

## 국립국어원 속성 기반 감성 분석 말뭉치 2021 (버전 1.1)

- **자료명:** 국립국어원 속성 기반 감성 분석 말뭉치 2021
- **공개일**
  - (버전 1.0) 2022. 4. 1.
  - (버전 1.1) 2023. 3. 31.
  - “2022년 국립국어원 인공 지능 언어 능력 평가” 대회 과제로 활용되지 않은 ‘영화 및 여행’ 대상 자료 포함
- **자료 유형:** 텍스트
- **관련 사업:** 2021년 말뭉치 감성 분석 및 연구(2021)
- **자료 설명**
  - ※ 자세한 내용은 국립국어원 누리집 > 자료 > 연구·조사 자료 > ‘2021년 말뭉치 감성 분석 및 연구’ 사업 결과보고서 참조
- **내용**
  - 국립국어원 감성 분석 말뭉치 2020(1.0)과 동일한 문서에 속성 기반 감성 정보 부착
  - “SemEval-2016 Task 5: Aspect-Based Sentiment Analysis”의 방법론을 따라 말뭉치를 구축하였으며, 분석 대상 문서의 분량·구성을 고려하여 개체#속성 쌍을 정의

<표 1> 영화 도메인의 개체#속성 쌍

	일반	구성력	연기력	인지도	연출력
영화	0			0	0
시나리오	0	0			
배우	0		0	0	
캐릭터	0				0
감독	0			0	0
음향·음악	0			0	0
영상·시각 요소	0				0

<표 2> 제품 도메인의 개체#속성 쌍

	일반	가격	디자인	품질	편의성	다양성	인지도
제품 전체	0	0	0	0	0		0
본품	0		0	0	0	0	
패키지·구성품	0		0	0	0	0	
브랜드	0	0	0	0			0

<표 3> 여행 도메인의 개체#속성 쌍

	일반	가격	편의성	심미성	접근성	청결도	인지도
장소 전체	0			0	0	0	0
시설	0	0	0	0		0	
상품	0	0					0
서비스	0						

· 분량

- 이 말뭉치는 “국립국어원 감성 분석 말뭉치 2020(1.0)” 대상 문서 중에서 “2022년 국립국어원 인공 지능 언어 능력 평가 대회” 과제로 활용되지 않은 ‘영화’ 및 ‘여행’ 자료를 포함하여 공개 (1,360개 문서)

<표 4> ‘2021년 말뭉치 감성 분석 및 연구’ 사업의 말뭉치 구축 결과\*

분석한 문장 수(문서 수)

대상 자료	영화	제품			여행			합계
		화장품 세정제	전자 기기	제품 기타	숙소	식당	관광지	
감성 분석 말뭉치 2020(1.0)	3,273 (379)	2,939 (775)	700 (123)	2,559 (640)	265 (27)	266 (35)	712 (74)	10,714 (2,053)
추가 구축**	1,094 (69)	1,545 (67)	1,002 (68)	-	-	-	-	3,641 (204)
합계	4,367 (448)	3,941 (842)	2,245 (191)	2,559 (640)	265 (27)	266 (35)	712 (74)	14,355 (2,257)

\* 공개 자료와 비공개 자료 모두 포함한 총 구축 결과

\*\* 자료 활용성 제고를 위해 본 사업에서 추가 구축한 자료 규모

- 파일 형식: JSON(UTF-8 인코딩)

- 파일 수 및 크기: 파일 2개, 총 2.30MB(ZIP 압축)

· 인용:

(국문) 국립국어원(2022). 국립국어원 속성 기반 감성 분석 말뭉치 2021 (버전 1.1). URL: <https://kli.korean.go.kr/corpus>

(영문) National Institute of Korean Language (2022). NIKL Korean Aspect Based Sentiment Analysis Corpus 2021 (v.1.1) URL: <https://kli.korean.go.kr/corpus>

