

# Data – 초기 데이터

\* 데이터 : 생활법령 데이터

(데이터 출처 : 생활법령 사이트 → 책자형 → 책자형 생활법령)

## ▣ 단순폭행죄

- 사람의 신체에 대해 폭행을 가한 경우에는 2년 이하의 징역, 500만원 이하의 벌금, 구류 또는 과료에 처해집니다(「형법」 제260조제1항).
  - 다만, 폭행죄는 피해자의 명시적인 의사에 반해 공소를 제기할 수 없는 반의사불벌죄(反意思不罰罪)에 해당합니다(「형법」 제260조제3항).
- ※ 단순폭행죄와 같이 피해자의 명시적인 의사에 반해 처벌할 수 없는 사건의 경우, 처벌을 희망하는 의사표시의 철회는 제1심 판결의 선고 전까지 할 수 있으며, 처벌 희망 의사표시를 철회한 사람은 다시 고소할 수 없습니다(「형사소송법」 제232조제3항).

### ※ 법령용어해설

- 구류(拘留): 자유형의 일종으로 1일 이상 30일 미만의 기간동안 유치장에 가두는 것을 말합니다. 이는 형벌 중 가장 가벼운 것으로서 주로 경미한 범죄에 대해 과해집니다.
- 과료(科料): 벌금과 같이 재산형의 일종으로 범죄인에게 일정한 금액의 지급을 강제적으로 부담지우는 형벌을 말합니다. 그러나 과료는 벌금에 비해 그 금액이 적고 비교적 경미한 범죄인 경우에 부과됩니다.
- 반의사불벌죄(反意思不罰罪): 피해자가 처벌을 희망하지 않는다면 처벌을 할 수 없는 죄를 말합니다. 예컨대, 단순폭행죄·과실 상해죄·단순협박죄·명예훼손죄와 같은 경우에는 피해자가 처벌을 희망하지 않는다는 의사표시를 하거나 처벌을 희망하는 의사표시를 철회한 때에는 공소를 제기할 수 없으며, 기소 후인 때에는 법원은 공소기각의 판결을 선고 합니다.

<출처: 법령용어사례집, 법제처·한국법제연구원, 2003>

### ※ 폭행에 대한 판례

- 폭행에 해당한다고 본 판례
  - 안수기도 행위에 수반하는 신체적 행위가 단순히 손을 얹거나 약간 누르는 정도가 아니라 그것이 지나쳐서 가슴과 배를 반복하여 누르거나 때려 그로 인해 사망에 이른 것과 같은 정도의 것이라면 이는 사람의 신체에 대한 유형력의 행사로서 폭행에 해당합니다(대법원 1994. 8. 23. 선고 94도1484 판결).
  - 피해자의 신체에 공격적으로 폭언이나 욕설을 하거나 동시에 손발이나 물건을 휘두르거나 던지는 행위는 직접 피해자의 신체에 접촉하지 않았다고 하더라도 피해자에 대한 불법한 유형력의 행사로서 폭행에 해당될 수 있습니다(대법원 2003. 1. 10. 선고 2000도5716 판결).
- 폭행에 해당하지 않는다고 본 판례
  - 문을 부수고 건물로 들어가 폭언하면서 잠겨 있는 방문을 여러 번 발로 찬 행위는 피해자들의 신체에 대한 유형력의 행사로는 볼 수 없기 때문에 폭행에 해당하지 않습니다(대법원 1984. 2. 14. 선고 83도3186 판결).
  - 갑이 먼저 읊에게 덤벼들고, 뺨을 꼬집고, 주먹으로 쥐어박았기 때문에 피고인이 상대방을 부동켜안은 행위는 유형력의 행사인 폭행에 해당하지 않습니다(대법원 1977. 2. 8. 76도3758 판결).
  - 피해자가 시비를 걸려고 양팔을 잡는 것을 피하고자 몸을 틀어 뿌리친 것뿐인 행위는 폭행에 해당하지 않습니다(대법원 1985. 10. 8. 85도1915 판결).
  - 거리상 멀어져 있는 사람에게 전화기를 이용하여 전화하면서 고성을 내거나 그 전화 대화를 높음 후 듣게 하는 경우에는 특수한 방법으로 수화자의 청각기관을 자극하여 그 수화자로 하여금 고통스럽게 느끼게 할 정도의 음향을 이용하였다는 등의 특별한 사정이 없는 한 폭행에 해당하지 않습니다(대법원 2003. 1. 10. 선고 2000도5716 판결).

- title에는 소제목이 들어가 있으며, content에는 여러 문장을 담은 리스트로 들어가 있음.
- content의 내용 전체를 하나의 chunk로 보기에는 내용이 多 → chunking 필요

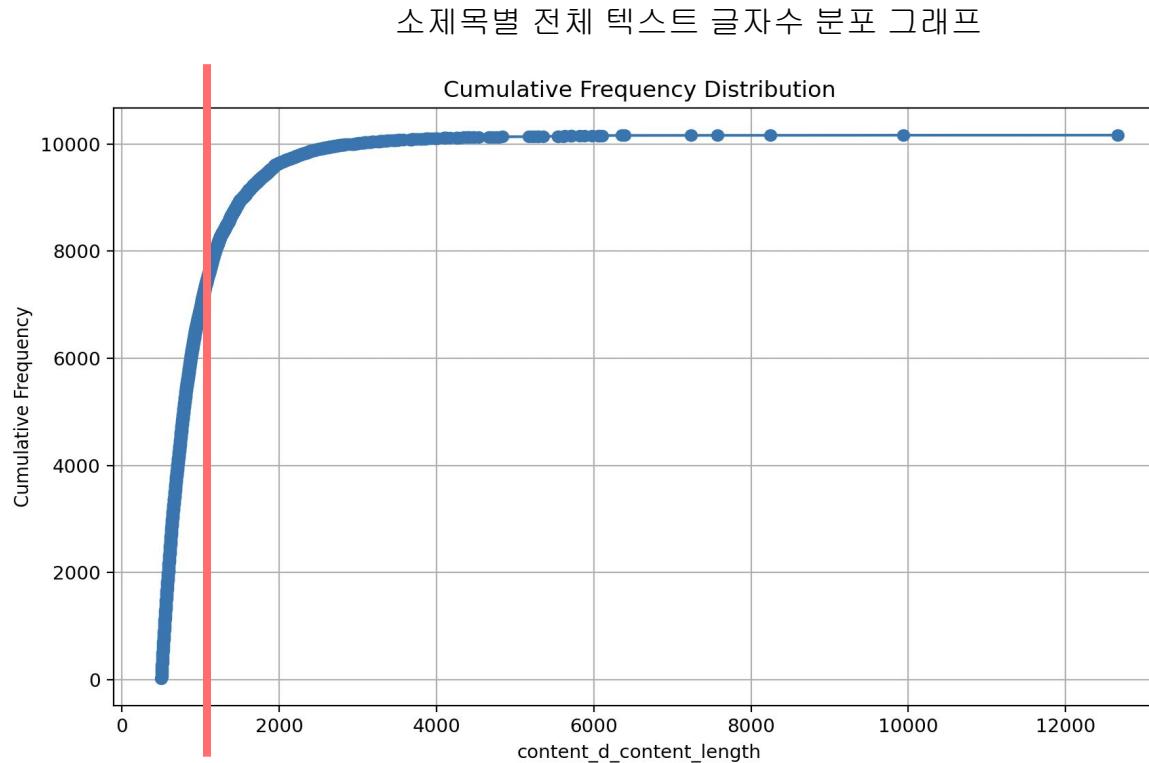
"title\":\"단순폭행죄\",

"content": [ \

    "사람의 신체에 대해 폭행을 가한 경우에는 2년 이하의 징역,  
     500만원 이하의 벌금, 구류 또는 과료에 처해집니다(「형법」 제260조제1항)."\",  
     "\\"다만, 폭행죄는 피해자의 명시적인 의사에 반해 공소를 제기할 수 없는  
     반의사불벌죄(反意思不罰罪)에 해당합니다(「형법」 제260조제3항)."\",  
     "\\"※ 단순폭행죄와 같이 피해자의 명시적인 의사에 반해 처벌할 수 없는  
     사건의 ... \"※ 법령용어해설\", \"구류(拘留): 자유형의 일종으로 1일 이상 30일  
     미만의 기간동안 유치장에 가두는 것을 말합니다. 이는 형벌 중 가장 가벼운  
     것으로서 주로 경미한 범죄에 대해 과해집니다.\", ... \"※ 폭행에 대한 판례\",  
     \"폭행에 해당한다고 본 판례\", \"안수기도 행위에 수반하는 신체적 행위가  
     단순히 손을 얹거나 약간 누르는 정도가 아니라 그것이 지나쳐서 가슴과 배를  
     반복하여 누르거나 때려 그로 인해 사망에 이른 것과 같은 정도의 것이라면  
     이는 사람의 신체에 대한 유형력의 행사로서 폭행에 해당합니다  
     (대법원 1994. 8. 23. 선고 94도1484 판결).\", ... ]

