

국립국어원 2021-01-14

발간등록번호
11-1371028-000864-01

2021년 회의록 요약 말뭉치 연구 분석

사업 책임자
김 한 샘

제 출 문

국립국어원장 귀하

국립국어원과 체결한 연구용역 계약에 따라 ‘2021년 회의록 요약 말뭉치 연구 분석’에 관한 연구 보고서를 작성하여 제출합니다.

■ 사업기간: 2021년 5월 ~ 2021년 12월

2021년 12월 21일

사업 책임자: 김한샘(연세대학교)

사업 수행자 연세대학교, 한국과학기술원, (주)답네츄럴, (주)테디썸

사업 책임자 김한샘

사업 참여자 주민재, 윤영민, 최기선, 전영민, 함영균, 김지성
한지윤, 오태환, 박서윤, 정승연, 박재완, 이규덕

<사업 수행자>

연세대학교 산학협력단 · 한국과학기술원 ·
주식회사 딥네츄럴 · 주식회사 테디썸

사업 책임자	김한샘(연세대학교 언어정보연구원 부교수)
사업 참여자	주민재(명지대학교 방목기초교육대학 부교수)
	윤영민(연세대학교 언어정보연구원 연구교수)
	한지윤(연세대학교 언어정보학협동과정 박사과정)
	오탈환(연세대학교 국어국문학과 박사과정)
	박서윤(연세대학교 언어정보학협동과정 박사과정)
	정승연(연세대학교 언어정보학협동과정 석사과정)
	최기선(한국과학기술원 전산학과 명예교수)
	김지성(한국과학기술원 전산학과 연수연구원)
	전영민((주)딥네츄럴)
	이규덕((주)딥네츄럴)
	함영균((주)테디썸)
	박재완((주)테디썸)

2021년 회의록 요약 말뭉치 연구 분석

본 사업은 논쟁적 대화 요약에 대한 대규모 말뭉치 구축을 위한 시범 연구로서, 국어 자원의 활용도와 가치를 제고하고 국내 표준화 및 참고 기반 자료가 될 수 있는 말뭉치를 구축하는 것을 목표로 한다. 논쟁적 대화 요약은 단순히 회의에서 이루어진 안건에 대한 핵심 정보만을 파악하는 것 이외에도 논증 채굴(Argument Mining) 및 의견 채굴(Opinion Mining) 등 다양한 응용 기술을 개발하기 위한 기초로써 활용 가치 및 연구 자원으로서의 효용성이 높다.

본 사업에서는 실제 세계에서 논쟁적 대화에 대한 요약 분석을 위한 논쟁적 대화 요약 분석 지침을 수립하고, 국회 소위원회 회의록 200건을 대상으로 하여 문서 내 중요 표현을 추출하고, 이를 바탕으로 추상적 요약문을 작성하여 지침의 효용성 및 실제 적용 방안을 검토한다. 본 사업의 주요 과업과 연구 성과는 다음과 같다.

▶ 논쟁적 대화 요약을 위한 지침 마련

본 과제에서 구축된 지침은 구어체의 대화로부터 문어체의 요약문을 작성하는 대화 요약을 넘어 논쟁적 대화에서의 입체적 정보를 상향식으로 요약하여 핵심 정보를 놓치지 않는 추상 요약문을 작성하는 대지침을 마련하였고, 세부 지침으로서 국회 회의록에 적용될 수 있는 방법을 마련하였다. 이를 통해 향후 대규모 논쟁적 대화 요약 말뭉치 구축을 위한 국내 표준화 및 참고 기반 자료가 될 수 있는 기틀을 확보하였다.

▶ 지침에 따른 추상 요약문 구축

본 과제에서는 개발된 지침을 사용하여 국회 소위원회 회의록 200건(전체 95,052 문장 및 1,560,255 어절, 회의록당 평균 475.26 문장 및 문장당 평균 16.415 어절)에 대해 이슈별로 중요 표현과 이슈 요약문(회의록 평균 4.42개 이슈, 이슈 요약문 평균 2.14 문장 · 45.28 어절)과 이슈의 세부 요약문(세부 요약문 평균 16.23개, 각 세부 요약문 평균 1.01 문장 · 23.27 어절)을 작성하고, 이를 토대로 회의록 전체에 대한 대표 요약문(평균 3.815 문장 · 83.875 어절)을 작성하였다. 본 과제를 통해 구축된 데이터는 논쟁적 대화 요약 모델에 대한 다양한 측면에서 평가지표를 제공할 것으로 기대된다.

주요어: 요약, 회의록 요약, 회의록 요약 말뭉치 구축

차례

제1장 서론

1. 사업의 목적 및 개요	2
2. 사업 범위	3
3. 사업의 필요성 및 기대 효과	4

제2장 회의록 요약 말뭉치 구축

1. 회의록 요약 말뭉치 구축 대상 회의록 선정	8
2. 회의록 전처리 및 데이터 형식	10
3. 회의록 요약 말뭉치 구축 과정	20
4. 회의록 요약 말뭉치 구축 결과	25

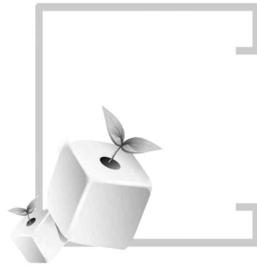
제3장 논쟁적 대화 요약 말뭉치 지침

1. 논쟁적 대화 요약 말뭉치 구축 지침	28
2. 논쟁적 대화 요약 예시 및 세부 분석	45
3. 검수 유의사항	61
4. 데이터 납품 형태	69

제4장 논쟁적 대화 요약 말뭉치 활용 방안

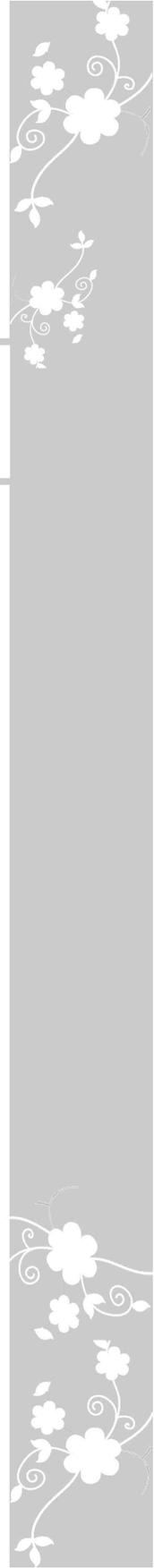
1. 회의록 요약 기술 개발	75
2. 회의록 요약 말뭉치 활용 방안	78

<부록 1> 분석 대상 회의록 목록	80
<부록 2> 논변 구조 분석 지침	85
<부록 3> 국회 회의록 요약 작업에 대한 전문가 의견	125
참고문헌	127
<Abstract>	128



제 1 장

서 론



1. 사업의 목적 및 개요

본 사업 “2021년 회의록 요약 말뭉치 연구 분석”은 논쟁적 대화 요약에 대한 대규모 말뭉치 구축을 위한 시범 연구로서, 국어 자원의 활용도와 가치를 제고하고 국내 표준화 및 참고 기반 자료가 될 수 있는 말뭉치를 구축하는 것을 목표로 한다.

논쟁적 대화는 사람들이 아이디어를 교환하고, 계획을 세우거나 정보를 공유하는 가장 일반적인 방법으로서, 학술적 회의, 기업의 회의, 국회 회의 등 실제 세계에서 존재하는 모든 회의에서 이루어지는 일종의 목적 지향적 대화라고 볼 수 있다. 이러한 논쟁적 대화는 서기에 의해서 수작업으로 작성되거나, 음성인식 등의 기술을 바탕으로 전사(transcription)되어 보관되고 있다.

기존의 요약 연구는 이렇게 전사되어 보관되는 비정형 데이터에서 필요한 정보만을 제공하는 텍스트 요약의 분야를 다루고 있다. 그러나 기존의 연구에서는 뉴스에 대한 요약데이터 중심의 말뭉치가 구축 및 공개되었으나, 관련 연구들은 뉴스, 논문, 도서 등에 대한 문서 요약에 집중되어 있다. 대화 요약의 연구에는 2020년 한국정보화진흥원에 의해 일상대화 및 토론 대화에 대한 요약데이터 구축 연구¹⁾가 시작 단계에 와 있으나 대화의 말차례(turn)가 4~5회의 짧은 대화를 주로 다루고 있으며, 대화 데이터 구축 과정에서 정형화된 프레임(예: 토론의 주제, 문제 제기, 찬성, 반론, 결론 등의 구성 요소가 사전에 정의된 상태로 구축된 가상의 대화 데이터)으로 구성된 하나의 주제에 대한 짧은 가공의 대화 데이터에 관한 요약 연구라는 점에서 한계가 있다.

본 사업에서는 실제 세계에서 논쟁적 대화 요약 분석에 필요한 지침 수립을 목표로 한다. 또한 해당 지침을 활용하여 국회 소위원회 회의록 200건에서 문서 내 중요 표현을 추출하고, 이를 바탕으로 추상적 요약문을 작성하여 지침의 효용성 및 실제 적용 방안 검토를 목표로 한다.

논쟁적 대화 요약 분석을 위한 지침 개발 및 회의록 요약 말뭉치를 구축하기 위해 설정한 하위 목표는 아래와 같다.

▶ 논쟁적 대화 요약을 위한 지침 마련
(세부 지침으로서 국회 회의록 요약 지침 마련)

▶ 지침에 따른 추상 요약문 구축
(국회 소위원회 회의록 200건)

1) <https://aihub.or.kr/aidata/30714>

2. 사업 범위

본 사업에서 수행한 과업의 범위는 아래와 같다.

과업의 범위	세부 과업 내용
회의록 요약 말뭉치 구축 방법론 수립	논쟁적 대화 요약 방법론 수립 논쟁적 대화 요약 분석 절차 수립
논쟁적 대화 요약을 위한 지침	논쟁적 대화 요약 분석의 원칙 논쟁적 대화의 중심어 선택 논쟁적 대화의 세부 요약문 작성 논쟁적 대화의 주제별 요약문 작성 국회 회의록 구축을 위한 세부 지침 마련
회의록 요약 말뭉치 시범 구축	국회 회의록 분석 말뭉치 형식 정의 국회 회의록 200건에 대한 정제 과정을 거쳐 원시 말뭉치를 구축하고 최종적으로 200건에 대한 중심어 선택 및 추상 요약문을 작성
회의록 요약 말뭉치 활용 방안 모색	논쟁적 대화 말뭉치의 산업적 활용 가치 검증

표 1 과업 범위 및 세부 과업 내용

3. 사업의 필요성 및 기대 효과

3.1. 요약 기술의 필요성과 관련 연구

요약 기술은 현실의 기업이나 기관 등에서 각종 문서 관리가 잘 이루어지지 못하고 데이터베이스 관리 시스템(DBMS, Database Management System)에 의존한 데이터의 관리와 탐색이 이루어지고 있다. 정보통신의 발달로 문서의 양이 급격히 늘어나는 상황에서, 문서를 직접 읽으며 정보를 찾는 전통적 방법은 시간과 자원을 많이 소모한다. 특히, 비정형 데이터가 정형 데이터에 비해 빠르게 증가하고 있는 상황²⁾에서 비정형 데이터에서 필요한 정보만을 제공하는 텍스트 요약 기술의 중요성은 점차 강조되고 있다.

해외의 텍스트 요약을 위한 연구 및 데이터로는 뉴스에 대한 요약데이터로서 씨엔엔/일상 메일 말뭉치(CNN/Daily Mail Corpus)(Nallapati 외, 2016), 기가워드 말뭉치(Gigaword Corpus)(Rush 외, 2015), 디유씨 2004 과업 1(DUC 2004 Task1³⁾), 문장 압축(Sentence Compression)(Filippova and Altun, 2013) 등의 말뭉치가 구축 및 공개되었으며, 웨비스-티엘디알-17 말뭉치(Webis-TLDR-17 Corpus)(Völske 외, 2017)처럼 소셜 미디어에 대한 요약이나 회사에서의 회의 대화에 대한 요약을 제공하는 에이엠아이 데이터 세트(AMI Dataset)(Carletta, 2006)이나 아이씨아이에스 데이터 세트(ICIS Dataset)(Boudin and Morin, 2013)도 구축되어 공개된 바 있다.

국내의 텍스트 요약을 위한 연구과제로는 2019년 국립국어원에서 13,167 문장의 신문 기사 요약 말뭉치 구축이 연구 및 시범 구축으로 진행되었고, 2020년 한국정보화진흥원에 의해 논문자료 요약 데이터, 도서자료 요약 데이터, 한국어 대화 요약 데이터 등이 구축되었다.

이러한 인공지능 연구 동향은 실제 기업의 시스템으로도 구축되어 서비스화되고 있다. 뉴스 기사를 최대 3문장 이내로 요약하는 네이버 요약봇, 카카오 다음뉴스 요약봇 등의 시스템이 서비스되고 있으며, 엔씨소프트의 페이지(PAIGE)와 같이 요약기능과 함께 문어체의 원문을 구어체로 문체 변환(Paraphrasing)하는 서비스 등, 요약 기술은 활발히 연구되고 있다.

2) IDC(2012-12). The Massive Growth in Unstructured Data.

3) <https://duc.nist.gov/duc2004/>

3.2. 논쟁적 대화 요약의 필요성과 어려움

기존의 텍스트 요약 연구는 저자에 의해서 잘 쓰인 뉴스, 논문, 도서자료 등에 대한 문서 요약(Document Summarization)에 집중되어 있으며, 대화 요약의 경우에는 2020년 한국정보화진흥원에 의해 일상대화 및 토론 대화에 대한 요약데이터 구축 연구가 시작 단계에 와 있으나, 대화 데이터 구축 과정에서 정형화된 프레임(예: 토론의 주제, 문제 제기, 찬성, 반론, 결론 등의 구성요소가 사전 정의된 상태로 구축된 대화 데이터)으로 구성된, 하나의 주제에 대한 짧은 가공의 대화 데이터에 관한 요약 연구라는 점에서 한계가 있다.

특히, 현실의 회의는 짧고 정형화된 대화가 아닌 긴 시간 동안 다양한 주제에 대해서 각 주제에 대한 자세한 설명들과 이에 대한 질의응답 등으로 구성되어 있어, (1) 회의 전체에 대한 구조를 분석(Document Structuring)하고, (2) 토론 주제들을 파악(Topic Modeling)하여, (3) 각각의 토론 주제에 대한 발화 간의 논변 구조(Argument Structure)를 파악하지 않으면 그 회의를 요약하는 것은 매우 어려운 일이다.

본 사업은 한 번의 회의가 A4 문서 기준 20쪽 내외로 구성된 긴 대화들로 구성된 국회 소위원회 회의록을 대상으로 하여 현실의 회의, 즉 논쟁적 대화에 대한 요약에 관한 연구이다. 기존의 해외 연구(예: AMI 데이터 세트)의 경우 회사에서의 회의(약 160개 말차례의 발화로 구성)에 대한 요약 데이터 세트를 구축한 바 있으나, 40여 개 법안에 대한 설명과 질의응답으로 구성된 20쪽 내외의 긴 국회 회의록과 같은 복잡한 데이터 세트에 대한 연구는 아직 국내외에서 수행된 바 없다.

이에 본 연구는 현실에서의 실제 회의, 즉 길고 다양한 주제에 대한 설명과 질의응답으로 구성된 회의에 대한 요약을 위한 분석 지침을 수립하고, 이러한 분석 지침의 검증의 측면에서 국회 소위원회 회의록 200건에 대한 시범 구축을 하는 것을 과제의 목표와 내용으로 한다.

3.3. 본 사업의 기대효과

▶ 비정형 회의록에 대한 메타데이터 추출 및 전처리 작업

본 사업에서는 HWP 파일에서 회의록에 해당하는 텍스트만을 추출하는 것이 아닌, 회의록에서 추출할 수 있는 메타데이터를 추출하여 회의록이 갖는 주요 정보들에 대해서도 JSON 형식으로 기록한다. 해당 작업을 통해 자동으로 회의록으로부터 메타데이터를 추출하는 시스템을 개발하여, 후속 연구에서 활용될 수 있도록 한다.

▶ 논변 구조 분석

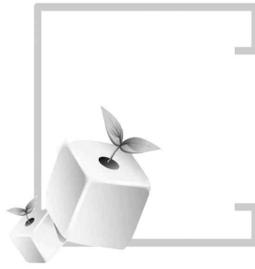
본 사업에서는 논쟁적 대화, 즉 각각의 주제에 대하여 질문과 답변, 지지와 부정, 요청과 설명 등의 다양한 속성의 발화들로 구성된 회의록에 대한 논변 구조(Argument Structure)를 분석할 수 있는 주석 지침을 만들어, 후속 연구에서 활용될 수 있도록 한다.

▶ 논쟁적 대화에 대한 요약 지침 구축

본 사업에서는 논쟁적 대화, 즉 회의록이라는 목적에 맞는 특정한 구조화된 요약문을 구축한다. 특히, 기존 연구인 뉴스 요약이나 일상 대화 요약이 주요 문장을 환언(paraphrasing)하는 방법이라는 것과 달리, 본 연구에서는 회의의 목적과 내용에 맞도록 구조화된 요약문(Abstract, Decisions, Problems/Issues, Actions 등)을 구축한다.

▶ 논쟁적 대화 요약의 활용도 제고

본 사업에서는 새로운 연구 분야인 논쟁적 대화 요약 말뭉치의 산업적/기술적 활용 방안을 제시한다. 본 과제에서 제안하는 방법론 및 지침에 따라 구축되는 논쟁적 대화 요약 말뭉치는 추후 회의 요약 기술 개발의 밑거름으로 삼을 수 있다.



제 2 장

회의록 요약 말뭉치 구축



1. 회의록 요약 말뭉치 구축 대상 회의록 선정

본 장에서는 회의록 요약 말뭉치 구축의 대상이 되는 원시 데이터인 회의록에 대한 개요를 기술한다. 본 사업에서 분석 대상이 된 회의록의 목록은 <부록 1> 분석 대상 회의록 목록에서 확인할 수 있다.

시범 구축 분석 대상 회의록의 요건은 다음과 같다.

- (1) 불필요한 논쟁이 과도하게 포함된 회의록 제외
- (2) 다양한 도메인을 가진 20쪽 내외의 회의록으로 구성

국회 회의록의 경우 회의의 안전에 대해 각 안전에 대한 정부 측의 의견 및 전문위원의 검토사항 및 의견, 해당 안전에 대한 참석 위원들의 질의응답 및 의견, 추가로 안전에 대한 동의 혹은 비동의 등이 이루어지는 구성으로 이루어져 있다. 본 사업팀의 검토 결과 일부 데이터의 경우 상당 부분 안전과 관계없는 말싸움(언쟁)으로 이루어져 있어 이를 전수 검사를 통해 제외하는 절차를 거쳤다. 예를 들어 아래와 같은 말싸움이 많은 회의록의 경우에는 구축 대상의 회의록에서 제외하였다. 아래 예시에서 발의자의 이름은 익명 처리(#)하였다.

<p>○### 실무자가 어느 선이에요? 사무관 선인가요, 과장 선인가요?</p> <p>○### 국장 정도 선까지라고 보시면 됩니다.</p> <p>○### 국장한테 보고된, 그 이상은 보고가 되지 않았고요?</p> <p>○### 저도 보긴 봤습니다마는 이거는 실무자 의견으로 그냥 공식의견이 아니라고 해라, 공식의견이 아니다, 전혀 우리가……</p> <p>○### 언론에 난 경위는, 그러면 발표는 안 했고요?</p> <p>○### 경위는 아마 그 보고서가 유출이 된 걸로, 기자가……</p> <p>○### 발표를 한 일은 없고요?</p> <p>○### 발표한 일은 전혀 없고요. 아마 기자가 그 보고서를 어디서 입수를 한 것 같아요. 그래서 그게 특종 비슷하게 돼 가지고……</p> <p>○### 그러면 그거 기강 문제네요.</p> <p>○### 그거 문제네요. 진짜 큰 문제네요.</p> <p>○### 전혀 기재부의 입장은 아니다?</p> <p>○### 예, 기재부의 입장은 아니라고 어제 분명히 언론에다가 보도 해명자료를 냈습니다.</p> <p>○### 자기 주장이 안 되니까 언론플레이 한 것같이 보이네요.</p>
--

표 2 구축 대상 회의록에서 제외된 사례. 안전과 관계없는 내용이 대다수를 차지하는 회의록

결과적으로 다양한 도메인에서 말싸움(언쟁)을 최소한으로 포함하고 있는 200개 회의록을 선정하였고, 이에 대한 통계는 아래와 같다.

분야	회의록 개수
국토	23
외통	22
국방	22
법사	20
환노	19
재정	16
여성	11
행안	11
기재	10
복지	9
교문	8
안행	8
교육	6
농해수	4
산자중기	3
정치개혁특위	3
미방	1
교과	1
문방	1
문체	1
사법개혁특위	1
합계	200

표 3 분석 말뭉치 통계 1: 분야별 회의록의 개수

항목	내용
회의록당 회의 참석자 인원	16.24 (명)
회의록당 실제 발화자 평균	12.45 (명)
회의록당 발화 개수 평균	475.26 (개)
발화 평균 길이 (글자 수)	68.8958 (개)
발화 평균 길이 (단어 수)	16.415 (개)

표 4 분석 말뭉치 통계 2: 회의록 문서 관련 통계

2. 회의록 전처리 및 데이터 형식

2.1. 회의록 전처리

본 장에서는 회의록 원시 말뭉치를 정제하는 전처리 작업 및 원시 말뭉치와 분석 말뭉치의 데이터 형식에 대해 기술한다.

국회 회의록 원시 데이터는 HWP 파일 형식으로 공개되어 있으며 본 사업팀은 해당 HWP 파일에 대해 메타데이터를 추출하고 이를 전처리하는 작업을 수행하였다. 국회 회의록은 아래와 같은 형식으로 작성되어 있다.

제377회-여성가족소위제1차(2020년5월6일) 1	4 제377회-여성가족소위제1차(2020년5월6일)
<p>제377회국회 (임시회)</p> <p>여성가족위원회회의록 (법안심사소위원회)</p> <p>제 1 호</p> <p>국 회 사 무 처</p> <p>일 시 2020년5월6일(수) 장 소 여성가족위원회소회의실</p> <p>의사일정</p> <ol style="list-style-type: none"> 1. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안 2. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안 3. 가족친화 사회환경의 조성 촉진에 관한 법률 일부개정법률안 4. 가족친화 사회환경의 조성 촉진에 관한 법률 일부개정법률안 5. 텔레그램을 통한 미성년자 대상 성범죄 처벌 강화 및 신상공개에 관한 청원 6. 아동·청소년의 성보호에 관한 법률 일부개정법률안(계속) 7. 아동·청소년의 성보호에 관한 법률 일부개정법률안 8. 아동·청소년의 성보호에 관한 법률 일부개정법률안 9. 아동·청소년의 성보호에 관한 법률 일부개정법률안 10. 아동·청소년의 성보호에 관한 법률 일부개정법률안 11. 아동·청소년의 성보호에 관한 법률 일부개정법률안 12. 아동·청소년의 성보호에 관한 법률 일부개정법률안 13. 아동·청소년의 성보호에 관한 법률 일부개정법률안 14. 성폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안 15. 양육비 이행확보 및 지원에 관한 법률 일부개정법률안(계속) 16. 양육비 이행확보 및 지원에 관한 법률 일부개정법률안(계속) 17. 일제하 일본군위안부 피해자에 대한 생활안정지원 및 기념사업 등에 관한 법률 일부개정법률안(계속) 18. 일제하 일본군위안부 피해자에 대한 생활안정지원 및 기념사업 등에 관한 법률 일부개정법률안(계속) 19. 일제하 일본군위안부 피해자에 대한 보호·지원 및 기념사업 등에 관한 법률 일부개정법률안 20. 일제하 일본군위안부 피해자에 대한 생활안정지원 및 기념사업 등에 관한 법률 일부개정법률안(계속) <p>상정된 안건</p> <ol style="list-style-type: none"> 1. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안(이명수 의원 대표발의).....2 2. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안(송희경 의원 대표발의).....2 3. 가족친화 사회환경의 조성 촉진에 관한 법률 일부개정법률안(이양수 의원 대표발의).....2 4. 가족친화 사회환경의 조성 촉진에 관한 법률 일부개정법률안(정부 제출).....2 5. 텔레그램을 통한 미성년자 대상 성범죄 처벌 강화 및 신상공개에 관한 청원(국민주 외 100,000인 국민동의로 제출).....2 6. 아동·청소년의 성보호에 관한 법률 일부개정법률안(윤후덕 의원 대표발의)(계속).....2 7. 아동·청소년의 성보호에 관한 법률 일부개정법률안(강창일 의원 대표발의).....2 8. 아동·청소년의 성보호에 관한 법률 일부개정법률안(신창원 의원 대표발의).....2 9. 아동·청소년의 성보호에 관한 법률 일부개정법률안(송희경 의원 대표발의).....2 10. 아동·청소년의 성보호에 관한 법률 일부개정법률안(박대술 의원 대표발의).....2 	<p>수정의견과 같이 의견을 제시하였습니다. 심사자료 4본입니다. 1쪽입니다. 가족친화지원센터의 지정위수 기준을 법률에 명시하고 지정위수하려는 경우에 청문을 실시하도록 하려는 조항이고 기타 자구 수정 사항 등이 있습니다. 국민의 권리와 의무에 변동을 가져오는 사항에 대하여는 법률에 규정하는 것이 타당한 입법조치로 판단이 됩니다. 조문별로 간단히 짚겠습니다. 2쪽입니다. 안 제10조에 대해서는 특별자치시장을 포함하는 안입니다. 3쪽입니다. 안 제15조와 16조에 대해서는 '받고자 하는 자'를 '받으려는 자'로, '지정요건'을 '지정기준'으로 수정하려는 안입니다. 4쪽입니다. 자구 수정 차원에서 '가족친화지원센터'를 '1항에 따라 지정된 가족친화지원센터'로 수정하는 안입니다. 안 4항에서는 바로 앞서 설명드린 지정위수 기준 상향 입법에 대해서 동법 시행령에 있는 지정위수 기준을 상향한 내용으로 타당한 입법조치로 보입니다. 5쪽에서는 제21조(청문) 조항에 바로 '가족친화지원센터의 지정위수' 조항을 포함하는 내용입니다. 이상입니다. ○소위원장 신용현 정부 측 의견 말씀해 주십시오. ○여성가족부차관 김희경 2건 모두 동의합니다. ○소위원장 신용현 위원님들 의견 말씀해 주십시오. ○신보라 위원 가족친화지원센터는 지금 어디에 지정이 되어 있는 것이지요? ○소위원장 신용현 배석하신 분이 말씀 주시고 말씀하십시오. ○여성가족부차관 김진영 여성가족부 가족정책관 김진영입니다. 지금 현재는 한가원에 지정돼 있습니다. 한국건강가정진흥원에 지정돼 있습니다. ○신보라 위원 그러면 중앙에 있는 건강가정지원센터가 지금.....</p> <p>○김현아 위원 전체 32개가 있고요. 이게 다 수용시킬 건데요? ○김현아 위원 생활시설. ○여성가족부차관 김희경 생활시설. ○김현아 위원 생활시설이요? ○여성가족부차관 김희경 예. ○김현아 위원 그러면 아까 두 번째 안건이 지금 상담 프로그램을 제공하겠다고 법률에서 개정을 하고 있는데 이 부분은 지금의 인력 갖고 할 수 있는 것입니까, 아니면 별도의 사람과 비용이 투입되어야 되는 것입니까? ○여성가족부차관 김희경 현재 폭력피해 이주여성 상담소를 전국에 늘려 가고 있는 상황입니다. ○김현아 위원 아니, 그러니까 늘려 가고 있는데 오는 여기 2항에 보면 전체적으로..... ○표창원 위원 아등으로 확대합니다. ○김현아 위원 예, 확대하고. 여러 가지 상담 프로그램 이런 것들을 지금 확대하는 걸로 되어 있지 않습니까? ○여성가족부차관 김희경 예, 아동 말씀하시는 것이지요? ○소위원장 신용현 예, 아동에 관한 것. 2항은 외국인은 아니고 아동에 관련된 것입니다. ○김현아 위원 예, 알겠습니다. ○소위원장 신용현 그러면 의견해도 되겠습니까? (「예」 하는 위원 있음) 그러면 의견하도록 하겠습니다. 의사일정 제1항 및 2항 2건의 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안은 본회의에 부의하지 아니하고 위원회 대안으로 의결하고자 하는데 이의 있으십니까? (「예」 하는 위원 있음) 이의가 없으므로 가결되었음을 선포합니다. 다음은 의사일정 제3항과 4항 2건의 가족친화 사회환경의 조성 촉진에 관한 법률 일부개정법률안에 대해서 수석전문위원님께서 주요 내용 설명해 주시기 바랍니다. ○수석전문위원 차인순 심사자료 3번 1쪽입니다. 주요 내용은 가족친화 사회환경 조성 기본계획이 수립된 경우 여성가족부장관은 지체 없이 상임위원회에 보고하고 공표하며 또 중앙행정기관과 지자체 장에게 알리도록 하는 내용입니다. 이러한 내용은 바람직한 입법조치로 보았습니다. 다만 이미 공개가 되고 있는 점을 감안하여</p>

표 5 국회 회의록 원시 데이터의 예

본 사업팀은 HWP 파일을 TXT 파일로 저장 후, 해당 회의록에서 다음의 메타데이터들을 추출하였다. (예시는 “377여성(법안심사)소위01(20.5.6.)” 회의록에서 가져옴)

- 회의록의 카테고리: “여성가족위원회회의록”
- 회의록의 제목: “제377회국회(임시회)여성가족위원회회의록(법안심사소위원회)제1호”
- 회의록 작성자: “국회사무처”
- 회의록의 상정된 안건:
 - > “1. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안(이명수 의원 대표발의)”
 - > “2. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안(송희경 의원 대표발의)”
 - > “3. 가족친화 사회환경의 조성 촉진에 관한 법률 일부개정법률안(이양수 의원 대표발의)”
 - > “4. 가족친화 사회환경의 조성 촉진에 관한 법률 일부개정법률안(정부 제출)”
 - > ...
- 회의록의 발화자 목록
 - > “신용현 소위원장”
 - > “김현아 위원”
 - > ...
- 회의록의 각 발화 및 발화에 대한 정보들(발화 id 및 발화자 정보)

각 발화는 다음과 같은 원칙에 의해 범위가 지정되고, id가 부여되었다.

(1) 발화의 범위

한 명의 발화가 여러 개의 문단(새로운 줄)으로 구성된 경우에는 각 문단을 독립적인 발화로 간주하였다. 그 예는 아래와 같다.

○송희경 위원 제가 법안을 낸 건데요, 두 번째 것 보면.
 지금 배우자가 대한민국 국민이 아닌 전체로 확대하는 것에 정부가 수용하시는 것이지요?

표 6 복수 문단으로 구성된 발화의 예

위 예시의 경우, 한 명의 발화자(송희경 위원)에 의해 발화가 이루어졌고 발화는 두 개의 문단으로 구성되어 있다. 이 경우, 발화는 두 개가 독립적으로 이루어졌다고 간주하였다. 이는 속기사가 해당 발화를 기재할 때, 의미상으로 분리가 필요하다는 필요성에 의해 기술했을 것임을 고려한 방식이다. 표 6의 예시에 대해 발화는 다음과 같이 구분되었다.

발화 id	발화자	발화
29	송희경	제가 법안을 낸 건데요, 두 번째 것 보면.
30	송희경	지금 배우자가 대한민국 국민이 아닌 전체로 확대하는 것에 정부가 수용하시는 것이지요?

표 7 복수 문단으로 구성된 발화에 대한 발화의 범위의 예

(2) 서기의 메모에 대한 노트(note) 처리

일부 발화의 경우, 실제 발화가 이루어지지 않았지만 서기의 메모가 하나의 발화처럼 기재된 경우가 있다.

예를 들어 “(「예」 하는 위원 있음)”, “(「아니오」 하는 위원 있음)”, “(웃음소리)” 등과 같은 경우이다. 이에 대한 예시는 아래와 같다.

○소위원장 신용현 그러면 의결해도 되겠습니까?
 (「예」 하는 위원 있음)
 그러면 의결하도록 하겠습니다.

표 8 서기의 메모가 포함된 발화의 예시

위 예시의 경우, 서기의 메모가 하나의 발화처럼 한 줄을 차지하고 있는 경우이다.

이 경우에는 “예”라는 발화의 발화자를 특정할 수 없기에 하나의 발화라고 볼 수 없지만, 회의의 진행 및 내용에 있어서 중요한 역할을 차지하고 있으므로 원시 말뭉치에서는 이를 데이터로 보존할 필요가 있다. 따라서 본 사업팀은 해당 서기의 메모에 대해서 발화에 대한 노트(note)로 간주하여 데이터로 보존하였다. 위 예시에 대한 원시 말뭉치의 실제 구축 결과는 다음과 같다.

```
{
    "id": "SBRW2100000040.1.1.57",
    "form": "그러면 의결해도 되겠습니까?",
    "original_form": "그러면 의결해도 되겠습니까?",
    "speaker_id": "신용현",
    "note": "(「예」 하는 위원 있음)"
},
```

표 9 서기의 메모에 대한 note 처리의 예

표 9의 결과에서는 발화 id(SBRW2100000040.1.1.57), 발화자의 id(신용현) 및 발화 문장(form, original_form)이 JSON 형식으로 변환되었고, 서기의 메모인 “(「예」 하는 위원 있음)”은 직전 발화의 노트로 처리되었다. 이처럼, 서기의 메모가 하나의 발화처럼 한 줄을 차지하고 있는 경우에는 일괄적으로 직전 발화의 노트로 처리하였다.

(3) 발화 내 서기의 삽입구에 대한 처리

실제로 발화자에 의해 발화가 이루어진 것은 아니지만 발화에 대한 부가적인 설명이 필요하여 서기가 기록한 삽입구(괄호로 삽입)가 회의록에 존재한다. 삽입구의 유형은 다음과 같다.

유형	내용	예시	삭제여부
1	부가 설명	성원이 되었으므로 제320회 국회(정기회) 환경노동위원회 제3차 예산결산기금심사소위원회를 개의회하겠습니다. 개성공단 기업 유동자산 피해지원 현황을 보시면 기존 (16년도)에는 실태조사 확인 피해의 70%, 22억 한도에서 1248억 원이... 지난 4월 2일 우리 위원회 해양법안심사소위에서 동 제정법안에 대해서 제1조(목적) 조항부터 제39조(과태료) 조	유지

		<p>항까지 그리고 5개의 부칙 조항까지 모두 심사를 마친 바 있습니다.</p> <p>총 투자비가 1조 3336억 원이고, 타인자본은 52.37%인 7000억, 자기자본은 1832억(13.71%)입니다.</p>	
2	한자어와 외래어 원어에 대응하는 한글 표기	<p>PRISM(정책연구관리시스템)에 연구용역 결과가 등록되어야 되는데 등록 안 된 것이 7건 있다, 이런 것들에 대해서는 저희가 앞으로 고쳐서 제도개선을 하도록 하겠습니다. 저희가 수용을 하고요.</p> <p>이번 개정안은 국민에 대한 법률 이해가능성을 높이기 위해서 어려운 한자식 표기인 조문을 한글로 개정하는 내용입니다. 2조3호에 있는 '備記(게기)된'이라는 표현을 기록하여 보게 한다는 의미는 '규정된' 의미로 사용되고 있는데 이걸 '규정된'으로 바꾸는 내용입니다. 타당한 입법이라고 생각합니다.</p>	삭제
3	외래어 원어 표기	<p>19페이지, 철도시설 영상기록장치(CCTV) 설치 의무화 관련입니다.</p> <p>최근에 에어부산이, 이동지역을 타맥(Tarmac)이라고 하는데, 타맥 딜레이(Tarmac delay) 건 관련해서 현재 규제가 아까...</p> <p>그것은 무슨 얘기냐 하면 이게 애당초 경쟁구조가 안 되기 때문에 공모를 해도 실효성이 없을 거예요. 인큐베이팅(Incubating)이 안 됐어요, 그걸 대행할 수 있는 주체들이. 그런 측면까지 같이 검토하세요, 다른 부처도 마찬가지로.</p>	삭제
4	한자어 한자 표기	<p>동포재단에서 '상'자 붙였어요, '사업 상(商)' 자를 붙여가지고. 비슷해요, 오는 사람들 비슷하고. 한 국가가 비슷한 사람들에</p> <p>한문을 아들 자(子) 자를 쓰잖아요.</p> <p>먼저 93페이지에 구(區)를 두지 않는 50만 이상의 대도시인데요,</p> <p>이것을 빨리 해 줘야 사실은 내(內)의 현재 여러 가지 고용 안정이라든가 근로조건 개선에 도움이 될 텐데.....</p>	삭제
5	행동 명시	<p>(웃음) 국회의 좋은 기능으로 저희들이 받아들일 것입니다.</p> <p>(웃음) 그다음 안건들은 그렇게 사람들 많거나 복잡하지는 않잖아요? 그래서 10번, 윤후덕 의원님의 수도권 광역급행철도(GTX) 파주연장 '건설에 관한 청원으로 넘어가겠습니다.</p> <p>(웃음) 그러면 청장님은 ADD 본부장 하다가 청장으로</p>	삭제

표 10 삽입구의 전체 유형

표 10의 삽입구의 유형 중, 유형 1의 사례인 “(정기회)”와 같은 부가 설명은 회의 진행에서 중요한 정보로 판단되며 직접 발화의 가능성이 있어 삭제하지 않았다. 유형 2, 3, 4, 5의 경우에는 직접 발화가 아니므로 삭제하였다. 이때, 상기의 원문 형태를 original_form에서는 그대로 유지하였고, 정제 후 발화인 form의 경우에 한하여 삭제를 진행하였다. 분석 작업 시에는 form을 분석 대상 문장으로 삼았다.

