



1. HTML5 Introduction & Syntax, Semantic Web

1. HTML

(HyperText Markup Language) : 웹페이지 기술하기 위한 마크업 언어

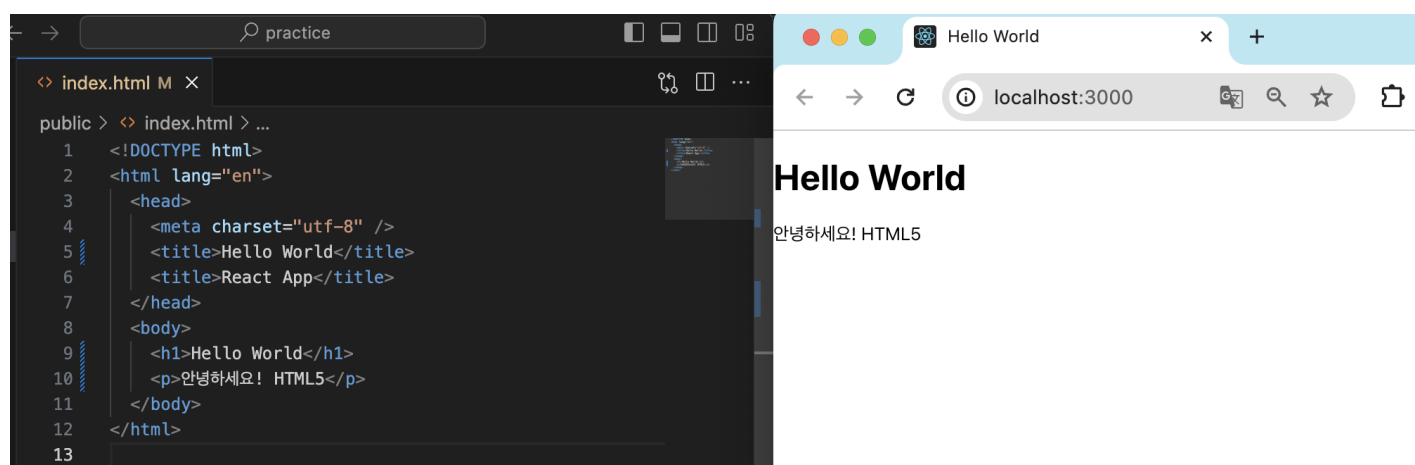
웹페이지의 **내용(content)**과 **구조(structure)**를 담당.

“HTML 태그” 를 통해 정보를 구조화

[HTML5 기능]

- 멀티미디어
 - 플래시와 같은 플러그인의 도움없이 비디오 및 오디오 기능을 자체적으로 지원
- 그래픽
 - SVG, 캔버스를 사용한 2차원 그래픽과 CSS3, WebGL을 사용한 3차원 그래픽을 지원
- 통신
 - 지금까지의 HTML은 단방향 통신만이 가능하였으나 HTML5는 서버와의 소켓 통신을 지원하므로 서버와의 양방향 통신이 가능
- 디바이스 접근
 - 카메라, 동작센서 등의 하드웨어 기능을 직접적으로 제어할 수 있음
- 오프라인 및 저장소
 - 오프라인 상태에서도 애플리케이션을 동작시킬 수 있음
- 시맨틱 태그
 - HTML 요소의 의미를 명확히 설명하는 시맨틱 태그를 도입하여 브라우저, 검색엔진, 개발자 모두에게 콘텐츠의 의미를 명확히 설명할 수 있음.
 - HTML 요소의 의미를 명확히 해석하고 그 데이터를 활용할 수 있는 시맨틱 웹을 실현 가능함
- CSS3
 - HTML5는 CSS3를 완벽하게 지원함

1-1) Hello HTML5



<!DOCTYPE html>

→ HTML5 문서는 반드시 <!DOCTYPE html>로 시작

→ 문서 형식(document type)을 HTML5로 지정

<html> </html>

→ 실제적 HTML document 기술

<head> </head>

→ document title, 외부 파일의 참조, 메타데이터의 설정 등 위치

→ 브라우저에 정보 표시 X

<body> </body>

→ 웹 브라우저에 출력되는 모든 요소

1-2) HTML5의 기본 문법

< 요소 (element)>

HTML document는 요소 (Element)들의 집합



시작태그 / 종료태그 / 태그 사이에 위치한 contents

- 요소의 중첩

- 요소는 중첩될 수 있음 (요소는 다른 요소를 포함 가능함)
- 부자관계 성립
- 중첩관계를 시각적으로 파악하기 쉽게 indent(들여쓰기) 활용
- "Hello HTML 예시"를 보면 html 요소는 body 요소를 포함하며 body 요소는 h1, p 요소를 포함한다