# Data – 데이터 분포

- 한 문서에서 소제목(s2)에 포함된 내용을 하나의 청크로 하기엔 텍스트의 길이가 긴 데이터들이 多



생활법령 데이터는 어느정도 통일된 구조를 가지는 데이터



구조 기반 청킹 로직 개발

# Data – 데이터 구조

## 기계식주차장 설치·철거 등

본문	100문 100답	카드뉴스	관련법령																																								
s1 <a href="#">기계식주차장 설치</a>																																											
s2 <a href="#">기계식주차장치의 개념</a>																																											
<p>“기계식주차장”은 노외주차장 및 부설주차장에 설치하는 주차설비로서 기계장치에 의하여 자동차를 주차할 장소로 이동시키기 위한 설비를 말합니다(「주차장법」 제2조제2호).</p>																																											
<p>• 기계식주차장의 설치 기준은 다음과 같습니다(「주차장법」 제19조의5 및 「주차장법 시행 규칙」 제16조의2제1항).</p> <p>➢ 기계식주차장 설치는 다음에 따라 자동차의 회전을 위한 공지(空地)(이하 “전면공지”라 함) 또는 자동차의 방향을 전환하기 위한 기계장치(이하 “방향전환장치”라 함)를 설치해야 합니다.</p> <ul style="list-style-type: none"> <li>✓ 중형 기계식주차장(길이 5.05m 이하, 너비 1.9m 이하, 높이 1.55m 이하, 무게 1,850kg 이하인 자동차를 주차할 수 있는 기계식주차장을 말함, 이하 같은) 너비 8.1m 이상, 길이 9.5m 이상의 전면공지 또는 지름 4m 이상의 방향전환장치와 그 방향전환장치에 접한 너비 1m 이상의 여유 공간</li> <li>✓ 대형 기계식주차장(길이 1.75m 이하, 너비 2.15m 이하, 높이 1.85m 이하, 무게 2,200kg 이하인 자동차를 주차할 수 있는 기계식주차장을 말함, 이하 같은) 너비 10m 이상, 길이 11m 이상의 전면공지 또는 지름 4.5m 이상의 방향전환장치와 그 방향전환장치에 접한 너비 1m 이상의 여유 공간</li> </ul> <p>➢ 기계식주차장의 내부에 방향전환장치를 설치한 경우와 2층 이상으로 주차구획이 배치되어 있고 출입구가 있는 층의 모든 주차구획을 기계식주차장치 출입구로 사용할 수 있는 기계식주차장의 경우에는 다음과 같이 설치합니다.</p> <ul style="list-style-type: none"> <li>✓ 자동차의 안전하고 원활한 통행을 확보하기 위하여 주차구획의 긴 변과 짧은 변 중 한 번 이상이 차로에 접하여 하며, 차로의 너비는 주차형식 및 출입구(차로식 또는 건축물식 주차장의 출입구를 포함)의 개수에 따라 다음 구분에 따른 기준 이상으로 해야 합니다(「주차장법 시행규칙」 제6조제1항제3호).</li> </ul>																																											
<table border="1"> <thead> <tr> <th colspan="3">이륜자동차전용 기계식주차장</th> </tr> <tr> <th rowspan="2">주차형식</th> <th colspan="2">차로의 너비</th> </tr> <tr> <th>출입구가 2개 이상인 경우</th> <th>출입구가 1개인 경우</th> </tr> </thead> <tbody> <tr> <td>평행주차</td> <td>2.25m</td> <td>3.5m</td> </tr> <tr> <td>직각주차</td> <td>4.0m</td> <td>4.0m</td> </tr> <tr> <td>45도 대향(對向)주차</td> <td>2.3m</td> <td>3.5m</td> </tr> </tbody> </table> <table border="1"> <thead> <tr> <th colspan="3">이륜자동차전용 외의 기계식주차장</th> </tr> <tr> <th rowspan="2">주차형식</th> <th colspan="2">차로의 너비</th> </tr> <tr> <th>출입구가 2개 이상인 경우</th> <th>출입구가 1개인 경우</th> </tr> </thead> <tbody> <tr> <td>평행주차</td> <td>3.3m</td> <td>5.0m</td> </tr> <tr> <td>직각주차</td> <td>6.0m</td> <td>6.0m</td> </tr> <tr> <td>60도 대향(對向)주차</td> <td>4.5m</td> <td>5.5m</td> </tr> <tr> <td>45도 대향주차</td> <td>3.5m</td> <td>5.0m</td> </tr> <tr> <td>교차주차</td> <td>3.5m</td> <td>5.0m</td> </tr> </tbody> </table> <p>➢ 보도와 차도의 구분이 있는 너비 12m 미만의 도로에 접하여 있는 부설주차장은 그 도로를 차로로 하여 주차단위구획을 배치할 수 있습니다. 이 경우 차로의 너비는 도로를 포함하여 6m 이상(평행주차형식의 경우에는 도로를 포함하여 4m 이상)으로 하며, 도로의 포함 범위는 중앙선까지로 하되, 중앙선이 없는 경우에는 도로 반대쪽 경계선까지로 합니다(「주차장법 시행규칙」 제11조제5항제2호).</p> <p>➢ 기계식주차장에는 도로에서 기계식주차장 출입구까지의 차로(이하 “진입로”라 함) 또는 전면공지와 접하는 장소에 자동차가 대기할 수 있는 장소(이하 “정류장”이라 함)를 설치해야 합니다. 이 경우 주차대수 20대를 초과하는 20대마다 한 대분의 정류장을 확보해야 하며, 정류장의 규모는 다음과 같습니다. 다만, 주차장의 풍구와 입구가 따로 설치되어 있으나 진입로의 너비가 6m 이상인 경우에는 충단경기도 6% 이하인 진입로의 길이 6m마다 한 대분의 정류장을 확보한 것으로 봅니다.</p> <ul style="list-style-type: none"> <li>✓ 중형 기계식주차장: 길이 5.05m 이상, 너비 1.9m 이상</li> <li>✓ 대형 기계식주차장: 길이 5.3m 이상, 너비 2.15m 이상</li> </ul> <p>➢ 기계식주차장에는 벽면으로부터 50cm 이내를 제외한 바닥면의 최소 조도를 다음과 같이 합니다.</p> <ul style="list-style-type: none"> <li>✓ 출입구: 최소 조도는 150.radians 이상</li> </ul> <p>• 특별사·광역사·특별자치도·시·군 또는 자치구는 지역실정에 고려된 구역을 정하여 기계식주차장의 설치대수, 기계식주차장치의 종류, 부설주차장과 주차대수 중 기계식주차장의 비율을 지방자치단체의 조례로 정할 수 있습니다(「주차장법」 제19조의5제2항).</p>				이륜자동차전용 기계식주차장			주차형식	차로의 너비		출입구가 2개 이상인 경우	출입구가 1개인 경우	평행주차	2.25m	3.5m	직각주차	4.0m	4.0m	45도 대향(對向)주차	2.3m	3.5m	이륜자동차전용 외의 기계식주차장			주차형식	차로의 너비		출입구가 2개 이상인 경우	출입구가 1개인 경우	평행주차	3.3m	5.0m	직각주차	6.0m	6.0m	60도 대향(對向)주차	4.5m	5.5m	45도 대향주차	3.5m	5.0m	교차주차	3.5m	5.0m
이륜자동차전용 기계식주차장																																											
주차형식	차로의 너비																																										
	출입구가 2개 이상인 경우	출입구가 1개인 경우																																									
평행주차	2.25m	3.5m																																									
직각주차	4.0m	4.0m																																									
45도 대향(對向)주차	2.3m	3.5m																																									
이륜자동차전용 외의 기계식주차장																																											
주차형식	차로의 너비																																										
	출입구가 2개 이상인 경우	출입구가 1개인 경우																																									
평행주차	3.3m	5.0m																																									
직각주차	6.0m	6.0m																																									
60도 대향(對向)주차	4.5m	5.5m																																									
45도 대향주차	3.5m	5.0m																																									
교차주차	3.5m	5.0m																																									
<p>※ 각 시·도의 조례는 &lt;국가법령정보센터(www.law.go.kr)-자치법규&gt;에서 확인하실 수 있습니다.</p>																																											

s3

s4 ~

table

- 생활법령 사이트에서 HTML 구조 기반으로 크롤링한 데이터 활용

- 원본 데이터의 구조에 따라 아래와 같이 한 문서 내의 텍스트를 구조 기반 문장 단위로 분리해서 사용

\* s2 는 소제목을 의미

\* s3 는 소제목에 포함되는 내용을 의미

\* s4 는 s3의 하위 내용을 의미

```
"doc_id":"7fe7951b8a55729461bf0bad62bd8137fd0da04a",
"full_title":"부설주차장 설치 · 운영>부설주차장 설치>기계식주차장 설치 등>기계식주차장 설치·철거 등",
"doc_title":"기계식주차장 설치·철거 등",
"chunks":[
    "s1_ 기계식주차장 설치",
    "s2_ 기계식주차장치의 개념",
    "s3_ \"기계식주차장치\"란 노외주차장 및 부설주차장에 설치하는 주차설비로서 기계장치에 의하여 자동차를 주차할 장소로 이동시키는 설비를 말합니다(「주차장법」 제2조제2호).",
    "s2_ 기계식주차장의 설치기준",
    "s3_ 기계식주차장치를 설치한 기계식주차장의 설치기준은 다음과 같습니다(「주차장법」 제19조의5 및 「주차장법 시행규칙」 제16조의2제1항).",
    "s4_ 기계식주차장치 출입구의 앞면에는 다음에 따라 자동차의 회전을 위한 공지(空地)(이하 \"전면공지\"라 함) 또는 자동차의 방향을 전환하기 위한 기계장치(이하 \"방향전환장치\"라 함)를 설치해야 합니다.",
    "doc_url":"https://www.easylaw.go.kr/CSP/CnpClsMain.laf?csmSeq=895&ccfNo=2&cciNo=3&cnpClsNo=1"}]
```

r2, t2 ~

# Chunking Logic

\* 앞서 데이터 구조에서 정의한 tag(s1, s2, s3, ...)를 바탕으로 작업

1

## 첨크 유형 정리

- s3까지만 있는 것
- s3 + s4
- s3 + s4 + reference
- s3 + table
- s3 + reference + table

