

## 1. Retrieved

```
<div>
<span id='author'
class='big'>
Author: John
Doe </span>
<span id='viewed'>
seen 42 times
</span>
</div>
```



## 2. Structured

```
{
  "author":
  "Author: John Doe",
  "viewed":
  "seen 42 times"
}
```



## 3. Cleaned

```
{
  "author":
  "john doe",
  "viewed":
  42,
}
```



## 4. Enriched

```
{
  "author":
  "john doe",
  "author_gender":
  "M",
  "viewed":
  42,
}
```