유형 2의 “備記(계기)”의 사례와 같이, 발화를 한자로 기재하고 이에 대한 한글을 함께 적은 경우에는 form에서 한자를 한글로 변환하고 삽입구를 삭제하였다.

추가로, 원시 말뭉치의 발화자 정보는 실제로 발화한 사람 이외에도, 회의에 참석한 사람을 모두 발화자의 목록으로 기록하여 원문 정보가 원시 말뭉치에서 손실되지 않도록 진행하였다.

○출석 위원(8인)			
김 현 아	송 옥 주	송 희 경	신 경 민
신 보 라	신 용 현	정 춘 숙	표 창 원
○출석 전문위원 및 입법심의관			
수 석 전 문 위 원		차 인 순	
입 법 심 의 관		김 수 욱	
○정부측 및 기타 참석자			
여성가족부			
차 관		김 희 경	
기 획 조 정 실 장		김 중 열	
청 소 년 가 족 정 책 실 장		윤 효 식	
여 성 정 책 국 장		이 건 정	
권 익 증 진 국 장		황 윤 정	
정 책 기 획 관		최 창 행	
청 소 년 정 책 관		심 민 철	
가 족 정 책 관		김 권 영	
한국여성인권진흥원장		박 봉 정	숙

표 11 국회 회의록 마지막 쪽에 기재된 회의 참석자 정보

표 11과 같이, 실제 발화자 이외에도 참석자에 대한 정보가 기재되어 있다. 본 사업 팀에서는 위와 같은 정보가 손실되지 않는 원시 말뭉치를 구축하기 위해 전처리를 통하여 모두 발화자의 목록에 추가하였다. 위 예시에서 “김희경”의 경우 직책이 “차관”으로 기재되어 있지만, 상위 정보에서 “여성가족부” 소속임을 확인할 수 있다. 이러한 경우에는 “김희경”의 직책을 “여성가족부차관”으로 기재하였다. “김종열”, “윤효식”, “이건정” 등의 참석자 역시 “여성가족부” 소속임을 명시하였다. 해당 작업의 경우 자동으로 진행될 수 있으나, 회의록의 데이터가 서로 다른 형식으로 기재되어 있는 사례가 있어, 이 경우에는 수작업으로 진행되었다.

```

{
    "id": "김종열",
    "occupation": "여성가족부기획조정실장"
},
{
    "id": "윤효식",
    "occupation": "여성가족부청소년가족정책실장"
},
{
    "id": "이건정",
    "occupation": "여성가족부여성정책국장"
},
}

```

표 12 회의 참석자 정보 원시 말뭉치의 예

2.2. 원시 말뭉치 데이터 형식

2.1장의 전처리 작업을 거쳐 원시 말뭉치 데이터는 아래와 같은 형식으로 구축되었다.

```
{
  "id": "SBRW2100000040",
  "metadata": {
    "title": "국립국어원 국회 회의록 원시 말뭉치 SBRW2100000040",
    "creator": "국립국어원",
    "distributor": "국립국어원",
    "year": "2021",
    "category": "회의록 > 국회소위원회 > 여성가족위원회회의록",
    "annotation_level": [
      "원시"
    ],
    "sampling": "본문 전체"
  },
  "document": {
    "id": "SBRW2100000040.1",
    "metadata": {
      "title": "제377회국회(임시회)여성가족위원회회의록(법안심사소위원회)제1호",
      "author": "국회사무처",
      "publisher": "국회사무처",
      "date": "20200506",
      "topic": "1. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안(이명수 의원 대표발의) / 2. 가정폭력방지 및 피해자보호 등에 관한 법률 일부개정법률안(송희경 의원 대표발의) ...",
      "speaker": [
        {
          "id": "신용현",
          "occupation": "소위원장"
        },
        {
          "id": "차인순",
          "occupation": "수석전문위원"
        }
      ]
    }
  }
}
```

```

...
    "utterance": [
      {
        "id": "SBRW2100000040.1.1.1",
        "form": "의석을 정돈하여 주시기 바랍니다.",
        "original_form": "의석을 정돈하여 주시기 바랍니다.",
        "speaker_id": "신용현",
        "note": "(09시58분 개의)"
      },
      {
        "id": "SBRW2100000040.1.1.2",
        "form": "성원이 되었으므로 제377회 국회(임시회) 제1차 여성가족위원회 법안심사소위원회를 개의하겠습니다.",
        "original_form": "성원이 되었으므로 제377회 국회(임시회) 제1차 여성가족위원회 법안심사소위원회를 개의하겠습니다.",
        "speaker_id": "신용현",
        "note": ""
      },
      {
        "id": "SBRW2100000040.1.1.3",
        "form": "바쁘신 일정 가운데에서도 20대 국회의 유종의 미를 거두기 위해서 회의에 참석해 주신 여러 위원님들께 감사의 말씀을 드립니다.",
        "original_form": "바쁘신 일정 가운데에서도 20대 국회의 유종의 미를 거두기 위해서 회의에 참석해 주신 여러 위원님들께 감사의 말씀을 드립니다.",
        "speaker_id": "신용현",
        "note": ""
      },
      {
        "id": "SBRW2100000040.1.1.4",
        "form": "법안 심사를 위해 김희경 여성가족부차관과 관계 공무원이 참석하셨습니다.",
        "original_form": "법안 심사를 위해 김희경 여성가족부차관과 관계 공무원이 참석하셨습니다.",
        "speaker_id": "신용현",
        "note": ""
      },
    ]
  }

```

표 13 원시 말뭉치 JSON 형식의 예

표 13의 원시 말뭉치의 예시에서 다음의 내용을 확인할 수 있다.

- 회의록 파일의 메타데이터 (root > metadata)
- 회의록 문서의 메타데이터 (root > document > metadata)
 - > 회의록의 발화자 정보 (root > document > metadata > speaker list)
- 회의록의 발화 목록 (root > document > utterance)
 - > 발화의 고유 아이디 (utterance > id)
 - > 정제 후 발화 (utterance > form)
 - > 정제 전 원문 발화 (utterance > original_form)
 - > 발화의 발화자 정보 (utterance > speaker_id)
 - > 발화의 노트 (utterance > note)

이 경우, 발화의 발화자 정보(speaker_id)는 회의록 문서 메타데이터의 발화자 정보 (speaker_list)의 id와 상호 호환된다.

3. 회의록 요약 말뭉치 구축 과정

3.1. 회의록 특징

□ 국회 회의록의 특징

회의는 개회, 안건 상정, 각 안건에 대한 의원 제안설명, 정부 제안설명, 전문위 검토 보고, 질의응답 및 산회 등의 여러 단계에 걸쳐 진행된다.

뉴스 요약이나 도서 자료 요약, 논문 요약의 경우 단일 문서를 다루기는 하지만 1~2 쪽의 비교적 짧은 원문에 대한 3~5문장의 요약문을 제공한다. 반면, 국회 회의록은 앞서 언급된 바와 같이 긴 회의록으로 회의의 각 단계가 구분되지 않고 시간 순서에 따라 발화들이 전사된 문서이다.

국회 회의록의 특징은 다음과 같다.

- 하나의 주제가 아닌, 여러 주제(안건)에 대한 논의
- 각 주제는 독립적 대화로 구성되지 않고 회의 전반에 걸쳐 논의

	뉴스, 도서, 논문 등	일상대화	국회 회의록
화자	1명(저자)	다자	다자
주제	1개	1개	다수 (안건별)
분량	약 1쪽	4~5개 말차례	20쪽 500개 말차례

표 14 요약 데이터 비교표

본 과제에서 대상으로 한 국회 소위원회 회의록은 약 20쪽의 문서로 구성되어 있다.

□ 논쟁적 대화의 특징

논쟁적 대화에 해당하는 회의록의 경우 한 명의 저자에 의해 하나의 주제에 대해 기술되는 뉴스, 도서와 같은 독백(Monologue)과는 그 성질이 다르다. 독백의 경우 논리적인 구조에 따라 기술한다는 점에서 문장, 혹은 텍스트의 일부 부분들은 수사 구조(Rhetorical Structure)에 의해 이해될 수 있다. 예를 들어, 아래 표 15와 같이 논변으로 주장을 지지하는 구조라고 볼 수 있다.

반면 대화의 경우에는 각 발화의 화행(Dialog Act)에 따라 그 의미가 달라지며, 각각

의 주제에 대해 동의하거나 비동의를 하는 여러 의견이 등장할 수 있다. 또한 각각의 이슈에 대해 특정한 주장이 등장할 수 있다. 예를 들어, 아래 표 15와 같이 진술(Statement) 혹은 특정 쟁점을 꺼내며 질문하는 이슈 제기(Open Issue)와 같은 역할이다. 그리고 각각의 발화는 여러 가지 수사 관계를 맺을 수 있다. 해당 주제에 대한 답변이거나 긍정적 혹은 부정적인 발언, 구체화하거나 요청하는 등의 다양한 역할이 있다. 특히 국회 회의록의 경우에는 40여 개의 법안에 대해 전문위원의 분석, 국회의원들에 의한 질의응답 등에 의해 더욱 복잡한 논변 구조가 나타난다. 따라서 이러한 복잡한 논변 구조에 대한 이해가 충분히 전제되지 않은 상태에서 핵심 정보를 모두 포함하는 요약문을 작성하는 것은 매우 어렵다.

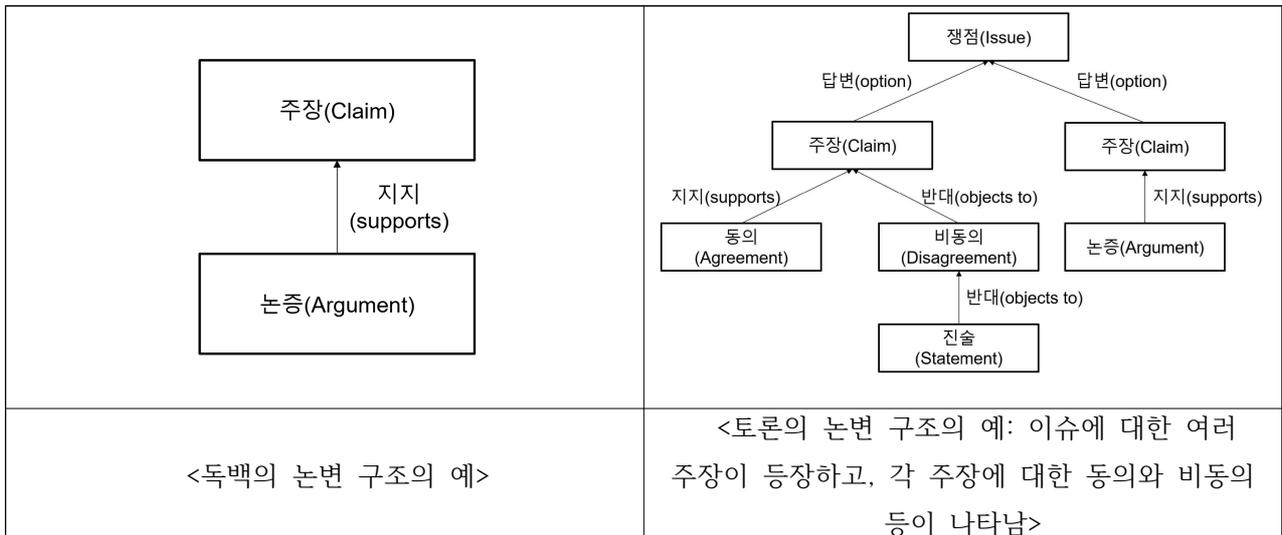


표 15 독백과 토론의 논변 구조의 비교표

□ 회의록 요약의 특징

일반적으로, 뉴스 요약 연구에서는 뉴스에서 가장 핵심이 되는 주제어와 주제 문장을 선정하고 해당 문장을 중심으로 한 1~5개의 요약 문장들을 작성 혹은 추출하는 연구를 수행하였다.

본 연구과제에서 다루는 국회 회의록의 경우에는 중요한 내용이 여러 발화에 걸쳐서 나타나고, 또한 다양한 안건과 쟁점에 대해 논의가 이루어지기 때문에 간단히 5개 문장 정도로 요약하기는 어렵다. 특히, 회의록의 목적과 특성에 적합한 구조의 요약이 선행되어야 한다. 예를 들어, AMI 말뭉치의 경우에는 약 163개 말차례(turn)로 구성된 회의에 대해서 약 23개의 요약 문장을 작성하였다.

따라서 회의록에 대한 요약문 작성은 핵심 문장 몇 개를 고르고 환언(paraphrasing)

하는 방식의 일반적인 요약문 작성 방법이 적용되기 어려우며 회의록의 목적과 특성에 맞는 요약문이 작성되어야 한다. 예를 들어, 회의에서 다룬 안건과 쟁점이 무엇이었는지, 그리고 각각의 안건과 쟁점에 관해 결정된 사항은 무엇인지, 회의에서 나왔던 이슈나 문제점들은 무엇이었는지, 그리고 다음 회의를 위해 수행해야 하는 것들은 무엇인지 등 여러 가지 요소가 포함된 요약문이 작성되어야 한다.

3.2. 추진 전략

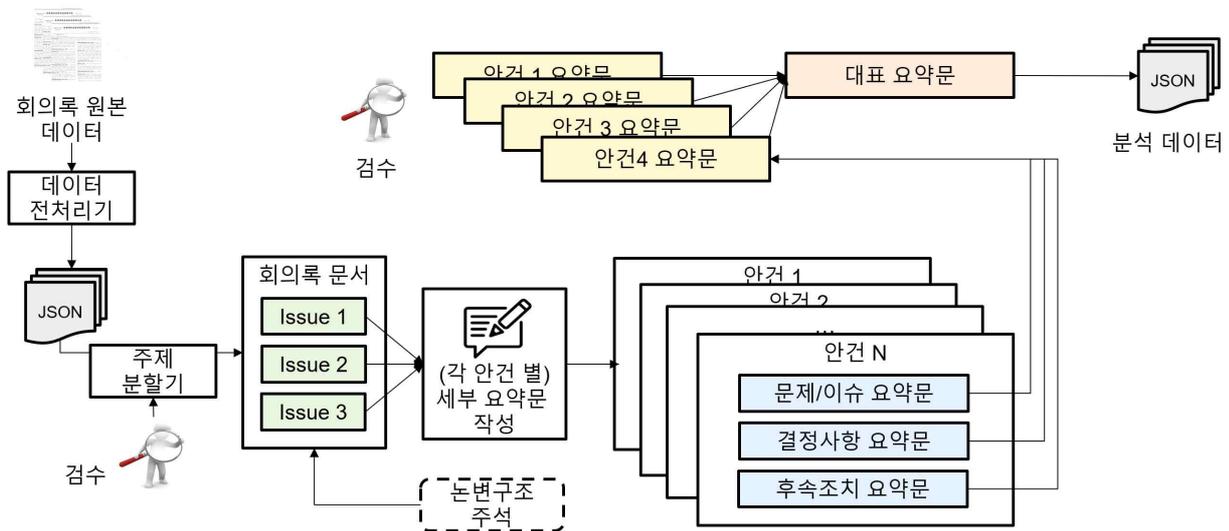


그림 1 추진 전략

위에서 논의된 바와 같이, 회의록은 여러 안건으로 구성되어 있으며, 각 안건에 관련된 쟁점들은 모두 회의록의 중요한 정보이다. 본 연구에서는 각 안건에 대해 논의된 이슈와 문제점, 결정사항, 후속조치 등에 관한 내용을 모두 포함하여 요약문을 작성하기 위해 상향식 요약문 구축 방법을 취한다.

3.2.1. 쟁점 분할에 의한 회의록 구조 파악

국회 회의록 문서 하나에는 여러 가지 안건에 대해 논의가 이루어진다. 또한 문서 안 대화는 비록 명확하게 구분되지는 않으나 대체로 안건 하나에 대해 논의가 이루어지고, 이후 다른 대화를 통해 다음 안건 논의가 시작되는 방법을 취한다. 각 안건에 관한 결정이나 후속조치 등은 논의 단계에서 언급되거나 회의의 마지막에서 결정되기도 한

다.

본 연구에서는 각 안전에 대해 집중적으로 요약 분석을 수행하기 위하여 회의록을 쟁점에 따라 나누어 회의록의 구조를 파악한다.

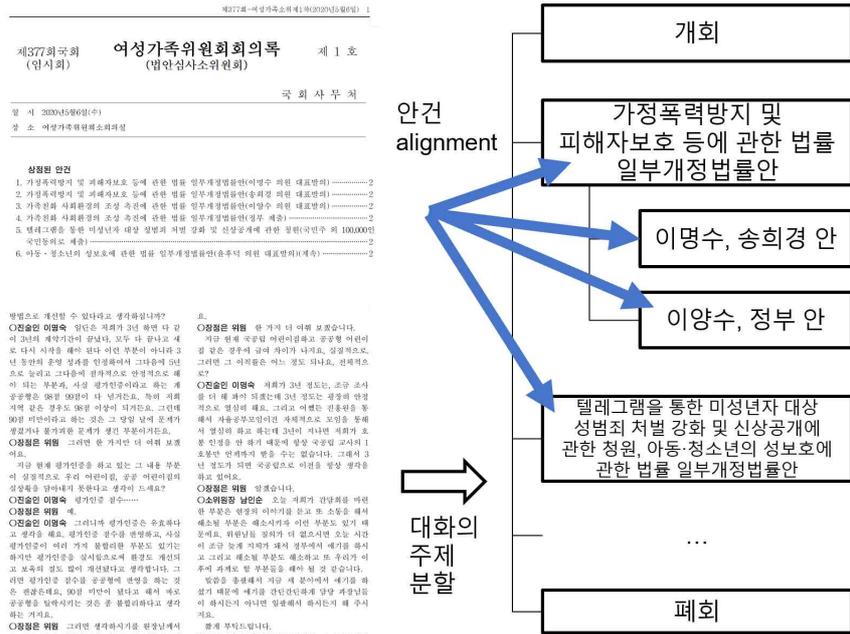


그림 2 주제 분할에 의한 회의록 구조 파악

특히 국회 회의록은 사회자(소위원장)의 의사 진행 발언으로 안전을 구분할 수 있으므로, 사업팀에서는 사회자의 발화를 중심으로 대화의 범위를 구분하는 방식을 취했다. 그리고 분할된 주제에서 중심어를 태깅하고, 이를 안전과 매칭하였다. 상세한 내용은 제3장의 지침에서 상술한다.

3.2.2. 상향식 대표 요약문 작성

추상 요약문은 논쟁의 중요한 정보를 파악하기 위하여 논쟁적 대화의 입체적 정보(문제/이슈, 결정사항, 후속조치 등)에 대한 각각의 세부 요약문을 작성하는 절차를 포함한다. 상향식 대표 요약문 작성의 절차는 다음과 같다.

- (1) 주제 분할된 대화의 중심어, 즉 쟁점에 대해 문제/이슈(Problems/Issues), 결정사항(Decisions), 후속조치(Actions)에 대한 각 세부 요약문을 작성한다.
- (2) 각 세부 요약문을 통합 및 중복을 제거하여 환언(paraphrasing) 방식으로 각 안전

에 대한 추상 요약문을 작성한다.

(3) 각 안건에 대한 요약문을 다시 통합 및 중복을 제거하여 환언하는 방식으로 회의록 전체에 대한 대표 요약문을 작성한다.

상향식 대표 요약문 작성 방법은 각 쟁점에서 논의되는 입체적 정보를 모두 반영한 요약문들을 바탕으로 대표 요약문을 작성하는 방식으로, 추상 요약문을 작성하는 작업자들은 여러 단계를 통해 회의에서 나타난 각 쟁점에 대한 중요한 정보들을 놓치지 않고 압축적인 요약문을 작성할 수 있다.

위에 적은 상향식 절차를 통해 회의록에 대한 대표 요약문과 각 안건에 대한 요약문, 그리고 각 안건에 대한 세부 요약문(문제/이슈제기, 결정사항, 후속조치 등)이 분석 및 평가의 결과가 된다.

요약문 작성에 대한 상세한 내용은 제3장의 지침에서 상술한다.

4. 회의록 요약 말뭉치 구축 결과

4.1. 원시 말뭉치 통계

항목	내용
회의록의 개수	200개
회의록 전체 문장	95,052 문장 ⁴⁾
회의록당 평균 문장	475.26 문장
회의록 전체 어절	- form 기준 1,560,255 어절 - original_form 기준 1,560,387 어절
회의록당 평균 어절	- form 기준 7,801.275 어절 - original_form 기준 7,801.935 어절

4.2. 분석 말뭉치 통계

항목		내용		
회의록의 개수		200개		
회의록 1개 기준	대표 요약문	평균 문장	3.815 문장	
		평균 어절	83.875 어절	
	이슈의 개수		4.42개	
	이슈 1개 기준	이슈 요약문	평균 문장	2.14 문장
			평균 어절	45.28 어절
		세부 요약문의 개수		평균 16.23개
	세부 요약문 1개 기준	평균 문장	1.01 문장	
평균 어절		23.27 어절		

4.3. 원시 말뭉치에 대한 요약문의 비율

항목	원시 말뭉치	요약(대표 요약문, 이슈 요약문, 세부 요약문)
회의록의 개수	200개	200개
회의록 전체 문장	95,052 문장	17,170 문장
회의록당 평균 문장	475.26 문장	85.85 문장
문장의 평균 어절 수	16.415 어절	22.75 어절
회의록 전체 어절(form 기준)	1,560,255 어절	390,713 어절
회의록당 평균 어절	(form 기준)	7,801.275 어절
	(original_form 기준)	7,801.935 어절
		1,953.565 어절

4.3. 작업 통계

작업자 정보

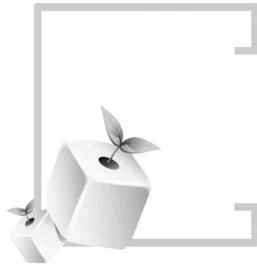
작업자	전공	연령대	주제 분할 작업 수량	요약문 분석 작업 수량	요약문 검수 참여	요약문 분석 평균 작업 시간
작업자 A	경영학	20대 중반	48	22	0	4

4) 원자료의 영향(공백인 줄 등)으로 형태가 비어 있는 form과 original_form도 있다. 이를 제거하면, 모두 95,035 문장이다.

작업자 B	컴퓨터공학	20대 중반	35	20	0	6
작업자 C	국어국문학	20대 중반	43	24	0	6
작업자 D	국어국문학	30대 초반	34	5	0	8
작업자 E	심리학	30대 초반	-	20	-	6
작업자 F	법학	30대 중반	-	17	-	7
작업자 G	일본문학	30대 초반	-	14	-	5
작업자 H	일본문학	30대 중반	-	14	-	5
작업자 I	스페인어	20대 초반	-	12	-	5
작업자 J	문예창작학	20대 후반	-	6	-	24
작업자 K	국어국문학	20대 후반	-	6	-	5

□ 작업 통계

구분	평균 소요시간(1인)	투입 작업자 수
주제분할 (작업)	1시간	4
주제분할 (검수)	0.5시간	4
요약문 (작업)	7시간	11
요약문 (검수)	3시간	4



제 3 장

논쟁적 대화 요약 말뭉치 지침



1. 논쟁적 대화 요약 말뭉치 구축 지침

1.1. 논쟁적 대화 요약 개요

1.1.1. 개요

본 장에서는 국립국어원 사업 '2021년 회의록 요약 말뭉치 연구 분석'에서 논쟁적 대화 요약 말뭉치를 구축하기 위한 지침을 기술하고, 해당 과제를 수행하기 위한 세부 분석 지침으로써 국회 회의록 요약에 관한 세부 지침을 추가한다. 본 지침에서는 논쟁적 대화의 요약문 예시로서 국회 회의록의 요약문을 위주로 사용하였으나, 일부 지침의 내용이 국회 회의록에 관한 세부 지침이면 따로 구분하였다.

요약 말뭉치는 원시 말뭉치에 대한 핵심적인 정보를 선택하고, 불필요한 정보를 삭제하며, 이를 종합하고 수사적 구조를 재구조화하는 과정을 거치며 구축된다. 기존의 문서 요약 말뭉치 분석 연구 및 지침은 위의 절차를 반영하여 문서(예: 신문 기사)에서 핵심 문장들을 찾고 이를 종합할 수 있는 요약 문장들을 작성하는 방식이 제안되었다. 이러한 방법을 통해 구축되는 요약 말뭉치는 정보를 전달하는 문서에서 중요한 정보를 선별하고, 이를 기술하는 요약문을 작성하는 것이 목적이다.

논쟁적 대화를 녹취한 문서(이하, 논쟁적 대화)는 화자(저자)가 한 명인 일반적인 문서와 달리 여러 화자가 발화를 주고받는 구조이다. 일반적으로 문서는 논리 구조를 통해 기술되는 방식을 취한다. 이와 다르게 논쟁적 대화에서는 대화에서 암묵적 혹은 명시적으로 쟁점이 존재하고 해당 쟁점에 대한 각 화자의 주장(문제제기, 견해, 결론, 입장 등)이 드러난다. 특히 일반적인 문서가 한 명의 저자(보고자 등)의 견해로 결론이 종합되고 작성되는 것과 달리 논쟁적 대화에서는 대화를 통해 합의점 혹은 결론이 도출된다.

따라서 논쟁적 대화는 일반적인 문서 요약과 달리 대화에서 제기된 쟁점은 무엇인지, 해당 쟁점에 대한 각 화자의 주장은 무엇인지, 그리고 논쟁적 대화를 통해 도출된 결정이나 해야 할 일은 무엇인지 등을 파악하여야 한다.

논쟁적 대화에서의 중요한 정보
(1) 대화에서 다루는 이슈가 무엇인가?
(2) 각 화자의 주장이 무엇인가?
(3) 회의를 통해 도출된 결론이 무엇인가?

표 16 논쟁적 대화에서의 중요한 정보

본 지침은 위의 논쟁적 대화에서 중요한 정보를 포함할 수 있는 요약 분석의 개념 (1.1.2)을 기술한다. 그리고 요약을 위한 방법론을 2장에서 기술하고, 세부 방법 및 예시를 3장에서 기술한다.

1.1.2. 논쟁적 대화 요약 분석의 개념

□ 논쟁적 대화

논쟁적 대화란 일상 대화와 달리 대화의 참가자(발화자)가 특정 주제에 대해 입장(stance)과 의견(opinion)이 있는 상태에서 이루어지는 대화로서, 대화의 주제가 되는 쟁점과 결정사항 등으로 구성되어 있으며 각 화자의 발화들은 자신의 주장에 대한 근거를 논증하는 방식으로 이루어진다. 대표적인 예로 회사에서의 프로젝트 미팅과 같은 회의 등을 생각해 볼 수 있다.

□ 논쟁적 대화 요약의 대상이 되는 핵심 정보

논쟁적 대화 요약 분석은 구어체로 이루어진 논쟁적 대화에서 나타난 핵심 정보를 포착하고 이를 문어체인 요약문의 형태로 작성하는 일련의 과정을 포함한다. 1.1장에서 논의된 바와 같이, 논쟁적 대화에서 중요한 정보는 대화에서 논의의 대상이 되는 쟁점이 무엇인지, 그리고 논쟁을 통해 결정된 사항이 무엇인지를 포함하여야 한다. 또한 한 명의 저자가 작성한 문서와 달리, 대화 참가자(발화자)의 의견이나 입장이 서로 다르게 나타나므로 각 화자의 주장이 무엇인지가 명시되어야 한다.

본 지침에서는 위의 정보를 포함하는 요약문 작성을 위해 해외 연구인 AMI Corpus⁵⁾의 요약문 구성을 참고하여 다음의 내용을 포함하는 요약문을 작성한다.

5) <https://groups.inf.ed.ac.uk/ami/corpus/>

주요 정보	내용
Problems/Issues	논쟁적 대화에서 논의의 대상이 되는 쟁점/문제/이슈
Decisions	논쟁적 대화를 통해 도출된 결정사항
Actions	논쟁적 대화를 통해 결정된 내용의 후속조치나 과업사항

표 17 요약문에 포함되는 주요 정보

위의 정보를 포함하기 위해, 작성되는 요약문은 “누가” 요약의 대상이 되는 “문제”를 제기하였는지, 이에 대해 각 화자의 “주장”은 무엇인지, 그리고 도출된 “결론”은 무엇인지 등을 포함하여야 한다.

□ 논쟁적 대화 요약의 방법 개요

논쟁적 대화는 하나의 안전에 대해서 심도 있는 논의를 하는 일도 있으나, 대개는 여러 안전에 관해서 대화가 이루어진다. 뉴스와 같은 문서를 요약할 때 여러 개의 뉴스가 혼재한 경우가 아닌 독립적인 뉴스를 요약의 대상으로 삼는 것처럼 논쟁적 대화 요약에서도 각각의 안전을 다루는 대화들을 상대로 요약을 수행하는 것이 중요한 정보를 집중적으로 요약할 수 있다. 본 지침에서는 대화가 만약 여러 가지 쟁점에 대해서 이루어지면 대화를 쟁점별로 나눈 후 일부를 선택하여 이에 대해 각각 요약문을 작성하는 방법을 채택하였다.

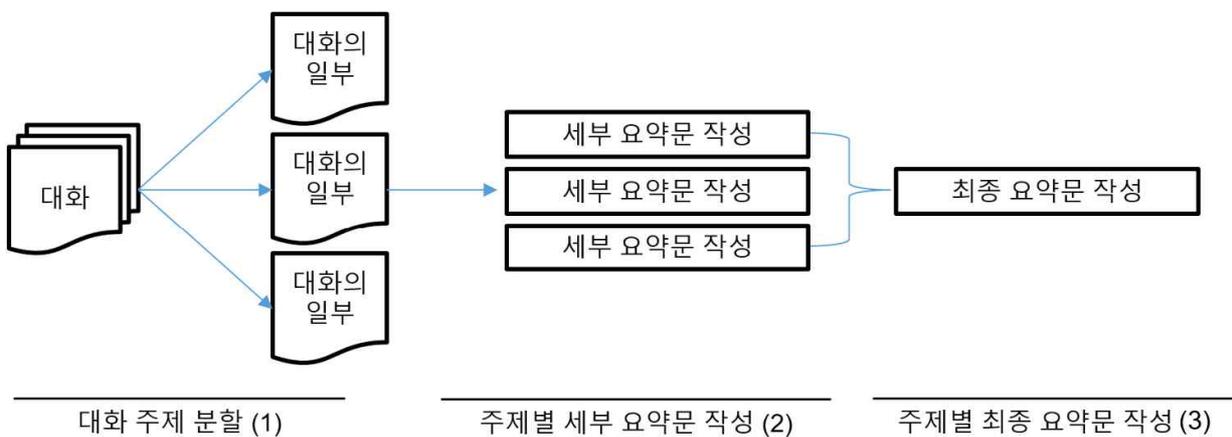


그림 3 논쟁적 대화 요약 절차

각각의 대화에 대한 요약문은 주요 정보를 모두 포함하여야 한다. 본 지침에서는 두 단계를 거쳐 요약문을 구축하도록 제안하였다. 먼저 논쟁적 대화에서의 논의의 대상이 되는 쟁점/문제/이슈에 대한 각각의 세부 요약문을 작성하고, 또한 논쟁적 대화에서 도출된 결정사항이나 후속조치, 과업사항에 대한 각각의 요약문을 작성한다. 이후 상기의 세부 요약문을 통합하여 핵심 정보를 선택하고 불필요한 정보를 제거하여 최종 요약문을 작성하는 방법을 채택하였다.

세부 요약문은 문제/이슈(Problems/Issues), 결정사항(Decisions), 후속조치(Actions)에 대한 요약문들로 구성되어 있으며, 최종 요약문은 복수의 요약문에 대해 통합된 요약문이다. 이러한 요약문의 구조는 아래와 같다.

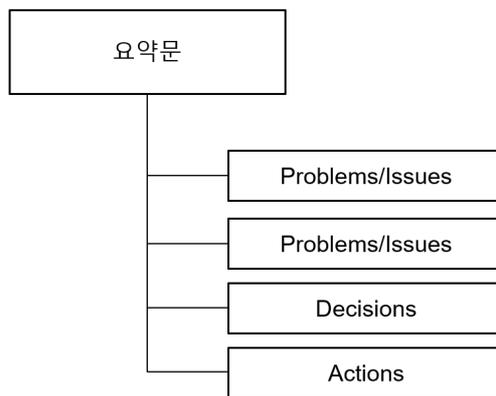


그림 4 논쟁적 대화 요약문의 구조
(하나의 안건에 대한 대화의 예시)

위 그림과 같이, 논쟁적 대화 요약은 하나의 쟁점에 대한 논쟁적 대화를 하나 이상의 문장으로 구성된 요약문을 작성하는 것이 목표이다. 해당 요약문은 여러 개의 세부 요약문으로 구성되며 각 세부 요약문은 문제/이슈, 결정사항, 후속조치의 각 유형에 대해 구축된다. 이러한 방법을 통해 해당 안건에 대해 이루어진 “문제”와 “주장”, “결론”을 모두 포함하는 요약문을 구축할 수 있다.

□ 논쟁적 대화 요약문 작성의 개요

논쟁적 대화에서 요약의 대상이 되는 중요한 정보는 1.1.1장에서 논의된 바와 같이 다음의 3가지이다: 대화에서 등장한 쟁점들, 각 화자의 주장, 회의를 통해 도출된 결론. 이러한 중요 정보를 요약문으로 작성하기 위해서는 다음의 4개 요소가 요약문에 포함되어야 한다.

- 누가 (누가 해당 문제를 제기했는가? 누가 어떤 의견을 제시했는가?)
- 무엇을 (어떤 내용을 이야기했는가? 해당 문제는 구체적으로 무엇인가?)
- 어떻게 (대화를 통해 도출된 사항은 무엇이고, 어떻게 결정되었는가?)
- 왜 (문제가 있는 이유는 무엇인가?)

이에 따라 요약문은 발화자 정보를 포함하고 각 문제/의견마다 개별적인 문장으로 작성되어야 한다. 또한 대화를 통해 쟁점이 명시되었다면, 이러한 쟁점은 반드시 포함되어야 한다. 이에 대한 구체적인 원칙과 절차는 2장에서 상술한다.

1.2. 논쟁적 대화 요약 방법론

1.2.1. 용어의 개념과 정의

□ 쟁점(issue)과 안건(topic)

본 지침에서는 쟁점(issue)과 안건(topic) 용어를 구분하여 사용한다. 본 지침에서 언급하는 쟁점은 논쟁적 대화에서 두 명 이상의 발화자가 논의를 수행한 논쟁의 대상이다. 안건은 회의록에서 사전에 명시된 혹은 사후에 기재된 회의의 안건을 의미하며 국회 회의록의 경우에는 “상정된 안건”이 안건에 해당한다. 본 지침에서 두 용어를 구분하여 사용하는 이유는 쟁점은 요약의 (후보) 대상이 되는 것이고, 안건은 국회 회의록에서 명시된 “상정된 안건”을 의미한다는 것을 명확히 하기 위함이다.