2

## 첨크 유형별 처리 방식

- chunk\_title
- s3가 1개인 경우
- s3가 여러 개인 경우
- s3 + s4
- s3 + s4 + reference
- table 처리

3

## 청킹 결과

# Chunking Logic – 1. 청크 유형 정리

유형		예시	
s3까지만 있는 것	s3가 한 개	<ul style="list-style-type: none"> <li>▪ s2 : 신고기관</li> <li>▪ s3 : 초록색 박스 내용</li> </ul> <p>▼ 예시 이미지</p> <p>■ 신고기관</p> <ul style="list-style-type: none"> <li>• 옥외집회 또는 시위의 신고는 관할 경찰서에 하면 되는데, 다만 옥외집회 또는 시위 장소가 두 곳 이상의 경찰서의 관할에 속하는 경우에는 관할 시·도경찰청에 제출해야 하고, 두 곳 이상 시·도경찰청의 관할에 속하는 경우에는 주최자를 관할하는 시·도경찰청에 제출해야 합니다(<a href="#">규제</a> 「집회 및 시위에 관한 법률」 제6조제1항 단서).</li> </ul> <p>■ 접수증 교부</p> <ul style="list-style-type: none"> <li>• 위 신고서를 제출하면 관할 경찰관서로부터 접수증(「집회 및 시위에 관한 법률 시행규칙」 별지 제3호서식)을 받습니다(<a href="#">규제</a> 「집회 및 시위에 관한 법률」 제6조제2항).</li> </ul>	<ul style="list-style-type: none"> <li>* s2 는 소제목을 의미</li> <li>* s3 는 소제목에 포함되는 내용을 의미</li> <li>* s4 는 s3의 하위 내용을 의미</li> </ul>
	s3가 여러 개	<ul style="list-style-type: none"> <li>▪ s2 : 금전공탁서 제출 및 공탁</li> <li>▪ s3 : 초록색 박스 내용 2 개</li> </ul> <p>▼ 예시 이미지</p> <p>■ 금전공탁서 제출 및 공탁</p> <ul style="list-style-type: none"> <li>• 법원으로부터 현금공탁명령을 받은 경우 금전공탁서(재판상의 보증)를 작성하여, 담보제공명령을 받은 법원의 공탁관에게 담보제공명령서와 공탁서 2부를 제출합니다(「공탁규칙」 제20조제1항).</li> <li>• 수리된 공탁서 중 1부를 받아 법원이 지정한 은행에 공탁금을 납부합니다(「공탁규칙」 제26조제1항).</li> </ul>	
s3 + s4	s4가 짧은 경우	<ul style="list-style-type: none"> <li>▪ s2 : 식중독의 예방</li> <li>▪ s3 : 초록색 박스 내 "식중독"이란?</li> <li>▪ s4 : 초록색 박스 내 s3 하단 내용</li> </ul> <p>▼ 예시 이미지</p> <p>■ 식중독의 예방</p> <ul style="list-style-type: none"> <li>• 캠핑을 갈 때 준비해야 할 물품 중 음식은 큰 비중을 차지하는 것들 중 하나입니다. 그러나 냉장보관을 해야 하는 음식을 잘못 보관해 가져간다거나, 캠핑을 하며 발견한 주위의 버섯이나 낙지를 한 생선 등을 잘못 먹는다거나 하면 식중독에 걸릴 수 있으므로 유의해야 합니다.</li> <li>• "식중독"이란 ?       <ul style="list-style-type: none"> <li>▶ 식품 섭취로 인체에 유해한 미생물 또는 유독물질에 의해 발생했거나 발생한 것으로 판단되는 감염성 질환 또는 독소형 질환을 말합니다(<a href="#">규제</a> 「식품위생법」 제2조제14호).</li> </ul> </li> </ul>	

# Chunking Logic – 1. 청크 유형 정리

<p>s4가 긴 경우</p> <ul style="list-style-type: none"> <li>▪ s2 : 수의사업의 구분</li> <li>▪ s3 : 법인세의 부과대상이 되는 수의사업</li> <li>▪ s4 : 초록색 박스 내 s3 하단 내용</li> </ul> <p>▼ 예시 이미지</p>	<p>* s2 는 소제목을 의미</p> <p>* s3 는 소제목에 포함되는 내용을 의미</p> <p>* s4 는 s3의 하위 내용을 의미</p>
--	--

# Chunking Logic – 1. 청크 유형 정리

reference가 있는 경우

s3 + s4 + reference(r, t)	reference가 한 개	<ul style="list-style-type: none"> <li>s3 : 이물질이 들어간 경우</li> <li>s4 : 초록색 박스 바로 위에 오는 문장</li> <li>r : 초록색 박스 내용</li> </ul> <p>▼ 예시 이미지</p>
	reference가 여러 개	<p>▼ 예시 이미지</p>