· 파일 명명 규칙

자리	1	2	3	4	5	6	7	8	9	10	11	12	13	14
속성	매체	장르	주석 단계	구축 연도	문서/문장	일련번호(8자리)								
정의값	E: 웹	X: 추출	SA: 감성	21: 2021년	1: 문서	2: 문장	0000001 ~ 9999999 (일곱 자리 일련번호)							
※ 예시: EXSA2112303090.json 2021년도에 구축한 문서 단위 감성 분석 말뭉치 파일 ※ 예시: EXSA2122303090.json 2021년도에 구축한 문장 단위 감성 분석 말뭉치 파일														

· 예시

```

· sentence_level
{
  "document_id": "EBRW2110000002.1",
  "domain": "영화",
  "metadata": {
    "title": "<온다> `나카시마 테츠야 신작, 관계의 빈틈을 노리는 그것'-23th 부천판타스틱영화제",
    "author": "doona09",
    "publisher": "네이버 블로그",
    "date": "20190630",
    "url": "http://blog.naver.com/PostView.nhn?blogId=doona90&logNo=221574210502"
  },
  "sentence": [
    {
      "sentence_id": 1,
      "sentence_form": "장cine 수다",
      "opinions": []
    },
    {
      "sentence_id": 2,
      "sentence_form": "'나카시마 테츠야 신작, 관계의 빈틈을 노리는 그것'",
      "opinions": []
    },
    {
      "sentence_id": 3,
      "sentence_form": "'-23th 부천판타스틱영화제'",
      "opinions": []
    },
    {
      "sentence_id": 4,
      "sentence_form": "doona09 · 2019. 6. 30. 14:56",
      "opinions": []
    }
  ]
}

```

```

        "sentence_id": 5,
        "sentence_form": "URL 복사 이웃추가 본문 기타 기능 번역보기",
        "opinions": []
    },
    {
        "sentence_id": 6,
        "sentence_form": "온다, 来る, It Comes, 2018, 나카시마 테츠야",
        "opinions": []
    },
    {
        "sentence_id": 7,
        "sentence_form": "제23회 부천판타스틱영화제에서 상영된 <온다>는 '사와무라 이치의 <보기왕이 온다>를 영화로 각색했습니다.",
        "opinions": []
    },
    {
        "sentence_id": 8,
        "sentence_form": "기대되었던 조합은 <고백>, <힘오스런 마츠코의 일상>의 '나카시마 테츠야가 연출했고, '츠마부키 사토시, '오카다 준이치, '쿠로키 하루, '마츠 다카코, '고마츠 나나등이 출연해 비주얼적 상상력으로 구현하겠다는 김작했습니다.",
        "opinions": []
    },
    {
        "sentence_id": 9,
        "sentence_form": "역시나 짐작대로네요!",
        "opinions": [
            {
                "opinion_id": 1,
                "category": "영화#일본",
                "opinion_target": "Null",
                "begin": 0,
                "end": 0,
                "polarity": "neutral"
            }
        ]
    },
    {
        "sentence_id": 10,
        "sentence_form": "영화 는 알 수 없는 존재의 저주를 그린다.",
        "opinions": []
    },
    {
        "sentence_id": 11,
        "sentence_form": "어릴 적 한 소녀와 함께한 숲속 놀이, 그 이후 돌아오지 않는 소녀.",
        "opinions": []
    },
    {
        "sentence_id": 12,
        "sentence_form": "그 소녀의 악몽에 시달리는 히데키(츠마부키 사토시)를 따라다니는 정체불명의 그것",
        "opinions": []
    },
    {
        "sentence_id": 13,
        "sentence_form": "성년이 된 히데키는 아내(쿠로키 하루)와 예쁜 딸 치사와 가정을 꾸렸습니다.",
        "opinions": []
    },
    },