□ 요약문(Abstract)

요약문은 논쟁적 대화를 포괄하는 요약문을 의미한다. 해당 요약문은 논쟁적 대화에서 다룬 쟁점(issue), 대화에서 이루어진 상호 간의 주장(claim), 그리고 대화를 통해 도출된 결정사항 등을 모두 포함한다. 대화에서 위의 내용 중 일부가 다루어지지 않았다면(예: 결정사항이 없는 경우) 요약문에 포함되지 않아도 무방하다.

요약문의 예시는 아래와 같다.

강은희 의원이 대표발의한 고등교육법 일부개정법률안에 관하여 정부는 2년의 유예기간을 달라고 요구하였다. 이에 1년이 적절하다는 의견과 국공립대학과 사립대학 간의 형평성을 고려한 유예 대책이 있는지에 대한 지적 등이 있었지만, 원안을 유지하면서 전문위원이 부대의견을 다는 것으로 합의되어 원안대로 의결됐다.

338교문(법안심사)소위01(15.12.23) 中

위 요약문은 특정 발화자가 제기한 쟁점(강은희 의원이 발의한 고등교육법 일부개정법률안)에 대해 상호 간의 주장(1년이 적절, 유예대책 등)을 기술하고, 결정사항(부대의견, 원안대로 의결)을 모두 포함하고 있다.

□ 문제/이슈(Problems/Issues)

문제/이슈(Problems/Issues)는 논쟁적 대화에서 나온 쟁점/문제/이슈 등을 기술한 세부 요약문이다. 각각의 문제/이슈에 대해서는 가능한 하나의 문장으로 작성되도록 하며, 한 명의 화자에게 복수의 이슈가 등장한 때도 각 문제/이슈를 하나의 문장으로 작성하도록 한다. 논쟁적 대화는 A화자가 무언가를 주장하고 해당 주장에 대해 B화자가 문제를 제기하는 방식으로 문제/이슈가 드러난다. 이에 대한 예시는 아래와 같다.

교육부차관 이영은 시간강사 처우개선 방안 추진에 있어, 시간 강사 처우 부분을 평가 지표로 넣는 등 대학들의 변화를 유도하는 한편, 평균 임금 상승을 위한 재정사업도 만들고 협의체를 구성해서 운영할 것이라고 정부의 의견을 밝혔다.
--

338교문(법안심사)소위01(15.12.23) 中

위 예시에서는 교육부차관 이영이 시간강사 처우개선이라는 쟁점에 대해서 어떠한 조처를 할 것인지를 언급하는 내용을 요약한 세부 요약문이다.

이러한 교육부차관의 발화에 대한 다른 화자의 의견 역시도 문제/이슈에 해당하는 세부 요약문이다. 이에 대한 예시는 아래와 같다.

박인숙 위원은 디테일로는 처우개선이 된 것으로 보이나 TF나 협의체 구성 플랜을 만들어서 2년 후에 해당 법에 대해 동일한 문제가 없도록 교육부에 요청했다.

338교문(법안심사)소위01(15.12.23) 中

문제/이슈에서는 위와 같이 논쟁적 대화에서 각 화자의 주요 발화를 중심으로 해당 화자의 주요 주장이나 문제제기 등을 요약의 대상으로 한다.

이러한 세부 요약문을 통해, 요약문에서는 논쟁적 대화에서 어떠한 내용이 논의되었는가를 요약된 정보로써 확인할 수 있다.

□ 결정사항(Decisions)

논쟁적 대화에서 문제점만 드러나는 일도 있지만, 그에 대해서 결론이 도출되는 경우도 존재한다. 결정사항(Decisions)은 이러한 결론 및 결정된 사항에 대한 세부 요약문이다. 이에 대한 예시는 아래와 같다.

소위원장 신성범은 고등교육법 일부개정법률안은 원안대로 의결할 것을 가결했다.
338교문(법안심사)소위01(15.12.23) 中

강은희 의원이 대표발의한 고등교육법 일부개정법률안에 대해서 여러 논쟁이 이루어졌고, 이에 대한 결론으로서 원안대로 의결하는 것으로 가결되었다는 정보가 결정사항에 해당하는 요약문으로 작성되었다. 위 예시와 같이 해당 법률안이 의결된 것은 소위원장의 가결을 통하여 이루어졌다는 정보를 포함하여 요약문이 작성된다.

이러한 세부 요약문을 통해, 요약문에서는 논쟁적 대화에서 어떠한 결정사항이 있었는가를 요약된 정보로써 확인할 수 있다.

□ 후속조치(Actions)

논쟁적 대화에서 결정된 사항 이외에 별도로 향후 후속조치나 과업사항에 대한 내용 역시도 논의될 수 있다. 후속조치(Actions)는 이러한 정보를 포함하는 세부 요약문으로서, 논쟁적 대화 참여자 혹은 관계자가 수행해야 할 과업에 해당한다.

교육부대학정책관 배성근은 유은혜 위원의 요청을 받아들여 강사 노조 2곳과 협의체 구성 후 의견 반영을 2곳 모두 할 것이다.
338교문(법안심사)소위01(15.12.23) 中

위 예시에서 해당 대화의 쟁점인 고등교육법 일부개정법률안이 의결되는 것과 별개로, 법률안에 대한 부대의견으로서 추가로 의견 반영이라는 후속조치를 수행한다는 내용을 의미한다.

이러한 세부 요약문을 통해, 요약문에서는 논쟁적 대화를 종료한 이후에 각 참여자가 수행해야 할 후속조치나 추가과업을 확인할 수 있다.

1.2.2. 요약문 작성의 원칙

□ 요약문 작성의 일반적 원칙

요약의 대상이 되는 중요한 정보는 다음과 같다.

- 대화에서 등장한 문제(Problems/Issues)는 무엇인가?
- 각 화자의 주장이 무엇인가?
- 회의를 통해 도출된 결론(Decisions, Actions)은 무엇인가?

중요한 정보를 요약하기 위하여서는 중요한 정보를 담고 있는 대화에 대하여, 누가, 무엇을, 어떻게, 왜의 4요소를 (되도록) 포함하여 요약문이 작성되어야 한다. 논쟁적 대화의 경우 “언제” 혹은 “어디서”는 회의 시간 및 회의 장소 등으로 명시되고 중복되는 경우가 많으므로 필요한 경우가 아니라면 포함하지 않는다.

요약문 작성의 세부 원칙은 다음과 같다.

	논쟁적 대화 지침	국회 회의록 요약 을 위한 세부지침
요약문에 발화 자 정보 포함 여 부	<ul style="list-style-type: none"> - 요약의 대상이 되는 중요한 정보를 누가 발언하였는가를 포함한다. - 여럿이 같은 의견을 낸 경우에는 발화 그룹을 명시한다. - 회의를 통해 결정되거나 동의된 사항은 발언자를 생략할 수 있다. (예: 본 회의를 통하여 A가 결정되었다) 	<ul style="list-style-type: none"> - 국회 회의록의 경우, 정부 측 의견에서 실무자의 의견과 정부의 의견을 구분하여 명시한다.
문장 형식	<ul style="list-style-type: none"> - 하나의 이슈는 하나의 문장으로 작성한다. - 한 명의 화자에게서 여러 이슈가 등장한 경우에는 각 이슈는 하나의 문장으로 작성한다. 	
대화의 주제(안 건)	<ul style="list-style-type: none"> - 논쟁적 대화에서 주제(안건)가 명시되었다면, 이러한 주제는 반드시 요약문에 포함되어야 한다. 	<ul style="list-style-type: none"> - 국회 회의록에서 소위원장(사회자)이 언급하는 안건 - 안건에 대한 동의/비동의 여부는 요약문에 반드시 작성

□ 문제/이슈(Problems/Issues) 요약문 작성

문제/이슈에서는 논쟁적 대화에서 제기된 문제들을 기술한다. 각각의 문제에 대해서는 가능한 한 하나의 문장으로 작성하며, 한 명의 화자에게서 복수의 문제가 등장한 경우에도 각각의 문제를 하나의 문장으로 작성한다.

문제/이슈 요약문은 다음의 세 가지 요소가 포함되도록 한다.

- 누가 문제를 제기하였는가? (다수 화자에 의한 문제/이슈면 생략 가능)
- 어떤 문제가 이야기되었는가?
- 근거가 드러난 경우, 해당 근거를 문장에 포함한다.

문제/이슈 요약문의 주요 형태는 다음의 유형에 해당한다.

- 요청사항, 문제 제기, 지적, 제안
- 사전 결정사항에 대해 동의하여 현장 결정한 경우
- 사전 결정사항에 대해 비동의하여 논쟁 후 결정한 경우 (혹은 다음에 결정)

<p>국회 회의록 요약을 위한 세부지침</p> <p>다음의 내용이 문제/이슈에 해당한다.</p> <ul style="list-style-type: none">- 우선순위 1: 안전(논의 대상)- 우선순위 2: 안전과 더불어 제기되는 쟁점/문제/지적사항 등- 우선순위 3: 해당 회의를 거쳐야만 하는 결정사항, 변화, 문제, 이슈 등 <p>전문위원발표와 같이 한 명의 화자에 의해 법안의 내용을 길게 설명하는 경우에는 다음의 방식으로 문제/이슈를 작성한다.</p> <ul style="list-style-type: none">- 뒤에 이어질 주제에 맞춰 쟁점별 1개 문장을 작성한다.- 문장이 너무 길어지면 세부 주제에 맞게 추가로 분할한다.- 법안심사의 경우 “조(항)”을 기준으로 1개 문장 작성도 고려한다.- 전문위원 발표와 같이 독립된 발화의 경우에는 아직 회의를 통해 결정된 사항이 없으므로 Decisions, Actions를 작성하지 않는다. <p>발표와 달리 논쟁인 경우 다음의 내용이 문제/이슈 요약의 대상이 된다.</p> <ul style="list-style-type: none">- 안전에 대한 의견 (지적, 제안, 요청)- 주요 안전과 별개인 의사진행발언은 요약의 대상이 되지 않는다.- 화자가 의견을 내기 위한 사전 질문(지적)의 경우에는 포함하지 않거나 뒤의 내용과 함께 요약의 대상이 된다.

□ 결정사항(Decisions) 요약문 작성

결정사항은 논쟁적 대화에서 결론이 도출된 경우에 이를 요약문으로 작성한다.

결정사항에 대한 요약문은 하나의 결정사항에 대해 하나의 문장으로 작성하도록 하며, 다음의 네 가지 요소가 포함되도록 한다.

- 누가 결정하였는가? (다수 화자에 의한 결정이면 생략 가능)
- 어떤 결정이 이루어졌는가?
- 근거가 드러난 경우, 해당 근거를 문장에 포함한다.
- 제안이나 요구사항이 전제된 쟁점이라면, 그 주체와 내용도 포함하여 작성한다.

결정사항 요약문의 주요 형태는 다음과 같다.

- 결정사항
- 합의된 사항
- 안전에 대한 동의
- 제안(요구사항) 등에 대한 동의

결정사항 요약문은 후속조치와 다음과 같은 차이점이 있다.

- 논쟁적 대화를 통해 결정된 사항은 결정사항에 해당한다. 이는 결론이 난 사항으로서 추가 논의가 없는 상황(문제가 해결된 상황)에 해당한다.
- 논쟁적 대화를 통해 진행되기로 된 과업사항은 후속조치에 해당한다. 결론이 나지 않아 추가 논의가 필요한 경우, 추가 제출, 재논의, 재검토 등이 여기(문제가 미해결된 상황)에 해당한다.

□ 후속조치(Actions) 요약문 작성

논쟁적 대화에서 결정사항은 대화를 통해 결정된 사항을 의미하는데, 이와 다르게 향후 조치가 결정된 사항이 있다. 즉 회의 참여자 혹은 관계자가 수행해야 할 과업이며 이는 후속조치 요약문에 해당한다.

후속조치 요약문은 하나의 결정 사항에 대해 하나의 문장으로 작성하도록 하며, 다음의 네 가지 요소가 포함되도록 한다.

- 누구에 의해서 과업이 지시되었는가?
- 어떤 과업인가?
- 누가 수행해야 하는 과업인가?
- 근거가 드러난 경우, 해당 근거를 문장에 포함하여야 한다.

후속조치의 주요 형태는 다음과 같다.

- 추가 논의사항
- 재논의 및 재검토가 필요한 사항

□ 쟁점별 요약문(Abstract)

전체 요약은, 쟁점별로 요약된 세부 요약문을 하나로 합치는 작업을 통해 수행된다.

- (1) 모든 내용을 논쟁적 대화의 시간 순서에 따라 취합한다.
- (2) 세부 요약문이 중복되는 내용이 있는 경우, 상위 개념으로 기술하여 통합하거나 불필요한 요약문은 제거한다.
- (3) 통합된 요약문을 다시 자연스러운 문장으로 작성한다.

요약문은 논쟁적 대화의 목적에 따라 간결하게 설명한 것으로, 대화에서 이루어진 주요 쟁점, 주요 논의점, 결정사항 등을 포함한다.

1.2.3. 요약 분석의 절차

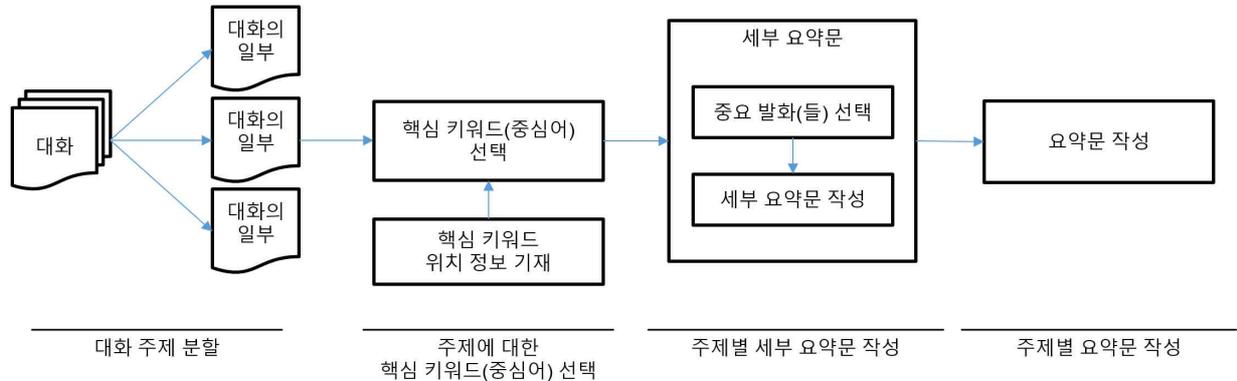


그림 5 요약 분석 절차 개념도

본 장에서는 요약 분석의 절차를 기술한다.

논쟁적 대화는 하나의 쟁점에 관한 대화로 이루어져 있거나, 혹은 하나의 회의에서 여러 개의 쟁점에 관한 대화들로 이루어져 있다. 1.1.2장에서 논의된 바와 같이, 쟁점에 대한 중요한 정보를 집중적으로 요약하기 위하여 각 대화들을 주제에 따라 나누는 절차를 거치고, 하나의 쟁점에 관한 대화를 수집하는 과정이 필요하다.

이후, 해당 대화에 대해서 다음의 절차가 이루어진다.

(1) 중심어(키워드, keyword) 선택

중심어란, 요약의 대상이 되는 대화를 대표할 수 있는 단어 혹은 명사구를 의미한다. 예를 들어 다음 대화에서 핵심어는 ‘평생교육법 일부개정법률안’이다.

○소위원장 조승래 좋습니다.
 그러면 이것도 의결을 잠시 보류하고요.
 의사일정 제3항을 심사하도록 하겠습니다.
 안민석 의원이 대표발의한 평생교육법 일부개정법률안을 심사하겠습니다.
 전문위원, 보고해 주시기 바랍니다.

○전문위원 조기열 심사 자료 2쪽을 참고해 주시기 바랍니다.
 개정안의 내용은 학점은행기관 형태의 평생교육시설에 대한 법적 근거를 마련해서 평생교육법 16조나 17조에 따른 국가나 지자체의 지원을 받을 수 있도록 하려는 내용입니다.
 ...

○전문위원 조기열 예, 평생교육시설이 좀 복잡하게 운영되고 있어서요, 운영되는 실태를 정확히 잘 몰라서 집행 부처의 의견을 좀 들으시면 좋을 것 같습니다.

367교육(법안심사)소위02(19.3.14.) 中

위 대화에서, 소위원장 조승래가 현재 대화의 쟁점인 ‘의사일정 제3항’에 관한 발화를 한다. 그리고 이후 해당 안전에 관한 대화가 이루어졌다.

안민석 의원이 대표발의한 평생교육법 일부개정법률안을 심사하겠습니다.

해당 발화를 통해 현재의 대화가 쟁점 “평생교육법 일부개정법률안”에 대해 이루어 짐을 알 수 있다. 이 경우, 현재의 대화에 대한 중심어로 “평생교육법 일부개정법률안”을 선택한다. 이때, 중심어는 어떤 발화에서 등장하였는지, 발화 내에서 어떠한 위치에 있는지를 분석 말뭉치에 기재하여야 한다. 이에 대한 데이터 형식의 예는 다음과 같다.

```
"issue": {  
  "keyword": "평생교육법 일부개정법률안",  
  "sentence_id": 50,  
  "begin": 14,  
  "end": 26  
},
```

이때 sentence_id는 중심어가 등장한 발화에 대한 고유 아이디 값, begin 및 end는 핵심 중심어의 해당 발화에서의 위치 정보(시작점, 끝점)를 의미한다. 이때 begin, end는 글자 수를 기준으로 계산되며 시작은 0이다.

(2) 세부 요약문 작성

세부 요약문 작성은 크게 두 가지의 절차를 거친다. 먼저, 해당 중심어에 관련된 대화 중 중요 발화를 선택하는 절차와 다음 해당 발화를 바탕으로 세부 요약문(들)을 작성하는 절차이다.

```
○소위원장 조승래 좋습니다.  
  그러면 이것도 의결을 잠시 보류하고요.  
  의사일정 제3항을 심사하도록 하겠습니다.  
  안민석 의원이 대표발의한 평생교육법 일부개정법률안을 심사하겠습니다.  
  전문위원, 보고해 주시기 바랍니다.  
○전문위원 조기열 심사 자료 2쪽을 참고해 주시기 바랍니다.  
...  
  검토의견으로서는, 현행법 제31조부터 제33조까지 유사한 형태들이 나열되고 있습니다. 학교형태
```

의 평생교육시설이라든지 사내대학형태, 원격대학형태의 평생교육시설이 나열되고 있는데요. 그쪽을 살펴보면 평생교육시설의 설치 주체라든지 등록 절차, 교육 대상, 교육비 부담 주체 등 상세한 규정이 들어 있습니다. 그런데 개정안은 그런 세부적인 규정이 없어서 현행법 체계와 좀 맞지 않는 측면이 있어서 수정·보완이 필요하다고 생각합니다.

...

○소위원장 조승래 차관, 정부 의견 말씀해 주십시오.

○교육부차관 박백범 전문위원님이 수정하자는 의견에 동의하면서, 저희들 대안으로는 제29조에 보면 ‘학교의 평생교육’이라는 게 있는데 지금 여기에는 안 나오지만, 그래서 제29조의2(학점은행기관의 평생교육) 해 가지고서 일부 문구 수정을 통해서 그 취지를 반영하고자 이렇게 수정 제안합니다.

...

○소위원장 조승래 전문위원은 어떻게 생각하십니까, 그 수정의견에 대해서?

○전문위원 조기열 사실 정확히 잘 모르겠습니다.

...

○소위원장 조승래 그러면 위원님들 혹시 이것 관련해서 의견 있으면 좀 말씀을 주시고요.

말씀이 없으시면 어차피 다음번 법안소위 또 하게 되니까요, 그전까지 발의하신 의원님과 전문위원 그리고 정부는 의견을 조율해서 대안을 좀 만들어서 주시면 그걸 우리가 존중해서, 위원님들 의견을 들어서 의결하는 걸로 하면 어떨까 싶습니다. 그렇게 하시지요?

(「예」 하는 위원 있음)

이것은 계속 심사하도록 하겠습니다.

...

367교육(법안심사)소위02(19.3.14.) 중

위 예시에서, 중심어 “평생교육법 일부개정법률안”에 관련된 대화가 이루어지는데, 이 중 세부 요약의 대상이 되는 문제/이슈, 결정사항, 후속조치에 관련된 발화를 모두 선택한다. 위 예시 중 음영에 해당한다.

세부 요약문은 이렇게 선택된 발화를 바탕으로 요약문을 작성한다. 예를 들어, 다음 발화의 경우 전문위원의 검토 의견에 대해서 전반적으로 동의하며, 일부 수정 의견을 제시하고 있다. 따라서 해당 발화는 해당 핵심 키워드(중심어)에 대한 문제/이슈에 해당하는 발화로, 다음과 같은 요약이 가능하다.

전문위원님이 수정하자는 의견에 동의하면서, 저희들 대안으로는 제29조에 보면 ‘학교의 평생교육’이라는 게 있는데 지금 여기에는 안 나오지만, 그래서 제29조의2(학점은행기관의 평생교육) 해 가지고서 일부 문구 수정을 통해서 그 취지를 반영하고자 이렇게 수정 제안합니다.

세부 요약문의 유형	세부 요약문
문제/이슈 (Problems/Issues)	교육부차관 박백범은 평생교육법 일부개정법률안 전문위원 검토의견을 수용하였으나 제29조의 2항 일부 문구 수정을 통해 반영하는 것을 제안했다.
367교육(법안심사)소위02(19.3.14.) 中	

또한, 대화의 마지막에서 소위원장 조승래의 발화를 통해 해당 쟁점인 “평생교육법 일부개정법률안”에 대해 정부가 의견을 조율하여 대안을 만들고 이에 대해 추후 의결하는 것으로 결정되었음을 알 수 있다.

... 그전까지 발의하신 의원님과 전문위원 그리고 정부는 의견을 조율해서 대안을 좀 만들어서 주시면 그걸 우리가 존중해서, 위원님들 의견을 들어서 의결하는 걸로 하면 어떨까 싶습니다. 그렇게 하시지요?

(「예」 하는 위원 있음)

이것은 계속 심사하도록 하겠습니다.

따라서 해당 발화는 해당 핵심 키워드(중심어)에 대한 결정사항에 해당하는 발화라고 볼 수 있으며 아래와 같이 요약 가능하다.

세부 요약문의 유형	세부 요약문
결정사항 (Decisions)	평생교육법 일부개정법률안은 개정안을 발의한 의원, 전문위원, 정부가 의견을 조율하여 대안을 완성하고 완성 시 의결하기로 했다.
367교육(법안심사)소위02(19.3.14.) 中	

각 안건에 대해 세부 요약문이 작성되며, 1.2.1장에서 논의된 요약문 작성의 원칙에 따라 각각의 쟁점, 결정사항 등은 독립적인 세부 요약문으로 작성되어 복수개의 세부 요약문들이 구축된다.

(3) 주제별 요약문 작성

상기의 과정을 거쳐 (1) 해당 주제에 대한 핵심 키워드(중심어)가 선정되었고, (2) 해당 주제에 대한 세부 요약문이 작성되었다. 다음 단계에서는 중심어와 세부 요약문을 토대로 해당 주제에 대한 통합된 요약문(Abstract)을 작성하는 과정이 이루어진다.

먼저 해당 안건인 “평생교육법 일부개정법률안”이 누구에 의해서 이루어졌는가를 명시하고(안민석 의원 발의), 이에 대해 이루어진 논쟁의 일반적 사항을 통합하는 과정을 거친다.

<p>요약문</p> <p>안민석 의원이 발의한 평생교육법 일부개정법률안과 관련하여 전문위원의 검토의견이 제출되었다. 이에 대해 개정안을 발의한 위원, 전문위원 그리고 정부가 의견을 조율하여 대안을 완성한 후에 의결하기로 했다.</p>
<p>367교육(법안심사)소위02(19.3.14.) 中</p>

상기 예시와 같이, 요약문은 쟁점, 그리고 쟁점을 언급한 사람, 쟁점에 대해 이루어진 논의, 쟁점에 대한 결정사항 등으로 구성되어 있다.

요약 분석의 절차를 종합하면 다음과 같다. 먼저, 대화에서 각 쟁점에 해당하는 대화들을 분리한다. 이후 해당 대화에 대해 쟁점을 대표할 수 있는 중심어를 선택하고, 이에 대한 발화 중 중요한 발화를 선택한다. 해당 발화들에 대해 세부 요약문이 작성한 후 이를 통해 논쟁적 대화에서 이루어진 안전과 논의들, 결정사항 등이 모두 포함된 요약문을 작성한다.

(4) 요약문의 참조(reference) 발화들 선택

본 지침 및 사업에서는 추상 요약문을 작성하는 작업과 동시에 논쟁적 대화에서 등장하는 모든 발화 중 어떤 발화를 참조하여 요약문을 작성하였는가를 명시하는 절차를 포함한다. 이는 해당 요약문이 올바르게 요약되었는가를 확인할 수 있는 근거자료를 구축한다는 목적과 함께 논쟁적 대화 요약 인공지능 모델 등을 개발할 때 어떤 발화가 중요한 발화인지를 파악할 수 있는 데이터로 활용될 수 있다는 것에 의의가 있다. 각 요약문에 대해 참조 발화를 기재한 데이터의 형식은 4장에서 확인할 수 있다.

1.2.4. 세부 지침: 국회 회의록 요약 분석 절차

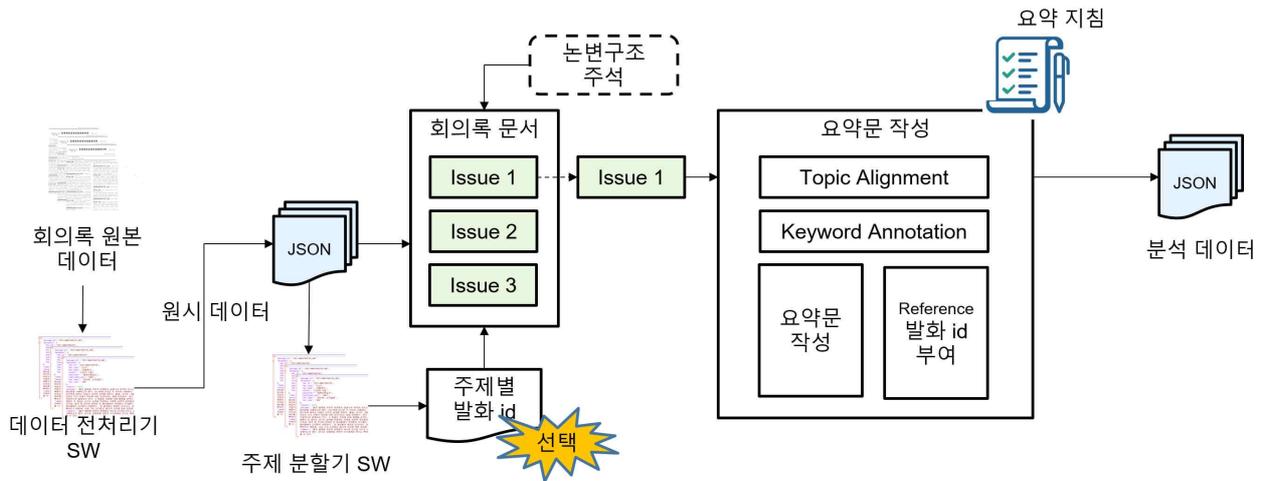


그림 6 국회 회의록 요약 분석 절차 개념도

본 장에서는 논쟁적 대화의 일종인 국회 회의록 요약 분석을 위한 세부 지침을 기술한다. 위 그림은 국회 회의록 요약 분석을 위한 절차이다.

국회 회의록은 논쟁적 대화가 전사된 말뭉치로써, 논쟁적 대화의 대표적 유형들을 살펴볼 수 있다는 점에서 유용도가 높다. 특히, 논쟁적 대화에서의 주요 정보인 안건(회의에서 상정된 안건)에 대한 다양한 논의(전문위원, 정부, 위원 등)가 이루어지며 각 안건에 대한 결정사항(의결/부결 등) 및 후속조치 등이 포함된 전형적인 논쟁적 대화를 담은 말뭉치이다.

본 사업에서 분석 대상으로 삼은 국회 회의록은 약 500개의 발화로 구성된 20쪽가량의 긴 대화록이며, 세로단 두 단의 한글 파일(HWP)로 작성되어 있다. 따라서 이에 대한 데이터 전처리가 필요하며, 긴 대화인 만큼 쟁점에 따라 대화를 자동으로 분할하는 소프트웨어를 사용하여 각 회의록 문서를 주제에 따라 분할하는 과정을 거친다. 국회 회의록에서는 쟁점 혹은 안건이라는 용어가 ‘국회 회의에서 상정된 안건’과 용어가 혼용될 수 있으니, 대화에서 논쟁의 대상은 쟁점(issue)이라는 용어를 사용하였고 국회의 “상정된 안건”은 안건이라는 용어를 사용하였다. 각 쟁점에 해당하는 대화를 분리한 뒤, 이에 대해 2.3장과 동일한 과정을 거쳐 요약 분석을 수행한다.

추가로, 국회 회의록의 경우에는 회의록 문서의 메타데이터로서 회의에 ‘상정된 안건’이 명시되어 있으므로, 중심어 작성에 앞서 해당 쟁점이 회의의 ‘상정된 안건’ 중 어느 안건에 해당하는가를 명시하는 절차가 추가되었다.

2. 논쟁적 대화 요약 예시 및 세부 분석

2.1. 중심어 선택

중심어란 대화의 주제를 나타내는 단어 혹은 구를 의미하며, 본 지침에서는 실제 발화에서 등장한 어휘로 한정한다. 이러한 중심어는 논쟁적 대화에서 논의의 대상이 되는 사안 혹은 안건으로서 해당 중심어에 대해 논의가 이루어진 대화가 요약의 대상이 된다. 실제 작업에서는 전체 대화 중 논의의 대상이 하나로 정해질 수 있는 부분 대화를 요약의 대상으로 하므로 중심어를 선택하는 작업은 중요하다.

□ 논쟁적 대화의 발화에서 중심어가 제기되는 경우

논쟁적 대화가 진행되면서 특정 화자에 의해서 대화의 논의 대상 혹은 안건이 제기되는 경우가 있다.

아래의 예시는 논쟁적 대화 중 발표 혹은 보고 과정에서 안건을 제기하는 경우이다.

...

이상 보고를 마치겠습니다.

○소위원장 이학재 정부 측 의견은 따로 들으실 필요 없을 것 같고 위원님들도 특별한 의견 없으시지요?

(「예」 하는 위원 있음)

다음 2항 보고하여 주시기 바랍니다.

○김상희 위원 결론은……

○소위원장 이학재 의결은 나중에……

○수석전문위원 임진대 입법 청원은……

○소위원장 이학재 입법 청원은 나중에 모아서 따로 보고를 해 주시고요.

○수석전문위원 임진대 예, 그렇게 하겠습니다.

○소위원장 이학재 4항 보고해 주시기 바랍니다.

○전문위원 박용수 4항 보고드리겠습니다.

사이버 따돌림 예방과 피해자 지원 관련 법률의 보완·강화에 관한 청원입니다.

청원인은 강호진입니다. 이인영·유은혜·인재근 의원님 소개로 작년 9월 달에 제출되었습니다.

주요 내용은 사이버 따돌림으로 인해 학생이 자살에까지 이르는 일이 빈번히 발생하는 등 사이버상에서의 집단적인 욕설, 따돌림, 언어폭력 등이 심각한 수준에 있으므로 실정법상 사

이바상의 각종 폭력에 대한 처벌이 쉽지 않기 때문에 사이버 언어폭력 등에 관한 처벌을 강화하는 입법을 청원하는 것입니다.

320교문(청원심사)소위01(13.11.18) 中

위 예시에서는 회의가 진행됨에 따라 다음 안건이 전문위원 박용수에 의해 제기되었다.

...

○전문위원 박용수 4항 보고드리겠습니다.

사이버 따돌림 예방과 피해자 지원 관련 법률의 보완.강화에 관한 청원입니다.

청원인은 강호진입니다. 이인영.유은혜.인재근 의원님 소개로 작년 9월 달에 제출되었습니다.

주요 내용은 사이버 따돌림으로 인해 학생이 자살에까지 이르는 일이 빈번히 발생하는 등 사이버상에서의 집단적인 욕설, 따돌림, 언어폭력 등이 심각한 수준에 있으므로 실정법상 사이버상의 각종 폭력에 대한 처벌이 쉽지 않기 때문에 사이버 언어폭력 등에 관한 처벌을 강화하는 입법을 청원하는 것입니다.

...

○교육부차관 나승일 없습니다.

○소위원장 이학재 위원님들도 다른 의견 없으시지요?

○김상희 위원 예, 없습니다.

○현영희 위원 잠깐만요. 사이버 따돌림 예방에 있어서 이미 우리가 학교폭력예방 및 대책에 관한 법률에 살인, 강간 등 강력범죄를 학생이 저질렀을 때에도 처벌규정을 따로 마련하고 있지 않고 있는데 지금 이것은 현재 학생들이 강력범죄를 범했을 경우에 형법에 따라서 처벌하고 있기 때문에 이것을 더 한 번 한다는 것은 좀 문제가 있지 않은가 이 생각이 드는데요.

○소위원장 이학재 그래서 이것은 정보통신망 이용촉진 및 정보보호 등에 관한 법률과 관련된 내용이기 때문에 미래창조과학방송통신위원회로 회부하기로 그렇게 하자는 의견이시니까요.

○현영희 위원 예, 거기로 보내는 게 좋을 것 같습니다.

○김상희 위원 미방위로 보내고, 법안은 또 거기서 심의하면 되니까요.

320교문(청원심사)소위01(13.11.18) 中

또한 전문을 보면 해당 키워드에 대해 추가적인 논의가 이루어지는 것을 볼 수 있다. 이에 따라 ‘사이버 따돌림 예방과 피해자 지원 관련 법률’은 안건으로 상정된 점, 그리고 추가적인 논의가 이루어진다는 점에서 중심어의 자격을 갖추었다고 볼 수 있다. 즉, 본 과제에 중심어는 논쟁적 대화에서 특정 화자에 의해 제기된 안건이면서 동시에 추가적인 논의가 이루어진 경우에만 논의 대상의 자격을 가질 수 있다.

따라서, 전문위원 박용수에 의해 제기된 안건 “사이버 따돌림 예방과 피해자 지원 관

련 법률의 보완·강화에 관한 청원”은 현영희 위원에 의해 논의가 진행되면서 중심어로서의 자격을 얻는다.

위 예시와 반대로, 특정 화자에 의해 논의 대상이 언급되었으나, 이와 관련한 논의가 전혀 이루어지지 않은 때도 있다.

...

○수석전문위원 임진대 예, 뒷부분은 보고를 생략하고요.

...

다음은 취업 후 학자금 상환 특별법 개정법률안에 관한 청원입니다.

청원인은 안진걸이라는 분인데 반값등록금 실현과 교육공공성 강화를 위한 국민본부 공동 집행위원장이라는 분이 되겠습니다. 정진후 의원님께서 소개하셨고 금년 8월 21일에 우리 위원회에 회부되어 왔습니다.

청원의 이유는 청년세대의 실업률과 비정규직이 증가하는 현실을 감안할 때 소득분위와 성적에 상관없이 취업 후에 학자금을 상환할 수 있도록 할 필요가 있다는 이유에서 취업 후 상환 학자금대출 대상에 대학원생을 포함시키고 취업 후 상환 학자금대출 자격 요건 중 가구 소득분위 및 다자녀가구 해당여부 부분 그리고 성적 석차 부분을 삭제하는 내용이 되겠습니다.

이 내용과 유사한 법률개정안이 있습니다. 유은혜 의원님께서 제출한 개정안과 유사한 내용이 되겠습니다.

...

320교문(청원심사)소위01(13.11.18) 中

위 예시에서 수석전문위원 임진대에 의해 언급된 안건인 “취업 후 학자금 상환 특별법 개정법률안에 관한 청원”에 대해서는 전문위원의 검토 의견이 이루어지지만, 이에 관련한 다른 대화 참가자의 질문이나 의견, 입장이 드러나지 않고 또한 이와 관련한 결정사항도 대화에서 이루어지지 않았다. 이럴 때, 해당 안건은 중심어로서의 자격이 없으며 또한 논쟁적 대화 요약의 대상이 되지 않는다.