- 빈 요소 (Empty element or Self-Closing element)

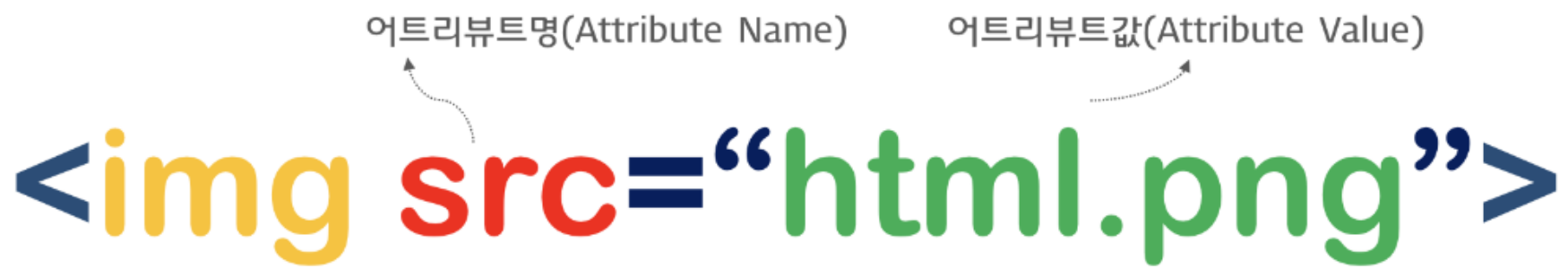
- content를 가질 수 없는 요소
- 어트리뷰트(Attribute)만 가질 수 있음

[대표적인 빈 요소]

- br
- hr
- img
- input
- link

- meta

< Attribute >



"어트리뷰트명 = 값"

```

```

- 요소의 성질, 특징을 정의하는 명세
- 요소에 추가적 정보 (이미지 파일의 경로, 크기 등)를 제공
- 시작 태그에 위치해야하며 이름과 값에 쌍을 이룬다

- 글로벌 Attribute

모든 HTML 요소가 공통으로 사용할 수 있는 어트리뷰트

Attribute	Description
id	유일한 식별자(id)를 요소에 지정한다. 중복 지정이 불가하다.
class	스타일시트에 정의된 class를 요소에 지정한다. 중복 지정이 가능하다.
hidden	css의 hidden과는 다르게 의미상으로도 브라우저에 노출되지 않게 된다.
lang	지정된 요소의 언어를 지정한다. 검색엔진의 크롤링 시 웹사이트의 언어를 인식할 수 있게 한다.
style	요소에 인라인 스타일을 지정한다.
tabindex	사용자가 키보드로 페이지를 내비게이션 시 이동 순서를 지정한다.
title	요소에 관한 제목을 지정한다.

< 주석 >

```
<!-- 주석은 화면에 표시되지 않는다. -->
```

<!-- 주석 표기 방법 -->

2. 시맨틱 웹 (Semantic Web) (배경지식)

대부분의 인터넷 사용자는 원하는 정보를 취득하기 위해 Google이나 Naver와 같은 검색사이트를 이용한다.

웹사이트는 검색 엔진에서의 노출이 매우 중요하다. 당연한 것이 검색엔진에 노출되지 않는 웹사이트에는 접속하는 이도 없기 때문이다. 검색 엔진은 이 시대의 가장 강력한 권력 중의 하나라고 말할 수도 있을 것이다.

SEO(검색엔진 최적화: Search Engine Optimization)같은 마케팅 도구를 사용하여 검색엔진이 본인의 웹사이트를 검색하기 알맞은 구조로 웹사이트를 조정하기도 하는데, 이것은 기본적으로 검색엔진이 웹사이트 정보를 어떻게 수집하는지 아는 것으로부터 시작된다.

크롤링 : 검색엔진은 로봇(Robot)이라는 프로그램을 이용해 매일 전세계의 웹사이트 정보를 수집한다.