```

```

{
  "sentence_id": 14,
  "sentence_form": "히데키는 결혼해서 남부럽지 않은 가정을 꾸리고 싶은 마음에 들떴고, 딸이아
태어나자 열심히 블로그에 소식을 업로드하며, 육아 모임도 나가며 정성을 쏟죠.",
  "opinions": []
},
{
  "sentence_id": 15,
  "sentence_form": "그러나 그것이 다시 단란한 가정을 찾아오며 일상을 잠식해 갑니다.",
  "opinions": []
},
{
  "sentence_id": 16,
  "sentence_form": "보기왕은 당신의 두려움을 먹고 산다",
  "opinions": []
},
{
  "sentence_id": 17,
  "sentence_form": "당연하게 생각했던 일도 누군기에게는 불편한 감정, 상처주는 안행일 수 있으며
편향된 시각과 이기심은 사람 사이의 관계를 멀어지게 하고 두려움을 양상하는지 알 수 있죠.",
  "opinions": []
},
{
  "sentence_id": 18,
  "sentence_form": "보기왕은 이런 관계의 빈틈을 찾아 찾아온답니다.",
  "opinions": []
},
{
  "sentence_id": 19,
  "sentence_form": "자신을 찾는 손님 방문을 전한 후배의 원인불명 병환, 전하나 계속되는 괴이한 일
의 반복",
  "opinions": []
},
{
  "sentence_id": 20,
  "sentence_form": "사달리다 못한 히데키는 민속학자인 친구에게 도움을 요청하고, 언론인마 영매사
를 소개받습니다.",
  "opinions": []
},
{
  "sentence_id": 21,
  "sentence_form": "과연 이 가족은 보기왕의 저주로부터 살아남을 수 있을까요?",
  "opinions": []
},
{
  "sentence_id": 22,
  "sentence_form": "캐릭터는 다들 말 못할 상처를 갖고 있다",
  "opinions": []
},
{
  "sentence_id": 23,
  "sentence_form": "영화와 소설은 세 사람의 시점으로 진행됩니다.",
  "opinions": []
},
{
  "sentence_id": 24,
  "sentence_form": "소설은 나눠있지만 영화는 정확히 자르지 않습니다.",
  "opinions": []
}

```

```

"opinions": []
},
{
  "sentence_id": 25,
  "sentence_form": "먼저 저주가 쓰인 히터키, 그의 아내, 영매의 남자친구의 이야기로 구성됩니다.",
  "opinions": []
},
{
  "sentence_id": 26,
  "sentence_form": "비주얼 리스트인 감독의 영화 답게 마치 뮤직비디오, CF를 보는 듯하며. 감각적이  
고 강렬한 색채, 빠른 화면 전환이 자칫 불친절하고, 과도하게 느껴질 수 있었습니다.",
  "opinions": [
    {
      "opinion_id": 1,
      "category": "영상/시각 요소#연출력",
      "opinion target": "색채",
      "begin": 51,
      "end": 53,
      "polarity": "positive"
    },
    {
      "opinion_id": 2,
      "category": "영상/시각 요소#연출력",
      "opinion target": "빠른 화면 전환",
      "begin": 55,
      "end": 63,
      "polarity": "negative"
    }
  ]
},
{
  "sentence_id": 27,
  "sentence_form": "고마흔 나나 비주얼 변신, 후반부에 알아 본. ",
  "opinions": []
},
{
  "sentence_id": 28,
  "sentence_form": "실체가 명확하지 않는 그것을 구현한다는 것 자체가 어불성설이며, 그것에게 당하  
는 인간의 나약함과 불안함을 영화 가득 채워 넣었습니다.",
  "opinions": []
},
{
  "sentence_id": 29,
  "sentence_form": "원작<보기왕이 온다>는 잊고 싶었던 기억, 내면 밑바닥의 공포를 건드립니다.",
  "opinions": []
},
{
  "sentence_id": 30,
  "sentence_form": "그 공포를 좀 먹고 자라난 마음은 어떤 수법으로 사람을 현혹시키기에 주목하는  
소설인데요.",
  "opinions": []
},
{
  "sentence_id": 31,
  "sentence_form": "일본의 구절 설화와 서양의 부기맨을 결합해 그럴듯한 이야기보따리를 풀어 냈니  
다.",
  "opinions": []
}