□ 논쟁적 대화의 진행자에 의해 안건이 명시되는 경우

발화자에 의해서 안건을 명시하는 경우 이외에도, 사회자가 회의의 순서를 따라 안건을 제시하는 때도 존재한다. 예를 들어 회사의 회의에서도 사회자가 안건을 언급하는 경우나, 국회 회의록에서 위원장이 안건을 언급하는 경우 등이 있다. 특히 이러한 경우 회의의 안건이 정해진 상태에서, 해당 안건을 사회자가 의사진행발언을 통해 언급하는 경우이다.

의사진행발언에 의해 중심어가 언급되는 예시는 아래와 같다.

...
심사된 안건
4. 자원의 절약과 재활용촉진에 관한 법률 일부개정법률안(최봉홍 의원 대표발의)
1. 먹는물관리법 일부개정법률안(이완영 의원 대표발의)
2. 물의 재이용 촉진 및 지원에 관한 법률 일부개정법률안(이낙연 의원 대표발의)
3. 환경범죄 등의 단속 및 가중처벌에 관한 법률 일부개정법률안(정부 제출)
...
4. 자원의 절약과 재활용촉진에 관한 법률 일부개정법률안(최봉홍 의원 대표발의) (14시13분)
○소위원장 김성태 그래서 오늘 안건 4번, 자원의 절약과 재활용촉진에 관한 법률 일부개정법률안을 순서를 일부 변경해서 의사일정 제4항을 먼저 상정하여 심사하도록 하겠습니다. 의사일정 제4항에 대해서 수석전문위원 설명해 주시기 바랍니다.
○수석전문위원 천병호 수석전문위원입니다. 심사자료 77쪽을 봐 주시기 바랍니다. 법안의 주요 내용을 말씀드리면, 첫째로, 지방자치단체가 폐자원 회수에 필요한 재원을 확보하고 효율적으로 운용·관리하기 위하여 재활용가능자원관리특별회계를 설치할 수 있도록 하고 있습니다.
...
320교문(청원심사)소위01(13.11.18) 中

위 예시에서는, 회의의 안건으로 “4. 자원의 절약과 재활용촉진에 관한 법률 일부개정법률안(최봉홍 의원 대표발의)”가 명시되어 있으며, 이에 대한 의사진행발언으로서 소위원장 김성태가 안건 4번에 대한 의사진행발언을 하며 “자원의 절약과 재활용촉진에 관한 법률 일부개정법률안”을 언급한다.

이는 논쟁적 대화의 범위 안에서 중심어를 명시한다는 원칙을 따른 것이다. 즉 발화로 언급된 중심어와 이와 관련된 발화만이 요약 대상이며, 비록 진행자에 의해 언급된 안건이라도, 실제 발화 안에 중심어가 존재하고 추가 논의가 이루어진다는 점에서 중심어로 인정할 수 있다. 이와는 반대로 일부 회의록에서는 발화로 이루어지지 않은 정보인 메타데이터에 회의 안건이 존재하고 실제 발화에는 존재하지 않을 때가 있는데, 이런 경우는 메타데이터 내 안건을 중심어로 인정하지 않는다.

위 예시에서는 중심어 “자원의 절약과 재활용촉진에 관한 법률 일부개정법률안”에 대해서 수석전문위원이 내용을 소개하고, 이에 관련하여 여러 의견이 오가고 이에 대한 결정사항 등이 논의되었다.

상기의 예시를 통해, 중심어는 다음의 조건에 의해 결정될 수 있다.

- 논쟁적 대화에서 논의의 대상이 되는 안건 혹은 쟁점
- 회의와 같은 논쟁적 대화에서 미리 주어진 안건
- 발화자의 발화 및 진행자의 의사진행발언에 의해 명시된 키워드

다음의 경우에는 중심어로 간주하지 않음.

- 논쟁적 대화에서 발화 이외의 외부 정보(속기사의 메모 등)
- 실제로 논의가 이루어지지 않은 안건 혹은 쟁점

▶ 세부지침: 국회 회의록에서의 중심어 선정

국회 회의록은 형식상 어떤 안건에 대한 논의(심사 등)가 이루어질 것인지를 미리 정하고 있다. 아래 예시에서 의사 일정은 회의가 시작하기 전에 계획된 안건을 의미하며, 심사된 안건은 실제 회의를 통해 심사가 진행된 안건을 의미한다. 중심어는 해당 심사된 안건이 발화자 혹은 소위원장에 의해 언급될 때 등장하게 되며, 이 중 일부에 대해서 논의가 이루어진다.

따라서 국회 회의록에서 중심어를 선정할 경우에는 국회 회의록의 1쪽에 기재된 “심사된 안건”이 논의의 대상이므로 이 안건이 발화자에 의해 언급된 경우, 이를 중심어로 선정하고 분석 말뭉치에서는 중심어와 해당 안건을 매칭하여 기록한다.

2.2. 세부 요약문: 문제/이슈(Problems/Issues)

문제/이슈는 논쟁적 대화에서 나온 쟁점, 문제, 이슈들에 대한 세부 요약문으로, 각 이슈는 가능한 하나의 문장으로 작성되며 한 명의 화자에게서 복수 개의 이슈가 등장한 경우에도 각각의 이슈를 하나씩의 문장으로 작성한다.

문제/이슈 요약문의 주요 형태는, 논쟁적 대화에서 나타나는 요청사항, 문제제기(지적), 제안 등에 해당한다.

...

1. 먹는물관리법 일부개정법률안(이완영 의원 대표발의)

(14시49분)

○소위원장 김성태 다음은 의사일정 제1항 이완영 의원이 대표발의한 먹는물관리법 일부개정법률안을 상정합니다.

○홍영표 위원 말씀 전에 회의진행사항 발언 하나만 말씀드리면, 먹는물관리법 일부개정법률안은 여기 지금 샘물 표지제 위반자에 대한 처벌조항 같은 것들이 지금 안 돼 있습니다.

그리고 또 허가제를 신고제로 안 하는 것에 저희가 이제 반대하고 있는데, 그래서 지금 사실 법안을 저희가 발의하려고 해서 이것도 다음에 병합해서, 왜냐하면 오늘 하게 되면 이걸 개정안을 또 만들어야 되거든요.

그래서 이 법안이 곧 나오니까 그것을 같이 다음에 제출해서 병합해서 해 주시면 어떻습니까?

○심상정 위원 그러세요. 이 안은 오늘 불가입니다, 저의 의견은.

○소위원장 김성태 수석전문위원 설명해 주시고, 정부 측 의견까지 듣고 판단하도록 하겠습니다.

수석전문위원 설명해 주십시오.

○수석전문위원 천병호 1쪽을 봐 주시기 바랍니다.

이완영 의원안의 주요 내용은,

첫째로, 다중이용시설에 정수기를 설치하는 경우에는 냉.온수기와 마찬가지로 오염 가능성이 높은 장소에 설치를 금지하고, 정수기를 주기적으로 청소, 소독하는 등 위생적 관리를 하도록 하고,

...

다음, 26쪽입니다.

현행법에서 '폐기처분 등'과 관련해서 '먹는샘물등, 수처리제 또는' 그렇게 되어 있습니다
마는 이 내용은 처음에 입법 불비에 의해서 정수기 등이 빠졌습니다. 그래서 '수처리제, 정수
기 또는 그 용기와 포장'으로 보완한 내용이 되겠습니다.

이상입니다.

...

○소위원장 김성태 그것은 의견을 제시해 주십시오.

○심상정 위원 기본적으로 사 먹는 물이 늘어나면 여러 가지 문제가 있는 거지요. 아까 원
칙적인 문제를 환원하신다고 했는데 이제 그것은 말도 안 되는 얘기고요, 제가 보기에.

그다음에 기본적으로 먹는물 판매가 확대되는 것은 바람직하지 않다는 게 저의 견해예요.
지하수하고 지표수 고갈 문제도 있고 또 하나는 페트병 환경오염 문제가 굉장히 심각하잖아
요. 그래서 이것은 근원적인, 기초 차이가 좀 있다 이렇게 생각이 됩니다.

그런데 민주당에서 어떤 안을 지금 준비하시는지 모르겠는데 이 안 자체에 대해서 가장
핵심적인 것은 원위치를 시켰지만 기본적으로 물 판매 확대와 관련해서 저는 확대하는 방
향으로의 어떤 개정은 반대한다는 말씀을 드립니다.

○소위원장 김성태 정부, 이 법의 시급성은 어떤 정도입니까?

○환경부차관 윤종수 이번 회기 안에 또 수정안이, 대안이 나온다고 하시니까 그러면……
그 내용을 잘 모르지만 저희들이 보기에 지금 여러 가지 지적을 해 주셔 가지고 그런 문제
에 대해서 다 수정했다고 생각하고 있습니다마는 흥 의원님실에서 또 나오면 병합해서, 다만
이번 회기를 지나가면 이게 또 너무 문제가 되기 때문에……

○심상정 위원 뭐가 문제가 돼요? 제가 보기에 정부는 오히려 수돗물 공급의 신뢰를 강
화시키는 데 신경을 써야지, 왜 생수시장 확대에 신경을 씩습니까?

○환경부차관 윤종수 그래서 병합해서 같이 처리하도록 하겠습니다.

○심상정 위원 제가 묻고 싶은 것이 그것이에요. 정부가 신경 써야 되는 것은 수돗물 공급
의 신뢰를 어떻게 강화할 거냐를 신경 써야지, 왜 생수시장 확대에 신경을 쓰냐 이 말입니다.

○소위원장 김성태 위원님들, 다른 의견 없습니까?

그러면 의사일정 제1항에 대해서도 보다 심도 있는 심사를 위해서 우리 소위원회에서 계
속 심사하도록 하겠습니다. 이의 없습니까?

(「예」 하는 위원 있음)

이의 없으시면 소위원회에서 계속 심사하도록 하겠습니다.

...

위 예시에서, 중심어 “먹는물관리법 일부개정법률안”에 대한 논의가 이루어졌음을 알 수 있는데, 먼저 수석전문위원 천병호가 해당 법률안에 관한 내용 전반을 소개하였다. 특히, 주요 내용과 함께 이에 대한 전문위원의 보완사항들을 중심으로 소개하였다. 해당 발화는 논쟁적 대화에서 특정 화자에 의한 문제 제기 및 제안의 내용을 포함하고 있으므로 요약의 대상이 된다.

해당 발화들은 전문위원이 중심어 “먹는물관리법 일부개정법률안”에 대해 가진 견해를 모두 포함하고 있으며 이에 대한 세부 요약문은 다음과 같다.

세부 요약문의 유형	세부 요약문
문제/이슈 (Problems/Issues)	전문위원은 먹는물관리법 일부개정법률안에 대해 주요 골자 및 수정 의견과 보완해야 할 사항들을 개괄하였다.
320교문(청월심사)소위01(13.11.18) 中	

문제/이슈 세부 요약문을 작성할 때에는 다음의 정보가 포함되어야 한다:

- 누가 문제를 제기하였는가 (다수 화자에 의한 문제/이슈인 경우 생략 가능)
- 어떤 문제가 이야기되었는가?
- 근거가 드러난 경우, 해당 근거를 문장에 포함한다.

이에 요약문은 전문위원(누가)이 먹는물관리법 일부개정법률안에 대한 내용 및 수정 의견, 보완사항 등(문제)을 발언하였음을 기준으로 작성된다.

또한 같은 예시에서, 전문위원의 의견에 대해 심상정 위원이 문제 제기를 했다. 심상정 위원은 해당 안을 심사하여 통과하는 것에 대해 반대입장을 취하고 있으며, 이에 대한 근거로서 여러 가지 문제점이 있음을 지적하였다.

○심상정 위원 기본적으로 사 먹는 물이 늘어나면 여러 가지 문제가 있는 거지요. 아까 원칙적인 문제를 환원하신다고 했는데 이제 그것은 말도 안 되는 얘기고요, 제가 보기에.

그다음에 기본적으로 먹는물 판매가 확대되는 것은 바람직하지 않다는 게 저의 견해예요. 지하수하고 지표수 고갈 문제도 있고 또 하나는 페트병 환경오염 문제가 굉장히 심각하잖아요. 그래서 이것은 근원적인, 기초 차이가 좀 있다 이렇게 생각이 됩니다.

그런데 민주당에서 어떤 안을 지금 준비하시는지 모르겠는데 이 안 자체에 대해서 가장 핵심적인 것은 원위치를 시켰지만 기본적으로 물 판매 확대와 관련해서 저는 확대하는 방향으로의 어떤 개정은 반대한다는 말씀을 드립니다.

위 발화에서 심상정 의원은 해당 법안에 대해서 반대한다는 견해를 명확히 표현하였고, 반대하는 근거로서 다양한 문제가 발생할 수 있음을 제시하였다.

세부 요약문의 유형	세부 요약문
문제/이슈 (Problems/Issues)	심상정 의원은 물을 구매하는 비율이 늘어나면 발생할 수 있는 문제가 많아진다는 점에서 해당 법안에 대해 반대하였다.
320교문(청원심사)소위01(13.11.18) 中	

위 요약문의 경우에는 심상정 의원이 해당 법안에 대해 반대하였다는 것과, 이에 대한 근거를 포함하여 요약문이 작성되었다.

또한 본문에서 심상정 의원은 정부가 생수시장 확대에 신경을 쓴다는 것 자체에 반대 입장을 취하고, 오히려 중요시해야 하는 것은 수돗물 공급의 신뢰를 강화하는 것이라고 발언하였다. 이는 위의 법안 반대 이유와는 달리 정부 견해 자체에 부정적인 의견을 표한 것으로 위의 내용과 다른 내용이다. 따라서 발화 중 정부에 대한 부정적 의견 표출 부분은 아래와 같이 다른 세부 요약문으로 작성되었다.

이에 대한 요약문은 다음과 같다.

세부 요약문의 유형	세부 요약문
문제/이슈 (Problems/Issues)	심상정 의원은 또한 해당 법안에 대해 정부가 수돗물 공급의 신뢰도를 올리기보다는 생수 시장 확대에 기여하는 것에 반대하였다.
320교문(청원심사)소위01(13.11.18) 中	

▶ 세부 지침: 국회 회의록에서의 문제/이슈

국회 회의록에서 법안에 대한 심사는 정부 측의 의견 및 전문위원의 의견, 그리고 각 위원의 질의응답과 의견제시가 이루어진다. 정부 측 또는 전문위원의 발화는 대체로 발표의 형식을 띠며, 위원들의 의견제시는 논쟁의 형식을 띤다.

전문위원의 발표 같은 경우 다음을 준수하여 요약문을 작성한다.

- 1) 뒤에 이어질 주제에 맞춰, 안전별로 1개의 문장을 작성한다.
- 2) 문장이 너무 길어지는 경우 세부 주제에 맞게 추가 분할을 수행한다.
 - 법안심사의 경우에는 '조(항)'을 기준으로 1개 문장 작성을 고려한다.

논쟁의 경우에는 안전에 대한 의견(지적, 제안, 요청) 등이 문제/이슈 요약문의 대상이 되며, 주요 안전과 별개인 의사진행발언의 경우에는 요약의 대상으로 삼지 않는다.

만약 여러 사람이 유사한 의견을 내면 통합하여 작성하며, 뒤에 진짜 의견을 내기 위한 사전 질문의 경우에는 포함하지 않거나 뒤의 내용과 연계하여 작성한다.

2.3. 세부 요약문: 결정사항(Decisions)

결정사항 세부 요약문은 논쟁적 대화를 통해 결정된 사항에 대한 것으로, 회의를 통해 이루어진 결정사항 및 합의된 사항이나 안전에 대한 동의/비동의 등을 포함한다.

□ 대화를 통해 결정되는 사항

...

○김상희 위원 김병국 씨하고 같은 입장을 갖고 있는 전문가들 의견이 다수 있다든가 그래서 그분들이 와서 좀 얘기를 한다든가 뭐가 그런 게 있어야지 그냥 억울하다고 여기에다 내놓은 것을 ‘그냥 한번 해 보지, 큰 비용이 들지 않는다면’ 이렇게 생각하는 건 좀 그럴 것 같고요.

일단 근거를 좀 주세요, 우리가 판단할 수 있는 자료를. 아무 그게 없으면 좀 곤란하지 않습니까? 보류하는 쪽으로……

○소위원장 이학재 이런 결정에 대해서 전문위원님 어떻게, 다른……

○입법심의회관 이승재 저희들은 저희가 혹시, 이게 법원 판결까지 난 사항이고 하니까 또 지금 김상희 위원님 말씀처럼……

○김상희 위원 뭐 근거가 있어야지요.

○입법심의회관 이승재 사정변경이 있는 것인지 전문가 의견도 청원소위를 통해서 한번 받아 보고 해서 현영희 위원님께도 보고를 드리고 그렇게 하겠습니다.

○소위원장 이학재 예, 보류하는 걸로.

다음은 처음에 미뤄 놔던 법안과 관련된 사항들에 대해서 일괄적으로 수석전문위원님이 설명해 주시기 바랍니다.

...

320교문(청원심사)소위01(13.11.18) 中

위 예시에서는 안전 “홍무18년 고려총통 국보지정 재심요청”에 대해서 보류하는 것이 결정되었다.

세부 요약문의 유형	세부 요약문
결정사항 (Decisions)	홍무18년 고려총통 국보지정 재심요청에 관한 청원은 보류되었다.
320교문(청원심사)소위01(13.11.18) 中	

□ 안전에 대한 동의/비동의로 결정되는 사항

...
○소위원장 김경협 더 이상 의견이 없으면 토론을 마치고 의결하도록 하겠습니다.
의사일정 제6항 아시아산림협력기구 설립에 관한 협정 비준동의안은 정부가 제출한 원안대로 의결하고자 하는데 이의 없으십니까?
(「예」 하는 위원 있음)
이의가 없으므로 가결되었음을 선포합니다.
...

346외통(법안심사)소위02(16.10.28.) 中

논쟁적 대화를 통해 도출된 결정사항에 대해 진행자가 대화 참여자의 동의/비동의 의견을 재확인하는 경우가 있다. 이러한 경우에도 안전에 대한 결정사항이 도출된 것으로 간주한다.

▶ 세부 지침: 국회 회의록에서의 결정사항

...
○소위원장 김성태 그러면 의사일정 제2항에 대해서는 보다 심도 있는 심사를 위하여 우리 소위원회에서 계속 심사하도록 하겠습니다. 이의 없습니까?
(「예」 하는 위원 있음)
...

320교문(청원심사)소위01(13.11.18) 中

위 예시에서는 의사일정 제2항에 대해서는 해당 소위원회에서 계속 심사하는 것으로 결정되었음을 알 수 있다. 국회 회의는 위원들의 동의나 비동의에 대해 「예」 혹은 「아니오」와 같은 대답을 듣는 방식으로 결정하거나 안전에 대한 의결 여부를 심사하는 방식을 취하며, 해당 대답들은 국회 회의록 서기에 의해서 메모와 같은 방식으로 이루어진다. 따라서 국회 회의록의 경우에는 발화자의 발화 이외에도 서기의 메모와 같은 추가적인 정보를 사용하여 요약문을 작성한다.

2.4. 세부 요약문: 후속조치(Actions)

후속조치 세부 요약문은 논쟁적 대화를 통해 도출된 후속조치 혹은 추가과업에 대한 요약문이다.

<p>...</p> <p>○소위원장 이학재 세 번이나 걸쳐서, 서울시문화재위원회까지 해서 문화재관리국, 문화재청 해 갖고 세 번이나 감정을 했던 사항인데 여기에서 결정을 그냥 하는 것보다……</p> <p>○김상희 위원 예, 판단 근거가 없어요.</p> <p>○현영희 위원 다시 보류를 하자 그 말이에요.</p> <p>○김상희 위원 저도 좀 수용할 수도 있겠는데 수용하려고 그러면 뭐 근거가 있어야 되지 않습니까? 그냥 좀 그런 것 같다 해 가지고 우리가 그렇게 하는 건 아닌 것 같고 조금 더 관련해서 보완해서 우리한테 보고를, 자료를 주셨으면 좋겠어요.</p> <p>...</p> <p>○입법심의회관 이승재 저희들은 저희가 혹시, 이게 법원 판결까지 난 사항이고 하니까 또 지금 김상희 위원님 말씀처럼……</p> <p>○김상희 위원 뭐 근거가 있어야지요.</p> <p>○입법심의회관 이승재 사정변경이 있는 것인지 전문가 의견도 청원소위를 통해서 한번 받아 보고 해서 현영희 위원님께도 보고를 드리고 그렇게 하겠습니다.</p> <p>○소위원장 이학재 예, 보류하는 걸로.</p> <p>...</p>	320교문(청원심사)소위01(13.11.18) 中
--	-----------------------------

위 대화에서, 위원들이 해당 사안에 관해서 결정을 내리기에는 근거자료가 부족하다는 것을 지적하였고, 이에 대해 입법심의회관은 전문가 의견을 청원소위를 통해 받아 현영희 위원에게 보고하겠다는 약속을 하였다. 이 대화를 통해 추측할 수 있는 것은 해당 회의가 끝난 이후에 입법심의회관은 후속조치로서 전문가 의견을 현영희 위원에게 제출할 것임을 알 수 있다. 따라서 이에 대한 요약문은 다음과 같이 작성될 수 있다.

세부 요약문의 유형	세부 요약문
후속조치 (Actions)	입법심의회관은 사정변경이 있는지 청원소위를 통해 받아보고 현영희 위원에게도 보고할 것이다.
320교문(청원심사)소위01(13.11.18) 中	

논쟁적 대화에서 향후 조치사항이 결정된 경우, 이것은 회의 참여자 혹은 관계자가 수행해야 할 과업이며 요약문에서는 결정사항에 해당한다. 또한 논쟁적 대화가 종료되는 시점에 후속조치에 대한 결정사항을 마지막에 논의하는 경우가 있다. 회의를 종료하는 시점에 회의에 관한 내용을 정리하는 과정이다.

...

의사일정 제1항, 제5항 및 제6항은 현재 기준으로 내용이 부당하게 되었거나 청원 취지가 달성되었으므로 본회의에 부의하지 않기로 하였습니다.

이의 없으시지요?

(「예」 하는 위원 있음)

가결되었음을 선포합니다.

다만 의사일정 제6항 사회복지법인 특수학교의 학교법인 전환 청원에 대하여 교육부는 사후에 처리 경과를 보고하여 주시기 바랍니다.

의사일정 제11항은 우리 소위원회에서 계속 심사하도록 하겠습니다.

이상으로 오늘 회의를 마치도록 하겠습니다.

위원님 여러분, 교육부 차관님, 문화재청 차장님을 비롯한 관계관 여러분, 수고 많으셨습니다.

산회를 선포합니다.

320교문(청원심사)소위01(13.11.18) 中

위와 같이, 회의의 의사진행발언의 후속조치사항에 대한 언급에서 결정사항 요약문을 작성할 수 있다. 위의 경우에는 결정사항과 후속조치가 모두 포함된 마무리 발언으로서, 후속조치의 경우에는 아래 발화를 통해 드러나 있다.

다만 의사일정 제6항 사회복지법인 특수학교의 학교법인 전환 청원에 대하여 교육부는 사후에 처리 경과를 보고하여 주시기 바랍니다.

위 발화에서처럼 교육부가 사후 처리 경과를 보고해야 함을 언급하고 있다. 이에 대한 요약문은 다음과 같이 작성될 수 있다. 후속조치사항은 “누가”, “무엇을”, “어떻게 (누구에게 제출할 것인가 등)”가 명시되어야 한다. 아래 요약문에서는 사후 처리 경과를 소위원회에 보고할 것임을 명시한 요약문이 작성되었다.

세부 요약문의 유형	세부 요약문
후속조치 (Actions)	의사일정 제6항 사회복지법인 특수학교의 학교법인 전환 청원에 대하여 교육부는 사후에 처리 경과를 소위원회에 보고할 것이다.
320교문(청원심사)소위01(13.11.18) 中	

2.5. 요약문 작성

요약문은 논쟁적 대화의 목적에 따라 간결하게 설명한 것으로, 대화에서 이루어진 주요 쟁점, 분위기, 주요 논의할 점 등을 포함한다.

전체 요약은, 주제별로 요약된 세부 요약문을 하나로 합치는 작업을 통해 수행된다. 요약문은 세부 요약문들을 시간 순서에 따라 취합하고, 세부 요약문이 중복되거나 불필요한 경우에는 통합하거나 제거하는 과정을 거친다. 이후 통합된 요약문을 자연스러운 문장으로 다시 작성(paraphrasing)한다.

예를 들어 세부 요약문이 다음과 같이 작성된 경우를 보자.

세부 요약문의 유형	세부 요약문
문제/이슈 (Problems/Issues)	전문위원은 김영춘 의원안인 도시교통정비 촉진법 일부개정법률안은 심의 절차를 강화하려는 취지로서 바람직하며 다만 교통영향평가서에 대한 심의의 원래 취지에 따라 승인관청 소속 심의를 거치도록 할 필요가 있다는 의견을 제시하였다.
결정사항 (Decisions)	국토교통부는 도시교통정비 촉진법 일부개정법률안에 관한 전문위원 검토의견을 모두 수용했다.
결정사항 (Decisions)	도시교통정비 촉진법 일부개정법률안은 심사 결과를 반영하여 수정한 부분은 수정한 대로, 기타 부분은 원안대로 의결되었다.
352국토(교통법안심사)소위01(17.7.5.) 中	

이 경우, 전문위원은 “도시교통정비 촉진법 일부개정법률안”에 대해서 여러 가지 의견을 제시하였고, 이에 대해서 국토교통부는 전문위원의 의견을 모두 수용하였으며, 위원들은 심사 결과를 반영하여 수정본과 그 외 원안에 대해서 의결이 이루어졌다. 이러한 정보들을 바탕으로, 작업자는 해당 안건에 대한 요약문을 다음과 같이 작성할 수 있다.

요약문
김영춘 의원이 대표발의한 도시교통정비 촉진법 일부개정법률안에 관하여 전문위원은 심의 관련 수정의견을 제시하였고, 정부 측이 이를 모두 수용하여 수정한 부분은 수정한 대로, 기타 부분은 원안대로 의결되었다.
320교문(청원심사)소위01(13.11.18) 中

해당 요약문은 김영춘 의원(누가)이 발의한 안건인 중심어 “도시교통정비 촉진법 일부개정법률안”에 대한 논의 사항과 이에 대한 결정사항들을 모두 포함한 요약문이다.

상기의 과정과 같이 요약문은 중심어와 함께 세부 요약문에서 요약된 이슈 및 논의 사항과 결정사항의 정보를 모두 포함한 요약문으로 작성될 수 있다.

3. 검수 유의사항

3.1. 환언(paraphrasing)

논쟁적 대화에 대한 요약은 대화의 내용을 포괄하는 요약 문장을 새로 작성하는 것으로, 핵심적인 정보를 선택하고 이를 이해하고 통합하는 과정을 거친다. 이후 원 대화의 구조를 반영하는 재구조화의 과정을 거쳐 요약문이 작성된다. 본 지침에서는 세부 요약을 작성할 때, 한 사람의 주장을 하나의 세부 요약으로 작성하기 때문에, 논쟁적 대화 속에서 한 사람의 발화를 바탕으로 요약한다.

작업자는 한 사람의 발화를 그대로 사용하여 요약문으로 작성할 경우, 대화의 내용에서 핵심적인 내용을 추출하고 재구성하는 것이 아니라 반복적으로 기술하기 때문에 이에 대하여 주의가 필요하다.

...

○박인순 위원 그러니까 디테일은, 처우개선이 확실히 된 것은 확실한데 저는 해마다 12월 마지막 날에 앉아서 이런 똑같은 얘기를, 2년 후에도 또 똑같이 할까 봐, 이 법 자체를..... 1년 얘기도 그래서 나온 것 아니에요? 아까 설훈 위원님 말씀도 다 맞거든요. 그러니까 타임테이블을 해서 내년 초에 TF 뭘 만들든지 협의체를 만들든지 뭘 만들어서 어떻게 어떻게 하겠다 플랜을 해서, 연말에 똑같은 얘기를 안 할 수 있게 그것을 확실히 하는 게 중요한, 그게 법안소위에서 할 일이거든요.

처우는 확실히 개선됐고 강사실을 마련해 주고 그런 것은 디테일이고, 그것을 좀 확실히 해주세요. 다음에 또 2년 후 연말에, 만약에 2년 유예된다면 2년 후 연말에 또 똑같은 것을 안 하게 이렇게 이렇게 하겠다 그런 의지를 좀 교육부에서 확실히 해 주세요.

...

338교문(법안심사)소위01(15.12.23) 中

해당 발화의 경우, 박인순 위원이 디테일은 개선되었으나 TF 협의체 구성 플랜을 만들어 연말에 똑같은 얘기를 안 할 수 있게 해 달라는 요청을 하는 발화이다. 이때 똑같은 얘기를 안 한다는 것은 해당 법에 동일한 문제가 발생하지 않도록 요청하는 의미임을 알 수 있다.

환언 전	박인숙 위원은 디테일로는 처우개선이 된 것으로 보이나 TF나 협의체 구성 플랜을 만들어서 2년 후에 해당 법에 대해 똑같은 말을 하지 않도록 교육부에서 확실하게 해달라고 요청했다.
환언 후	박인숙 위원은 디테일로는 처우개선이 된 것으로 보이나 TF나 협의체 구성 플랜을 만들어서 2년 후에 해당 법에 대해 동일한 문제가 없도록 교육부에 요청했다.
338교문(법안심사)소위01(15.12.23) 中	

위 예시에서, “똑같은 말을 하지 않도록”이라는 것은 발화에서 그대로 나타난 표현으로 요약문의 문장으로 작성되기에는 부적합하다. 이는 해당 발화가 갖는 진짜 의미가 요약문에서 빠졌기 때문이다. 따라서 해당 표현은 “동일한 문제가 없도록”으로 환언하는 것이 적합하다.

다음 예시에서도 구어체로 인한 환언의 필요성이 드러난다.

...	○소위원장 신성범 제가 다시 한 번, 위원님들, 특히 야당 위원님들 중심으로 해서 2년간 유예를 해 주면, 1년이 아니라 2년간 주장하는 근거가 2년 동안 정말 시간강사의 처우개선과 신분보장이라는 입법취지에 맞게 정부에서 해낼 수 있는가에 대한 약간의 의구심들을 갖고 있는 것 같아요. 그래서 그 부분에 대한 것을 아까 설명했지만 좀 요약해서 이야기를 해 봐요.
...	338교문(법안심사)소위01(15.12.23) 中

위 예시에서, “입법취지에 맞게 정부에서 해낼 수 있는가에 대한 약간의 의구심”을 갖는다는 것은 대화에서 나타나는 구어체 표현으로서 요약문의 형식에는 적합하지 않다. 이러한 발화를 요약문에 그대로 사용하는 것은 적합하지 않고, 해당 발화가 갖는 진짜 의미가 요약문에서 드러나도록 환언 과정을 거쳐야 한다.

환언 전	소위원장 신성범은 2년간 유예를 할 경우, 시간강사의 처우개선과 신분보장이라는 입법 취지에 맞게 정부에서 해낼 수 있는가에 대해 위원들의 의구심이 있으니 야당 위원 중심으로 질의할 것을 요청했다.
환언 후	소위원장 신성범은 시간강사의 처우개선과 신분보장을 정부에서 유예기간 동안 해낼 수 있는가에 대해 야당 위원 중심으로 질의할 것을 요청했다.
338교문(법안심사)소위01(15.12.23) 中	

...
○여성부기획조정실장 김태석 저희도 예비비를 추진했었는데 기재부와 협의하는 가운데 절감 예산 중에 사용이 어려운 부분도 있기 때문에 그런 예산을 우선적으로 쓰도록 기재부와 협의가 됐습니다.
...
284여성(예산결산기금심사)소위01 中

위 예시의 경우, 발화의 특성상 불필요한 내용이 중복하여 표현된 예시이다. 기재부와 협의한다는 내용이 중복되어서 나타났다.

환언 전	여성부기획조정실장 김태석은 예비비를 추진했으나 기재부와 협의하면서 절감 예산 중 사용이 어려운 예산을 우선 쓰도록 기재부와 협의가 되었었다고 해명했다.
환언 후	여성부기획조정실장 김태석은 예비비를 추진했으나 절감 예산 중 사용이 어려운 예산을 우선 쓰도록 기재부와 협의가 되었었다고 해명했다.
338교문(법안심사)소위01(15.12.23) 中	

따라서 위 요약문과 같이, 발화에서 불필요하게 중복되는 내용은 삭제하고 필요한 정보만을 요약문으로 작성할 필요가 있다.

이처럼, 논쟁적 대화는 구어체로 작성되어 중복되는 내용을 포함하여 발언하기도 하며, 또한 실제로 의도한 내용과 달리 대화의 형식으로 발언하기도 한다는 점에 유의하여, 논쟁적 대화에서 의도된 정보와 내용이 함축된 요약문을 작성하여야 한다.

3.2. 화자의 명시

논쟁적 대화에 참여한 발화자들은 각자 주제에 대한 견해나 의견이 존재하며, 이에 대한 요약문은 내용에 대한 요약뿐만 아니라 화자에 대해 명확하게 명시할 필요가 있다. 특히 발화자의 직책이나 역할이 중요한 경우에는 더욱 그러하다.

<p>...</p> <p>○이태규 위원 그러니까 그 다자조약이 커버하는 부분하고 각 국가 간에 체결한 부분하고 사법공조의 효력이나 이런 측면에서는 어떤 게 훨씬 더 문제 해결에 도움이 되나요?</p> <p>○외교부제2차관 조태열 양자가 먼저, 특별협정에 해당하니까 양자조약이 먼저 적용이 되겠지요.</p> <p>○이태규 위원 그러면 이 양자조약이 큰 문제가 없으면 계속해서 체결을 내가 볼 때는 신속하게 더 확대시켜 나가야 된다고 보는데 이 부분이 지연되거나 이게 잘 안 되는 특별한 이유가 있는 겁니까?</p> <p>...</p>	<p>346외통(법안심사)소위02(16.10.28.) 中</p>
---	-------------------------------------

상기 대화에서 정보만 파악한다면 다자조약과 양자조약 중 양자조약이 먼저 적용이 된다는 것이 핵심 내용이다. 그러나 위 대화에서는 이태규 위원이 해당 정보에 대해 질문했다는 점, 그리고 답변한 사람이 외교부제2차관 조태열이라는 점 역시도 중요한 정보이다. 추가로, 양자조약을 확대할 필요성이 있는데 이것이 지연되는 이유가 무엇인지 이태규 위원이 질의했다는 것 역시 중요한 정보이다.

<p>화자를 명시하지 않은 요약문</p>	<p>다자조약보다 양자조약이 특별협정으로 우선 적용되고 사법 공조의 효력이 높다는 점에서 양자조약을 확대해나가는 방침이 바람직한데, 이태규 위원은 양자조약을 확대하는 과정에서 지연 등 문제가 발생하는 이유에 대해 질의하였다.</p>
<p>화자를 명시한 요약문</p>	<p>외교부제2차관 조태열은 국가 간 체결 부분과 사법공조의 효력 부분에서 양자조약이 효력이 높고 특별협정으로 우선 적용된다고 밝혔고, 이에 이태규 의원은 양자조약을 확대해나가는 방침이 바람직하다고 보았고, 양자조약을 확대하는 과정에서 지연 등 문제가 발생하는 이유에 대해 질의하였다.</p>
<p>346외통(법안심사)소위02(16.10.28.) 中</p>	

위와 같이, 화자를 명시하지 않은 요약문은 “다자조약보다 양자조약이 특별협정으로 우선 적용되고 사법 공조의 효력이 높다는 점에서 양자조약을 확대해나가는 방침이 바람직한다”라는 의견이 외교부 제2차관 조태열의 발화임이 누락되어 있다.

따라서 화자를 명시한 요약문에서 볼 수 있는 것처럼, 근거가 되는 발화를 누가 발언하였는지, 또한 바람직하다고 생각한 사람은 누구인지, 그리고 질의한 사람이 누구인지 등에 대한 정보를 모두 명확히 기재하여야 한다.

이와 마찬가지로 화자에 대한 정보도 모든 화자의 정보를 명시하여야 한다.