table이 있는 경우

s3 + table	table 한 개	▼ 예시 이미지																																				
		<p>▼ 상근예비역의 입영순서는 입영희망시기가 빠른 사람이 1순위입니다.</p> <p>· 지방병무청장은 상근예비역소집 대상자로 선발된 사람에 대해 군소요 제기 시-구-읍-면별로 다음의 순위로 입영순서를 결정합니다. [1] 「법역병 시행령」 제33조제1항제2항 및 「현역병 입영업무 규정」(방무청훈령 제2071호, 2024. 6. 발령-시행) 제45조 제1항].</p> <table border="1"> <thead> <tr> <th>순위</th><th>대상</th><th>비고</th></tr> </thead> <tbody> <tr> <td>1순위</td><td>입영희망시기가 빠른 사람</td><td>순위가 동일한 때에는 생년 월일이 빠른 순으로 함</td></tr> <tr> <td>2순위</td><td>입영희망시기를 선택하지 않은 사람</td><td></td></tr> </tbody> </table> <p>· 지방병무청장은 상근예비역소집 대상자로 선발된 사람에게 입영할 때에는 가까운 입영일에 입영하게 할 수 있습니다(「법역병 입영업무 규정」 제45조제2항).</p> <p>· 지방병무청장은 현역병 입영일자로 선발된 사람 중에서 수령 등의 사용료 상근예비역소집 대상자로 추가 선발한 사람에 대해서는 선발전의 입영일자 및 입영부대에 상근예비역소집 대상자로 입영하게 해야 합니다(「현역병 입영업무 규정」 제45조제3항).</p> <p>※ 다만, 그 입영일자 및 입영부대에 상근예비역으로 입영하게 할 수 있음</p> <p>· 지방병무청장은 해당연도에 상근예비역소집 대상자의 현역병입영일자 연기(기자기 포함) 사용료가 해소되었으나, 해당 연도에 상근예비역 입영계정이 종료되어 입영할 수 없는 경우에는 다음 연도에 상근예비역소집 대상자로 입영하게 해야 합니다(「현역병 입영업무 규정」 제45조제4항).</p>	순위	대상	비고	1순위	입영희망시기가 빠른 사람	순위가 동일한 때에는 생년 월일이 빠른 순으로 함	2순위	입영희망시기를 선택하지 않은 사람																												
순위	대상	비고																																				
1순위	입영희망시기가 빠른 사람	순위가 동일한 때에는 생년 월일이 빠른 순으로 함																																				
2순위	입영희망시기를 선택하지 않은 사람																																					
	table 여러 개	<p>▼ 예시 이미지</p> <p>▼ 이용객의 귀책사유로 계약을 취소한 경우</p> <p>· 이용객에게 사용이 생겨서 속박·식박·이용계약을 해제하는 경우에는 다음의 기준에 따른 금액을 환급받을 수 있습니다(「소비자 분쟁해결기준」 제4조 및 별표 2 26. 속박).</p> <p>▶ 성수기인 경우</p> <table border="1"> <thead> <tr> <th>유형</th><th>원금기준</th></tr> </thead> <tbody> <tr> <td>사용예정일 10일 전까지 취소 또는 계약체결일 취소</td><td>계약금 원금</td></tr> <tr> <td>사용예정일 7일 전까지 취소</td><td>총요금의 10% 공제후 원금</td></tr> <tr> <td>성수기 주 중 사용예정일 5일 전까지 취소</td><td>총요금의 30% 공제후 원금</td></tr> <tr> <td>사용예정일 3일 전까지 취소</td><td>총요금의 50% 공제후 원금</td></tr> <tr> <td>사용예정일 1일 전까지 또는 사용예정일 당일 취소</td><td>총요금의 80% 공제후 원금</td></tr> <tr> <td>사용예정일 10일 전까지 취소 또는 계약체결일 취소</td><td>계약금 원금</td></tr> <tr> <td>사용예정일 7일 전까지 취소</td><td>총요금의 20% 공제후 원금</td></tr> <tr> <td>성수기 주 말 사용예정일 5일 전까지 취소</td><td>총요금의 40% 공제후 원금</td></tr> <tr> <td>사용예정일 3일 전까지 취소</td><td>총요금의 60% 공제후 원금</td></tr> <tr> <td>사용예정일 1일 전까지 또는 사용예정일 당일 취소</td><td>총요금의 90% 공제후 원금</td></tr> </tbody> </table> <p>▶ 비수기인 경우</p> <table border="1"> <thead> <tr> <th>유형</th><th>배상기준</th></tr> </thead> <tbody> <tr> <td>사용예정일 2일 전까지 취소</td><td>계약금 원금</td></tr> <tr> <td>사용예정일 1일 전까지 취소</td><td>총요금의 10% 공제후 원금</td></tr> <tr> <td>사용예정일 당일 취소 또는 연락 없이 불참</td><td>총요금의 20% 공제 후 원금</td></tr> <tr> <td>사용예정일 2일 전까지 취소</td><td>계약금 원금</td></tr> <tr> <td>사용예정일 1일 전까지 취소</td><td>총요금의 20% 공제후 원금</td></tr> <tr> <td>사용예정일 당일 취소 또는 연락 없이 불참</td><td>총요금의 30% 공제 후 원금</td></tr> </tbody> </table>	유형	원금기준	사용예정일 10일 전까지 취소 또는 계약체결일 취소	계약금 원금	사용예정일 7일 전까지 취소	총요금의 10% 공제후 원금	성수기 주 중 사용예정일 5일 전까지 취소	총요금의 30% 공제후 원금	사용예정일 3일 전까지 취소	총요금의 50% 공제후 원금	사용예정일 1일 전까지 또는 사용예정일 당일 취소	총요금의 80% 공제후 원금	사용예정일 10일 전까지 취소 또는 계약체결일 취소	계약금 원금	사용예정일 7일 전까지 취소	총요금의 20% 공제후 원금	성수기 주 말 사용예정일 5일 전까지 취소	총요금의 40% 공제후 원금	사용예정일 3일 전까지 취소	총요금의 60% 공제후 원금	사용예정일 1일 전까지 또는 사용예정일 당일 취소	총요금의 90% 공제후 원금	유형	배상기준	사용예정일 2일 전까지 취소	계약금 원금	사용예정일 1일 전까지 취소	총요금의 10% 공제후 원금	사용예정일 당일 취소 또는 연락 없이 불참	총요금의 20% 공제 후 원금	사용예정일 2일 전까지 취소	계약금 원금	사용예정일 1일 전까지 취소	총요금의 20% 공제후 원금	사용예정일 당일 취소 또는 연락 없이 불참	총요금의 30% 공제 후 원금
유형	원금기준																																					
사용예정일 10일 전까지 취소 또는 계약체결일 취소	계약금 원금																																					
사용예정일 7일 전까지 취소	총요금의 10% 공제후 원금																																					
성수기 주 중 사용예정일 5일 전까지 취소	총요금의 30% 공제후 원금																																					
사용예정일 3일 전까지 취소	총요금의 50% 공제후 원금																																					
사용예정일 1일 전까지 또는 사용예정일 당일 취소	총요금의 80% 공제후 원금																																					
사용예정일 10일 전까지 취소 또는 계약체결일 취소	계약금 원금																																					
사용예정일 7일 전까지 취소	총요금의 20% 공제후 원금																																					
성수기 주 말 사용예정일 5일 전까지 취소	총요금의 40% 공제후 원금																																					
사용예정일 3일 전까지 취소	총요금의 60% 공제후 원금																																					
사용예정일 1일 전까지 또는 사용예정일 당일 취소	총요금의 90% 공제후 원금																																					
유형	배상기준																																					
사용예정일 2일 전까지 취소	계약금 원금																																					
사용예정일 1일 전까지 취소	총요금의 10% 공제후 원금																																					
사용예정일 당일 취소 또는 연락 없이 불참	총요금의 20% 공제 후 원금																																					
사용예정일 2일 전까지 취소	계약금 원금																																					
사용예정일 1일 전까지 취소	총요금의 20% 공제후 원금																																					
사용예정일 당일 취소 또는 연락 없이 불참	총요금의 30% 공제 후 원금																																					

# Chunking Logic – 1. 청크 유형 정리

s3 + reference + table

## ▼ 예시 이미지

### ▣ 물품이용의 유형과 내용

물품 이용의 유형	내용
물품의 대부	지방자치단체의 장은 대부를 목적으로 한 물품 또는 대부하여도 지방자치단체의 사업 또는 사무에 지장이 없는 물품은 대부할 수 있습니다(「공유재산 및 물품 관리법」 제74조제1항).
물품의 매각	지방자치단체의 장은 매각을 목적으로 한 물품이나 불용의 결정을 받은 물품(이하 “불용품”이라 함)의 경우에는 매각할 수 있습니다(「공유재산 및 물품 관리법」 제76조제1항).  물품을 매각할 때에는 일반입찰에 부쳐야 합니다. 다만 특례가 인정되는 경우에는 경매나 수의계약으로 매각할 수 있습니다(「공유재산 및 물품 관리법」 제76조제2항 참조).  지방자치단체의 장은 물품으로서 매각되지 않은 것은 매각의 특례를 둘 수 있습니다(「공유재산 및 물품 관리법」 제76조제3항 참조).
물품의 양여	지방자치단체의 장은 용도가 지정된 국고보조금이나 기부금에 의하여 취득한 물품으로서 불용품은 그 용도에 따라 양여할 수 있습니다(「공유재산 및 물품 관리법」 제78조제1항).
물품의 교환	지방자치단체는 물품을 효율적으로 관리하기 위하여 특히 필요한 경우에는 그 물품을 그 지방자치단체 외의 자가 소유한 같은 종류의 물품과 교환할 수 있습니다(「공유재산 및 물품 관리법」 제79조제1항).

### ※ 용어해설

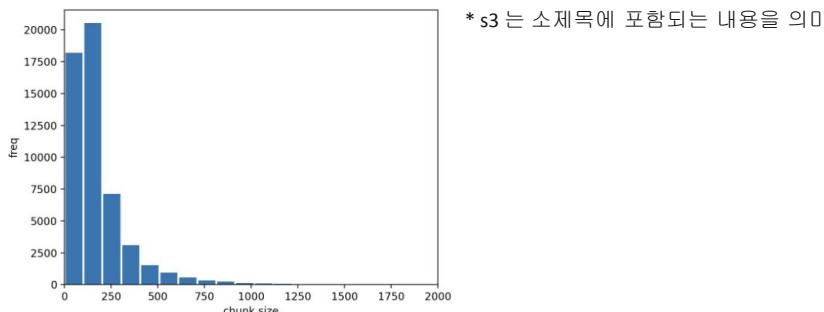
» “불용품(不用品)”이란 지방자치단체의 장은 소관 물품 중 사용할 필요가 없거나 사용할 수 없는 물품이 있으면 그 물품에 대하여 불용의 결정을 하여야 하며 이에 따라 불용의 결정을 한 물품을 불용품이라고 합니다(「공유재산 및 물품 관리법」 제75조제1항 및 제2항).