인덱싱 : 검색 사이트 이용자가 검색할 만한 키워드를 미리 예상하여 검색 키워드에 대응하는 인덱스(색인)을 만들어 둔다. (검색엔진의 인덱서가 이를 수행함)

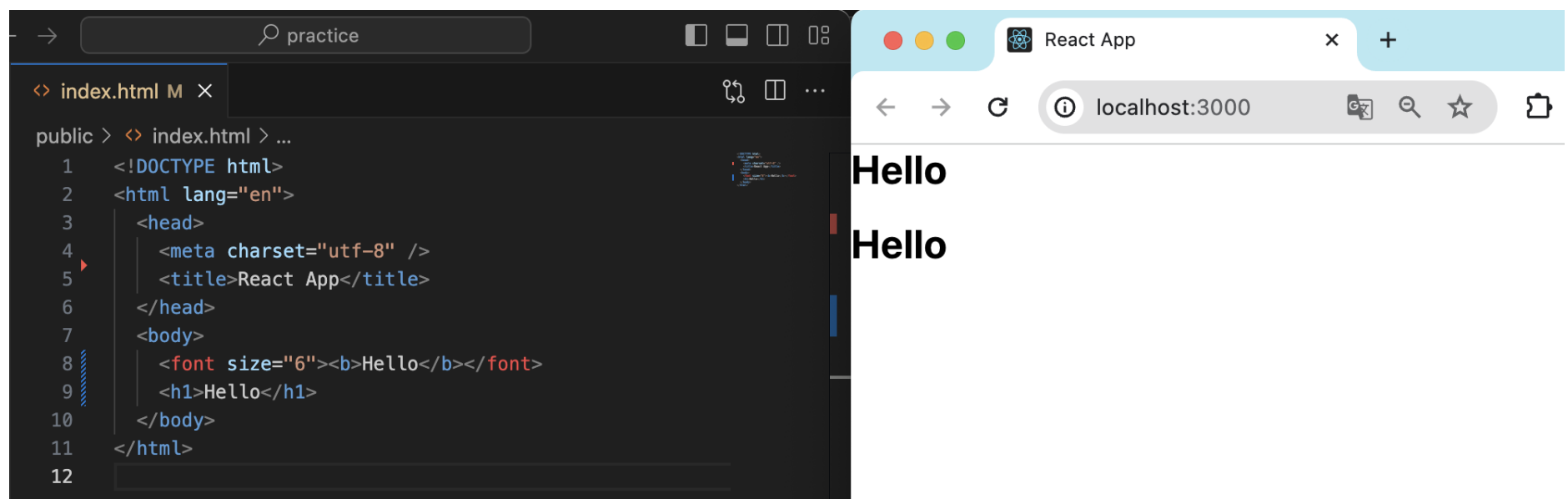
인덱스를 생성할 때 사용되는 정보는 검색 로봇이 수집한 정보인데 결국 웹사이트의 HTML 코드이다.

즉, 검색 엔진은 HTML 코드만으로 그 의미를 인지하여야 하는데

이때

시맨틱 요소(Semantic element)를 해석하게 된다.

HTML로 작성된 문서는 컴퓨터가 해석할 수 있는 메타데이터와 사람이 사용하는 자연어 문장이 뒤섞여 있다.



위 코드의 1행과 2행은 브라우저에서 동일한 외형을 갖는다. h1 태그의 디폴트 스타일이 1행과 같기 때문이다.

[1행의 요소]

```
<font size="6"><b>Hello</b></font>
```

의미론적으로 어떤 의미도 가지고 있지 않다. 즉, 의도가 명확하지 않다.

개발자가 의도한 요소의 의미를 명확하게 나타내지 않고, 다만 폰트 크기와 볼드체를 지정하는 메타데이터만을 브라우저에게 알리고 있다.

[2행의 요소]

```
<h1>Hello</h1>
```

header(제목) 중 가장 상위 레벨이라는 의미를 내포하고 있어서 **개발자가 의도한 요소의 의미가 명확히 드러나고 있다. 이것은 코드의 가독성을 높이고 유지보수를 쉽게한다.**

검색엔진은 대체로 h1 요소 내의 콘텐츠를 웹문서의 중요한 제목으로 인식하고 인덱스에 포함시킬 확률이 높다.

또한 사람도 h1 요소 내의 콘텐츠가 제목임을 인식할 수 있다. 시맨틱 요소로 구성되어 있는 웹페이지는 검색엔진에 보다 의미론적으로 문서 정보를 전달할 수 있고 검색엔진 또한 시맨틱 요소를 이용하여 보다 효과적인 크롤링과 인덱싱이 가능해졌다.

