

국립국어원 대화 맥락 추론 말뭉치 2023 (버전 1.0)

- **자료명:** 국립국어원 대화 맥락 추론 말뭉치 2023
- **공개일**
 - (버전 1.0) 2024. 3. 29.
- **자료 유형:** 텍스트
- **관련 사업:** 2023년 대화 맥락 추론 및 부적절 발언 탐지 말뭉치 연구 분석(2023)
- **자료 설명**
 - ※ 자세한 내용은 국립국어원 누리집 > 자료 > 연구·조사 자료 > ‘2023년 대화 맥락 추론 및 부적절 발언 탐지 말뭉치 연구 분석’ 사업 보고서 참고
- **내용**
 - 메신저 대화 자료에서 나타나는 발화를 대상으로 대화 맥락이나 상식, 세계 지식 등에 의거하여 ‘원인’, ‘후행 사건’, ‘전제 조건’, ‘내적 동기’, ‘감정 반응’의 5가지 유형별 추론문을 작성하여 구성한 말뭉치.
- **분량**
 - 문서 4,000건(추론문 20,000개)
- **파일 형식:** JSON(UTF-8 인코딩)
- **파일 수 및 크기:** 텍스트 파일 1개, 총 3.95MB(ZIP 파일 기준)
- **인용:**
 - (국문) 국립국어원(2024). 국립국어원 대화 맥락 추론 말뭉치 2023(버전 1.0).
URL: <https://corpus.korean.go.kr>
 - (영문) National Institute of Korean Language (2024). NIKL Dialogue Contextual Inference Corpus 2023 (v.1.0). URL: <https://corpus.korean.go.kr>

· 예시

```
{
  "id": "MDRW2100006683",
  "metadata": {
    "title": "국립국어원 온라인 대화 말뭉치 MDRW2100006683",
    "creator": "국립국어원",
    "distributor": "국립국어원",
    "year": "2021",
    "category": "온라인 대화 > 2인 대화",
    "annotation_level": "원시",
    "sampling": "실시간 대화"
  },
  "document_list": [
    {
      "id": "MDRW2100006683.1",
      "metadata": {
        "title": "온라인 대화",
        "author": "개인 대화 참여자",
        "publisher": "심심이",
        "date": "20210908",
        "topic": "날씨와 계절",
        "speaker": [
          {
            "id": "1",
            "age": "30대",
            "occupation": "기타",
            "sex": "남성",
            "birthplace": "서울",
            "pricipal_residence": "서울",
            "current_residence": "서울",
            "device": "스마트폰",
            "keyboard": "천지인"
          },
          {
            "id": "2",
            "age": "30대",
            "occupation": "가정 주부",
            "sex": "여성",
            "birthplace": "서울",
            "pricipal_residence": "서울",
            "current_residence": "서울",
            "device": "스마트폰",
            "keyboard": "천지인"
          }
        ]
      },
      "setting": {
        "relation": "기타>연인",
        "intimacy": 5,
        "contact_frequency": "(거의) 매일 연락한다."
      }
    }
  ]
}
```

```

    },
    "utterance": [
      {
        "id": "MDRW2100006683.1.1",
        "form": "요즘 왜이렇게 춥냐",
        "original_form": "요즘 왜이렇게 춥냐",
        "speaker_id": "2",
        "time": "20210908 06:23"
      },
      {
        "id": "MDRW2100006683.1.2",
        "form": "비와서 그런거 아닐까? 나도 잘 때 좀 춥더라",
        "original_form": "비와서 그런거 아닐까? 나도 잘 때 좀 춥더라",
        "speaker_id": "1",
        "time": "20210908 06:23"
      },
      {
        "id": "MDRW2100006683.1.3",
        "form": "나도 자다가 추워서 깨고 그래 ㅎㅎ",
        "original_form": "나도 자다가 추워서 깨고 그래 ㅎㅎ",
        "speaker_id": "2",
        "time": "20210908 06:24"
      },
      {
        "id": "MDRW2100006683.1.4",
        "form": "이불 잘 덮고 자",
        "original_form": "이불 잘 덮고 자",
        "speaker_id": "1",
        "time": "20210908 06:24"
      },
      {
        "id": "MDRW2100006683.1.5",
        "form": "이불 잘 덮지 근데 얇아서 그런가봐 바꿔야겠어",
        "original_form": "이불 잘 덮지 근데 얇아서 그런가봐 바꿔야겠어",
        "speaker_id": "2",
        "time": "20210908 06:24"
      },
      {
        "id": "MDRW2100006683.1.6",
        "form": "이제 비 그치면 본격적으로 가을 아닐까 싶다",
        "original_form": "이제 비 그치면 본격적으로 가을 아닐까 싶다",
        "speaker_id": "1",
        "time": "20210908 06:25"
      },
      {
        "id": "MDRW2100006683.1.7",
        "form": "난 이미 비오기 전부터 가을 같던데",
        "original_form": "난 이미 비오기 전부터 가을 같던데",
        "speaker_id": "2",