# Chunking Logic – 2. 청크 유형별 처리 방식

- 청크 분리 기준 : 글자수

- 텍스트 길이가 길 때 자르는 기준 : 글자수 **700** 이상이면 cut

s3로 텍스트를 끊었을 때 글자수 분포 그래프



글자수	예시
1000	<p>신고절차 및 첨부서류 건축물의 건축신고를 하려는 사람은 건축·대수선·용도변경(변경)신고서에 다음의 서류를 첨부해서 특별자치시장·특별자치도지사 또는 시장·군수·구청장에게 제출(전자문서로 제출하는 것을 포함)해야 합니다. 이경우 특별자치시장·특별자치도지사 또는 시장·군수·구청장은 행정정보의 공동이용을 통해「건축법 시행규칙」제12조제1항제4호의 서류 중 토지등기사항증명서를 확인해야 합니다(「건축법」제14조제1항, 「건축법 시행규칙」제12조및별지 제6호서식). 1. 「건축법 시행규칙」별표 2 중 배치도·평면도(총별도 작성된 것만 해당)·입면도 및 단면도. 다만, 다음 각 목의 경우에는 각 목의 구분에 따른 도서를 말합니다. 가. 연면적의 합계가 100제곱미터를 초과하는 「건축법 시행령」별표 1제1호의 단독주택을 건축하는 경우: 「건축법 시행규칙」별표 2의 설계도서 중 건축계획서·배치도·평면도·입면도·단면도 및 구조도(구조내력상 주요한 부분의 평면 및 단면을 표시한 것만 해당) 나. 「건축법」제23조제4항에 따른 표준설계도서에 따라 건축하는 경우: 건축계획서 및 배치도 다. 「건축법」제10조에 따른 사전결정을 받은 경우: 평면도 2. 「건축법」제11조제5항 각 호에 따른 허가 등을 받거나 신고를 하기 위하여 해당법령에서 제출하도록 의무화하고 있는 신청서 및 구비서류( 해당사항이 있는 경우로 한정) 3. 건축할 대지의 범위에 관한 서류 4. 건축할 대지의 소유 또는 그 사용에 관한 권리증명하는 서류. 다만, 건축할 대지에 포함된 국유지·공유지에 대해서는 허가권자가 해당 토지의 관리청과 협의하여 그 관리청이 토지를 건축주에게 매각하거나 양여할 것을 확인한 서류로 그 토지의 소유에 관한 권리를 증명하는 서류를 갈음할 수 있습니다. 특별자치시장·특별자치도지사 또는 시장·군수·구청장이 건축신고서를 받은 때에는 그 기재내용을 확인한 후 그 신고의 내용에 따라건축신고필증을 신고인에게 교부해야 합니다(「건축법」제14조제1항, 「건축법 시행규칙」제12조제3항 및별지 제7호서식).</p>
634	<p>조합을 설립하려면 토지등소유자의 4분의 3 이상 및 토지면적의 2분의 1 이상의 토지소유자의 동의를 받아야 합니다(「도시 및 주거환경정비법」제35조제2항). 토지등소유자의 동의를 받으려는 경우에는 조합설립 동의서(「도시 및 주거환경정비법 시행규칙」별지 제6호서식)에 다음의 사항을 포함하여 동의를 받아야 합니다(「도시 및 주거환경정비법 시행령」제30조제1항·제2항 및「도시 및 주거환경정비법 시행규칙」제8조제3항). 건설되는 건축물의 설계의 개요 정비사업비 정비사업비의 분담기준 사업 완료 후 소유권의 귀속에 관한 사항 조합 정관 ※ 조합의 정관에 기재해야 하는 사항에 대한 자세한 내용은 「도시 및 주거환경정비법」제40조제1항에서 확인할 수 있습니다. 추진위원회는 조합설립에 필요한 동의를 받기 전에 다음의 정보를 토지등소유자에게 제공해야 합니다(「도시 및 주거환경정비법」제35조제10항 및「도시 및 주거환경정비법 시행령」제32조). 토지등소유자별 분담액 및 산출근거 그 밖에 추정 분담금의 산출 등과 관련하여 특별시·광역시·특별자치시·도·특별자치도 또는 「지방자치법」제175조에 따른 서울특별시·광역시 및 특별자치시를 제외한 인구 50만 이상 대도시의 조례로 정하는 정보 ※ “토지등소유자 동의방법” 더 알아보기</p>

- 설정 이유 :

글자수가 700 이하인 데이터들이 하나로 보기 좋으며, 아래와 같은 사례 내용의 텍스트도 “(판단)”에 해당하는 내용이 대략 600자 정도

"chunk":

"2. 변질된 음료로 인한 손해배상 청구 (사건개요) 신청인은 2008. 10. 15. 편의점에서 피신청인이 판매하는 비닐팩 음료인 '제로칼로리 \*\*맛(150g)' 제품 4개(이하 '이 사건 음료'라고 함)를 구입한 후 점심시간에 1개를 마시려고 개봉하였는데 비닐팩 마개 주위에 검게 탄 흔적이 있었고, 투명색인 비닐팩 내용물 전체가 검은색으로 변색되어 있어 당시의 충격과 불쾌함으로 인한 정신적 손해배상 등을 요구함. (판단) 이 사건 제품의 변질은 포장재 불량으로 인한 공기 흡입 때문인 것으로 확인되었는바 제품을 유통·판매하는 피신청인은 제품을 판매함에 있어 필요한 주의의무를 다하지 못하였으므로 신청인이 변질된 제품으로 인해 입은 손해를 배상할 의무가 있다. 재산적 손해배상액: 신청인이 함께 구입한 이 사건 음료 4개 중 1개만 변질되었지만, 나머지 3개 음료는 이 물질 발견으로 인한 불쾌감 때문에 마시기 어려워므로 피신청인은 변질된 제품을 구매함으로서 신청인이 입은 재산적 손해인 제품 4개의 구입대금에 상당하는 금 4,800원을 배상하는 것이 타당하다. 정신적 손해 배상액: 일반적으로 타인의 불법행위로 인하여 재산권이 침해되는 경우 재산적 손해의 배상에 의하여 정신적 고통도 충분히 회복되지만(서울고등법원 1992. 10. 30. 선고 92나23102 판결 참조), 음료의 변질 사고는 혐오 물질을 발견하였을 당시의 정신적 충격이 매우 크므로, 피신청인은 이 사건 음료의 변질 사고로 인한 정신적 손해배상의 책임을 보상해야 한다. 따라서 정신적 손해배상의 범위에 대하여는 신청인이 이 사건 제품을 오랫동안 음용했다는 사실과 제품의 변질 정도 등을 종합적으로 고려한 금 200,000원으로 정하는 것이 상당하다."

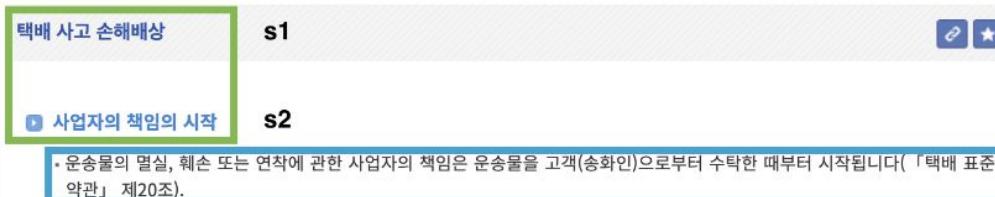
# Chunking Logic – 2. 청크 유형별 처리 방식

## 0. chunk\_title

\* s1(대제목), s2(소제목)을 의미

### ▲ title 처리

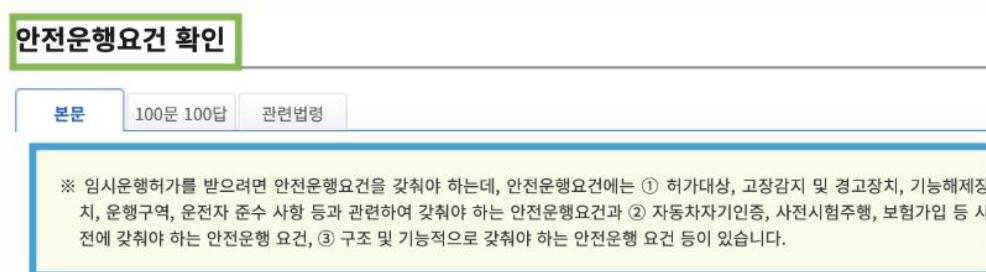
- s1, s2가 있는 경우
  - chunk\_title : s1 + s2 + ({idx}) (\* idx는 내용이 나뉘지는 경우 한 문서에 있는 내용임을 나타내기 위해 추가)



- CHUNK : {"chunk\_title" : "택배 사고 손해배상 : 사업자의 책임의 시작", "chunk\_content" : "운송물의 멸실, 훼손 또는 ..."}

- s1, s2가 없는 경우

- 예시



- CHUNK : {"chunk\_title" : "안전운행요건 확인", "chunk\_content" : "※ 임시운행허가를 받으려면 안전운행요건을 갖춰야 하는데, ..."}

# Chunking Logic – 2. 청크 유형별 처리 방식

## 1. s3 가 1개인 경우 (기본)

(아래에서 길이는 s3에 해당하는 파란색 박스 내의 텍스트의 길이를 의미)

\* s3(sentence)는 소제목에 포함되는 내용을 의미

- a. 길이가 짧은 경우 → 하나의 chunk

결과 예시

택배 사고 손해배상	<b>s1</b>	
	<b>s2</b>	
사업자의 책임의 시작	<b>s3</b>	

• 운송물의 멸실, 훼손 또는 연착에 관한 사업자의 책임은 운송물을 고객(송화인)으로부터 수탁한 때부터 시작됩니다(「택배 표준 약관」 제20조).

CHUNK :

```
{ chunk_title : s1 + s2(초록색 박스 텍스트),  
chunk_content : 파란색 박스}
```

예시)

```
{ chunk_title : “택배 사고 손해배상 : 사업자의 책임의 시작”,  
chunk_content : “운송물의 멸실, 훼손 또는 연착에 관한 ...”}
```

- b. 길이가 긴 경우 → 글자수(700)를 기준으로 내용을 나눈다.

데이터 중 b에 해당하는 경우는 없는 것으로 확인.

