

Schema.org 이용하여 HTML 데이터 구조화하기

개발자들이 정의한 각종 데이터를 최종 결과물인 HTML 문서를 통해서 출력하지만 이를 다시 데이터형으로 돌리기에는 많은 어려움이 있기 때문에 HTML 문서에서 데이터를 구분해 낼 수 있는 방법으로 [마이크로데이터](#)이나 [마이크로포맷](#) 또는 [RDFa](#)를 이용할 수 있습니다. RDFa는 XHTML에나 잘 어울릴 법한 모양새이고 사용법이 다소 까다롭다는 점에서 개인적으로는 마이크로포맷이나 마이크로데이터를 선호합니다. 마이크로포맷은 별도의 네임스페이스 선언을 필요로하지 않고 class와 rel 속성만을 이용하여 구조화된 데이터를 표현할 수 있지만, 단계가 있거나 관계가 있는 데이터를 구조화하기에는 부족한 면이 있습니다. 마이크로데이터는 RDFa와 마이크로포맷의 중간쯤으로 볼 수 있습니다. 자세한 사용 방법은 구글이 문서([microdata](#), [microformats](#), [RDFa](#))를 잘 만들어 뒀습니다.

이 얘기는 수년 전부터 거론되었지만 웹사이트 소유주에게 돌아가는 이렇다 할 가치가 없었기 때문에 지들끼리 말만 많았던 HTML5 스펙이죠. 최근 구글은 이 구조화된 데이터를 이용하여 저자나 인물을 표시하거나 제품이나 장소 등의 평점 및 리뷰, 앨범의 트랙 정보 이벤트 등을 검색결과에 노출하고 있으며, 이것을 [리치 스니펫](#)이라 부르고 있습니다. 구글은 리치 스니펫을 제공하는 방법으로는 마이크로데이터를 권장하고 있으며 이는 데이터가 어떤 종류인지를 정의하는 스키마를 필요로합니다. 대표적으로 [data-vocabulary.org](#)또는 [schema.org](#)에서 제공하는 데이터 스키마를 선언하는데 이를 잘 이용하면 관계형 데이터 구조화가 가능해 집니다.

구글은 웹마스터 도구를 통해 [구조화된 데이터 테스트 도구](#)를 제공합니다. 다음은 schema.org의 데이터 스키마를 기반으로 이 블로그의 데이터를 구조화하고 테스트 도구를 통해 얻은 [결과](#)입니다. 제법 쓸만한 데이터가 만들어 지더라고요. 조금 더 발전하면 ATOM이나 RSS를 별도로 제공하는게 무의미해질지도 모르겠네요.

Item

- type: <http://schema.org/blog>
- property:
 - url: [Firejune](#)
 - name: [Firejune](#)
 - description: The Web is still changing.
 - version: 2.19
 - blogpost: Item 1
 - author: Item 2

Item 1

- type: <http://schema.org/article>
- property:
 - datepublished: 2012-12-19
 - name: [Node.JS용 Scribd 모듈 배포](#)
 - url: [Node.JS용 Scribd 모듈 배포](#)
 - author: [파이어준](#)
 - interactioncount: [Reactions4](#)

- articlebody: 역시나, 기업 요구사항은 파일 교환보다는 문서 관리 쪽으로 많이 치우쳐져 있더군요. 그래서 문서관련 기능을 강화하기 위해 매시업할 수 있는...
- keywords: [자료 - 웹개발](#)
- keywords: [Node.JS](#)
- keywords: [Scribd](#)
- keywords: [HTML5](#)
- keywords: [자바스크립트](#)
- keywords: [프로젝트](#)
- keywords: [GitHub](#)
- keywords: [NPM](#)
- keywords: [매시업](#)
- comment: Item 3
- comment: Item 4
- comment: Item 5
- comment: Item 6

Item 3

- type: <http://schema.org/comment>
- property:
 - datecreated: 2012-12-20T08:55:50+09:00
 - image: http://a0.twimg.com/profile_images/2399594250/8dp14wildkk8xبت8ownl_normal.png
 - url: [beyonditblog](#)
 - name: [beyonditblog](#)

- text: Node.JS용 Scribd 모듈 배포: 역시나, 기업 요구사항은 파일 교환 보다는 문서 관리쪽으로 많이 치우쳐져 있더군요. 그래서 문서관련 기능을 강화하기...

Item 4

- type: <http://schema.org/comment>
- property:
 - datecreated: 2012-12-20T09:24:06+09:00
 - image: http://a0.twimg.com/profile_images/1123405595/IT_NEWS_normal.jpg
 - url: [All IT News](#)
 - name: [All IT News](#)
 - text: text: [경준호] Node.JS용 Scribd 모듈 배포: 역시나, 기업 요구사항은 파일 교환 보다는 문서 관리쪽으로 많이 치우쳐져 있더군요. 그래서 문서관련 기능을...

Item 5

- type: <http://schema.org/comment>
- property:
 - datecreated: 2012-12-22T00:50:28+09:00
 - image: http://a0.twimg.com/profile_images/2514809179/stcr8k585g4tkcch9pgw_normal.jpeg
 - url: [nanhapark](#)

- name: [nanhapark](#)
- text: Node.JS용 Scribd 모듈 배포
[#nodeqa](http://t.co/qpzDlro0)

Item 6

- type: <http://schema.org/comment>
- property:
 - datecreated: 2012-12-22T01:20:16+09:00
 - image: http://a0.twimg.com/profile_images/1032276925/imazine_normal.jpg
 - url: [iMaZiNe80](#)
 - name: [iMaZiNe80](#)
 - text: RT @nanhapark: Node.JS용 Scribd 모듈 배포
[#nodeqa](http://t.co/qpzDlro0)

Item 2

- type: <http://schema.org/person>
- property:
 - image:
<https://firejune.com/attach/image/272778.jpeg>
 - name: Joon Kyoung
 - jobtitle: web front-end development
 - worksfor: Spark & Associates Corp
 - url: [@firejune](#)

- url: [Blog](#)
- url: [More information](#)