```

	<pre> "time": "20210908 06:25" }, { "id": "MDRW2100006683.1.8", "form": "그래? 나는 낮에는 여전히 덥던데", "original_form": "그래? 나는 낮에는 여전히 덥던데", "speaker_id": "1", "time": "20210908 06:26" }, { "id": "MDRW2100006683.1.9", "form": "그렇긴한데 나는 이제 긴바지입어ㅋㅋㅋ", "original_form": "그렇긴한데 나는 이제 긴바지입어ㅋㅋㅋ", "speaker_id": "2", "time": "20210908 06:26" }, { "id": "MDRW2100006683.1.10", "form": "별써? 난 아직도 반바지에 반팔인데도 낮에 움직이면 땀 나던데", "original_form": "별써? 난 아직도 반바지에 반팔인데도 낮에 움 직이면 땀나던데", "speaker_id": "1", "time": "20210908 06:27" }, { "id": "MDRW2100006683.1.11", "form": "헉 땀이 나다니ㅋㅋ 일교차 심하니까 너무 얇게 입고다니 지는 마", "original_form": "헉 땀이 나다니ㅋㅋ 일교차 심하니까 너무 얇게 입고다니지는 마", "speaker_id": "2", "time": "20210908 06:28" }, { "id": "MDRW2100006683.1.12", "form": "그래 근데 나는 새벽에 잘 때 빼고는 밤에는 바람 많이 불어서 시원하더라고", "original_form": "그래 근데 나는 새벽에 잘 때 빼고는 밤에는 바람 많이 불어서 시원하더라고", "speaker_id": "1", "time": "20210908 06:28" }, { "id": "MDRW2100006683.1.13", "form": "몸에 열이 많은듯", "original_form": "몸에 열이 많은듯", "speaker_id": "1", "time": "20210908 06:28" }, </pre>
--	---

```

    {
      "id": "MDRW2100006683.1.14",
      "form": "맞아 그래서 겨울에도 막 반팔에 패딩입고 ㅋㅋㅋ 덥다고
하자나 어디 북극이나 남극 가야지나 춥다고 할 사람임ㅋㅋ",
      "original_form": "맞아 그래서 겨울에도 막 반팔에 패딩입고 ㅋㅋ
ㅋ 덥다고하자나 어디 북극이나 남극 가야지나 춥다고 할 사람임ㅋㅋ",
      "speaker_id": "2",
      "time": "20210908 06:29"
    },
    {
      "id": "MDRW2100006683.1.15",
      "form": "ㅋㅋ그럴지도모르지 빨리 겨울이나 왔으면 좋겠다",
      "original_form": "ㅋㅋ그럴지도모르지 빨리 겨울이나 왔으면 좋겠
다",
      "speaker_id": "1",
      "time": "20210908 06:30"
    }
  ],
  "inference": {
    "cause": {
      "output": "가을비가 내리고 날씨가 추워졌다.",
      "reference": [
        "MDRW2100006683.1.3"
      ]
    },
    "subsequent event": {
      "output": "화자2는 겨울 이불을 꺼낼 것이다.",
      "reference": [
        "MDRW2100006683.1.3"
      ]
    },
    "prerequisite": {
      "output": "화자1은 추위를 많이 타지 않는다.",
      "reference": [
        "MDRW2100006683.1.4"
      ]
    },
    "motivation": {
      "output": "화자2는 따뜻하게 자고 싶다.",
      "reference": [
        "MDRW2100006683.1.5"
      ]
    },
    "reaction": {
      "output": "화자1은 추위를 많이 타는 화자2가 안쓰럽다.",
      "reference": [
        "MDRW2100006683.1.9"
      ]
    }
  }
}

```

--

- ※ 'cause': 원인 추론문
- 'subsequent event': 후행 사건 추론문
- 'prerequisite': 전제 조건 추론문
- 'motivation': 내적 동기 추론문
- 'reaction': 감정 반응 추론문

· 자료 내용 문의: 02-2669-9638