# Chunking Logic – 2. 청크 유형별 처리 방식

## 2. s3 가 여러개인 경우

(아래의 길이는 s3에 해당하는 파란색 박스 내의 각 텍스트의 길이를 의미)

\* s3(sentence)는 소제목에 포함되는 내용을 의미

### a. 길이가 짧은 경우 → s2로 묶는다.

: s2(소제목)에 포함된 전체 내용이 글자수 기준 700을 넘지 않는 경우를 의미

#### s2

##### 민사조정기일의 통지

- s3** ■ 민사조정 신청을 하면 분쟁의 당사자에게 조정기일이 통지됩니다(「민사조정법」 제15조제1항).
- s3** ■ 분쟁 당사자는 조정기일에 출석해서 조정담당판사 앞에서 진술을 해야 합니다. 다만, 특별한 사정이 있는 경우에는 조정담당판사의 허가를 얻어 대리인을 출석시키거나 보조인을 동반할 수 있습니다(「민사조정규칙」 제6조제1항).
- s3** ■ 조정담당판사는 상당하다고 인정하는 때에는 당사자의 신청을 받거나 동의를 얻어 비디오 등 중계장치에 의한 중계시설을 통하여거나 인터넷 화상장치를 이용하여 조정기일을 열 수 있습니다(「민사조정규칙」 제6조의2제1항).

#### CHUNK :

```
{ chunk_title : "민사조정기일의 통지",
  chunk_content : "민사조정 신청을 하면 ... 분쟁 당사자는...
  조정담당판사는 상당하다고 인정하는 ..."}
```

# Chunking Logic – 2. 청크 유형별 처리 방식

## 2. s3 가 여러개인 경우

(아래의 길이는 s3에 해당하는 파란색 박스 내의 각 텍스트의 길이를 의미)

### b. 길이가 긴 경우 → s3로 묶는다.

\* “s3로 묶는다”는 의미는 s3의 tag를 가지는 sentence가 하나의 청크로 생성된다는 것

: s2(소제목)에 포함된 전체 내용이 글자수 기준 700을 넘는 경우를 의미

s2	<b>운송장에 운송물의 가액을 기재하지 않은 경우</b>	
s3	운송장에 운송물의 가액을 기재하지 않은 경우에는 손해배상한도액은 50만원으로 하되, 운송물의 가액에 따라 할증요금을 지급하는 경우의 손해배상한도액은 각 운송가액 구간별 운송물의 최고가액으로 합니다(「택배 표준약관」 제22조제3항).	→ CHUNK 0
s3	운송물이 전부 멸실된 때에는 인도예정일의 인도예정장소에서의 운송물 가액을 기준으로 산정한 손해액 또는 고객(송화인)이 입증한 운송물의 손해액(영수증 등)을 지급합니다(「택배 표준약관」 제22조제3항제1호).	→ CHUNK 1
s3	운송물이 일부 멸실된 때에는 인도일의 인도장소에서의 운송물 가액을 기준으로 산정한 손해액 또는 고객(송화인)이 입증한 운송물의 손해액(영수증 등)을 지급합니다(「택배 표준약관」 제22조제3항제2호).	→ CHUNK 2
s3	운송물이 훼손된 때에는 ① 수선이 가능한 경우는 실수선 비용(A/S비용), ② 수선이 불가능한 경우는 인도일의 인도장소에서의 운송물 가액을 기준으로 산정한 손해액 또는 고객(송화인)이 입증한 운송물의 손해액(영수증 등)을 지급합니다(「택배 표준약관」 제22조제3항제3호).	→ CHUNK 3
s3	운송물이 연착되고 일부 멸실 및 훼손되지 않은 때에는 ① 수선이 가능한 경우는 실수선 비용(A/S비용), ② 수선이 불가능한 경우는 인도일의 인도장소에서의 운송물 가액을 기준으로 산정한 손해액 또는 고객(송화인)이 입증한 운송물의 손해액(영수증 등)을 지급합니다(「택배 표준약관」 제22조제3항제4호).	→ CHUNK 4
s3	운송물이 연착되고 일부 멸실 또는 훼손된 때에는 ① 일부 멸실된 때에는 인도예정일의 인도장소에서의 운송물 가액을 기준으로 산정한 손해액 또는 고객(송화인)이 입증한 운송물의 손해액(영수증 등)을, ② 수선이 가능한 경우는 수선해 주고, ③ 수선이 불가능한 경우는 인도예정일의 인도장소에서의 운송물 가액을 기준으로 산정한 손해액 또는 고객(송화인)이 입증한 운송물의 손해액(영수증 등)을 지급합니다(「택배 표준약관」 제22조제3항제5호).	
		{chunk_title : “택배 사고 손해배상 : 운송장에 운송물의 가액을 기재하지 않은 경우 1/6” chunk_content : “운송장에 운송물의 가액을 기재하지 않은 ...”}
		{chunk_title : “택배 사고 손해배상 : 운송장에 운송물의 가액을 기재하지 않은 경우 2/6” chunk_content : “운송물이전부 멸실된 때에는 인도예정일 ...”}
		{chunk_title : “택배 사고 손해배상 : 운송장에 운송물의 가액을 기재하지 않은 경우 3/6” chunk_content : “운송물이일부 멸실된 때에는 인도일의 ...”}
		{chunk_title : “택배 사고 손해배상 : 운송장에 운송물의 가액을 기재하지 않은 경우 4/6” chunk_content : “운송물이훼손된 때에는 ① ...”}
		{chunk_title : “택배 사고 손해배상 : 운송장에 운송물의 가액을 기재하지 않은 경우 5/6” chunk_content : “운송물이연착되고 일부 멸실 ...”}

# Chunking Logic – 2. 청크 유형별 처리 방식

## 3. s3 + s4

(s4는 s3의 하위 내용으로 s3보다 들어쓰기 된 내용을 의미)

\* s3(sentence)는 소제목에 포함되는 내용을 의미

s4(sentence)는 s3의 하위 내용을 의미

### a. s4의 길이가 짧은 경우 → s3로 묶는다.

: s3(sentence)를 포함한 s3의 하위 내용(s4)이 전체 텍스트가 글자수 기준 700을 넘지 않는 경우를 의미

#### ▶ 대부료의 납입기한

- s3** ▶ 물품 대부료의 납입기한은 다음의 구분에 따릅니다(「공유재산 및 물품 관리법 시행령」 제74조제3항).
- s4** ▶ 물품 대부기간이 1년 미만인 경우 : 대부일부터 1개월 이내
- s4** ▶ 물품 대부기간이 1년 이상인 경우 : 대부일이 속한 해의 다음 해부터 매년 그 대부일에 해당하는 날부터 1개월 이내

#### CHUNK :

```
{ chunk_title : "대부료의 납입기한",
  chunk_content : "물품 대부료의 납입기한은 ... 물품 대부
  기간이 1년 미만인 ... 물품 대부 기간이
  1년 이상인 ..."}
```

# Chunking Logic – 2. 청크 유형별 처리 방식

## 3. s3 + s4

(s4는 s3의 하위 내용으로 s3보다 들어쓰기 된 내용을 의미)

\* s3(sentence)는 소제목에 포함되는 내용을 의미  
s4(sentence)는 s3의 하위 내용을 의미

### b. 길이가 긴 경우 → s3로 묶은 후, 길이를 기준으로 다시 나눈다.

: s3(sentence)를 포함한 s3의 하위 내용(s4)이 전체 텍스트가 글자수 기준 700을 넘는 경우를 의미

세부 과정	
s3	1. s3로 묶는다. (이미지의 파란색 박스 단위를 의미)
s4	2. 1에서 묶은 청크의 전체 텍스트가 글자수 700을 넘어가는 경우 (이미지의 첫번째 파란색 박스를 의미)
s4	3. 2에 해당하는 청크를 다시 글자수를 기준으로 나눈다. (이미지의 빨간색 박스 단위를 의미)
CHUNK 0	<p>{chunk_title : “음식물류 폐기물을 배출제한 1/4”</p> <p>chunk_content : "음식물류 폐기물을 다량으로 배출하는 ... “}</p>
CHUNK 1	<p>{chunk_title : “음식물류 폐기물을 배출제한 2/4”</p> <p>chunk_content : "2. 식품접객업 중 사업장 규모가 200제곱 ... “}</p>
CHUNK 2	<p>{chunk_title : “음식물류 폐기물을 배출제한 3/4”</p> <p>chunk_content : "음식물류 폐기물을 배출하는 음식물류 폐기물의 발생 ... “}</p>
CHUNK 3	<p>{chunk_title : “음식물류 폐기물을 배출제한 4/4”</p> <p>chunk_content : "음식물류 폐기물을 배출하는 음식물류 폐기물을 스스로 ... “}</p>

# Chunking Logic – 2. 청크 유형별 처리 방식

## 4. s3 + s4 + reference

(reference는 아래 이미지에서 하늘색 박스에 해당하는 내용을 의미)

- reference의 tag는 r2, t2, ... 와 같이 s가 아닌 r, t로 나타냄.
- tag를 s(sentence)로 치환한 후, 1 ~ 3번 과정으로 처리

\* r,t(reference) tag를 s로 치환하는 이유

: reference의 내용이 결국에는 한 문서에서 s3와 같은 역할을 하는 것으로 볼 수 있음.

### 물놀이 전 주의사항

s3

• 상수원보호구역에서의 물놀이 금지

▶ 환경부장관은 상수원의 확보와 수질 보전을 위해 필요하다고 인정되는 지역을 상수원보호구역으로 지정할 수 있습니다(규제 「수도법」 제7조제1항).