즉, 시맨틱 태그란 브라우저, 검색엔진, 개발자 모두에게 콘텐츠의 의미를 명확히 설명하는 역할을 한다.

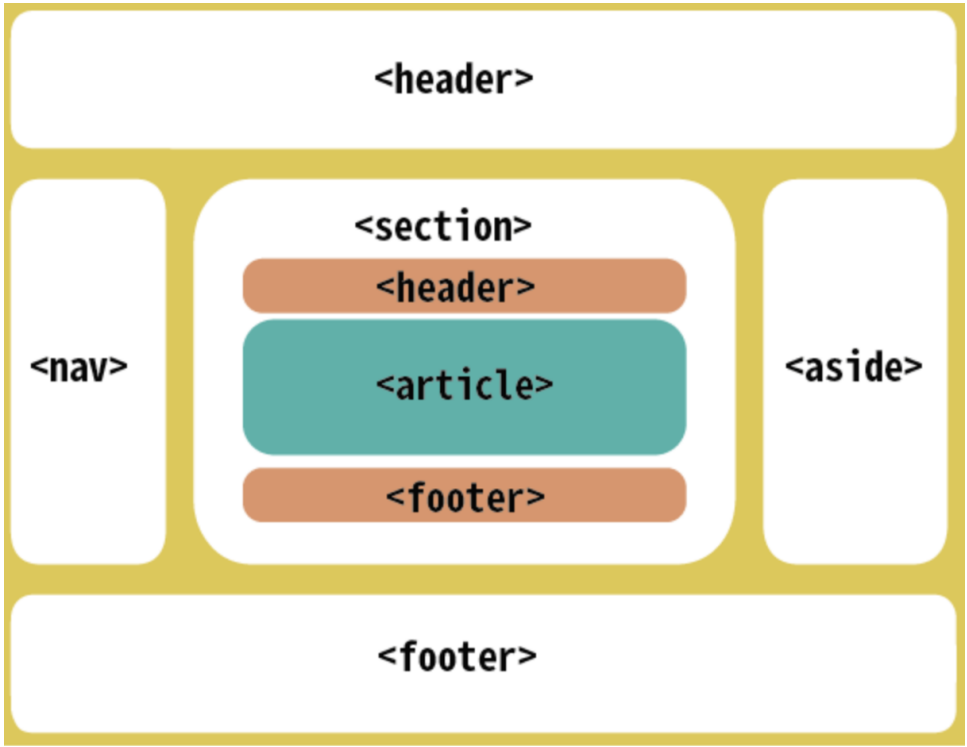
시맨틱 태그에 의해 컴퓨터가 HTML 요소의 의미를 보다 명확히 해석하고 그 데이터를 활용할 수 있는 시맨틱 웹이 실현될 수 있다.

시맨틱 웹이란 웹에 존재하는 수많은 웹페이지들에 메타데이터(Metadata)를 부여하여, 기존의 잡다한 데이터 집합이었던 웹페이지를 ‘의미’와 ‘관련성’을 가지는 거대한 데이터베이스로 구축하고자 하는 발상이다.

cf 1) HTML5에서 새롭게 추가된 Semantic tag

<code><header></code>	일반적으로 페이지나 해당 섹션의 가장 윗부분에 위치하며, 사이트의 제목이 보통 들어갑니다. (선택적으로 상단바나 검색창 배치 가능)
<code><nav></code>	내비게이션(navigation)의 약자로, 일반적으로 상단바 등 사이트를 안내하는 요소에 사용됩니다. (보통은 안에 <code></code> 을 넣어 목록 형태로 사용)
<code><main></code>	문서의 주된 콘텐츠를 표시
<code><section></code>	문제 전체적인 내용과 관련되어 있는 콘텐츠들의 집합으로, html에서 특정한 구역을 의미
<code><article></code>	독립적이고 자체 포함된 콘텐츠를 지정
<code><aside></code>	본문의 주요 부분을 표시하고 남은 부분을 설명하는 요소입니다. (특별한 일이 아니면 사이드바나 광고창 등 중요하지 않은 부분에 사용)
<code><footer></code>	일반적으로 페이지나 해당 파트의 가장 아랫부분에 위치하며 사이트의 라이선스, 주소, 연락처 등을 넣을 때 사용

cf 2)



HTML로 웹페이지를 만들때 위와 같은 구조의 시맨틱 요소를 활용한다