<p>화자를 일부만 명시한 요약문</p>	<p>병무청 차장 김노운은 예산불용액이 과다하다는 지적을 수용한다고 답했고, 이에 안규백 위원과 송영근 위원 등은 외국 출장 기회의 확대 등을 통해 선진국의 병무 행정을 배우고 좀 더 적극적으로 사업을 수행해 나갈 것을 주문했다.</p>
<p>화자를 모두 명시한 요약문</p>	<p>병무청 차장 김노운은 예산불용액이 과다하다는 지적을 수용한다고 답했고, 이에 안규백 위원, 송영근 위원, 심성찬 위원은 외국 출장 기회의 확대 등을 통해 선진국의 병무 행정을 배우고 좀 더 적극적으로 사업을 수행해 나갈 것을 주문했다.</p>
<p>334국방(예산결산심사)소위01(15.7.2) 中</p>	

3.3. 세부 지침: 국회 회의록 요약에서 유의할 사항

□ 회의 진행 발언

해당 발화는 국회 회의에서 소위원장 등의 진행자가 발언하는 의사진행에 관련된 발화로서, 요약의 대상이 되지 않는다.

○신성범 소위원장 좌석을 정돈하여 주시기 바랍니다.

성원이 되었으므로 제338회 국회 제1차 법안심사소위원회를 개의하겠습니다.

오늘은 교육부 소관 법안을 먼저 심사한 후에 문화체육관광부 소관 법안을 심사하도록 하겠습니다.

그리고 안내말씀 드립니다.

지난 12월 11일 자로 새정치민주연합 소속 유기홍 위원님이 사임하시고 유은혜 위원님이 새로 보임해 오셨습니다.

유은혜 위원님 인사말씀해 주시기 바랍니다.

○유은혜 위원 전반기에 법안소위였다가 이렇게 후반기, 막바지에 다시 법안소위를 하게 됐는데요. 여러 위원님들 많이 도와주시고 가르쳐 주시고, 제가 열심히 하겠습니다.

유사하게, 꼭 소위원장 등의 진행자가 발언하는 의사 진행발화가 아니더라도, 회의 참석자가 발화하는 의사 진행발화는 요약의 대상이 되지 않는다.

...

○김희선 위원 우선 기간, 전문위원이 한 것하고 발의한 법안 내용하고 다른 부분하고, 특위 구성에 대해서 정부 의견을 한번 들어 보시지요.

...

284여성(예산결산기금심사)소위01 中

□ 중심어와 상정된 안건의 연결

요약문은 중심어가 해당 대화에 대한 주제를 의미하는데, 국회 회의록의 경우에는 법률안에 대한 정확한 이름이 명시될 필요가 있다. 예를 들어, 다음과 같은 발화에서 명시된 중심어는 “고등교육법 일부개정법률안”으로서 이후의 대화는 해당 안건에 대한 회의가 진행된다.

<p>...</p> <p>○소위원장 신성범 의사일정 제1항 강은희 의원이 대표발의한 고등교육법 일부개정법률안을 심사하겠습니다.</p> <p>...</p>
338교문(법안심사)소위01(15.12.23) 中

위 발화에서 명시된 중심어는 “고등교육법 일부개정법률안”으로서 이후의 대화는 해당 안건에 대한 회의가 진행된다. 그리고 해당 법률안은 국회 회의록에서 “상정된 안건”으로 명시되어 있다.

심사된 안건

- 1. 고등교육법 일부개정법률안(강은희 의원 대표발의) 1
- 2. 교육환경 보호에 관한 법률안(유은혜 의원 대표발의)(계속) 1
- 3. 학교보건법 일부개정법률안(유은혜 의원 대표발의)(계속) 1
- 4. 문학진흥법안(도종환 의원 대표발의)(계속) 9
- 5. 관광숙박시설 확충을 위한 특별법 일부개정법률안(이에리사 의원 대표발의)(계속) 9
- 6. 공공디자인문화 진흥법안(이중훈 의원 대표발의)(계속) 9

위의 경우, 해당 중심어가 심사된 안건 중 어느 것에 해당하는지를 명시해 줄 필요가 있다. 따라서 실제 구축되는 분석 말뭉치에서는 해당 안건이 topic으로 명시되어 JSON 형식의 파일에 기록되어야 한다.

이에 대한 예시는 다음과 같다.

<pre> "issue": { "topic": "1. 고등교육법 일부개정법률안(강은희 의원 대표발의)", "keyword": "고등교육법 일부개정법률안", "sentence_id": SBRW2100000038.1.1.18, "begin": 22, "end": 34 }, "summary": { </pre>
--

"abstract": {

"form": "강은희 의원이 대표발의한 고등교육법 일부개정법률안에 관하여 정부는 2년의 유예기간을 달라고 요구하였다. 이에 1년이 적절하다는 의견과 국공립대학과 사립대학 간의 형평성을 고려한 유예 대책이 있는지에 대한 지적 등이 있었지만, 원안을 유지하면서 전문위원이 부대의견을 다는 것으로 합의되어 원안대로 의결됐다."

위와 같이 해당 요약문은 국회 회의록의 안건인 1. 고등교육법 일부개정법률안(강은희 의원 대표발의) 에 대하여 이루어진 대화에 대한 요약문이다. 해당 요약문의 중심어는 “고등교육법 일부개정법률안”이며, 이에 대한 요약문은 아래와 같이 요약되었다.

"강은희 의원이 대표 발의한 고등교육법 일부개정법률안에 관하여 정부는 2년의 유예기간을 달라고 요구하였다. 이에 1년이 적절하다는 의견과 국공립대학과 사립대학 간의 형평성을 고려한 유예대책이 있는지에 대한 지적 등이 있었지만, 원안을 유지하면서 전문위원이 부대의견을 다는 것으로 합의되어 원안대로 의결됐다."

이처럼, 국회 회의록처럼 회의의 안건이 회의록에 명시되어 있는 경우 상기와 같이 회의록의 안건과 중심어를 매핑하는 작업을 통해 대화의 주제가 무엇인가를 명시하도록 한다.

만약, 특정 중심어를 갖는 대화일지라도 상정된 안건을 여러 개를 다룰 때는 다음의 방식으로 안건(topic)을 선택한다.

- (1) 상정된 안건에 관련된 대화인 경우, 관련 안건들 모두 안건(topic)으로 매핑한다.
- (2) 상정된 안건과 무관한 대화인 경우, 해당 대화는 요약의 대상이 되지 않는다.

4. 데이터 납품 형태

4.1. 메타데이터

```
{
  "id": "SBSC2102112100",
  "metadata": {
    "title": "국립국어원 국회 회의록 요약 분석 말뭉치 SBSC2102112100",
    "creator": "국립국어원",
    "distributor": "국립국어원",
    "year": "2021",
    "category": "회의록 > 국회소위원회",
    "annotation_level": "요약 분석",
    "sampling": "본문 전체"
  },
  "document": [
    {
      "id": "SBRW2100000001.1",
      "metadata": {
        "title": "第284回國會(定期會)女性委員會會議錄(豫算決算基金審查小委員會)第1號",
        "author": "國會事務處",
        "publisher": "國會事務處",
        "date": "20090925",
        "topic": "1. 2008회계연도 세입세출결산 > 가. 여성부 소관 / 2. 2008회계연도 기  
금결산 > 가. 여성발전기금",
        "speaker": [
          {
            "id": "김상희",
            "occupation": "소위원장",
            "original_id": "김상희"
          },
          ...
        ]
      }
    }
  ],
}
```

위 예시는 분석 말뭉치의 메타데이터의 정보를 보여주는 예시이다. 분석 말뭉치는 200개 회의록에 대해 1개의 파일로 저장되어 납품되었으며, 파일에 대한 메타데이터 및 각 문서에 대한 메타데이터 정보를 기록하고 있다.

- 분석 말뭉치 파일의 메타데이터 (root > metadata)
 - 분석 문서의 메타데이터 (root > document > metadata)
 - > 회의록의 제목, 저자, 날짜 등 회의록에 기재된 정보들
 - > 회의록의 발화자 정보 (root > document > metadata > speaker)
- * 이때, 화자(speaker)의 original_id는 한자-한글 변환 적용 전 원문 그대로의 이름(즉, 한자 또는 한글)이며, id는 한자-한글 변환 적용 후의 이름(즉, 모두 한글)이다.

4.2. 발화 정보

각 문서의 발화에 대한 정보는 sentence에 모두 기록되어 있으며, 각 발화를 띄어쓰기로 토큰화한 word의 위치 정보 및 문자열을 모두 기록하여 납품되었다. 각 발화에 대한 id값은 원시 말뭉치의 발화 id와 같이 적용되었다.

```
"sentence": [  
  {  
    "id": "SBRW2100000001.1.1.1",  
    "form": "회의를 시작하도록 하겠습니다.",  
    "word": [  
      {  
        "id": 0,  
        "form": "회의를",  
        "begin": 0,  
        "end": 3  
      },  
      {  
        "id": 1,  
        "form": "시작하도록",  
        "begin": 4,  
        "end": 9  
      },  
      {  
        "id": 2,  
        "form": "합니다",  
        "begin": 10,  
        "end": 15  
      }  
    ]  
  }  
]
```

```

        "id": 2,
        "form": "하겠습니다.",
        "begin": 10,
        "end": 16
    }
]
},

```

4.3. 요약 정보

각 회의록 문서에 대한 요약문은 크게 두 가지로 구분되는데, 하나는 main_summary와 다른 하나는 issue_summary이다.

main_summary는 회의록 전체에 대한 요약문으로서 회의에서 다루어진 안건들에 대한 요약 정보가 포함되어 있으며, 이에 대한 의결 사항 등이 추상 요약되어있다.

issue_summary는 회의록 내 쟁점별 요약문으로서 특정 중심어에 대한 요약문과 세부 요약문들의 집합으로 구성되어 있다. 아래의 예시에서는 중심어(keyword) “2008회계연도 여성부 소관 세입세출결산”이 회의의 안건 “1. 2008회계연도 세입세출결산 > 가. 여성부 소관”에 관련된 것이며, 이에 대한 주제 요약문이 “summary > abstract > form”에 기록되어 있다. 그리고 해당 요약문을 작성할 때 참고한 발화들의 목록이 “summary > abstract > reference”에 나열되어 있다.

또한 해당 주제 요약문에 대한 세부 요약문(Problems/Issues, Decisions, Actions)들 역시 “summary > sub”에 리스트로 들어가 있으며, 각 세부 요약문의 유형(type) 및 참고한 발화의 목록 역시 같은 형태로 저장되어있다.

```

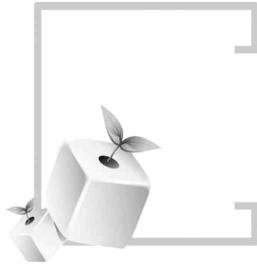
{
  "id": "SBRW2100000001.1",
  "SC": {
    "main_summary": "본 회의는 제284회 정기국회 여성위원회 제1차 예산결산기금
심사소위원회로 진행되었으며, 이·전용 과다 및 예산 10% 절감 부당 부분에 대해서는 주의, 이·전용액
으로 청사 임차료 충당 문제는 주의, 건국 60주년 기념사업 추진 부적절은 시정, 연구용역 연도 말 발
주 및 과다 수의계약 지양은 주의와 부대의견을 달고, 성별영향평가 운영방식 개선 및 활용도 제고는
제도개선, 직업교육훈련의 실적 및 운영방식 개선은 제도개선, 여성주간사업 축소집행 문제는 주의, 생
애주기별 양성평등의식 교육 예산 부족 및 소극적 집행 개선은 주의, 폭력피해여성 주거지원사업 임의
적 축소집행은 시정, 성매매 예방 및 교육사업 임의적 축소집행은 시정, 일본군위안부 피해자 기념사업
예산의 목적 외 사용 및 축소집행은 시정으로 요구사항이 결정됐다. 따라서 시정요구사항은 시정 4건,
주의 5건, 제도개선 2건, 부대의견 1건이 되며, 2008회계연도 여성부 소관 세입세출결산, 2008회계연
도 여성발전기금 결산에 대하여는 각각 소위원회에서 합의한 시정요구사항을 반영하여 정부 원안대로
의결했다.",
    "issue_summary": [
      {
        "issue": {
          "topic": "1. 2008회계연도 세입세출결산 > 가. 여성부 소관",
          "keyword": "2008회계연도 여성부 소관 세입세출결산",
          "sentence_id": "SBRW2100000001.1.1.5",
          "begin": 9,
          "end": 31
        },
        "summary": {
          "abstract": {
            "form": "본 회의에서 여성부의 2008회계연도 세입세출결산에 관
해 세입 징수 결정된 금액의 정산 잔액을 조속히 반납 조치가 필요하다고 했고, 여성부는 다 납부가
되었고 수납률 제고를 하겠다고 답변하여 해당 사안은 넘어갔다. 그리고 이·전용 과다 예산 10% 절감
부당에 관해서 위원들은 '시정'으로 의견을 모았지만, 다시 결정하기로 하고 넘어갔고, 이·전용액으로
여성부가 청사 임차료를 충당한 문제에 대해서는 여성부의 해명을 듣고 불가피한 전용이었다고 보고 '
주의'를 주기로 했다.",
            "reference": [
              "SBRW2100000001.1.1.17",
              "SBRW2100000001.1.1.18",
              "SBRW2100000001.1.1.21",
              ...
            ]
          }
        }
      }
    ]
  }
}

```

```

    },
    "sub": [
      {
        "type": "Problems/Issues",
        "form": "전문위원은 세입 수납률 제고 조치 필요에 관한 사항에 관하여 세입 징수 결정 금액 중 징수율 감소로 인한 정산 잔액 반납 조치가 필요하다고 발언했다. 그리고 시정요구수준을 어떤 식으로 할지 위원들의 의견을 요청했다.",
        "reference": [
          "SBRW2100000001.1.1.17",
          "SBRW2100000001.1.1.18"
        ]
      },
      {
        "type": "Problems/Issues",
        "form": "여성부기획조정실장 김태석은 세입 수납률 제고 조치에 대해 서울시, 강원도 등 미수납액이 있고 현재는 다 납부가 되었다고 밝히며, 세입고지서 발부 및 독촉으로 수납률 제고를 하겠다고 발언했다.",
        "reference": [
          "SBRW2100000001.1.1.21"
        ]
      }
    ]
  },

```



제 4 장

논쟁적 대화 요약 말뭉치 활용 방안



1. 회의록 요약 기술 개발

회의는 사람들이 아이디어를 교환하고, 정보를 공유하며, 계획을 수립하는 가장 흔한 방법이다. 최근의 음성 인식 기술이 진보하면서, 회의록(meeting transcripts)을 자동으로 작성하는 기술들이 많이 개발되었다. 회의록을 자동으로 작성하는 솔루션의 현황은 다음과 같다.

□ 네이버 클로바

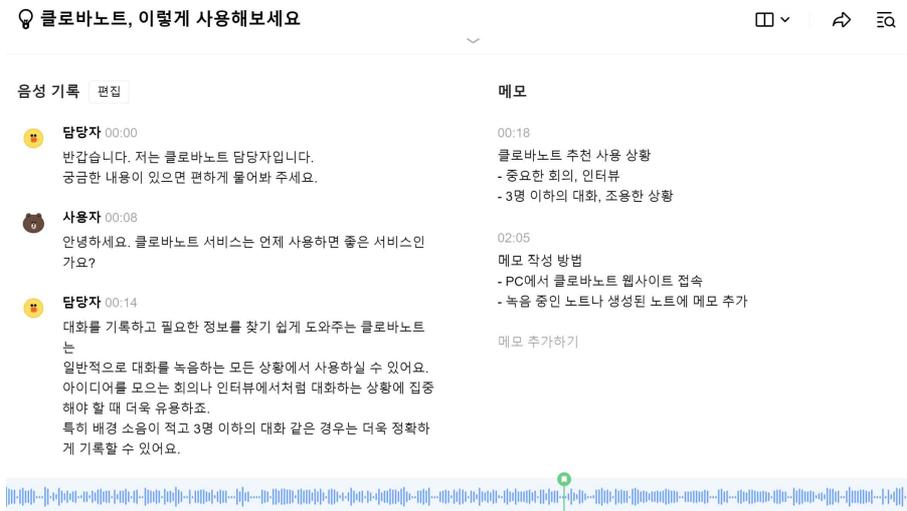


그림 7 네이버 클로바 화면 및 예시

□ 마음 AI

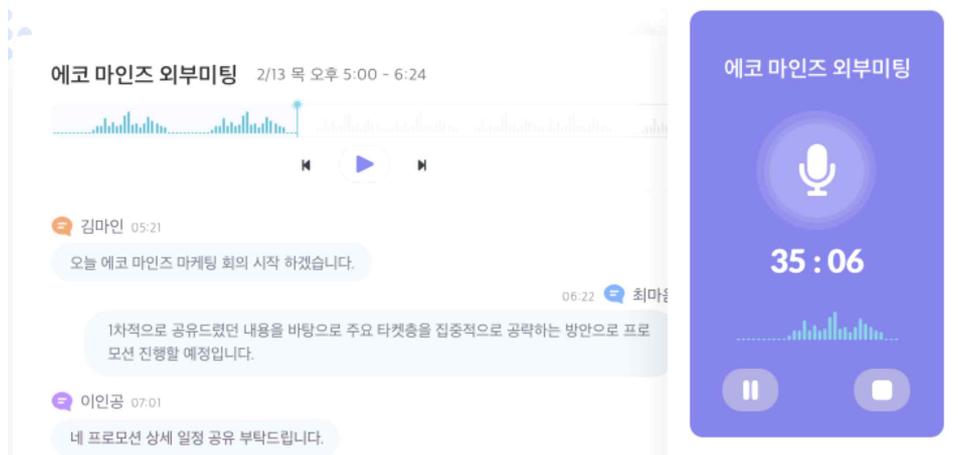


그림 8 마음 AI 화면 및 예시

네이버 클로바 노트 서비스나 마음 AI의 경우 회의에 대한 전사된 기록을 제공하며, 이러한 회의의 콘텐츠의 핵심 내용을 요약하여 제공하고자 하는 연구도 최근 수행되기 시작했다.

□ LG전자의 AI 기반 회의 녹취록 요약 경진대회

LG전자는 수기 회의록 작성의 부담을 절감하기 위해 회의에서 논의된 핵심 안건을 추상 요약하는 경진대회를 2021년 9월에 주최한 바 있다.⁶⁾

□ 회의록 요약에 위한 transformer 모델

해외 연구(Zhu 외, 2020)에서는 문서 요약에 사용되는 transformer 모델을 대화의 구조에 맞게 변형하여 문장 단위의 encoding 구조를 확장하여 발화의 순서에 따른 맥락을 해석할 수 있는 구조화된 transformer 모델을 제안한 바 있다.

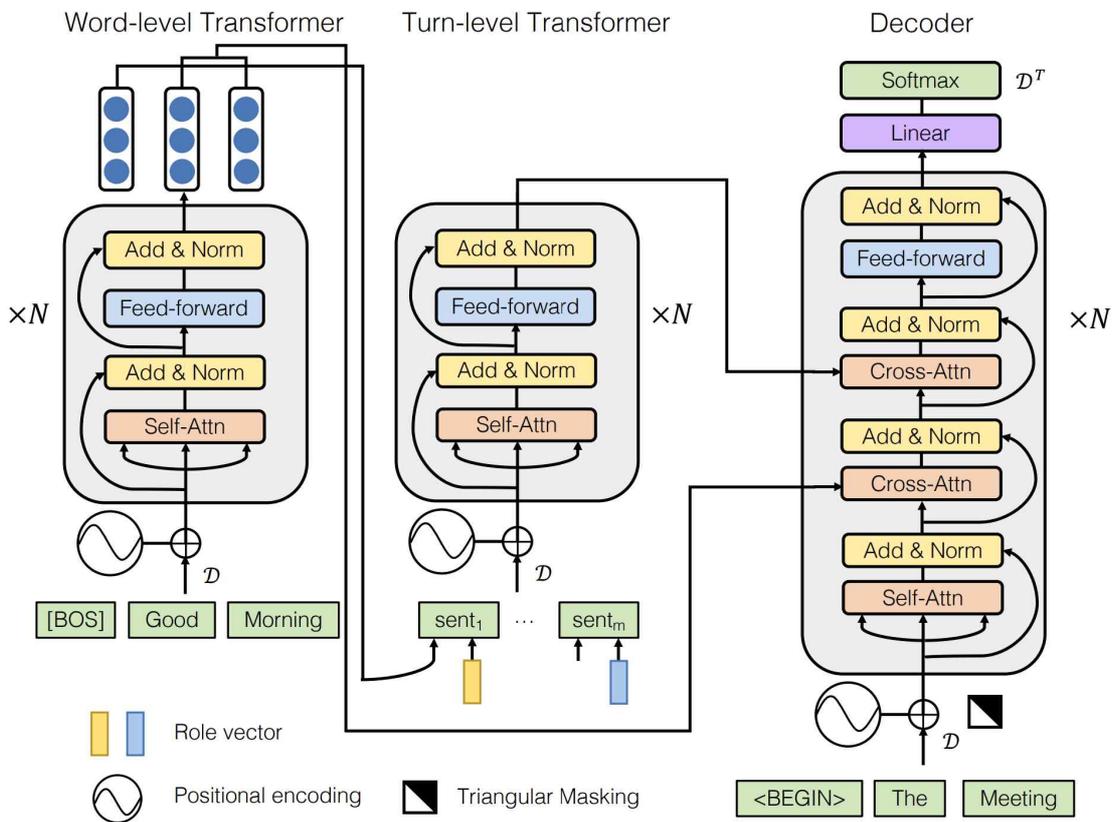


그림 9 회의록 요약에 위한 구조화된 transformer 구조

6) <https://dacon.io/competitions/official/235813/overview/description>

상기 논의된 바와 같이, 회의록 요약에 관한 연구가 최근 활발히 이루어지고 있으며, 해당 연구의 흐름을 살펴보면 다음과 같다.

▶ 문서 요약을 넘어 대화 요약으로

최근의 요약 연구는 뉴스, 도서에 치중되어 이루어졌는데 뉴스나 도서의 경우에는 한 명의 전문가에 의해 하나의 주제에 대해 논리적으로 잘 쓰인 문서이다. 반면 회의록의 경우에는 실제 발화를 바탕으로 작성된 문서로서 대화에서 다룬 안건, 대화에서 이루어진 각 화자의 의견이나 주제에 대한 결정사항 등의 다양한 입체적 정보를 포함하고 있다. 최근의 연구에서 문서 요약 AI는 대화 요약에 대해 상대적으로 낮은 성능을 보이고 있어 고품질의 논쟁적 대화 요약 말뭉치 구축의 필요성이 대두되고 있다.

▶ 정보 요약을 넘어 의견 채굴로

최근의 요약 연구는 뉴스나 도서의 핵심 정보 요약에 초점을 맞추고 있다. 회의록의 경우에는 회의의 전체적인 결론 이외에도 각 화자가 가진 견해와 감성 등도 중요한 정보로 간주된다. 예를 들어, 국회 회의록에서도 하나의 안건에 대해서 정부 측의 입장과 위원회의 입장이 상충하는 경우, 어떤 측면에서 상충했는지, 각자의 주장은 무엇인지도 중요한 정보가 된다. 최근의 연구는 논증 채굴의 측면에서 다양한 관점으로 회의록에 관한 분석 연구가 이루어지고 있다.

7) <https://2021.argmining.org/index.html>

2. 회의록 요약 말뭉치 활용 방안

본 과제에서 구축한 회의록 요약 말뭉치 분석 지침과 말뭉치는 최근 주목받는 회의록 요약을 위한 기초 연구로서 중요한 자원으로 활용될 수 있다. 회의록 요약 및 논증 채굴 모델을 위한 성능 향상의 기반이 됨과 동시에 다양한 분야의 발전을 견인할 것으로 기대한다.

1) 대규모 논쟁적 대화 요약 말뭉치 구축을 위한 지침

본 과제에서 구축된 지침은 구어체의 대화로부터 문어체의 요약문을 작성한다는 대화 요약을 넘어, 논쟁적 대화에서의 입체적 중요 정보를 분석하는 방법(논변 구조) 및 입체적 정보를 상향식으로 요약하여 핵심 정보를 놓치지 않는 추상 요약문을 작성하는 일련의 과정을 상세히 지침으로 기술하였다. 특히 기술적으로 어려운 추상 요약 지침을 개발하는 것은 비용과 시간 측면에서 비용이 높은 작업이므로 본 과제를 통해 구축된 지침은 이후 산학계에서 또 다른 말뭉치를 구축하거나 확장할 때 중요한 지침으로 활용될 수 있다.

2) 모델 성능 평가용 데이터

영어의 경우 AMI Corpus 등 데이터가 공개되어 있으나 한국어의 경우 논쟁적 대화를 위한 고품질의 말뭉치가 거의 없으며, 일상 대화 및 짧은 대화에 대한 1~2 문장의 간략한 요약데이터만이 존재한다. 본 과제를 통해 구축된 말뭉치는 회의록에 대한 대표 요약문 이외에도, 각 주제(안건)에 대한 요약문과 함께 문제/이슈, 결정사항, 후속조치에 해당하는 세부 요약문을 작성하여 대화 요약 모델에 대한 다양한 측면에서의 평가가 가능하여 향후 모델의 성능을 개선하기 위한 다양한 지표를 제공해 줄 수 있다.

3) 응용 기술 개발을 위한 기반 마련

논쟁적 대화 요약 분석은 단순히 회의에서 이루어진 안건에 대한 핵심 정보만을 파악하는 것 이외에도, 각 화자의 입장에 대한 의견 채굴 및 감성분석 등의 추가적인 연구가 가능하다. 또한 논증 채굴 관점에서 회의에서 이루어진 결정사항이나 각 발화 간의 논변 구조 등을 분석하는 데 활용될 수 있다. 본 사업을 통해 구축된 말뭉치는 상기의 정보를 분석하는 데 필요한 요약 정보를 모두 제공하므로 향후 다양한 응용 기술을 개발하기 위한 기초 말뭉치의 역할을 수행할 수 있다.

4) 논변 구조 분석 및 표준화

본 과제에서 제출된 논변 구조 분석 지침은 아직 국내에서 연구가 거의 없는 분야로서, 논쟁적 대화에서 각 화자의 발화가 갖는 역할과 각 발화 간의 관계를 분석할 수 있는 연구 방법론과 체계를 제공한다. 논변 구조는 논쟁적 대화에서의 논증 채굴을 위한 핵심적인 스키마로서 향후 활용 가치 및 연구 자원으로서의 효용성이 높다.

국립국어원과 본 과제 연구진은 본 과제를 통해 구축된 논변 구조 분석 지침의 스키마에 대한 표준화 작업을 진행하고 있으며, 2022년 1분기에 KS표준으로 기고하여 산업 표준심의회 전문가 심사 안건으로 상정할 예정이다.

<부록 1> 분석 대상 회의록 목록

파일명	대상 회의록	분야
SBRW2100000001	284여성(예산결산기금심사)소위01	여성
SBRW2100000002	311법사(법안심사제1)소위01(12.11.15)	법사
SBRW2100000003	311환노(법안심사)소위01(12.9.17)	환노
SBRW2100000004	320교문(청원심사)소위01(13.11.18)	교문
SBRW2100000005	320환노(예산결산기금심사)소위03(13.11.28)	환노
SBRW2100000006	321교문(예산결산기금심사)소위03(13.12.16)	교문
SBRW2100000007	321국토(교통법안심사)소위01(13.12.12)	국토
SBRW2100000008	322국토(민투mrg대책)소위01(14.2.26)	국토
SBRW2100000009	322국토(철도산업발전)소위01(14.2.7)	국토
SBRW2100000010	322국토(철도산업발전)소위02(14.2.17)	국토
SBRW2100000011	323국토(철도산업발전)소위02(14.4.10)	국토
SBRW2100000012	323안행(법안심사)소위03(14.4.15)	안행
SBRW2100000013	323환노(노사정)소위03(14.4.17)	환노
SBRW2100000014	326국토(국토법안심사)소위01(14.7.9)	국토
SBRW2100000015	329여성(법안심사)소위02(14.11.25)	여성
SBRW2100000016	329외통(법안심사)소위04(14.12.5)	외통
SBRW2100000017	332미방(법안심사)소위02(15.4.23)	미방
SBRW2100000018	334국방(예산결산심사)소위01(15.7.2)	국방
SBRW2100000019	334복지(법안심사)소위05(15.7.6)	복지
SBRW2100000020	337국방(국방운영개선)소위01(15.11.20)	국방
SBRW2100000021	337복지(보육제도개선)소위02(15.11.25)	복지
SBRW2100000022	337복지(보육제도개선)소위04(15.11.30)	복지
SBRW2100000023	338교문(법안심사)소위01(15.12.23)	교문
SBRW2100000024	346국방(예산결산심사)소위03(16.11.3.)(비공개포함)	국방
SBRW2100000025	346외통(법안심사)소위02(16.10.28.)	외통
SBRW2100000026	347안행(안전및선거법심사)소위01(16.12.20.)	안행
SBRW2100000027	347안행(행정및인사법심사)소위01(16.12.21.)	안행
SBRW2100000028	352국토(교통법안심사)소위01(17.7.5.)	국토
SBRW2100000029	354외통(법안심사)소위04(17.11.22.)	외통
SBRW2100000030	354행안(안전및선거법심사)소위03(17.11.30.)	행안
SBRW2100000031	354환노(고용노동)소위03(17.9.28.)	환노
SBRW2100000032	364환노(예산결산기금심사)소위04(18.11.22.)	환노
SBRW2100000033	367교육(법안심사)소위02(19.3.14.)	교육
SBRW2100000034	369외통(신북방)소위01(19.7.18.)	외통
SBRW2100000035	370교육(예산결산기금심사)소위01(19.8.22.)	교육
SBRW2100000036	370행안(법안심사)소위01(19.8.26.)(공청회)	행안
SBRW2100000037	371행안(법안심사)소위03(19.10.1.)	행안
SBRW2100000038	371환노(예산결산기금심사)소위01(19.10.30.)	환노
SBRW2100000039	376기재(조세)소위01(20.3.11.)	기재
SBRW2100000040	377여성(법안심사)소위01(20.5.6.)	여성
SBRW2100000041	311국토(예산결산기금심사)소위02(12.11.7)	국토
SBRW2100000042	311재정(예산결산기금심사)소위09(12.11.20)	재정
SBRW2100000043	311행안(법안심사)소위02(12.9.19)	행안
SBRW2100000044	311교과(법안심사)소위03(12.9.24)	교과
SBRW2100000045	311법사(예산결산심사)소위03(12.11.21)	법사
SBRW2100000046	311법사(예산결산심사)소위04(12.12.3)	법사
SBRW2100000047	311외통(법안심사)소위03(12.11.19)	외통

SBRW2100000048	311재정(경제재정)소위03(12.9.24)	재정
SBRW2100000049	311재정(예산결산기금심사)소위08(12.11.15)	재정
SBRW2100000050	313문방(법안심사)소위01(13.2.15)	문방
SBRW2100000051	315국방(예산결산심사)소위01(13.4.23)	국방
SBRW2100000052	315국토(교통법안심사)소위01(13.4.23)	국토
SBRW2100000053	315여성(유엔협약이행정책심사)소위01(13.4.23)	여성
SBRW2100000054	315재정(예산결산기금심사)소위03(13.5.3)	재정
SBRW2100000055	315재정(조세)소위02(13.4.18)	재정
SBRW2100000056	315재정(조세개혁)소위01(13.4.24)	재정
SBRW2100000057	315환노(법안심사)소위02(13.4.17)(공청회)	환노
SBRW2100000058	316법사(법안심사제1)소위04(13.6.28)	법사
SBRW2100000059	316교문(법안심사)소위01(13.6.24)	교문
SBRW2100000060	316국방(국방운영개선)소위01(13.6.4)	국방
SBRW2100000061	317여성(일본군위안부대책)소위01(13.7.12)	여성
SBRW2100000062	320국토(교통법안심사)소위01(13.10.8)	국토
SBRW2100000063	320국토(교통법안심사)소위05(13.12.9)	국토
SBRW2100000064	320국토(예산결산기금심사)소위03(13.12.5)	국토
SBRW2100000065	320국토(청원심사)소위01(13.10.8)	국토
SBRW2100000066	320안행(법안심사)소위04(13.12.10)	안행
SBRW2100000067	320재정(예산결산기금심사)소위03(13.12.10)	재정
SBRW2100000068	321교문(법안심사)소위01(13.12.19)	교문
SBRW2100000069	321국방(법률안심사)소위02(13.12.30)	국방
SBRW2100000070	321국토(국토법안심사)소위01(13.12.26)	국토
SBRW2100000071	321법사(예산결산심사)소위02(13.12.12)	법사
SBRW2100000072	321복지(예산결산심사)소위02(13.12.12)	복지
SBRW2100000073	322국방(법률안심사)소위03(14.2.26)	국방
SBRW2100000074	322법사(법안심사제1)소위02(14.2.25)	법사
SBRW2100000075	322외통(법안심사)소위02(14.2.20)	외통
SBRW2100000076	322재정(경제재정)소위01(14.2.17)	재정
SBRW2100000077	323국방(법률안심사)소위02(14.4.18)	국방
SBRW2100000078	323국토(철도산업발전)소위03(14.4.16)	국토
SBRW2100000079	323국토(청원심사)소위01(14.4.11)	국토
SBRW2100000080	323복지(법안심사)소위03(14.4.16)(공청회)	복지
SBRW2100000081	323환노(노사정)소위01(14.4.7)	환노
SBRW2100000082	324국방(법률안심사)소위01(14.4.21)	국방
SBRW2100000083	324법사(법안심사제2)소위01(14.4.22)	법사
SBRW2100000084	324재정(조세개혁)소위01(14.4.22)	재정
SBRW2100000085	324환노(법안심사)소위01(14.4.21)	환노
SBRW2100000086	326국방(법률안심사)소위01(14.7.17)	국방
SBRW2100000087	326법사(법안심사제1)소위02(14.7.11)	법사
SBRW2100000088	326법사(법안심사제2)소위01(14.7.15)	법사
SBRW2100000089	326재정(2013회계연도결산심사)소위01(14.7.14)	재정
SBRW2100000090	327복지(예산결산심사)소위02(14.7.24)	복지
SBRW2100000091	329환노(법안심사)소위06(14.12.8)	환노
SBRW2100000092	329국방(예산결산심사)소위01(14.11.7)	국방
SBRW2100000093	329법사(법안심사제1)소위01(14.11.6)	법사
SBRW2100000094	329재정(예산결산기금심사)소위03(14.11.11)	재정
SBRW2100000095	330국토(국토법안심사)소위01(14.12.23)	국토
SBRW2100000096	330법사(법안심사제1)소위02(15.1.8)	법사
SBRW2100000097	330외통(법안심사)소위01(15.1.6)	외통

