. IETF(Internet Engineering Task Force)에서 2005년에 제안.
<http://tools.ietf.org/html/rfc3987>