▶ 상수원보호구역에서는 수영, 목욕, 야영 또는 수면을 이용한 레저행위가 금지되므로(규제 「수도법」 제7조제3항제2호 및 규제 「수도법 시행령」 제12조제1항제2호 및 제3호), 물놀이를 하려는 곳이 상수원보호구역은 아닌지 확인하세요.

※ 상수원보호구역에서는 물놀이 뿐만 아니라 세탁, 야외 취사행위 등도 금지된다는 사실을 잊지 마세요.

s3

• 오염된 하천 등에서의 물놀이 금지

▶ 하천 등이 오염되어 수영 등의 물놀이를 하는 경우 사람의 건강이나 생활에 미치는 피해가 크다고 인정되는 경우 관할 구역 내 주민 등 이해관계자에게 해당 하천 등에서 수영 등의 행위를 자제하도록 안내하는 등의 조치가 취해지므로(「물환경보전법」 제21조의2제1항 및 제2항 참조), 물놀이를 하려는 곳에 이런 안내문이 있는지 확인하세요.

s3

※ 물에 들어가기 전 준비사항

▶ 물에 들어가기 전에는 반드시 준비운동을 하고, 물놀이는 물 깊이가 배꼽 아래까지 오는 곳에서 합니다.

▶ 물에 처음 들어가기 전 심장에서 먼 부분부터(다리, 팔, 얼굴, 가슴 등의 순서) 물을 적신 후 들어가세요.

▶ 물놀이 중 비가 오거나 천둥, 번개가 치면 물에서 즉시 나와야 합니다.

▶ 음식물을 먹은 직후 바로 물에 들어가거나 음주 후에 물에 들어가지 않도록 합시다.

▶ 물놀이 도중 소리를 듣고 피부가 땅기거나 몸이 떨리고 입술이 파래지면 물놀이를 중단하고 옷이나 타월 등으로 몸을 따뜻하게 감싸고 휴식을 취합시다.

▶ 수영을 할 때 껌을 씹거나 음식물을 먹으면 질식의 위험이 있으므로 주의합시다.

< 행정안전부 어린이, 안전 배움터, 여름철 안전 >

CHUNK 0

{chunk\_title : “물놀이 전 주의사항 1/3”  
chunk\_content : “상수원보호구역에서의 물놀이 금지...”}

CHUNK 1

{chunk\_title : “물놀이 전 주의사항 2/3”  
chunk\_content : “오염된 하천 등에서의 물놀이 금지...”}

CHUNK 2

{chunk\_title : “물놀이 전 주의사항 3/3”  
chunk\_content : “※ 물에 들어가기 전 준비사항 ...”}

# Chunking Logic – 2. 청크 유형별 처리 방식

## 4. s3 + s4 + reference

- 문제 : 다음과 같이 과하게 잘리는 문제가 발생.

가압류신청 진술서	
채권자는 가압류 신청과 관련하여 다음 사실을 진술합니다. 다음의 진술과 관련하여 고의로 누락하거나 허위로 진술한 내용이 발견된 경우에는, 그로 인하여 보정명령 없이 신청이 기각되거나 가압류이의절차에서 불이익을 받을 것임을 잘 알고 있습니다.	
s3	20...
s3	채권자(대리인) _____ (날인 또는 서명)
s3	◇ 다음 ◇
1. 피보전권리와 관련하여	
가. 채무자가 신청서에 기재한 청구채권을 인정하고 있습니까?	
» 예	
» 아니오 → 채무자의 주장의 요지 :	
나. 채무자가 청구채권과 관련하여 오히려 채권자로부터 받을 채권을 가지고 있다고 주장하고 있습니까?	
» 예 → 채무자의 주장의 요지 :	
» 아니오	
다. 채권자가 신청서에 기재한 청구금액은 본안소송에서 승소할 수 있는 금액으로 적정하게 산출된 것입니까? (과도한 가압류로 인해 채무자가 손해를 입으면 배상하여야 함)	
» 예	
» 아니오	

### 문제 발생 원인 :

reference가 시작되는 sentence의 tag는 보통 r2 인데,

reference의 중간 내용에 포함된 sentence가 tag를 r2를 갖는 경우가 있다.

\* 빨간색 박스 내의 텍스트들이 각각 s3 tag로 치환되어 아래 결과와 같이 각각 하나의 청크로 분리되는 문제

가압류신청 진술서	20...
작성 가압류신청 진술서의 누락 및 허위 작성 5/11	
가압류신청 진술서	채권자(대리인) _____ (날인 또는 서명)
작성 가압류신청 진술서의 누락 및 허위 작성 6/11	
가압류신청 진술서	◇ 다음 ◇
작성 가압류신청 진술서의 누락 및 허위 작성 7/11	

# Chunking Logic – 2. 청크 유형별 처리 방식

## 4. s3 + s4 + reference

- **개선** : 내용이 적당하게 묶이도록 tag 치환 규칙을 재설정.

\* 치환 규칙

: reference에서 내용이 시작되는 텍스트만 tag를 s3로 설정하고, 이후 텍스트는 s8로 치환하여 S3의 하위 내용으로 처리

### ▣ 가압류신청 진술서의 누락 및 허위 작성

**s3**

· 가압류신청 진술서를 첨부하지 않거나, 고의로 진술 사항을 누락하거나 허위로 진술한 내용이 발견된 경우에는 특별한 사정이 없는 한 보정명령 없이 신청이 기각될 수 있습니다(「보전처분 신청사건의 사무처리요령」 제3조).

※ 채무자가 여러 명인 경우에는 각 사람별로 진술서를 작성해야 합니다.

**s3**

가압류신청 진술서

채권자는 가압류 신청과 관련하여 다음 사실을 진술합니다. 다음의 진술과 관련하여 고의로 누락하거나 허위로 진술한 내용이 발견된 경우에는, 그로 인하여 보정명령 없이 신청이 기각되거나 가압류이의절차에서 불이익을 받을 것임을 잘 알고 있습니다.

20...

채권자(대리인) \_\_\_\_\_ (날인 또는 서명)

◇ 다음 ◇

**s8**

**s8**

**s8**

**CHUNK 0**

{chunk\_title : “가압류신청 진술서의 누락 및 허위 작성 1/n”  
chunk\_content : “ 가압류신청 진술서를 첨부하지 않거나,  
... ”}

**CHUNK 1**

{chunk\_title : “가압류신청 진술서의 누락 및 허위 작성 2/n”  
chunk\_content : “ ※ 채무자가 여러 명인 경우에는 각 ... ”}

**CHUNK 2**

{chunk\_title : “가압류신청 진술서의 누락 및 허위 작성 3/n”  
chunk\_content : “ 가압류신청 진술서 채권자는 ... ”}

# Chunking Logic – 2. 청크 유형별 처리 방식

## 5. table 처리

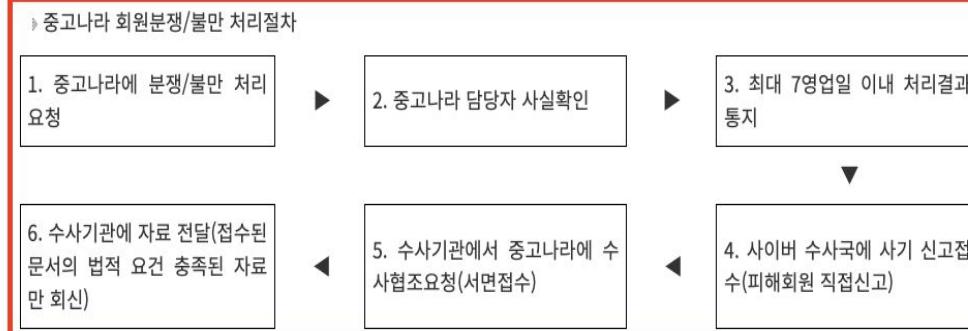
- 처리 방식 : table 데이터 바로 앞에 텍스트가 나오는 경우, 텍스트에 `table_id("{doc_id}_table{idx}")` 을 붙여서 한 문장으로 생성
- 이유 : 테이블 앞에 나오는 텍스트가 테이블을 설명하는 경우가 대부분일 것 같다.

### 온라인 직거래플랫폼을 통한 피해 신고

s3

- 온라인 직거래 피해를 입은 사람이 경찰서 등에 신고하는 방법 이외에도 온라인 직거래 플랫폼을 통해 피해를 신고하거나 분쟁을 처리할 수 있습니다(중고나라, 회원분쟁/불만처리절차 및 당근, 거래문제참조).

s4



위의 방식대로 처리 후, 로직 1 ~ 4번 방식 적용한 결과

CHUNK	
	<pre>{   chunk_title : "온라인 직거래플랫폼을 통한 피해 신고"   chunk_content : "온라인 직거래 피해를 입은 사람이 ... 중고나라 회원분쟁/불만 처리절차"   d01b0401404f4265cab736c1818fb34c9244b4eb_table0 당근 거래문제   (사기를 당했을 경우)   d01b0401404f4265cab736c1818fb34c9244b4eb_table1 ※ 온라인 플랫폼을 통한 피해 신고 또한 ..." }</pre>

! 이후, 테이블 데이터가 처리되면 `table_id` 자리에 테이블 내용으로 대체됨.