SBRW2100000098	331법사(법안심사제2)소위01(15.2.24)	법사
SBRW2100000099	332교문(법안심사)소위02(15.4.30)	교문
SBRW2100000100	332법사(법안심사제1)소위02(15.5.1)	법사
SBRW2100000101	332법사(법안심사제1)소위03(15.5.4)	법사
SBRW2100000102	332안행(법안심사)소위04(15.4.30)	안행
SBRW2100000103	332외통(법안심사)소위01(15.4.28)	외통
SBRW2100000104	332재정(조세)소위02(15.4.27)	재정
SBRW2100000105	332재정(조세)소위03(15.4.29)	재정
SBRW2100000106	332환노(법안심사)소위02(15.4.28)	환노
SBRW2100000107	334국방(법률안심사)소위03(15.7.1)	국방
SBRW2100000108	334국방(예산결산심사)소위04(15.7.7)(비공개포함)	국방
SBRW2100000109	334안행(예산결산기금심사)소위02(15.7.3)	안행
SBRW2100000110	334외통(예산결산기금심사)소위01(15.7.2)	외통
SBRW2100000111	334외통(예산결산기금심사)소위02(15.7.3)	외통
SBRW2100000112	334재정(조세)소위01(15.7.2)	재정
SBRW2100000113	335법사(법안심사제2)소위01(15.7.21)	법사
SBRW2100000114	337국방(예산결산심사)소위01(15.10.20)	국방
SBRW2100000115	337국방(청원심사)소위01(15.11.18)	국방
SBRW2100000116	337법사(예산결산심사)소위03(15.10.28)	법사
SBRW2100000117	337복지(보육제도개선)소위03(15.11.27)	복지
SBRW2100000118	337외통(법안심사)소위01(15.10.29)	외통
SBRW2100000119	337재정(예산결산기금심사)소위02(15.10.23)(비공개포함)	재정
SBRW2100000120	337재정(조세)소위11(15.11.30)	재정
SBRW2100000121	340국방(법률안심사)소위01(16.2.16)	국방
SBRW2100000122	343외통(예산결산기금심사)소위01(폐)(16.7.12)	외통
SBRW2100000123	343외통(예산결산기금심사)소위02(폐)(16.7.13)	외통
SBRW2100000124	346교문(예산결산기금심사)소위08(16.11.16.)	교문
SBRW2100000125	346기재(경제재정)소위01(16.11.8.)	기재
SBRW2100000126	346기재(예산결산기금심사)소위02(16.11.4.)	기재
SBRW2100000127	346기재(조세)소위08(16.11.28.)	기재
SBRW2100000128	346복지(예산결산심사)소위04(16.10.31.)	복지
SBRW2100000129	348법사(법안심사제1)소위01(17.1.18.)	법사
SBRW2100000130	349국방(법률안심사)소위02(17.2.13.)(공청회)	국방
SBRW2100000131	349환노(고용노동)소위02(17.3.2.)	환노
SBRW2100000132	350국방(법률안심사)소위01(17.3.14.)	국방
SBRW2100000133	350국토(국토법안심사)소위01(17.3.23.)	국토
SBRW2100000134	350기재(경제재정)소위01(17.3.23.)	기재
SBRW2100000135	350안행(안전및선거법심사)소위01(17.3.24.)	안행
SBRW2100000136	350안행(행정및인사법심사)소위01(17.3.23.)	안행
SBRW2100000137	353기재(조세)소위01(17.8.22.)	기재
SBRW2100000138	354여성(법안심사)소위01(17.9.18.)(공청회)	여성
SBRW2100000139	354여성(예산결산심사)소위02(17.11.20.)	여성
SBRW2100000140	354외통(법안심사)소위03(17.11.21.)	외통
SBRW2100000141	354외통(법안심사)소위05(17.11.29.)	외통
SBRW2100000142	354환노(고용노동)소위05(17.11.28.)	환노
SBRW2100000143	355교문(교육법안심사)소위01(17.12.19.)	교문
SBRW2100000144	356여성(국제규약점검)소위01(18.2.13.)	여성
SBRW2100000145	358행안(안전및선거법심사)소위01(18.3.27.)	행안
SBRW2100000146	359환노(고용노동)소위01(18.4.11.)(비공개포함)	환노
SBRW2100000147	360국토(교통법안심사)소위01(18.5.24.)	국토

SBRW2100000148	363기재(경제재정)소위02(18.8.29.)	기재
SBRW2100000149	363법사(법안심사제1)소위01(18.8.28.)	법사
SBRW2100000150	363복지(예산결산심사)소위02(18.8.28.)	복지
SBRW2100000151	364국방(병역특례제도개선)소위02(18.11.29.)	국방
SBRW2100000152	364국토(국토법안심사)소위04(18.12.07.)(공청회)	국토
SBRW2100000153	364문체(법안심사)소위02(18.9.12.)	문체
SBRW2100000154	364법사(법안심사제2)소위01(18.11.12.)	법사
SBRW2100000155	364여성(법안심사)소위02(18.11.9.)	여성
SBRW2100000156	364외통(법안심사)소위04(18.12.6.)	외통
SBRW2100000157	364외통(청원심사)소위01(18.11.22.)	외통
SBRW2100000158	365행안(법안심사)소위01(18.12.26.)	행안
SBRW2100000159	367국방(병역특례제도개선)소위01(19.3.25.)	국방
SBRW2100000160	367기재(조세)소위01(19.3.28.)	기재
SBRW2100000161	367행안(법안심사)소위01(19.3.8.)	행안
SBRW2100000162	367환노(고용노동)소위02(19.3.19.)	환노
SBRW2100000163	368행안(제천화재관련평가)소위01(폐)(19.5.9.)	행안
SBRW2100000164	369교육(법안심사)소위01(19.6.26.)	교육
SBRW2100000165	369교육(법안심사)소위03(19.7.11.)	교육
SBRW2100000166	369국토(교통법안심사)소위02(19.7.16.)	국토
SBRW2100000167	369기재(예산결산기금심사)소위01(19.7.15.)	기재
SBRW2100000168	369여성(법안심사)소위02(19.6.27.)	여성
SBRW2100000169	369행안(제천화재관련평가)소위02(19.7.11.)	행안
SBRW2100000170	370국토(교통법안심사)소위01(19.8.22.)	국토
SBRW2100000171	370환노(환경)소위01(19.8.21.)	환노
SBRW2100000172	371국방(법률안심사)소위02(19.9.20.)	국방
SBRW2100000173	371국방(법률안심사)소위05(19.11.18.)	국방
SBRW2100000174	371외통(미주)소위01(19.11.12.)	외통
SBRW2100000175	371외통(법안심사)소위02(19.11.21.)	외통
SBRW2100000176	371외통(신남방)소위01(19.11.13.)	외통
SBRW2100000177	371외통(예산.결산.기금심사)소위01(19.9.17.)	외통
SBRW2100000178	371외통(유럽.아프리카)소위01(19.11.18.)	외통
SBRW2100000179	371환노(예산결산기금심사)소위05(19.11.7.)	환노
SBRW2100000180	372환노(환경)소위02(19.12.12.)	환노
SBRW2100000181	376기재(예산결산기금심사)소위01(20.3.11.)	기재
SBRW2100000182	376행안(법안심사)소위01(20.3.4.)	행안
SBRW2100000183	376환노(예산결산기금심사)소위01(20.3.10.)	환노
SBRW2100000184	377법사(법안심사제1)소위01(20.4.28.)	법사
SBRW2100000185	377행안(예산.결산및기금심사)소위01(20.4.28.)	행안
SBRW2100000186	355여성(법안심사)소위01(17.12.15.)	여성
SBRW2100000187	364산자중기(산업통상자원특허)소위02(18.11.20.)	산자중기
SBRW2100000188	364정치개혁특(정치개혁제2)소위02(18.12.6.)	정치개혁특
SBRW2100000189	365농해수(농림축산식품법안심사)소위01(19.1.14.)	농해수
SBRW2100000190	365사법개혁특(법원.법조)소위03(18.12.20.)	사법개혁특
SBRW2100000191	365정치개혁특(정치개혁제2)소위07(폐)(19.1.17.)	정치개혁특
SBRW2100000192	368교육(법안심사)소위03(폐)(19.6.12.)	교육
SBRW2100000193	368정치개혁특(정치개혁제2)소위08(19.4.11.)	정치개혁특
SBRW2100000194	369농해수(해양수산법안심사)소위01(19.7.9.)	농해수
SBRW2100000195	370농해수(농림축산식품법안심사)소위01(19.8.26.)	농해수
SBRW2100000196	371교육(법안심사)소위03(19.10.30.)	교육
SBRW2100000197	371농해수(농림축산식품법안심사)소위01(19.9.25.)	농해수

SBRW2100000198	371산자중기(중소벤처기업)소위01(19.9.26.)	산자중기
SBRW2100000199	371산자중기(중소벤처기업)소위04(19.11.28.)	산자중기
SBRW2100000200	377국토(교통법안심사)소위01(20.5.6.)	국토

[분야 정보]

법사: 법제사법위원회

교문: 교육문화체육관광위원회

환노: 환경노동위원회

국토: 국토교통위원회

안행: 안전행정위원회

여성: 여성가족위원회

외통: 외교통일위원회

미방: 미래창조방송통신위원회

국방: 국방위원회

안행: 안전행정위원회

복지: 보건복지위원회

교육: 교육위원회

행안: 행정안전위원회

기재: 기획재정위원회

환노: 환경노동위원회

교과: 교육과학기술위원회

재정: 기획재정위원회

산자중기: 산업통상자원중소벤처기업위원회

논변 구조 주석 체계

2021.12.21.

초안: 연세대학교, 한국과학기술원, (주)딥네츄럴, (주)테디섬

검토: 국립국어원

1. 논변 구조 주석 체계

1.1 개요

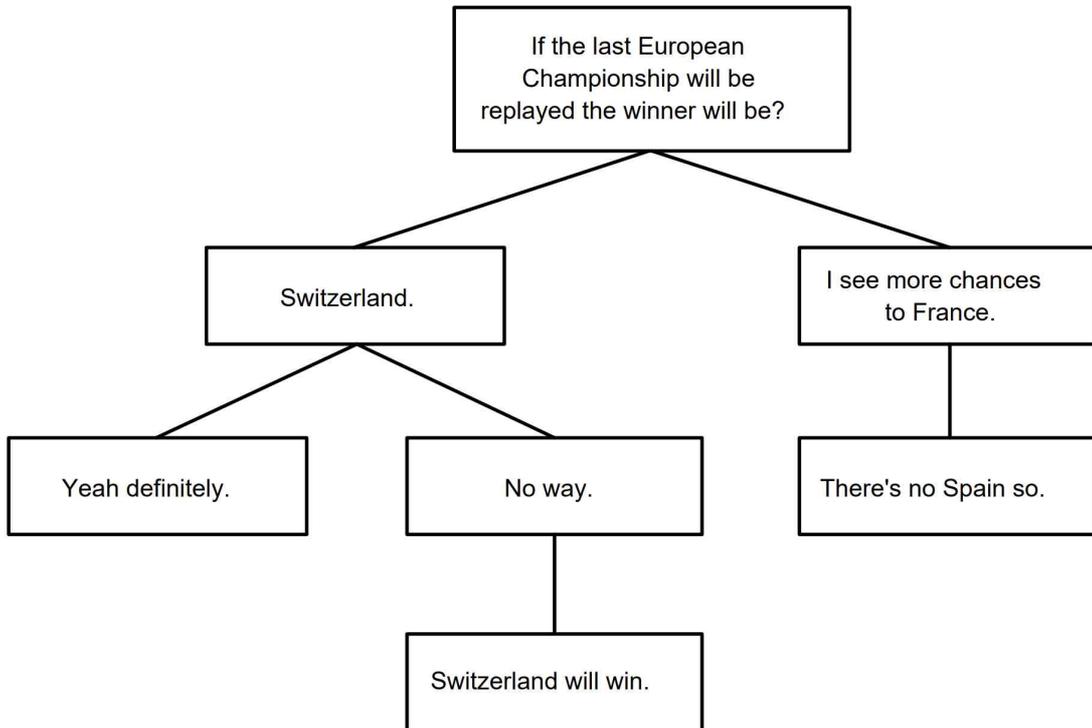
본 문서에서는 논쟁적 대화(예: 회의록)에서 나타나는 논변(논쟁, Argumentation)을 구조화하기 위한 논변 구조(argument structure)의 주석 체계(1장)와 주석 예시(2장)를 다룬다.

논변이란 어떤 의견에 대해서 자신의 견해를 서술하고 주장하는 것으로, 대화에서는 대화의 참가자가 주제에 대한 의견을 논하고 진술하는 것을 의미한다. 본 문서에서 언급하는 논변 구조는 대화 속에서 이루어지는 의견을 주고받는 담화의 구조(discourse structure)를 논변의 측면에서 분석하기 위한 일종의 스키마이다. 예를 들어, 회의의 참가자들이 각자의 의견(opinion)에 대해 동의하거나 비동의를 하는 상황이 논변 구조의 분석 대상이 된다. 다음은 논변 구조의 예시이다⁸⁾:

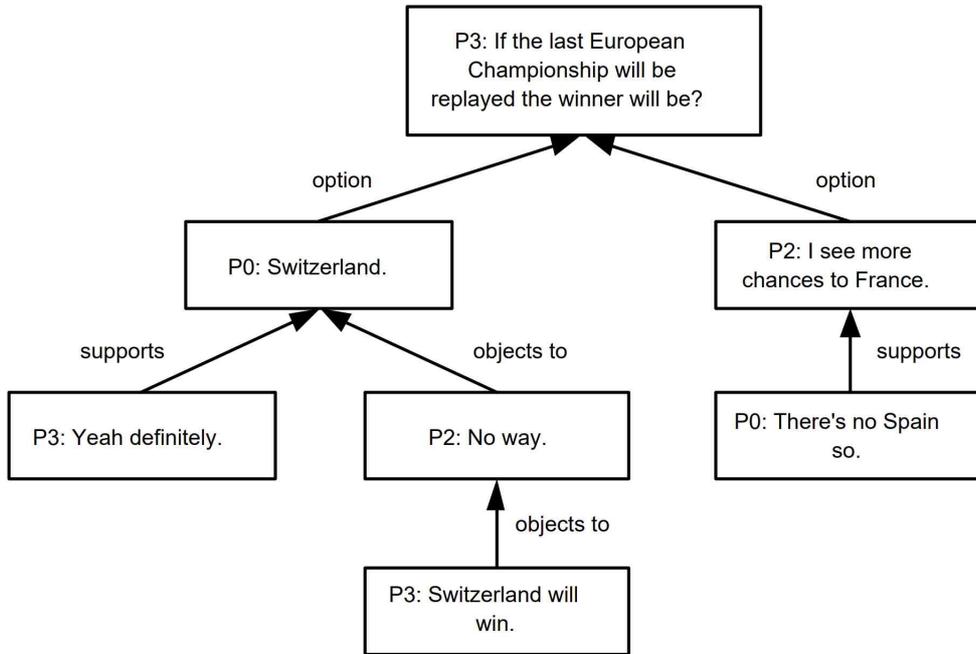
8) 본 문서에 등장하는 예시의 일부는 다음 논문에서 인용하였다: "Structuring argumentation in meetings Visualizing the argument structure", Eric van der Weijden, Master Thesis, 2005

id	화자	발화
1	P3	If the last European championship will be replaced the winner will be?
2	P0	Switzerland.
3	P0	Yeah definitely.
4	P0	No way.
5	P0	Switzerland will win.
6	P0	I see more chances to France.
7	P0	There's no Spain so.

논변 구조는 위와 같은 대화에 대해 구조화하여 시각화할 수 있는 스키마를 의미한다. 이 경우 발화 간의 관계를 그래프로 그린다면 아래와 같을 것이다.



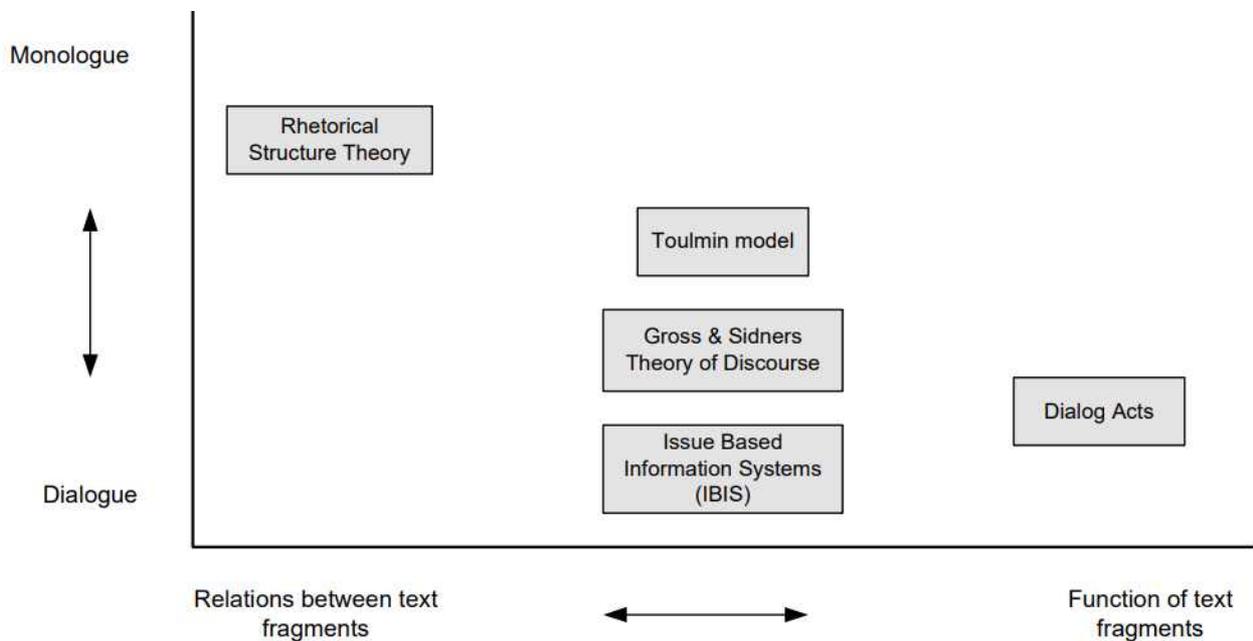
위와 같은 그림은 논쟁에서 이루어진 각각의 발화를 트리 구조(tree structure)로 표현한 것으로, 각각의 발화는 독립적으로 이루어진 것이 아니라 앞선 발화에 대한 대답, 질문, 문제제기, 동의, 비동의 등의 다양한 의미로 결합되어 있다. 논변 구조는 이렇게 대화에서 이루어진 각각의 발화를 그래프의 노드(node)로, 그리고 각 발화 간의 관계(relation)를 명시하는 것을 목적으로 한다. 위 그림에 대해, 각각의 발화의 관계를 추가한다면 아래와 같은 그림이 될 수 있다.



위 그림과 같이, 하나의 논변 구조는 주어진 문제(발화 id 1)에 대한 의견제시나 답변들이 있으며 이 경우에는 Option의 관계로 주석되었다. 또한 각각의 답변에 대해 지지(support)하거나 반대(objects to)하는 경우에도 주석되었다. 이와 같이 논변 구조의 구성 요소는 크게 논의의 요소(argument element)와 논의의 관계(argument relation)로 구성되며, 각 논의 요소는 발화 단위로 태깅되며 각 논의 요소는 논의 관계로 태깅된다. 본 문서에서 다루는 논변 구조는 상기와 같은 조건하에서 기술될 것이다.

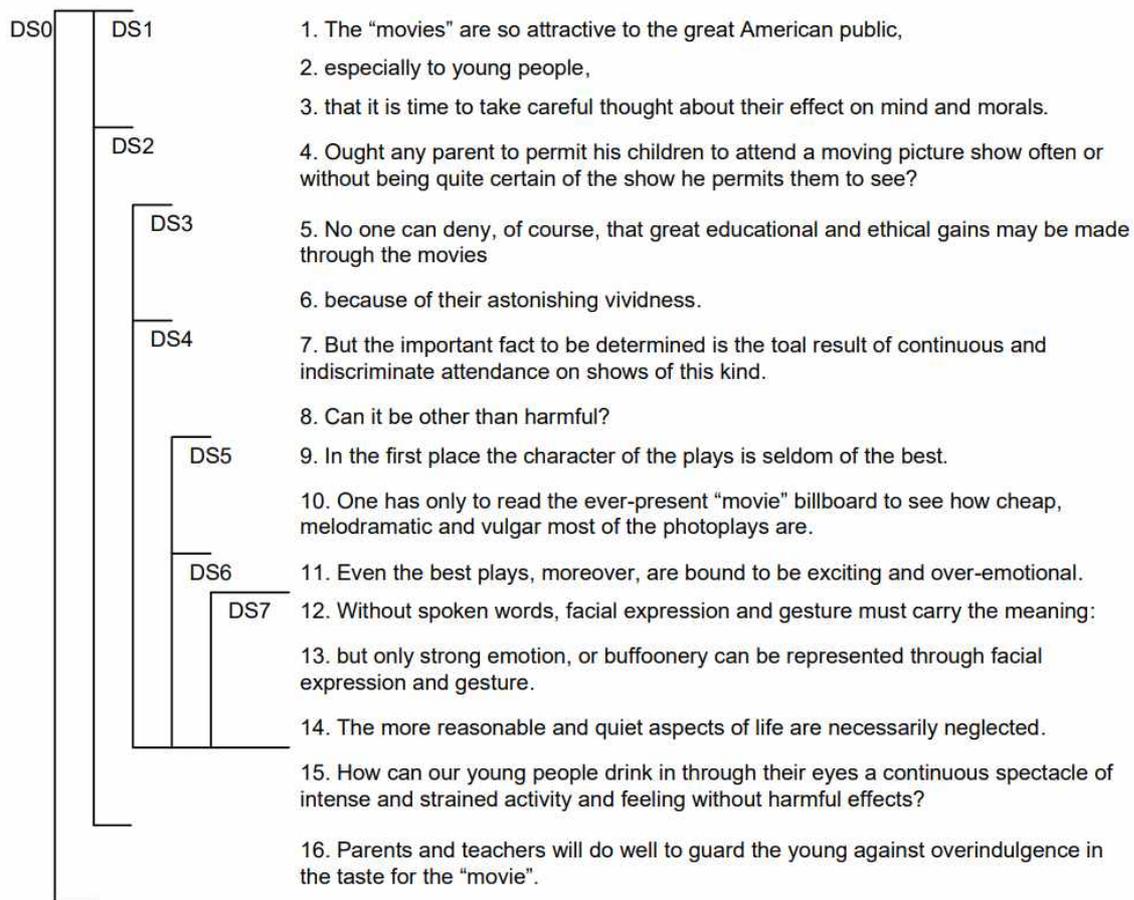
1.2 관련 연구

기존의 연구에서는 이러한 논변 구조를 분석하기 위하여 다양한 연구가 수행되었다. 그러나 대부분 도서, 논문과 같은 문서에 대한 수사 구조(Rhetorical Structure) 연구에 치중되어 있다. 반면 대화의 경우에는 발화의 의도와 내용을 이해하는 연구(Natural Language Understanding, NLU)가 진행되고, 맥락에서의 발화의 화행(Dialog Act)을 분석하는 연구가 수행된 바 있다.

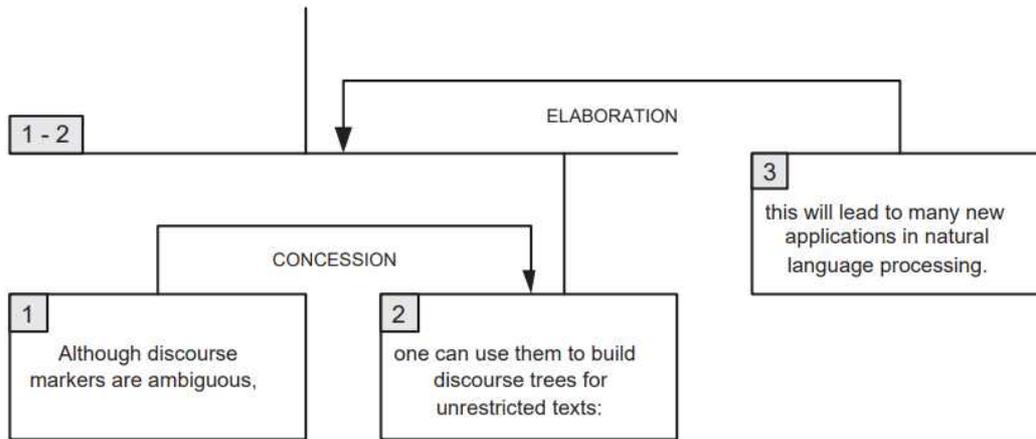


<텍스트 구조 이해를 위한 이론 및 연구 위치>

예를 들어, 담화(Discourse)의 측면에서는 Grosz, Sidner (1987)의 Discourse Theory의 연구가 대표적인데, 이는 담화에서 나타나는 언어학적 구조, 발화의 의도 구조, 발화가 주목하는 상태 등을 분석하는 것에 목적이 있다. 아래 예시는 담화 이론(Discourse Theory)을 시각화한 예시이다. 아래 예시와 같이, 하나의 담화는 세부적인 부분으로 나누어지고, 각각의 부분은 전체 구조하에서 관계를 갖는 구조이다. 이 경우 수사 구조의 경우에는 보다 스키마와 구조에 집중하여 각각의 발화들이 갖는 관계를 구조화하는 것에 초점을 맞추었다.

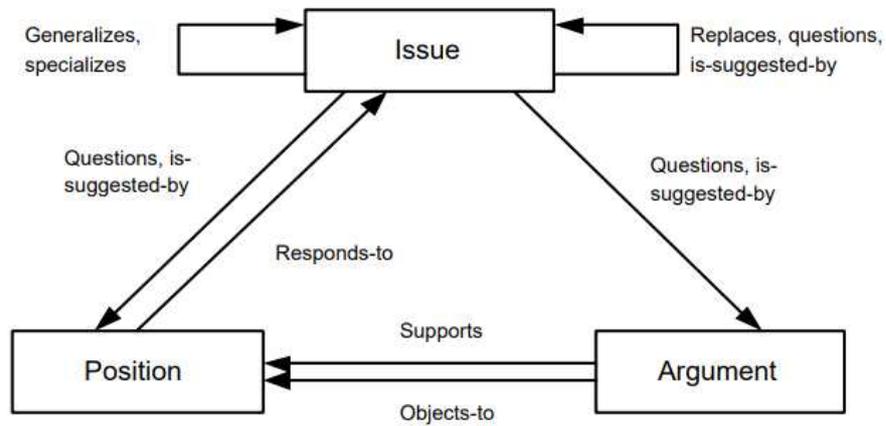


<Discourse Structure의 예. 각 발화 간의 관계를 주석하였다. B.J. Grosz and C.L. Sidner, 1986, "Attention, intentions, and the structure of discourse".>



<Rhetoric Structure의 예. 각 발화 간의 관계를 주석하고, 각 역할을 레이블링하였다. D. Marcu, 1997, "The rhetorical parsing of natural language texts">

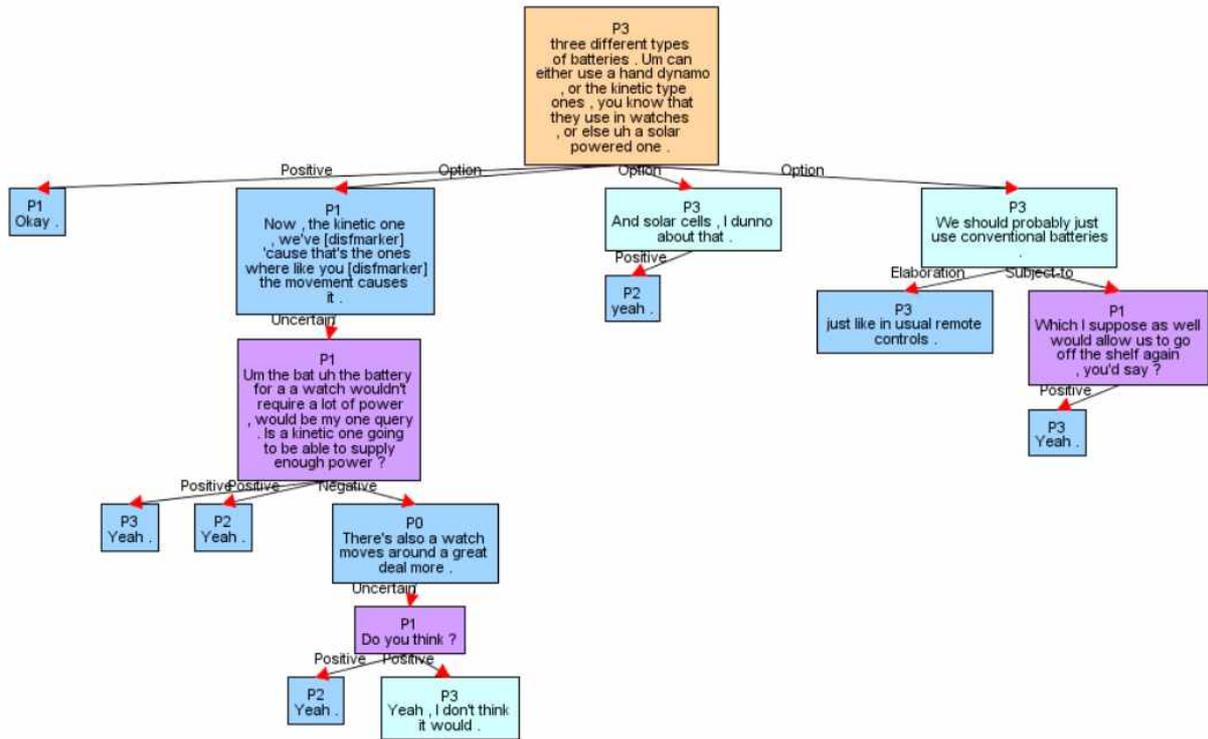
본 문서에서는 논쟁적 대화, 즉 각각의 주제에 대하여 질문과 답변, 지지와 부정, 요청과 설명 등의 다양한 속성의 발화들로 구성된 회의를 다루며, 해당 연구만으로는 충분히 논변 구조를 분석하기에는 불충분하다. 본 연구에서는 다음 두 가지의 논변 구조 스키마를 참고하여 하나의 통일된, 그리고 표준화된 논변 구조 스키마를 제안하고 이를 바탕으로 한 주석 지침을 만들고자 한다. 참고할 하나는 Isenmann, Reuter (1997)의 토론 구조를 바탕으로 한 IBIS(Issue based Information Systems)의 논변 구조이다. 해당 모델은 이슈에 대해 일반화/구체화 및 질문과 제안, 응답, 지지, 반박 등으로 구성된 모델로서 이에 대한 예시는 아래와 같다. 아래 예시와 같이, 각각의 논변 요소들은 상호 간에 지지하거나 이슈에 대해 설명, 반박하는 등의 구조로 되어 있다.



<IBIS 주석의 예.

https://en.wikipedia.org/wiki/Issue-based_information_system>

다른 참고 연구는, Twente Argument Schema(TAS)이다. TAS에서는 Dialog Act의 Accept, Reject, Response, Suggest 등의 구체화된 논변 관계 레이블을 단순화하여 보다 일반적으로 적용될 수 있는 레이블(예: Positive, Negative)을 사용하였다는 특징이 있다. 이를 통해 회의에서 나타나는 논쟁에 대하여 논변 구조를 파악하고 주석하기에 용이하다는 장점이 있다. 이에 대한 예시는 아래와 같다.

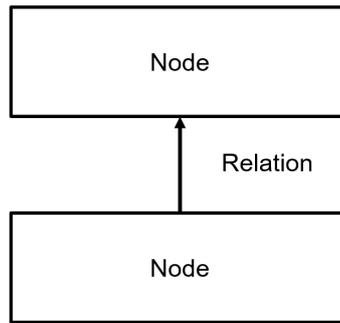


<TAS 주석의 예>

본 연구에서는 IBIS의 구체화된 스키마와 TAS의 일반화된 스키마를 참고하여, 국회 회의록에 적용될 수 있는 주석 체계를 구축하는 것을 목표로 하여 논쟁적 대화 요약을 위한 분석 자료로서 논변 구조를 주석하는 지침을 만드는 것을 목표로 한다.

1.3 논변 구조 주석 체계

논변 구조는 1.1장에서 논의된 바와 같이, 다음과 같은 주석 체계를 갖는다.

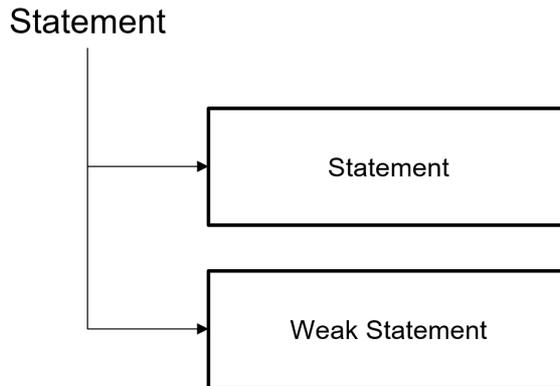


본 문서에서 논의하는 논변 구조는 다음과 같은 특징이 있다.

(1) 대화 내 모든 발화가 논변 구조에 속하지 않으며, 특정 사안에 대해 논의가 이루어지는 발화들이 논변 구조의 노드가 되며, 또한 해당 발화들 사이에 논변 관계가 있는 경우에만 논변 구조에 속한다. 다시 말해, 한 명의 화자에 의한 일방적인 주장/연설 등은 담화(discourse) 혹은 화행(dialog act)의 분석이 될 수 있으나 논변 구조의 구성 요소에 해당하지 않는다.

(2) 상기의 원칙을 따라, 동일 화자의 연속적인 발화들은 서로 관계를 갖지 않는다. 오직 다른 화자와 논의가 이루어지는 발화만이 논변 구조의 노드가 되며, 이에 따라 해당 발화와 다른 화자의 발화만이 논변 관계를 가질 수 있다.

□ 노드 레이블의 유형 1. Statement

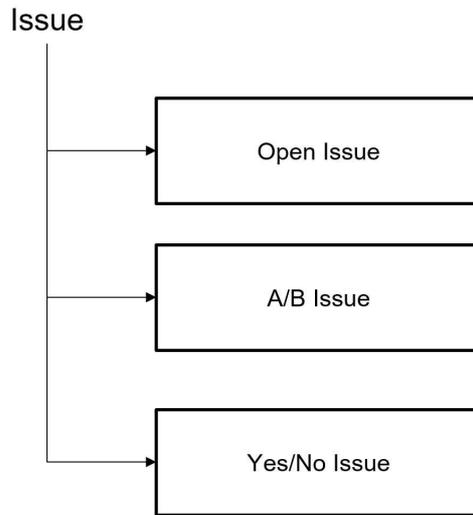


기본적으로 논쟁적 발화에서 문제를 제기하는 Issue 관련 발화를 제외한 모든 발화는 Statement와 Weak Statement로 구성된다. Statement는 논쟁적 대화에서 논변 구조에 속하는 약한 발화를 제외한 모든 발화이며, Weak Statement는 논변 구조에 속하는 약한 발화이다. 약한 발화는 간단한 답변 정도에 해당한다.

Statement 예시	
A/B Issue	○손숙미 위원 그 부분은 매해 반복이 되는 것 같거든요. 연구자가 프로젝트 계획서를 늦게 내서 그런 겁니까, 아니면 평가가 지연돼서 그런 것입니까?
Statement	○김태석 여성부기획조정실장 작년에 조금 늦었던 취업지원통합사이트 구축 부분은 원래 보육업무가 있을 때 고객만족센터를 설치하려던 예산이 보육업무가 넘어감으로써 민원이 적어져서 수요가 없어짐에 따라서 새로운 사업을 하게 되어서 조금 늦어지게 됐고, 또한 그것은 지난해 경단법 시행이 12월 말에 됨으로써 일정이 좀 늦어지게 됐습니다.
Statement	○김태석 여성부기획조정실장 특별히 지난해에는 새 정부 들어서면서 전체적으로 조금 늦게 시작된 측면이 있습니다.
O p e n Issue	○손숙미 위원 보통 연구기간은 얼마나 잡습니까?

Weak Statement 예시
좋아요.
아마도요.
네, 저도요.
아닙니다.
예, 그렇습니다.
예.

□ 노드 레이블의 유형 2. Issue



Issue 발화는 논쟁적 대화에서 문제를 제기하는 발화를 의미한다.

- Open Issue 발화는 논쟁적 대화에서 특정 이슈를 제기하는 질문 혹은 발화이다. A/B Issue, Yes/No issue와 같이 제약조건 없이 가능한 모든 대답이 해결책이 될 수 있는 경우에 국한된다.

Open Issue 예시
주의보다 더한 것을 받아야 되는데, 신 전문위원 잘 지적해 가지고 오늘 이렇게, 이걸 주의를 훨씬 뛰어넘는 강한 페널티를 받고 그다음에 법에 따라서 해야 될 거라고 생각하는데 그래도 주의 정도를 주신 것 보면 많이 참작해서 주신 것 같은데, 김해영 위원님 여러 가지 또 많이 배려하신 것 같은데 제가 볼 때는 이걸 주의보다 더 강한 페널티를 줘야 된다 이런 지적을 할 수 있는데, 아무튼 지금 시정요구유형 주의까지 받았다고 하셨으니까 하되 영거주춤하게 그러지 말고 아주 쿨하게 이 부분은 잘못됐다 인정하고 다시 법률에 근거해서 하든지 아니면 법이 개정될 때 그때 맞춰서 하는 것이 옳다 이런 판단을 합니다. 어떠세요?
그다음에 유예하면 1년 동안 이것 해결할 수 있습니까? 또 똑같은 것 들고 나올까 봐 지금 걱정하는 것도 있는 거거든요. 그런 것을 한번 얘기를 해 보시라고요, 계획을 제출하든가.
아니, 교육부에서 국정교과서에 대해서 권한을 가지고 고칠 수 있다 그 얘기하고 교육부에서 교과서를 고치기 위해서 관제 여론을 만들고 책임 집필자의 동의 없이 이 교과서가 고쳐지도록 할 수 있다라는 게 같은 얘기입니까?

그러나 모든 질문 형식을 띠다고 해서, 모든 질문이 Open Issue 발화에 해당하지는 않는다. 예를 들어 아래의 경우에는 Open Issue가 아닌 다른 노드 레이블로 분류된다.

예시	Node Label
아니, 우리 자료하고 다른 내용을 지금 얘기하시려는 거예요, 아니면 내용 중에 일부 승인된 부분에 별도 설명을 하시려고 하는 거예요?	A/B Issue
상담사라는 것을 제가 잘 몰라서 물어봤더니 집중기에 단기 상담사가 지금 필요하다라는 것 아닙니까?	Yes/No Issue
여러 가지 여론이 뭐가 들어왔습니까? 여러 가지 여론이 뭐가 들어와요?	Statement

- A/B Issue 발화는 답이 명시적으로 열거된 이슈를 제기하는 발화이다.

A/B Issue 예시
아니, 우리 자료하고 다른 내용을 지금 얘기하시려는 거예요, 아니면 내용 중에 일부 승인된 부분에 별도 설명을 하시려고 하는 거예요?

- Yes/No Issue 발화는 쟁점과 관련이 있으면서 예/아니오로 답할 수 있는 이슈에 해당한다. Yes/No Issue 발화에 대한 답변은 Positive/Negative/Uncertain의 관계로만 연결될 수 있다.

ID	예시	node label	relation label	relation ID
1.5.2.1	그러면 그 입장 차이들이 있는데 협의체 구성할 때 이 두 노조들 각각 다 협의체 구성해서 논의하실 건가요?	YES/NO Issue		
1.5.2.2	예, 그렇습니다.	Weak Statement	Positive	1.5.2.1

ID	예시	node label	relation label	relation ID
1.5.3.1	21 층 이상 건물의 경우는 지금은 광역단체장의 승인인데 교육감의 승인하고 승인을 2 개 받아야 되는 것은 맞는 것입니까? 그러니까 그 영향평가 대상이 되면 교육감의 승인을 받아야 되는 것 아니에요?	YES/NO Issue		
1.5.3.2	예, 환경평가의 승인을 받도록 되어 있습니다.	Statement	Positive	1.5.3.1

□ Relation 레이블의 유형 1. Positive

뒷 발화가 앞 발화를 뒷받침, 지지, 긍정하는 경우에 사용된다.

ID	예시	node label	relation label	relation ID
2.1.3.1	전문위원님 아까 1년으로 말씀하셨는데 그게 굳이 2년까지 걸려서 해야 될 일인지도 좀 의문이고(...) 근본적인 문제들을 어떻게 해결할 것인가에 대한 답이 좀 나와야 될 것 같아요. 그래야지 이게 의견도 접근이 되는 것 아니겠습니까?	Open Issue		
2.1.3.2	위원님 말씀이 맞습니다. 그래서 저희가 두 가지 측면에서 지금 유예기간 동안에 준비를 하려고 합니다.	Statement	Positive	2.1.3.1

□ Relation 레이블의 유형 2. Negative

뒷 발화가 앞 발화를 반박하는 경우에 사용된다.

ID	예시	node label	relation label	relation ID
2.2.2.1	그래서 저희가 대안을 낸다면, 전문위원님 말씀하고 위원님 말씀 받아서 대안을 낸다면 여기 시정요구사항에 이렇게 되어 있는 것을 '교육부는 현 단계에서는 국가진로교육센터의 성인진로교육 관련 예산 집행을 최소화하고 향후 성인진로교육 지원의 법적 근거를 마련하도록 할 것' 이런 정도만 해 주신다면 저희들이 법을 고쳐 가면서 하면 되겠다는 말씀을 드리겠습니다.	Statement		
2.2.2.2	법에 어긋나는 부분을 최소화시키고 이런다는 것 자체가 모순이 되잖아요. 이 부분은 주의를 받고 빨리 시정해서 다음에 법이 마련되면 그 근거법에 따라서 하는 게 옳은 것 아니에요?	Open Issue	Negative	2.2.2.1

□ Relation 레이블의 유형 3. Uncertain

Yes/No Issue에 대한 답변 중 Positive 또는 Negative로 주석하기 어려운 경우에만 사용된다. 그 이외의 경우에는 사용하지 않는다.

ID	예시	node label	relation label	relation ID
2.3.1.1	그런데 이 부분에 대해서 법적인 검토를 해 왔는데 여성발전기본법상 여성정책기본계획의 변경 사항과 관련해서는 시점은 맞았다고 그랬지요? 어쨌든 이것을 변경하고 나서 이렇게 했다고 한 것이지요? 맞습니까?	Yes/No Issue		
2.3.1.2	그것은 확인을 해 보겠습니다. 시점은 12월에 동시에 이루어졌기 때문에.....	Statement	Uncertain	2.3.1.1

□ Relation 레이블의 유형 4. Specialization

뒷 발화가 앞 발화의 전문화된 발화일 경우 사용. 이슈를 재해석하여 앞 발화에 대한 답변을 제한하거나 다른 답변(해결책)을 유도하는 경우.

ID	예시	node label	relation label	relation ID
2.4.1.1	몰도바의 수도는 어디인가?	Open Issue		
2.4.1.2	가장 몰도바인처럼 들리는 것은 무엇인가?	Open Issue	Specialization	2.4.1.1

□ Relation 레이블의 유형 5. Elaboration

발화에 대한 추가정보가 제공된 경우.

ID	예시	node label	relation label	relation ID
2.5.2.1	상세보고라는 게 뭐예요?	Statement		
2.5.2.2	저희가 협의체 구성하겠다는 부분하고요, 그다음에 프라임사업은 다음 주에 발표를 합니다. 프라임사업도 그렇고 이번 주 화요일 날, 엇그제 발표했던 인문학 강화사업, 그러니까 코어사업에서도 평가지표에 기본적으로 시간강사들에 대한 평균 임금 부분이 지표로 들어갔습니다. 그래서 그런 여러 가지 평가지표들을 통해서 대학들의 변화를 유도하는 부분들이 있고요.	Statement	Elaboration	2.5.2.1

□ Relation 레이블의 유형 6. Request

발화에 대한 자세한 정보를 요청하거나 다시 말해주기를 요청하는 경우.

ID	예시	node label	relation label	relation ID
2.6.2.1	저희가 조금 이따 상세히 보고드리는 부분에 대해서 위원님들께서 한번 판단해 보시고 특위 구성에 대해서는 다시 한 번 논의해 주셨으면 하고 생각합니다.	Statement		
2.6.2.2	상세보고라는게 뭐예요?	Statement	Request	2.6.2.1

□ Relation 레이블의 유형 7. Option

앞 발화에 대한 답변, 옵션, 솔루션이거나 아이디어를 제공하는 경우 사용.

ID	예시	node label	relation label	relation ID
2.7.2.1	그래서 저희들이 도종환 의원님의 수정을 수용을 하되 명확히 저희 정부 입장은 이것을 영으로 받아서 영에서 법인으로 명시하겠다는 것을 사전에 오해를 풀기 위해서 명시를 한다는 전제하에 의원님의 수정안을 수용을 하겠습니다.	Statement		
2.7.2.2	지난번에 차관님이 미국의 스미소니언 박물관의 사례 같은 것을 들면서, 재정을 안정적으로 계속 지원하는 그런 사례 같은 것들을 참조하면서 책임을 질 수 있도록 하겠다고 말씀하셨는데 미국처럼 정말 기부문화가 활성화되고 자체 모금도 막 성공적으로 되고 보잉 같은 기업이나 개인들이 우리 돈으로 한 300억씩 수백억씩 막 이렇게 기부하는 그런 문화가 아닌 상황에서 문학에 대해서, 문학관에 대해서 누가 그렇게 많이 기부할 것인가를 생각해 보면 문화부가, 정부가 좀 더 적극적으로 책임지는 자세로 운영을 하겠다고 하면 저도 시행령에 법안을 담는 것 이런 것 등을 동의할 수 있다는 말씀을 드립니다.	Statement	Option	2.7.2.1

2. 논변 구조 주석 가이드라인

2.1 NODE label

2.1.1 Statement

논쟁적 발화에서 논변 구조에 속하는, 약한 발화를 제외한 모든 발화

- 1) 회의진행 발언의 경우 앞뒤 발화와 relation label을 달지 않음.
- 2) 회의록 앞부분의 보고, 의견은 대부분 Statement로 분류함.
 - 이 경우에는 뒷 발화에 따라 relation label을 주석할 수 있음.

예시
<p>국·공립대의 경우에는 저희가 실제로 재정을 투입해서 올린 부분입니다. 그래서 재정을 투입해서 올리는 게 가능하기 때문에 그것을 좀 선도적으로 국·공립대가 먼저 가고 그것을 따라오기를 기대했는데 아직은 부족합니다. 국·공립대학이 지금 8만 원 수준이고요. 사립대가 지금 5만 원 수준입니다. 그래서 그 갭을 좀 줄일 수 있도록 저희가 평가지표에도 넣고 기본적인 처우 이러한 부분에 상세한 재정사업도 만들고 협의체를 구성해서 운영을 하겠다는 말씀을 일단 보고드리겠습니다.</p>
<p>(...)그래서 1만 3000원이 증액이 되고 국립대 시간강사 지원하는 예산이 올해도 1060억 해서 예산이 반영돼 있습니다. 그래서 국립대학은 유예기간 동안에 많이 개선이 되었고요. 사립대학은 저희 정부가 재정지원을 할 수 없기 때문에 지금 차관님 말씀하신 대로 정부 각종 재정사업을 통한 유도지표를 통해서 저희가 실질적인 처우개선을 유도해 왔습니다. 그래서 사립대학도 그동안 4만 원 수준에서 지금 차관님 말씀하신 대로 5만 원까지 1만 원이 올라가 있는 상태입니다. 가령 예를 들어 드리면 이번에 대학구조개혁 평가지표에도 사립대학으로 하여금 유도하기 위해서 시간강사 처우개선 부분을 지표로 반영했습니다. 이미 그렇게 돼 있고 앞으로 신설되는 코어나 프라임 등을 통해서도 저희가 계속 유도지표를 넣는다면 사립대학들도 계속적으로 강사 처우개선을 할 것으로 보고 있습니다. 코어하고 프라임 신규 사업이 추진되는 기간을 앞으로 2년으로 볼 경우에 저희가 2년을 주시사하는 그런 부탁을 드리는 겁니다.</p>
<p>위원님, 저희가 안 그래도 로드맵을, 2년을 주실 경우에 생각하고 있는 타임라인을 별도로 보고 올리도록 하겠습니다. 그래서 저희가 준비하고 있는 것은 협의체를 바로 구성해 가지고요. 이번에는 어떻게 보면 옛날의 유예하고 달리 저희가 정부 대책을 확실히 만든다는 조건하에 유예를 시켜 주시는 거니까요. 저희가 바로 협의체를 구성해서 6월 중에, 상반기 안에는 대책 보고를 올릴 수 있도록 준비를 하겠습니다.</p>
<p>그러니까요. 그래서 그 제도개선을 큰 틀에서 보시고 그 큰 틀 속에서 제대로 된 평가, 제대로 된 인센티브 그리고 그것이 페널티 중심이 아니라 교원사회에 활력을 불어넣는 제대로 된 인센티브제도가 될 수 있도록 그렇게 제도개선을 하는 방향으로 교육부가 설계를 하시라는 겁니다.</p>
<p>그러면 여기에 대해서 시정요구 수준을 어떤 식으로 선택하실 것인지 위원님들께서 의견을 주시면 감사하겠습니다.</p>
<p>지금 조직적 개입이라는 말씀을 쓰셨는데요. 직권으로 수정하는 것은 위임 전결상에 따</p>

<p>라서 누가 국장 지시에 의해서 하든 아니면 과장 지시에 의해서 하든 잘못된 것을 고치는 것은 항상 있을 수 있는 일입니다.</p>
<p>좌석을 정돈해주시기 바랍니다.</p>
<p>정부 측 발언하세요.</p>
<p>이것과 관련해서 여성부에서 의견이 있으십니까?</p>
<p>박기풍 차관, 정부 측 의견 있으면 말씀해 주시기 바랍니다.</p>

*Statement 간에 Relation label을 주석하는 경우

발화자	ID	예시	node label	relation label	relation ID
A	1	이게 과하다는 것이지 부당한 것과는 다르잖아요.	Statement		
B	2	왜냐하면 예산 10%를 일률적으로 절감을 해 가지고, 국회에서 예산을 심의해서 확정된 것을 부처에서 마음대로 한 것이거든요.	Statement	Elaboration	1
C	3	그것을 일률적으로 했다고 얘기할 수는 없는 것이고 부처에서 나름대로 판단해서 한 것인데 우리가 일률적인 사항이라고 판단하기에는 사실 쉽지 않습니다.	Statement	Negative	2
B	4	정부에서 10%라고 하는 일률적인 적용을 했다는 겁니다.	Statement	Negative	3
A	5	제가 보기에는 이것은 부적절까지는 해당이 되는데 위법이나 부당한 것은 아니라는 것이지요.	Statement	Negative	4

3) Relation label을 주석하는 경우에는 가장 가까운 발화와 연결함.

회의진행 발화인 4, 5번 발화를 제외하면 6번 발화와 관계된 발화는 1~3번 발화이기 때문에 그중 가장 마지막 발화인 3번과 연결함.

발화자	ID	예시	node label	relation label	relation ID
A	1	첫 번째로 세입 수납률 제고조치 필요에 관한 사항입니다.	Statement		
A	2	이것은 세입 징수 결정된 금액 6억 5400 만 원 중 수납액은 57.5%로서 징수율이 감소해서 정산 잔액이 조속히 반납될 수 있도록 조치해 달라는 내용입니다.	Statement		
A	3	그러면 여기에 대해서 시정요구 수준을 어떤 식으로 선택하실 것인지 위원님들께서 의견을 주시면 감사하겠습니다.	Statement		

B	4	위원님께서 말씀해 주시지요.	Statement		
C	5	이것과 관련해서 여성부에서 의견이 있으십니까?	Statement		
D	6	이것과 관련해서는 서울시가 3건이 되고 그 외 강원도 등의 미수납액입니다. 지자체 추경예산에 미편입이 되었는데 지금 현재는 모두 다 납부가 됐습니다. 앞으로도 세입고지서 발부 및 독촉을 지속적으로 해 가지고 수납률이 제고될 수 있도록 하겠습니다.	Statement	Option	3

발화자	ID	예시	node label	relation label	relation ID
A	1	시정은 '법령상·사실상 시정이 가능할 것'이라고 했거든요. 그러니까 사실상 시정이 가능한 것이지요.	Statement		
B	2	이·전용 문제는 좀 짚고 넘어가야 됩니다.	Statement	Negative	1
A	3	참고로 어제 환노위에서 노동부와 환경부 결산을 했는데 이것은 시정으로 했습니다.	Statement	Elaboration	1

2.1.2 Weak Statement

논쟁적 대화에서 논변 구조에 속하는 약한 발화

예시
좋아요.
아마도요.
네, 저도요.
아닙니다.
예, 그렇습니다.
예.
예, 주의를 받겠습니다.

2.1.3 Open Issue

논쟁적 대화에서 특정 이슈를 제기하는(쟁점과 관련된) 질문 혹은 발화. A/B Issue, Yes/No issue와 같이 제약조건 없이 가능한 모든 대답이 해결책이 될 수 있는 경우에 한함.

*간단한 질문(또는 발화)이더라도 쟁점과 관련된 질문 또는 이슈를 제기하는 발화일 경우 open issue로 분류함.

예시
주의보다 더한 것을 받아야 되는데, 신 전문위원 잘 지적해 가지고 오늘 이렇게, 이견 주의를 훨씬 뛰어넘는 강한 페널티를 받고 그다음에 법에 따라서 해야 될 거라고 생각하는데 그래도 주의 정도를 주신 것 보면 많이 참작해서 주신 것 같은데, 김해영 위원님 여러 가지 또 많이 배려하신 것 같은데 제가 볼 때는 이견 주의보다 더 강한 페널티를 줘야 된다 이런 지적을 할 수 있는데, 아무튼 지금 시정요구유형 주의까지 받았다고 하셨으니까 하되 엉거주춤하게 그러지 말고 아주 쿨하게 이 부분은 잘못됐다 인정하고 다시 법률에 근거해서 하든지 아니면 법이 개정될 때 그때 맞춰서 하는 것이 옳다 이런 판단을 합니다. 어떠세요?
그다음에 유예하면 1년 동안 이것 해결할 수 있습니까? 또 똑같은 것 들고 나올까 봐 지금 걱정하는 것도 있는 거거든요. 그런 것을 한번 얘기를 해 보시라고요, 계획을 제출하든가.
아니, 교육부에서 국정교과서에 대해서 권한을 가지고 고칠 수 있다 그 얘기하고 교육부에서 교과서를 고치기 위해서 관제 여론을 만들고 책임 집필자의 동의 없이 이 교과서가 고쳐지도록 할 수 있다라는 게 같은 얘기입니까?
법에 어긋나는 부분을 최소화시키고 이런다는 것 자체가 모순이 되잖아요. 이 부분은 주의를 받고 빨리 시정해서 다음에 법이 마련되면 그 근거법에 따라서 하는 게 옳은 것 아니에요?
그런데 국가재정법 위반 여부나 이런 부분을 판단해서 해야 되는데 지금 국립대학 회계법 시행규칙 그다음에 국가재정법 제46조 이런 부분들이..... 국가재정법보다 더 우위에 있나요, 아니면 어떤가요?
주의가 아니라 시정으로 해야하는 것 아니에요?

***예외사항**

아래 예시는 질문형식을 띠고 있으나 Open issue에 해당하지 않는 경우.

예시	node label
아니, 우리 자료하고 다른 내용을 지금 얘기하시려는 거예요, 아니면 내용 중에 일부 승인된 부분에 별도 설명을 하시려고 하는 거예요?	A/B Issue
상당사라는 것을 제가 잘 몰라서 물어봤더니 집중기에 단기 상당사가 지금 필요하다는 것 아닙니까?	YES/NO Issue
여러 가지 여론이 뭐가 들어왔습니까? 여러 가지 여론이 뭐가 들어와요?	Statement

2.1.4 A/B Issue

답이 명시적으로 열거된 이슈.

TAS-annotation-manual 예시
가장 똑똑한 동물은 무엇인가? A, B 또는 C?

예시
아니, 우리 자료하고 다른 내용을 지금 얘기하시려는 거예요, 아니면 내용 중에 일부 승인된 부분에 별도 설명을 하시려고 하는 거예요?

2.1.5 YES/NO Issue

쟁점과 관련이 있으면서 예/아니오로 답할 수 있는 이슈

*YES/NO Issue의 다음 발화는 Positive / Negative / Uncertain으로 분류하며 질문의 의도와 상관없이 예/아니오로 답할 수 있는 질문을 했다면 답변에 따라 예=positive, 아니오=negative로 주석함.

2.1.5.1

TAS-annotation-manual 예시
그 이름은 모두 가짜인가?

2.1.5.2

ID	예시	node label	relation label	relation ID
1.5.2.1	그러면 그 입장 차이들이 있는데 협의체 구성할 때 이 두 노조들 각각 다 협의체 구성해서 논의하실 건가요?	YES/NO Issue		
1.5.2.2	예, 그렇습니다.	Weak Statement	Positive	1.5.2.1

2.1.5.3

ID	예시	node label	relation label	relation ID
1.5.3.1	21층 이상 건물의 경우는 지금은 광역단체장의 승인인데 교육감의 승인하고 승인을 2개 받아야 되는 것은 맞는 것입니까? 그러니까 그 영향평가 대상이 되면 교육감의 승인을 받아야 되는 것 아니에요?	YES/NO Issue		
1.5.3.2	예, 환경평가의 승인을 받도록 되어 있습니다.	Statement	Positive	1.5.3.1

2.1.5.4

ID	예시	node label	relation label	relation ID
1.5.4.1	상담사라는 것을 제가 잘 몰라서 물어봤더니 집중기에 단기 상담사가 지금 필요하다는 것 아닙니까?	YES/NO Issue		
1.5.4.2	예, 정규 상담사가 약 350여 명 정도 있고요.	Statement	Positive	1.5.4.1

2.1.5.5

ID	예시	node label	relation label	relation ID
1.5.5.1	정부 예산에서 편성되어서 올라온 것도 아니지요? 아니, 그러니까 기재부에서 잘린 거지요?	YES/NO Issue		
1.5.5.2	최종 확정되는 시기가 조금 늦어져서 그것은 부득이하게 저희가 반영을 못 시켰습니다.	Statement	Positive	1.5.5.1

2.1.5.6

ID	예시	node label	relation label	relation ID
1.5.6.1	그러니까 이것을 오히려 제도화하는 게 맞는 거 아니냐 이거예요. 만약에 지난해만 그런 것이 아니고 계속해서 이런 일이 있었다면 한시적으로 신청 상담사가 필요하다 그러면 이것을 두는 것이 맞는 것 아닌가요?	YES/NO Issue		
1.5.6.2	예, 그래서 정규 상담사는 약 350여 명이 있고요 시기적으로 몰릴 때는 몇 달 동안 임시 상담사가 50명 정도 필요합니다. 그래서 이것은 저희들이 내역에 반영하도록 하겠습니다.	Statement	Positive	1.5.6.1

*쟁점과 관련이 있으면서 예/아니오로 답했을 때 어색하면 open issue로 분류함.

예시	node label	relation label
그러니까 소급이 안 되는 부분이, 지금 공사 중인 것뿐만 아니라 이미 승인받은 것은 어떻게 돼요?	Open issue	
적용되지 않습니다.	Statement	Option

*형식은 예/아니오로 대답할 수 있는 질문이지만 실제로 예/아니오 대답을 듣기

위한 발화가 아닌 경우.

*2번 발화의 경우 쟁점과 관련이 있으면서 본인의 의견을 어필하기 위한 질문이므로 Open Issue로 주석함.

발화자	ID	예시	node label	relation label	relation ID
A	1	참고로 어제 환노위에서 노동부와 환경부 결산을 했는데 이것은 시정으로 했습니다.	Statement		
B	2	이것을 한 게 위법입니까? 시정으로 하려면 위법 또는 부당한 사실이어야 되는데.....	Open Issue	Negative	1

2.2 RELATION LABEL

2.2.1 Positive

뒷 발화가 앞 발화를 뒷받침, 지지, 긍정하는 경우 사용.

2.2.1.1

ID	예시	node label	relation label	relation ID
2.1.1.1	나는 개미에게 투표한다.	Statement		
2.1.1.2	네 저도요.	Weak Statement	Positive	2.1.1.1

2.2.1.2

ID	예시	node label	relation label	relation ID
2.1.2.1	지난번에 차관님이 미국의 스미소니언 박물관의 사례 같은 것을 들면서, 재정을 안정적으로 계속 지원하는 그런 사례 같은 것들을 참조하면서 책임을 질 수 있도록 하겠다고 말씀하셨는데 미국처럼 정말 기부문화가 활성화되고 자체 모금도 막 성공적으로 되고 보잉 같은 기업이나 개인들이 우리 돈으로 한 300억씩 수백억씩 막 이렇게 기부하는 그런 문화가 아닌 상황에서 문학에 대해서, 문학관에 대해서 누가 그렇게 많이 기부할 것인가를 생각해 보면 문화부가, 정부가 좀 더 적극적으로 책임지는 자세로 운영을 하겠다고 하면 저도 시행령에 법안을 담은 것 이런 것 등을 동의할 수 있다는 말씀을 드립니다.	Statement		
2.1.2.2	예, 위원님 우려 말씀 주신 것 저희들도	Statement	Positive	2.1.2.1

	100% 공감을 하고요. 저희 문체부의 입장은 국립문학관이 문화융성의 주요한 견인차 역할을 꼭 해야 되고 또 그럴 당위성이 있다라는 것을 저희들도 100% 공감을 하고 있습니다.			
--	---	--	--	--

2.2.1.3

ID	예시	node label	relation label	relation ID
2.1.3.1	전문위원님 아까 1년으로 말씀하셨는데 그게 굳이 2년까지 걸려서 해야 될 일인지도 좀 의문이고(...)근본적인 문제들을 어떻게 해결할 것인가에 대한 답이 좀 나와야 될 것 같아요. 그래야지 이게 의견도 접근이 되는 것 아니겠습니까?	Open Issue		
2.1.3.2	위원님 말씀이 맞습니다. 그래서 저희가 두 가지 측면에서 지금 유예기간 동안에 준비를 하려고 합니다.	Statement	Positive	2.1.3.1

2.2.1.4

ID	예시	node label	relation label	relation ID
2.1.4.1	교육부 자꾸 그런 식으로 전부 다 얼렁뚱땅 해 오고 그러지 마세요. (...)이것은 시정하고 제도개선 쪽으로 아울러 같이 가야 될 것이라고 생각돼요.	Statement		
2.1.4.2	예, 그렇게 하겠습니다.	Weak Statement	Positive	2.1.4.1

2.2.1.5

ID	예시	node label	relation label	relation ID
2.1.5.1	필요하다고 그러면 이것을 제도화하는 것이 맞지 이것을 가지고 무슨 경고로 하고 시정으로 하느냐, 주의로 하느냐 하는 게 이게 맞아요?	Open Issue		
2.1.5.2	예산 내역에 반영해서 저희들이 제도화하도록 하겠습니다.	Statement	Option	2.1.5.1

2.2.1.6

ID	예시	node label	relation label	relation ID
2.1.6.1	(...)국회에는 없기 때문에 교육부를 통해서 되지 않을까 하는 생각이 듭니다.	Statement		

2.1.6.2	맞는 말씀입니다. 저희들도 해석하기에 국가 재정법이 있습니다만 특별법으로서 국립대학 회계법이 있고 저희들은 그 절차에 맞게 했다고 생각은 하고 있습니다.	Statement	Positive	2.1.6.1
---------	---	-----------	----------	---------

2.2.1.7

ID	예시	node label	relation label	relation ID
2.1.7.1	그러면 그 입장 차이들이 있는데 협의체 구성할 때 이 두 노조들 각각 다 협의체 구성해서 논의하실 건가요?	YES/NO Issue		
2.1.7.2	예, 그렇습니다.	Weak Statement	Positive	2.1.7.1

2.2.1.8

ID	예시	node label	relation label	relation ID
2.1.8.1	21 층 이상 건물의 경우는 지금은 광역단체장의 승인인데 교육감의 승인하고 승인을 2개 받아야 되는 것은 맞는 것입니까? 그러니까 그 영향평가 대상이 되면 교육감의 승인을 받아야 되는 것 아니에요?	YES/NO Issue		
2.1.8.2	예, 환경평가의 승인을 받도록 되어 있습니다.	Statement	Positive	2.1.8.1

2.2.1.9

ID	예시	node label	relation label	relation ID
2.1.9.1	상담사라는 것을 제가 잘 몰라서 물어봤더니 집중기에 단기 상담사가 지금 필요하다는 것 아닙니까?	YES/NO Issue		
2.1.9.2	예, 정규 상담사가 약 350여 명 정도 있고요.	Statement	Positive	2.1.9.1

2.2.1.10

ID	예시	node label	relation label	relation ID
2.1.10.1	정부 예산에서 편성되어서 올라온 것도 아니지요? 아니, 그러니까 기재부에서 잘린 거지요?	YES/NO Issue		
2.1.10.2	최종 확정되는 시기가 조금 늦어져서 그것은 부득이하게 저희가 반영을 못 시켰습니다.	Statement	Positive	2.1.10.1

2.2.1.11

ID	예시	node label	relation label	relation ID
2.1.11.1	그러니까 사업 내용이 과다하게 불용이 된 다거나 추계가 잘못됐다거나 이런 부분에서 삭감을 하는 것이지 일률적으로 사업비를 삭감을 하는 것은 아니잖아요?	YES/NO Issue		
2.1.11.2	일률적인 것은 아닙니다. 사업 내용에 따라서 했는데 그것도 삭감은 아니고 절감해 가지고 다른 용도로 사용을 한 겁니다.	Statement	Positive	2.1.11.1

*예/아니오로 대답한 뒤에 의견을 추가로 제시한 경우 Elaboration이 아니라 Positive, Negative로 분류함.

예/아니오로 대답한 경우에는 앞 발화가 높은 확률로 yes/no issue이기 때문에 질문-대답의 관계에 중점을 두고 주석함. 예/아니오 라고만 대답해도 되는데 추가적으로 의견을 제시한 것으로 보고 대답에 따라 positive 또는 negative로 주석함.

2.2.1.12

ID	예시	node label	relation label	relation ID
2.1.12.1	그러니까 이것을 오히려 제도화하는 게 맞는 거 아니냐 이거예요. 만약에 지난해만 그런 것이 아니고 계속해서 이런 일이 있었다면 한시적으로 신청 상담사가 필요하다 그러면 이것을 두는 것이 맞는 것 아닌가요?	YES/NO Issue		
2.1.12.2	예, 그래서 정규 상담사는 약 350여 명이 있고요 시기적으로 몰릴 때는 몇 달 동안 임시 상담사가 50명 정도 필요합니다. 그래서 이것은 저희들이 내역에 반영하도록 하겠습니다.	Statement	Positive	2.1.12.1

2.2.1.13

ID	예시	node label	relation label	relation ID
2.1.13.1	그렇게 모호하게 말씀하시면 안 되지요. 이것이 단체장의 권한을 침해하거나 뭔가 이 법 때문에 변경사항이 생기거나 하는 것이 아니지 않습니까?	YES/NO Issue		
2.1.13.2	예, 그렇습니다. 지금 다만 종래 해 왔던 것은 학교용지 선정 시에 환경영향평가가 있었고 그다음에 도시정비사업 그것이 있었는데 이번에 제정법을 하면서 추가된 것은 건축물에 관련된 사항이 추가되었기 때문에 말씀드린 것입니다.	Statement	Positive	2.1.13.1

2.2.2 Negative

뒷 발화가 앞 발화를 반박하는 경우 사용.

2.2.2.1

ID	예시	node label	relation label	relation ID
2.2.1.1	지금 폐기 쪽을 주장하시는 분들은 지금 말씀하신 대로 또 유예해 봐야 다른 대안이 있겠느냐 하는 그런 비관적 측면에서 폐기를 말씀하시는 거고요. 유예를 더 강하게 많이 원하십니다. 만약에 폐기할 경우에는 2011년도에 그 자살사건 이후에 그 당시 원점으로 돌아가는 것은 아니냐..... 그래서 그동안 유예하는 동안에 전혀 의미가 없었던 것은 아니다	Statement		
2.2.1.2	아주 부족하나마 개선되고 있다는 말씀을 하셨지만 이 법에 문제 제기가 되고 우리가 시간강사와 관련해서 이렇게 근본적인 해결책을 요구하고 있는 것 아닙니까? (...)전문위원님 아까 1년으로 말씀하셨는데 그게 굳이 2년까지 걸려서 해야 될 일인지도 좀 의문이고, 1년으로 정한다고 하면 역순으로 타임스케줄을 짜서 각각의 이견이 있는 당사자들과 이견들을 좁히면서 여기에서 제기됐던 근본적인 문제들을 어떻게 해결할 것인가에 대한 답이 좀 나와야 될 것 같아요. 그래야지 이게 의견도 접근이 되는 것 아니겠습니까?	Open Issue	Negative	2.2.1.1

2.2.2.2

ID	예시	node label	relation label	relation ID
2.2.2.1	그래서 저희가 대안을 낸다면, 전문위원님 말씀하고 위원님 말씀 받아서 대안을 낸다면 여기 시정요구사항에 이렇게 되어 있는 것을 '교육부는 현 단계에서는 국가진로교육센터의 성인진로교육 관련 예산 집행을 최소화하고 향후 성인진로교육 지원의 법적 근거를 마련하도록 할 것' 이런 정도만 해 주신다면 저희들이 법을 고쳐 가면서 하면 되겠다는 말씀을 드리겠습니다.	Statement		
2.2.2.2	법에 어긋나는 부분을 최소화시키고 이런다는 것 자체가 모순이 되잖아요. 이 부분은 주의를 받고 빨리 시정해서 다음에 법이 마련되면 그 근거법에 따라서 하는 게 옳은 것 아니에요?	Open Issue	Negative	2.2.2.1

2.2.2.3

ID	예시	node label	relation label	relation ID
2.2.3.1	학교 밖 학생도 학생이잖아요.	Statement		
2.2.3.2	학생 밖 청소년은 엄밀하게는 학생은 아닙니다.	Statement	Negative	2.2.3.1
2.2.3.3	우리가 법률에 명시된 학생이 학교에 다니는, 학교에 적응을 두고 있는 사람만 학생이라고 좁게 생각할 수도 있지만 또래에 있는 것도 우리가 학교 밖 학생이라고 본다면 하면 그 부분까지는 우리가 용인하고, 여기 보니 성인평생학습자 성인들에 대한 부분이 있기 때문에 그 부분은 학생으로 보기 어렵지 않느냐 이런 관점입니다.	Statement	Negative	2.2.3.2

2.2.2.4

ID	예시	node label	relation label	relation ID
2.2.4.1	그러니까요. 그런데 그게 실무자 선에서 이루어질 수 있는 정도의 행위 양태가 아니다. 워선에서까지의 아주 조직적 개입이 있을 수 있다라는 문제 제기도 계속 되고 있는 거예요.	Statement		
2.2.4.2	지금 조직적 개입이라는 말씀을 쓰셨는데요. 직권으로 수정하는 것은 위임 전결상에 따라서 누가 국장 지시에 의해서 하든 아니면 과장 지시에 의해서 하든 잘못된 것을 고치는 것은 항상 있을 수 있는 일입니다.	Statement	Negative	2.2.4.1
2.2.4.3	아니, 교육부에서 국정교과서에 대해서 권한을 가지고 고칠 수 있다 그 얘기하고 교육부에서 교과서를 고치기 위해서 관제 여론을 만들고 책임 집필자의 동의 없이 이 교과서가 고쳐지도록 할 수 있다라는 게 같은 얘기입니까?	Open Issue	Negative	2.2.4.2
2.2.4.4	책임 집필자는 동의가 필요한 게 아니고 오히려 계약 의무를 해태한 겁니다. 계약에서는 정부가 수정 지시를 하면 성실하게 따라야 될 의무가 있고요. 그것은 박용조 교수도 언론 인터뷰를 통해서 인정을 했습니다.	Statement	Negative	2.2.4.3

2.2.2.5

ID	예시	node label	relation label	relation ID
2.2.5.3	아닙니다.	weak statement		
2.2.5.4	아니, 제가 얘기하는 것은 교육과정 자체의 부당한 일들이라고 하는 부분은 포괄적 책임을 교육부가 져야 된다는 거예요. 그러면 '그	Open Issue	Negative	2.2.5.3

	건 출판사가 했기 때문에 나는 모른다' 그렇게 해서 변명하고 넘어갈 소지는 아니라는 거예요. 전체적으로 6학년 1학기 국정교과서를 세상에 내보낼 때 그 총체적인 부분은 교육부가 처음부터, 1부터 말변까지 전부 다 들여다보고 부당한 일이 있는지 없는지에 대한 부분은 교육부가 책임져야지 그렇게 무책임하게 '우리 교육부는 모른다' 그렇게 할 수 있는 거예요?			
--	--	--	--	--

2.2.2.6

ID	예시	node label	relation label	relation ID
2.2.6.1	예, 정규 상담사가 약 350 여 명 정도 있고요.	Statement		
2.2.6.2	필요하다고 그러면 이것을 제도화하는 것이 맞지 이것을 가지고 무슨 경고로 하고 시정으로 하느냐, 주의로 하느냐 하는 게 이게 맞아요?	Open Issue	Negative	2.2.6.1

2.2.2.7

ID	예시	node label	relation label	relation ID
2.2.7.1	그러면 교육부는 뭐하고 있었어요, 지금까지 그것을 제도개선 안 하고? 이런 것이 매번 반복적으로 일어난다면 이 부분에 당연히 지적이 들어갈 것이고 그런데 왜 아직까지 그것을 안 고치고 지금까지 계속 이렇게 해 왔어요?	Open Issue	Negative	
2.2.7.2	그동안은 저희들이 기존 예산 중에서 활용해 왔습니다. 그중에 서류처리원이라는 항목이 있어서 거기에서 같이 썼었는데 이것은 좀 더 세분화해서 명확하게 표현하도록 하겠습니다.	Statement	Option	2.2.7.2
2.2.7.3	교육부 자꾸 그런 식으로 전부 다 얼렁뚱땅 해 오고 그러지 마세요. 이것을 한번 그랬으면 존경하는 신 위원 말씀처럼 제도개선 하겠지만 반복해서 계속 이런 식으로 해 온다는 자체가 이것은 나태예요, 직무의 해태이고. 이것은 시정하고 제도개선 쪽으로 아울러 같이 가야 될 것이라고 생각돼요.	Statement	Negative	2.2.7.3

2.2.3 Uncertain

Yes/No Issue에 대한 답변 중 Positive 또는 Negative로 주석하기 어려운 경우에만 주석. 나머지는 주석하지 않음.