# Chunking Logic – 3. chunking

알아두면 유용한, 저작권 필수사항

본문 100문 100답 카드뉴스 관련법령

다른 사람의 저작권을 침해하지 않고 콘텐츠를 제작하는 방법



## ▣ 사전에 저작권 허락 받고 이용하기

- 저작재산권자는 다른 사람에게 그 저작물의 이용을 허락할 수 있습니다. 이에 따라 허락을 받은 사람은 허락받은 이용 방법 및 조건의 범위 안에서 그 저작물을 이용할 수 있습니다(「저작권법」 제46조제1항 및 제2항).

※ “저작권”이란 저작자가 자신의 저작물을 배타적·독점적으로 이용할 권리로서 저작재산권과 저작인격권으로 구성되고(「저작권법」 제10조제1항), “저작재산권”이란 저작자의 경제적 이익을 보호하기 위한 권리입니다[한국저작권위원회(<https://www.copyright.or.kr>)-저작권상담센터-상담사례집].

## ▣ 저작권 이용허락은 누구에게 받아야 할까?

- 저작권을 침해하지 않으려면 저작권자 또는 저작물에 따라서 저작권자 뿐만 아니라 저작인접권자에 대한 이용허락을 받아야 합니다(「저작권법」 제46조 및 제88조).

※ “저작인접권”이란 저작권을 전제로 하여 이와 별개의 독립된 저작물의 내용을 일반공중에게 전달하는 매체로서의 실연·음반·방송에 대해서 원저작물에 준하는 일종의 정신적 가치를 인정·보호하기 위해 그 행위자인 실연자·음반제작자·방송사업자에게 저작권과 유사하게 주어지는 권리입니다[「저작권법」 제2조제4호·제6호·제9호 및 국가법령정보센터(<https://www.law.go.kr>) 법령용어사전 참조].

## ▣ 권리자를 찾을 수 없다면 ‘법정허락’을 신청하세요.

- “법정허락제도”란 다음의 어느 하나에 해당하는 경우 문화체육관광부장관의 승인을 받은 후 문화체육관광부장관이 정하는 기준에 따른 보상금을 지급하거나 공탁하고 이를 이용할 수 있는 제도를 말합니다(규제) 「저작권법」 제50조제1항, 제51조 및 제52조 참고).

- 누구든지 상당한 노력을 기울였어도 공표된 저작물의 저작재산권자나 그의 거소를 알 수 없어 그 저작물의 이용허락을 받을 수 없는 경우
- 공표된 저작물을 공익을 위한 필요에 따라 방송하려는 방송사업자가 그 저작재산권자와 협의하였으나 협의가 성립되지 아니하는 경우
- 상업용 음반이 우리나라에서 처음으로 판매되어 3년이 지난 경우 그 음반에 녹음된 저작물을 녹음하여 다른 상업용 음반을 제작하려는 자가 그 저작재산권자와 협의하였으나 협의가 성립되지 않은 경우

※ 저작권자를 알기 어려운 경우에는 <한국저작권위원회 권리자 찾기(<https://www.findcopyright.or.kr>)>를 이용하시면 저작권자 조회 및 법정허락 승인 신청 서비스를 이용할 수 있습니다.

## CHUNK 0

{chunk\_title : “다른 사람의 저작권을 침해하지 않고 콘텐츠를 제작하는 방법 : 사전에 저작권 허락 받고 이용하기”  
 chunk\_content : “저작재산권자는 다른 사람에게 그 저작물의 이용을 허락할 수 있습니다. 이에 따라 허락을 받은 사람은 허락받은 이용 방법 및 조건의 범위 안에서 그 저작물을 이용할 수 있습니다(「저작권법」 제46조제1항 및 제2항).”}  
 ※ “저작권”이란 저작자가 자신의 저작물을 배타적·독점적으로 이용할 권리로서 저작재산권과 저작인격권으로 구성되고(「저작권법」 제10조제1항), “저작재산권”이란 저작자의 경제적 이익을 보호하기 위한 권리입니다[한국저작권위원회(<https://www.copyright.or.kr>)-저작권상담센터-상담사례집].

## CHUNK 1

{chunk\_title : “다른 사람의 저작권을 침해하지 않고 콘텐츠를 제작하는 방법 : 저작권 이용허락은 누구에게 받아야 할까?”  
 chunk\_content : “저작권을 침해하지 않으려면 저작권자 또는 저작물에 따라서 저작인접권자에 대한 이용허락을 받아야 합니다(「저작권법」 제46조 및 제88조).”}  
 ※ “저작인접권”이란 저작권을 전제로 하여 이와 별개의 돋립된 저작물의 내용을 일반공중에게 전달하는 매체로서의 실연·음반·방송에 대해서 원저작물에 준하는 일종의 정신적 가치를 인정·보호하기 위해 그 행위자인 실연자·음반제작자·방송사업자에게 저작권과 유사하게 주어지는 권리입니다[「저작권법」 제2조제4호·제6호·제9호 및 국가법령정보센터(<https://www.law.go.kr>) 법령용어사전 참조].

## CHUNK 2

{chunk\_title : “다른 사람의 저작권을 침해하지 않고 콘텐츠를 제작하는 방법 : 권리자를 찾을 수 없다면 ‘법정허락’을 신청하세요.”  
 chunk\_content : ““법정허락제도”란 다음의 어느 하나에 해당하는 경우...”}

# Chunking Logic – 3. chunking

내 저작권이 침해되면, 이렇게 대응하세요.



## ■ 저작권이 침해되면, 대응할 수 있는 법적조치와 처벌

- 저작권이 침해되면 온라인 플랫폼 사업자에 게시 중단 요청, 민사 소송 제기, 형사고소 등을 할 수 있습니다.

※ 민사책임과 별도로 형사책임도 처벌 받을 수 있습니다.

- 저작권이 침해되면, 다음에 해당하는 법적 조치를 취할 수 있습니다(「저작권법」 제123조제1항, 제125조제1항, 제125조의2제1항·제3항, 제136조제1항제1호·제2항제1호, 제138조제2호 및 「민법」 제741조).

<민사 조치>

› 침해 정지 청구

› 침해 예방 및 손해배상 담보 청구

› 손해배상 청구

› 법정손해배상 청구

› 부당이득반환청구

<형사 처벌>

› 저작재산권, 그 밖에 「저작권법」에 따라 보호되는 재산적 권리를 복제, 공연, 공중송신, 전시, 배포, 대여 및 2차적저작물 작성의 방법으로 침해한 자는 5년 이하의 징역 또는 5천만원 이하의 벌금에 처해짐

› 저작인격권 또는 실연자의 인격권을 침해하여 저작자 또는 실연자의 명예를 훼손한 자는 3년 이하의 징역 또는 3천만원 이하의 벌금에 처해짐

› 출처 명시를 하지 않은 자는 5백만원 이하의 벌금에 처해짐

※ “저작인격권”이란 저작물과 관련하여 저작자의 명예와 인격적 이익을 보호하기 위한 권리입니다[한국저작권위원회 (<https://www.copyright.or.kr>)-저작권상담센터-상담사례집].

※ 저작권 침해 신고 시 온라인 플랫폼 사업자가 해야 할 조치

- 온라인 플랫폼 사업자는 복제·전송의 중단요구를 받은 경우에는 즉시 그 저작물등(저작물, 실연·음반·방송 또는 데이터베이스)의 복제·전송을 중단시키고 자신의 권리가 침해됨을 주장하는 자에게 그 사실을 통보해야 합니다(「저작권법」 제103조제1항·제2항 참고).

※ 저작권 문제에 따른 저작권 상담 및 분쟁 조정 서비스는 <한국저작권위원회 홈페이지 (<https://www.copyright.or.kr>)>에서 이용할 수 있습니다.

※ 그 밖에 저작권 이용 및 보호에 대한 자세한 내용은 이 사이트 『저작권보호』에서 확인할 수 있습니다.

## CHUNK 3

{chunk\_title : “내 저작권이 침해되면, 이렇게 대응하세요. : 저작권이 침해되면, 대응할 수 있는 법적조치와 처벌 1/4”}

chunk\_content : ” 저작권이 침해되면 온라인 플랫폼 사업자에 게시 중단 요청 ... ”}

## CHUNK 4

{chunk\_title : “내 저작권이 침해되면, 이렇게 대응하세요. : 저작권이 침해되면, 대응할 수 있는 법적조치와 처벌 2/4”}

chunk\_content : “저작권이 침해되면, 다음에 해당하는 법적 조치를 취할 수 있습니다 (「저작권법」 ... ”)

## CHUNK 5

{chunk\_title : “내 저작권이 침해되면, 이렇게 대응하세요. : 저작권이 침해되면, 대응할 수 있는 법적조치와 처벌 3/4”}

chunk\_content : “※ “저작인격권”이란 저작물과 관련하여 저작자의 명예와 인격적 이익 ... ”}

## CHUNK 6

{chunk\_title : “내 저작권이 침해되면, 이렇게 대응하세요. : 저작권이 침해되면, 대응할 수 있는 법적조치와 처벌 4/4”}

chunk\_content : ”※ 저작권 문제에 따른 저작권 상담 및 분쟁 조정 서비스는 ... ”}