ID	예시	node label	relation label	relation ID
2.3.1.1	그런데 이 부분에 대해서 법적인 검토를 해 왔는데 여성발전기본법상 여성정책기본계획의 변경 사항과 관련해서는 시점은 맞았다고 그랬지요? 어쨌든 이것을 변경하고 나서 이렇게 했다고 한 것이지요? 맞습니까?	Yes/No Issue		2.3.1.1
2.3.1.2	그것은 확인을 해 보겠습니다. 시점은 12월에 동시에 이루어졌기 때문에.....	Statement	Uncertain	

ID	예시	node label	relation label	relation ID
2.3.2.1	우리가 청원소위에서 이렇게 심사해 가지고 통과해 주면 보건복지부에서 반대하는 것 아니에요? 될 가망성이 있는 거예요? 없어요?	Yes/No Issue		2.3.2.1
2.3.2.2	글쎄요.	Statement	Uncertain	

2.2.4 Specialization

뒷 발화가 앞 발화의 전문화된 발화일 경우 사용. 이슈를 재해석하여 앞 발화에 대한 답변을 제한하거나 다른 답변(해결책)을 유도하는 경우.

2.2.4.1

ID	예시	node label	relation label	relation ID
2.4.1.1	몰도바의 수도는 어디인가?	Open Issue		2.4.1.1
2.4.1.2	가장 몰도바인처럼 들리는 것은 무엇인가?	Open Issue	Specialization	

2.2.5 Elaboration

발화에 대한 추가정보가 제공된 경우.

2.2.5.1

ID	TAS-annotation-manual 예시	node label	relation label	relation ID
----	--------------------------	------------	----------------	-------------

2.5.1.1	개미는 우리 정원을 바꿀 수 있다.	Statement		
2.5.1.2	그들은 공원 전체를 바꿀 수도 있다.	Statement	Elaboration	2.5.1.1

2.2.5.2

ID	예시	node label	relation label	relation ID
2.5.2.1	상세보고라는 게 뭐예요?	Statement		
2.5.2.2	저희가 협의체 구성하겠다는 부분하고요, 그다음에 프라임사업은 다음 주에 발표를 합니다. 프라임사업도 그렇고 이번 주 화요일 날, 엇그제 발표했던 인문학 강화사업, 그러니까 코어사업에서도 평가지표에 기본적으로 시간강사들에 대한 평균 임금 부분이 지표로 들어갔습니다. 그래서 그런 여러 가지 평가지표들을 통해서 대학들의 변화를 유도하는 부분들이 있고요.	Statement	Elaboration	2.5.2.1

2.2.5.3

ID	예시	node label	relation label	relation ID
2.5.3.1	그 부분에 대해서 광역단체장이나 안행부가 아무런 이의를 제기하지 않았다는 것 확실해요? 차관이 얘기해 보세요, 차관이. 국장의 말은 도저히 이해가 안 됩니다.	Statement		
2.5.3.2	그러니까 실제로 공문으로 받은 것은 아니고요. 일단 유선상으로 국토부와 행자부 쪽에 확인을 했습니다. 그리고 광역자치단체는 기초자치단체의 사항이기 때문에 그쪽의 의견을 좀 수렴하라 그래서 그쪽의 의견도 실제 유선상으로 확인된 것은 큰 이견은 없다 이렇게 들어왔다고 합니다.	Statement	Elaboration	2.5.3.1

2.2.5.4

ID	예시	node label	relation label	relation ID
2.5.4.1	그다음에 유예하면 1년 동안 이것 해결할 수 있습니까? 또 똑같은 것 들고 나올까 봐 지금 걱정하는 것도 있는 거거든요. 그런 것을 한번 얘기를 해 보시라고요, 계획을 제출하든가.	Open Issue		
2.5.4.2	그래서 지금 당장 저희들이 필요하다 그래 가지고 직접적인 민원을 받아서 한 게 말씀드린 것처럼 지자체 11개에 업체만	Statement	Elaboration	2.5.4.1

	<p>해도 18개 이상입니다. 그리고 그런 업체들이 최소한의 행정 절차를 사업계획 승인부터 해서 도시계획 변경까지 하기 위해서는 1년의 기간은 필요하다는 것이 국토부와의 협의 결과입니다. 그래서 일단 그 협의에 따라서 이게 수용이 된다면 말씀드린 것처럼 그렇게, 지금 용적률 혜택을 보는 것이 전체적으로 봐서 23% 정도 밖에 안 되기 때문에 그렇게 큰 주차장이라든지 용적률로 인한 객실 증가라든지 이런 것으로 인해서 도심 복잡화라든지 이런 부분에 있어서는 큰 영향은 없을 것이다. 그래서 그 부분에 대해서는 일단은 동의가 된 것이기 때문에 그 이후의 것에 대해서는 그 결과를 보고 다시 판단을 해야 된다. 지금 저희 생각으로는 다시 그것을 하겠다 그런 생각은 없는 상황입니다.</p>			
--	---	--	--	--

2.2.5.5

ID	예시	node label	relation label	relation ID
2.5.5.1	<p>이게 출판사에서 이렇게 협의를 거쳐서 이런 부분에 대해서 저자들이 이런 부분을 고치겠다라고 해서 교육부가 그걸 승인한 거지요, 교육부가 직권 권고해서 수정한 부분이 문제가 된 게 아니라. 지금 이학재 위원이 지적하신 부분은.....</p>	Statement		
2.5.5.2	<p>조금 과정을 말씀드리면 저희들이 여러 군데에서 수정이 필요하다는 요청도 있었고 민원도 있었고 또 언론 보도도 있었습니다. 그래서 직권 수정을 하는 가운데 집필자의 의견도 묻고 또 다른 전문가의 의견도 묻고 여러 군데에 의견을 물었습니다. 그래서 직권 수정을 했다는 말씀을 드리겠습니다.</p>	Statement	Elaboration	2.5.5.1

2.2.5.6

ID	예시	node label	relation label	relation ID
2.5.6.1	<p>(...)이것은 기재부와 법제처의 의견 확인이 좀 필요하지 않느냐, 그래서 시정에 그칠 것이 아니라 이런 부분은 소위 징계위에 회부해야 된다, 징계감이다 저는 그렇게 판단하고 있는데 이 부분은 일단 조금 더 논의를 계속해 나갑시다.</p>	Statement		
2.5.6.2	<p>사실 위원장님이 지적하신 것처럼 국가재정이 들어가는 데 대해서는 일반법인 국가재정법이 원칙적으로 다 적용되어야 합니</p>	Statement	Elaboration	2.5.6.1

	다. 그런데 이 부분에 대해서는 유권해석 기관은 중앙행정기관의 장이 유권해석을 의뢰하는 경우에는 법제처에서 유권해석을 할 수 있습니다. 그런데 저희가 국회에서 법제처에 바로 요구할 수는 없고 이 건에 대해서 저희가 고민한 결과 교육부에서 법제처에 법령해석과 관련해서 유권해석을 의뢰할 수는 있을 것입니다.(...)			
--	---	--	--	--

2.2.5.7

ID	예시	node label	relation label	relation ID
2.5.7.1	그런데 손실액이 적지 않은 것 같아요.	Statement		
2.5.7.2	전년도 18년도의 손실액은 8600억, 한 1조 정도가 되고 있습니다. 그런데 현재 시점에서는 실질적으로 저희가 그 전년도 대비해서 전년도에 총 한 4000억 정도가 손실이 났지만 현재는 1조 2000억 정도를 이익을 내고 있습니다. 불요불급하게 단년도만 보게 되면 손실이지만 중장기적으로 가게 되면 괜찮습니다, 저희가 지금.	Statement	Elaboration	2.5.7.1

2.2.5.8

ID	예시	node label	relation label	relation ID
2.5.8.1	그다음에 글로벌교육 교류사업 관련해서 세계교육회의 준비 이 부분에 대해서는 논의가 필요하다고 한 이유가 있는데, 이 부분은 정진후 위원님이 조금 설명해 주셨으면 좋겠습니다.	Open Issue		
2.5.8.2	이것은 2015년 교육회의, 회의를 준비하는데(...) 세계 여러 나라에 설치된 한국대사관에 항의 활동, 회의 개최나 회의 참여 문제를 포함해서 항의 활동까지를 결의했거든요. 그렇기 때문에 저는 이 예산에 대해서는 추후에 편성해도 늦지 않다고 보고 이번 예산에서는 제외하고 삭감할 것을 다시 한 번 제안합니다.	Statement	Elaboration	2.5.8.1

2.2.5.9

ID	예시	node label	relation label	relation ID
2.5.9.1	주의보다 더한 것을 받아야 되는데, 신전문위원 잘 지적해 가지고 오늘 이렇게, 이건 주의를 훨씬 뛰어넘는 강한 폐	Open Issue		

	<p>널티를 받고 그다음에 법에 따라서 해야 될 거라고 생각하는데 그래도 주의 정도를 주신 것 보면 많이 참작해서 주신 것 같은데, 김해영 위원님 여러 가지 또 많이 배려하신 것 같은데 제가 볼 때는 이 건 주의보다 더 강한 페널티를 줘야 된다 이런 지적을 할 수 있는데, 아무튼 지금 시정요구유형 주의까지 받겠다고 하셨으니까 하되 엉거주춤하게 그러지 말고 아주 쿨하게 이 부분은 잘못됐다 인정하고 다시 법률에 근거해서 하든지 아니면 법이 개정될 때 그때 맞춰서 하는 것이 옳다 이런 판단을 합니다. 어때세요?</p>			
2.5.9.2	<p>그런데 저희들 이미 정책연구를 시행하고 있습니다. 그래서 권고하는 쪽으로 해 주시면 저희들이, 그리고 진로교육법에는 분명히 학생이라고 되어 있지만 예를 든다면 학교 밖 청소년도 있습니다. 그런 아이들도 진로교육 해 주고 있거든요.</p>	Statement	Elaboration	2.5.9.1

2.2.6 Request

발화에 대한 자세한 정보를 요청하거나 다시 말해주기를 요청하는 경우.

2.2.6.1

ID	예시	node label	relation label	relation ID
2.6.1.1	개미는 우리 정원을 바꿀 수 있다.	Statement		
2.6.1.2	정원을 바꾼다는 게 무슨 말인가?	Open Issue	Request	2.6.1.1

2.2.6.2

ID	예시	node label	relation label	relation ID
2.6.2.1	저희가 조금 이따 상세히 보고드리는 부분에 대해서 위원님들께서 한번 판단해 보시고 특위 구성에 대해서는 다시 한번 논의해 주셨으면 하고 생각합니다.	Statement		
2.6.2.2	상세보고라는 게 뭐예요?	Statement	Request	2.6.2.1

2.2.6.3

ID	예시	node label	relation	relation
----	----	------------	----------	----------

			label	ID
2.6.3.1	국·공립대의 경우에는 저희가 실제로 재정 투입해서 올린 부분입니다. 그래서 재정을 투입해서 올리는 게 가능하기 때문에 그것을 좀 선도적으로 국·공립대가 먼저 가고 그것을 따라오기를 기대했는데 아직은 부족합니다. 국·공립대학이 지금 8만 원 수준이고요. 사립대가 지금 5만 원 수준입니다. 그래서 그 갭을 좀 줄일 수 있도록 저희가 평가지표에도 넣고 기본적인 처우 이러한 부분에 상세한 재정사업도 만들고 협의체를 구성해서 운영을 하겠다는 말씀을 일단 보고드리겠습니다.	Statement		
2.6.3.2	(...)1년이 아니라 2년간 주장하는 근거가 2년 동안 정말 시간강사의 처우개선과 신분보장이라는 입법취지에 맞게 정부에서 해낼 수 있는가에 대한 약간의 의구심들을 갖고 있는 것 같아요. 그래서 그 부분에 대한 것을 아까 설명했지만 좀 요약해서 이야기를 해 봐요.	Statement	Request	2.6.3.1

2.2.6.4

ID	예시	node label	relation label	relation ID
2.6.4.1	예, 환경평가의 승인을 받도록 되어 있습니다.	Statement		
2.6.4.2	그 부분에 대해서 광역단체장이나 안행부가 아무런 이의를 제기하지 않았다는 것 확실해요? 차관이 얘기해 보세요, 차관이. 국장의 말은 도저히 이해가 안 됩니다.	Statement	Request	2.6.4.1

*Request는 자세한 정보를 요청하거나 다시 말해주기를 요청할 때에 주석하므로, 위치럼 질문 형식으로 발언하더라도 구체적이거나 추가적인 정보를 요청하는 발화는 Request로 분류함.

2.2.6.5

ID	예시	node label	relation label	relation ID
2.6.5.1	예, 이것은 정부 예산 10% 절감 추진 방침에 따라서 저희들도.....	Statement		
2.6.5.2	그러니까 사업비를 깎으라는 얘기는 아니고 인건비나 경상비에서 깎으라고 하지 않았습니까?	Yes/No Issue	Request	2.6.5.1
2.6.5.3	사업비도 같이 포함된.....	Statement	Negative	2.6.5.2
2.6.5.4	사업비도 포함이 되어 있는 게 확실합니까?	Yes/No Issue	Request	2.6.5.3

*아래와 같이 자료 등을 ‘요청’하는 경우에는 request 로 주석하지 않으며, 구체적으로 말하기, 반복해서 말하기 등 발화적 측면에서의 요청인 경우에만 주석함.

2.2.6.6

ID	발화자	예시	node label	relation label	relation ID
2.6.6.1	A	그런데 이미 심의위원회에서 돈을 주지 않았습니까? 방향은 이미 그전에 결정했을 것 아니에요? 그러니까 절차상 문제가 있다는 거예요. 이렇게 일을 하면 안 됩니다.	Statement		
2.6.6.2	A	그래서 그 경과를 제출해 주세요. 그래서 만약 이것이 위법적인 요소가 있다고 하면 징계요구를 하도록 하겠습니다.	Statement		
2.6.6.3	B	일단 시정으로 하고 부대의견을 넣을 수 있지 않아요?	Yes/No Issue		

2.2.7 Option

앞 발화에 대한 답변, 옵션, 솔루션이거나 아이디어를 제공하는 경우 사용.

2.2.7.1

ID	예시	node label	relation label	relation ID
2.7.1.1	물도바의 수도는 어디인가?	Open Issue		
2.7.1.2	치시나우	Statement	Option	2.7.1.1

2.2.7.2

ID	예시	node label	relation label	relation ID
2.7.2.1	그래서 저희들이 도종환 의원님의 수정을 수용을 하되 명확히 저희 정부 입장은 이것을 영으로 받아서 영에서 법인으로 명시하겠다는 것을 사전에 오해를 풀기 위해서 명시를 한다는 전제하에 의원님의 수정안을 수용을 하겠습니다.	Statement		
2.7.2.2	지난번에 차관님이 미국의 스미소니언 박물관의 사례 같은 것을 들면서, 재정을 안정적으로 계속 지원하는 그런 사례 같은 것들을 참조하면서 책임을 질 수 있도록 하겠다고 말씀하셨는데 미국처럼 정말 기부문화가 활성화되고 자체 모금도 막 성공적으로 되고 보잉 같은 기업이나 개인들이 우리 돈으로 한 300억씩 수백억씩 막 이렇게 기부하는 그런 문화가 아닌 상	Statement	Option	2.7.2.1

	<p>황에서 문학에 대해서, 문학관에 대해서 누가 그렇게 많이 기부할 것인가를 생각해 보면 문화부가, 정부가 좀 더 적극적으로 책임지는 자세로 운영을 하겠다고 하면 저도 시행령에 법안을 담는 것 이런 것 등을 동의할 수 있다는 말씀을 드립니다.</p>			
--	--	--	--	--

2.2.7.3

ID	예시	node label	relation label	relation ID
2.7.3.1	<p>(...)그래서 위원님 우려하시는 부분 저희들이 마음에 정확히 담고 이것이 반드시 문화융성의 주요한 진원지가 될 수 있도록 끌어 나가면서 또한 위원님이 문제 제기하신 정부의 지원 부분, 연구기능이랄지 이런 것이 축소가 되지 않도록..... 결국 재정당국의 동의가 필요한데요. 저희 문체부에서 기재부를 적극 잘 설득을 해서 이것이 소속기관 못지않게 잘 활성화가 될 수 있도록, 또 그리고 공적기능을 담당하는 재정 지원 부분은 결코 축소가 되지 않도록 저희가 부처 간에 협의를 잘 해서 절대 위원님들 우려하시지 않도록 잘 끌어가도록 노력을 하겠습니다.</p>	Statement		
2.7.3.2	<p>산하기관 또 하나 만들었기 때문에 퇴직 공무원 낙하산인사 식으로 이렇게 운영되면 안 된다는 점을 조금 더 명확하게 하는 것이 필요하다고 생각해요.</p>	Statement	Option	2.7.3.1

2.2.7.4

ID	예시	node label	relation label	relation ID
2.7.4.1	<p>다만 최근의 급변하는 기술의 발전과 사회의 변화에 비추어 봤을 때 학생뿐만 아니라 성인평생학습도 그 중요성이 날로 커지고 있습니다. 때문에 이 부분에 대해서는 빠른 시간 내에 진로교육법을 개정해서 예산에 반영하는 것이 맞다고 봅니다.</p>	Statement		
2.7.4.2	<p>그래서 저희가 대안을 낸다면, 전문위원님 말씀하고 위원님 말씀 받아서 대안을 낸다면 여기 시정요구사항에 이렇게 되어 있는 것을 '교육부는 현 단계에서는 국가진로교육센터의 성인진로교육 관련 예산 집행을 최소화하고 향후 성인진로교육 지원의 법적 근거를 마련하도록 할 것' 이런 정도만 해 주신다면 저희들이 법을 고쳐 가면서 하면 되겠다는 말씀을 드리겠습니다.</p>	Statement	Option	2.7.4.1

2.2.7.5

ID	예시	node label	relation label	relation ID
2.7.5.1	그렇게 지금 저도 믿고 싶고요. 학령인구가 감소가 된다는가 또 학교들이 통폐합이 되면서 선생님들 숫자도 감소하면서 기금의 축적이 좀 적어지지 않겠어요?	Open Issue		
2.7.5.2	그래서 저희가 전략적 자산 배분이나 롤링을 통해 가지고 수익을 다변화할 수 있도록 노력을 하고 있습니다.	Statement	Option	2.7.5.1

*아래 예시의 두 번째 발화와 같이 두 가지 LABEL 을 모두 주석할 수 있는 경우(수정 수용(option-조건부-), 수용(positive)이 포함된 발화)에는 더 큰 의미인 option(답변, 조건부)으로 주석함.

2.2.7.6

ID	예시	node label	relation label	relation ID
2.7.6.1	정부 측 의견 있으면 말씀해 주세요.	Statement		
2.7.6.2	개정안에 대해서는 수용을 하고 전문위원 검토의견에 대해서는 수정 수용을 하고자 합니다.	Statement	Option	

2.2.8 Option Exclusion

앞 발화의 Node Label 이 Issue 이고, Issue 의 대답에 해당하는 발화가 답변, 옵션, 솔루션이 아닌 경우

ID	TAS-annotation-manual 예시	node label	relation label	relation ID
2.8.1	가장 똑똑한 동물은 무엇인가?	Open Issue		
2.8.2	개미를 뛰어난 개인으로 보지 않을 것이다. 그 자체로는 아무것도 아니지 않은가?	Statement	Option Exclusion	2.8.1

2.3. 예외

예외에 해당하는 발화의 경우 주석하지 않음

2.3.1 Noise

회의와 관련 없는 말다툼일 경우 주석을 수행하지 않음.

2.3.2 Relation Label

1) 회의 진행 발언이거나 앞 발화와 뒷 발화의 관계가 없다고 판단되는 경우

회의진행 발언 또는 회의내용과 관계없는(쟁점을 해결하기 위한 내용이 아닌) 발화의 경우 relation label 및 ID를 작성하지 않음. 따라서 아래 예시의 경우 relation label을 주석하지 않음.

ID	예시	node label	relation label	relation ID
1	검찰국장 발언하세요	Statement		
2	검찰국장입니다	Statement		

ID	예시	node label	relation label	relation ID
1	일단 넘어가고 필요하면 나중에 다시 하지요.	Statement		
2	다음은 이·전용액으로 청사 임차료 총 당한 문제입니다.	Weak Statement		

ID	예시	node label	relation label	relation ID
1	의석을 정돈해 주시기 바랍니다.	Statement		
2	성원이 되었으므로 제316회 국회(임시회) 제1차 법안심사소위원회를 개의하겠습니다.	Statement		
3	의사일정 제1항부터 제8항까지 강동원 의원·강은희 의원·조영순 의원·김영환 의원·장병완 의원·김성태 의원·박주서 의원이 각각 대표발의하고 정부에서 제출한 고용정책 기본법 일부개정법률안 8건을 일괄상정합니다.	Statement		
4	8건의 법률안에 대하여 전문위원 설명해 주시기 바랍니다.	Statement		
5	전문위원 설명 전에 제가 의사진행발언을 좀 하겠습니다.	Statement		
6	예.	Weak Statement		
7	제가 이미 오늘 법안심사소위의 안건 심사 순서를 놓고 문제 제기를 했습니다. 이미 위원장님께서도 말씀을 드렸는데, 법안심사소위에서 안건 심사 순서는 대단히 중요합니다.	Statement		

2) 한 사람의 발화

한 사람의 발화 간에는 relation label을 주석하지 않으며, 각각의 내용이(ID별로) 다른 것이 아닌 이상 맨 위의 발화에만 주석을 붙임. 따라서 맨 위의 발화인 2번에 relation label 및 ID를 작성하고 3번 발화의 경우 relation label 및 ID를 작성하지 않음.

발화자	ID	예시	node label	relation label	relation ID
A	1	그렇지요? 그러니까 부처에서 말씀을 똑바로 하셔야지요.	Statement		
B	2	그러니까 국회에서 예산을 책정해서 집행해야 되는데 갑자기 전 부처에서 똑같이 10%를 절감해서 하도록 사업 조정을 하다 보니까 그게 잘 검토되어 있지 않은 상태에서 사업을 조정해서 불용이 막 생긴 겁니다. 그래서 아무리 예산절감을 하더라도 이렇게 주먹구구식으로 하지 말라는 것이지요.	Statement	Negative	1
B	3	그런 측면에서 이번에는 국회가 정부에게 강력하게 시정요구를 해서 다음부터는 국회의 예산심의권을 침해하지 말라는 겁니다.	Statement		

발화자	ID	예시	node label	relation label	relation ID
A	1	정부 측 의견 얘기해 주십시오.	Statement		
B	2	전문위원 검토의견처럼 박성호 의원안은 수정수용, 또 배기운 의원안은 일부 수정수용, 그리고 조현룡 의원안은 수용하는 입장입니다	Statement	Option	(1번 앞의 발화)
B	3	수정수용이라고 말씀드렸던 것은 아까 박성호 의원안은 미터기 등 택시 유사표시 금지를 운수종사자 준수사항에만 규정을 하고 있고 배기운 의원안은 운송사업자 준수사항으로 되어 있어서 그것을 같이 묶어서 볼 필요가 있기 때문에 말씀드린 거고요. 또 조현룡 의원님은 국민 편익 제고라든지 지방분권 촉진 차원에서, 또 이게 이양사무로 되어 있습니다. 그래서 필요한 사항이라는 점에서 수용합니다.	Statement		

*한 사람의 발화 간에는 relation label을 주석하지 않지만, 아래와 같이 한 사람의 발화이면서 앞 발화에 대해 각각 다른 내용으로 이야기하는 경우에는 relation label을 주석함.

발화자	ID	예시	node label	relation label	relation ID
A	1	그러면 일단 대기 수요들이 사업계획 승인을 받고 그다음에 도시계획을 통해 가지고 이 법에 의한 혜택은 받을 수 있도록 하는 거고요. 그게 최소한 1년은 해 줘야지 대기 수요들이 풀릴 수 있다고 보는 겁니다.	Statement		
B	2	아니, 그래서 문체부가 지금 대기 수요도 있고 하나까 이 부분에 있어서는 유예를 좀 하고, 내가 묻는 거는 대기 수요 있는 것은 이미 다 알려져 있었을 거고 그 사이에 이 법을 유예하기 이전에 뭐 다른 조치를 한 게 있는가와 그런 것에 대해서 좀 보고를 하시고요.	Open Issue	Request	1
B	3	그다음에 유예하면 1년 동안 이것 해결할 수 있습니까? 또 똑같은 것 들고 나올까 봐 지금 걱정하는 것도 있는 거거든요. 그런 것을 한번 얘기를 해 보시라고요, 계획을 제출하든가.	Open Issue	Request	1

<부록 3> 국회 회의록 요약 작업에 대한 전문가 의견

전문가 의견 1: 강윤주(한국일보 기자)
<p>- 주제 분할 시, 논의되는 법안의 구체적 내용을 제목에 적시해줬으면 합니다. 국회엔 수많은 법률안이 발의됩니다. 지금 이 순간도 마찬가지로입니다. 그러나 이 모든 법률안이 세상에 전혀 없던 제정안인 경우는 거의 드뭅니다. 보통은 기존에 있던 법률안의 내용을(자구만 수정하는 경우도 있습니다) 바꿔 내는 개정안인 경우가 많습니다. 따라서 주제 분할 시, 가장 큰 구분을 차지하는 법안의 구체적 내용을 간단히 정리해주는 게 필요해 보입니다. 그 자체로 내용 요약이 되는 효과도 얻을 수 있습니다.</p> <p>- 법안을 '주어'로 놓고 작성하는 것도 하나의 방법입니다. 위원장 법안 소개 > 의원들의 동의 여부 등 의견 확인 > 위원장 결론 > 위원장 후속 법안 소개 이런 식의 진행인데, 해당 회의록에선 위원장의 안건 발의와 다음 안건 발의로 넘어갈 때 2번씩 같은 구분이 반복됩니다. 불필요해 보입니다. 연장선상에서 위원장의 일상적인 회의 진행 발언에 대해선 별도의 카테고리(▲하위내용-진행발언 등)를 만들어 표시하고 넘어가주는 게 좋을 것 같습니다. 전문위원의 법안 소개 발언도 ▲하위내용-법안소개로 처리해도 무방해 보입니다. 발언의 주목도를 낮추는 것도 일의 속도를 높이는 요소라 생각합니다.</p> <p>- 법안의 쟁점, 심사의 결론을 한번씩 짚고 넘어가줬으면 좋겠습니다. 논쟁적인 법안을 중심으로 의견교환이 진행되는 경우, 토론 상황을 한눈에 파악할 수 있도록 개별 쟁점을 정리해주는 것도 방법일 것 같습니다. 의원들이 어떤 부분에서 문제를 제기하고 있는지 쟁점별로 한번씩 정리해주고 넘어가는 것도 필요해 보입니다. 마찬가지로, 위원장이 다음 법안으로 넘어가기 전 보통 마무리 발언을 하고 넘어갑니다. 이때 해당 법안에 대한 결론이 어찌 났는지를 정리하면 좋을 것 같습니다.</p> <p>- ▲추후 논의 사항 ▲정부 요구 사항 등을 적고 구체적으로 왜 합의가 안 됐고, 뭐가 더 필요한지 보완 조치를 담으면 될 것 같습니다. 외통위 소위 통일경제특구의 경우, 관련 제반 법률 미비로 법사위 논의가 불투명하니 이 부분을 보완하라는 취지의 언급들이 있었습니다. 이런 부분을 반영해 결론을 마무리 짓는 게 필요해 보입니다.</p> <p>- 국회 법안심사소위는 법률안 심사의 최전선에 있는 회의체입니다. 소위에서 어느 정도 여야가 합의를 이루고 정부와 조율 끝에 상임위원회에 안건으로 회부됩니다. 그런 만큼 디테일하게 격론이 맞붙을 수밖에 없는 자리입니다. 쟁점 분할 작업에서도, 맥락을 살피며 쟁점을 정리하고 기술하는 작업이 구체적으로 필요해 보이는 이유입니다. 개별 법안마다 쟁점은 판이하게 다를 수밖에 없습니다. 이걸 일괄적으로 표준화하기는 어렵습니다. 그러나 최종 목적이 요약문 작성인 만큼, 이 작업을 수월하게 진행하기 위해선 쟁점 분할에서 단순 법안 구분보다는 쟁점을 사전에 한번씩 정리해주고 가는 게 필요하다는 생각입니다.</p>

전문가 의견 2: 박상준(한국일보 기자)
<p>국회 회의록은 매우 민감하다. 여러 이해 관계자들의 입장이 서로 갈리는 경우가 많고 단어 하나 문장 하나로 인해서 갈등이 빚어지는 경우도 많다. 때문에 그 어떤 회의 내용보다도 요약을 시도할 때 신중해야 한다.</p> <p>국회 회의록의 특성을 분석해서 한 가지 잣대로만 요약을 하기 보다는 내용에 따라 복수의 기</p>

준을 적용해서 요약할 시도해 보는 것도 고려해 볼 만 하다. 회의 진행 순서, 소위원장에 의한 회의 진행 등 정형화 된 대화 진행 내용을 요약할 때와 토론과 논쟁으로 이뤄지는 부분을 요약할 때는 다른 기준을 가지고 접근해 보는 것이다.

특히 토론과 논쟁 부분을 요약할 때 글자 수 줄이는 작업만 할지 아니면 내용을 해체한 뒤 재구성하는 것까지 염두에 둘지 확실히 해야 할 것으로 보인다.

풀어 쓴 기존 회의록을 표 위주로 정리한 것으로 그쳐서는 안 된다. 현재 각각의 화자 별로 분류하는 방식에서 벗어나 소주제 별로 묶는 방식으로 진행하면 어떨까 생각이 든다. 그리고 다시 그 소주제 별로 묶는 것을 다시 대주제로 묶는 시도를 해보는 것도 필요하지 않을까 한다.

전문가 의견 3: 이서희(한국일보 기자)

전체적인 지침과 세부 지침은 쟁점을 추출한다는 목적에 맞춰 구성돼 있는 것으로 판단함. 다만 첫째 ‘주제 중복’ 가능성이 많은 데이터를 가지고 더 검증할 필요가 있어 보임. 둘째 쟁점에 따라 세부적인 가지 수가 많아질 수 있고, 이를 구분해야 요약의 의미가 명확해질 수 있음. 이에 대한 검토가 필요해 보임.

데이터 검토 및 구축 과정 또한 국회 회의록의 특징을 반영하여 전체적으로 구성돼 있음. 세부 내용에 있어서만 검증이 필요해 보임.

첫째 Problem/Issues에서 다수 화자에 의한 내용은 화자를 생략하게 돼 있는데, 다수 화자라고 하더라도 누가 말했는지(위원장, 간사 등)에 따라 의미와 실행 내용이 달라질 수 있으므로 이에 대한 검토가 필요함.

둘째 Actions에서 법 개정 사안인지, 시행령 개정 사안인지 등 층위에 따라 수행 주체가 달라질 수 있기 때문에 이에 대해 Decision에서 명시하는 것이 국회 회의록 특성에 맞는 구축 방법이 될지 검토하는 것이 타당해 보임.

셋째 국회 회의의 경우 여러 회기 내 회의록 간의 중복이나 관련 가능성이 크기 때문에 이에 대해 링크를 거는 등 보완점을 검토해볼 필요가 있음.

넷째, 단순 회의록을 요약하는 것을 넘어 수용자 관점에서 살펴보면, 회의 이후 결과(법안 본회의 통과 여부, 정책 시행 여부) 등이 실질적인 내용이 될 수 있으므로 의안정보시스템 등 연계를 통해 참고 데이터를 보완하는 것도 필요해보임.

참고문헌

- Nallapati, Ramesh and Zhou, Bo-Wen and Cicero Nogueira dos santos, Caglar Gulcehre and Xiang, Bing (2016), Abstractive Text Summarization Using Sequence-to-Sequence RNNs and Beyond, CoNLL 2016.
- Rush, Alexander M. and Chopra, Sumit and Weston, Jason (2015), A Neural Attention Model for Abstractive Sentence Summarization, Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing.
- Filippova, Katja and Altun, Yasemin (2013), Overcoming the Lack of Parallel Data in Sentence Compression, Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing (EMNLP '13).
- Völske, Michael and Potthast, Martin and Syed, Shahbaz and Stein, Benno (2017), TL;DR: Mining Reddit to Learn Automatic Summarization, Proceedings of the Workshop on New Frontiers in Summarization.
- Carletta, J. (2006) Announcing the AMI Meeting Corpus. The ELRA Newsletter 11(1), January-March, p. 3-5.
- Boudin, F., and Morin, E. (2013). Keyphrase extraction for n-best reranking in multi-sentence compression. In North American Chapter of the Association for Computational Linguistics (NAACL).
- Zhu, Chenguang and Xu, Ruochen and Zeng, Michael and Huang, Xuedong (2020), A Hierarchical Network for Abstractive Meeting Summarization with Cross-Domain Pretraining, EMNLP 2020.

<Abstract>

Corpus Research and Analysis 2021: Meeting Minutes Summarization

This project is a pilot study to build a large-scale corpus for argumentative dialog summarization. This project aims to enhance the utilization and value of Korean language resources and to build foundational data that can be standardized in Korea.

The argumentative dialogue summary corpus is an effective resource for not only the identification task of important information but also various application technologies such as the argument mining and opinion mining.

In this project, we build a guideline for abstractive summarization of argumentative dialogues. We also build abstractive summaries with manually extracted keywords for 200 subcommittee minutes of the National Assembly of the Republic of Korea based on the guideline. We review the utility and practical application of the guideline. The main tasks and research results of this project are as follows.

- ▶ To build a guideline for abstractive summarization of argumentative dialogues

This project built the guideline that is not only for the task to summarize dialog in written style but also for the task to abstractedly summarize the argumentative dialogue in a bottom-up manner. The guideline includes a detailed guideline to be applied to the minutes of the National Assembly. The guideline will serve the standardization of dialogue summarization and reference materials to build a large-scale argumentative dialog summary corpus in the

future.

- ▶ To build an abstractive summary corpus based on the guideline

The developed guideline was applied to the 200 minutes of the National Assembly subcommittee (95,052 sentences and 1,560,255 *eojeols* 'words' in total, 475.26 sentences and 7,801.275 *eojeols* per minutes on average). The summary corpus includes (1) keywords, (2) detailed sub-summaries ("sub" in the summarization corpus: 16.23 sub-summaries per issue, 1.01 sentences and 23.27 *eojeols* per "sub" on average), and (3) entire summaries for each issue ("abstract" in the corpus: 4.42 issues per minutes, 2.14 sentences and 45.28 *eojeols* per "abstract" on average) and a final summary of each minute ("main_summary" in the corpus: 3.815 sentences and 83.875 *eojeols* per final summary on average). The project outcomes are expected to be used for evaluation indicators in various aspects of the abstractive summarization models for argumentative dialogues.

Keywords: Summarization, Meeting Summarization, Building a Meeting Summary Corpus

Project Director: Kim Hansaem(Yonsei University)

<기획·연구>

국립국어원 이승재 언어정보과장

국립국어원 서셋별 학예연구사

국립국어원 서혜진 연구원

<사업 참여자>

사업 책임자 김한샘

사업 참여자 주민재, 윤영민, 최기선, 전영민, 함영균, 김지성, 한지윤,
오태환, 박서윤, 정승연, 박재완, 이규덕

발행인: 국립국어원장

발행처: 국립국어원

서울시 강서구 금남화로 154

전화 02-2669-9775, 전송 02-2669-9727

인쇄일: 2021년 12월 21일

발행일: 2021년 12월 21일

인 쇄: 연세대학교 POD센터

※ 이 책은 국립국어원의 용역비로 수행한 ‘2021년 회의록 요약 말뭉치 연구 분석’
사업의 결과물을 발간한 것입니다.