```

```

    },
    {
      "sentence_id": 32,
      "sentence_form": "원작 《보기왕이 온다》 원작은 호러, 심령의 느낌이 강했다면 영화는 오컬트적 요소가 큼니다.",
      "opinions": []
    },
    {
      "sentence_id": 33,
      "sentence_form": "후반부 거대한 곳판은 비주얼의 극치를 달리며 누군기에게 환율을 누군기에게 피로감을 선사할 것 입니다.",
      "opinions": [
        {
          "opinion_id": 1,
          "category": "영상/시각 요소#연출력",
          "opinion target": "후반부 거대한 곳판",
          "begin": 0,
          "end": 10,
          "polarity": "neutral"
        }
      ]
    },
    {
      "sentence_id": 34,
      "sentence_form": "이 부분은 <곡상>이 떠오르기도 했으며, 한국 무속인도 언뜻 보입니다.",
      "opinions": []
    },
    {
      "sentence_id": 35,
      "sentence_form": "[제23회 부천판타스틱영화제] 온다",
      "opinions": []
    },
    {
      "sentence_id": 36,
      "sentence_form": "감독 나카시마 테츠야",
      "opinions": []
    },
    {
      "sentence_id": 37,
      "sentence_form": "출연 오카다 준이치, 쿠로키 하루, 고마츠 나나, 미츠 다카코, 츠부키 시토시",
      "opinions": []
    },
    {
      "sentence_id": 38,
      "sentence_form": "개봉 미개봉",
      "opinions": []
    },
    {
      "sentence_id": 39,
      "sentence_form": "평점: ★★★",
      "opinions": []
    },
    {
      "sentence_id": 40,
      "sentence_form": "한 줄 평: 원작의 지나친 변주",
      "opinions": [

```

```

        "opinion_id": 1,
        "category": "영화#연출력",
        "opinion target": "Null",
        "begin": 0,
        "end": 0,
        "polarity": "negative"
    }
  ]
},

```

· document\_level

```

{
  "document_id": "EBRW2110000002.1",
  "domain": "영화",
  "metadata": {
    "title": "<온다> `나카시마 테츠야` 신작, 관계의 빈틈을 노리는 그것!-23th 부천판타스틱영화제",
    "author": "doona09",
    "publisher": "네이버 블로그",
    "date": "20190630",
    "url": "http://blog.naver.com/PostView.nhn?blogId=doona90&logNo=221574210502"
  },
  "sentence": [
    {
      "sentence_id": 1,
      "sentence_form": "장ane 수다"
    },
    {
      "sentence_id": 2,
      "sentence_form": "`나카시마 테츠야` 신작, 관계의 빈틈을 노리는 그것!"
    },
    {
      "sentence_id": 3,
      "sentence_form": "-23th 부천판타스틱영화제"
    },
    {
      "sentence_id": 4,
      "sentence_form": "doona09 · 2019. 6. 30. 14:56"
    },
    {
      "sentence_id": 5,
      "sentence_form": "URL 복사 이웃추가 본문 기타 기능 번역보기"
    },
    {
      "sentence_id": 6,
      "sentence_form": "온다, 来る, It Comes, 2018, 나카시마 테츠야"
    },
    {
      "sentence_id": 7,
      "sentence_form": "제23회 부천판타스틱영화제에서 상영된 <온다>는 사와무라 이치의 《보기왕이 온다》를 영화로 각색했습니다."
    },
    {
      "sentence_id": 8,
      "sentence_form": "기대되었던 조합은 <고백>, <혐오스런 마츠코의 일생>의 `나카시마 테츠야`가 연출했고, `츠미부키 사토시`, `오카다 준이치`, `쿠로키 하루`, `마츠 다카코`, `고마츠 나나`등이 출연해 비주얼적 상상력으로 구현하

```



```

"갓다라고 짐작했습니다."
},
{
  "sentence_id": 9,
  "sentence_form": "역시나 짐작대로네요!"
},
{
  "sentence_id": 10,
  "sentence_form": "영화 는 알 수 없는 존재의 저주를 그린다."
},
{
  "sentence_id": 11,
  "sentence_form": "어릴 적 한 소녀와 함께한 숲속 놀이, 그 이후 돌아오지 않는 소녀."
},
{
  "sentence_id": 12,
  "sentence_form": "그 소녀의 악몽에 시달리는 히데키(츠마부키 사토시)를 따라다니는 정체불명의 그
것"
},
{
  "sentence_id": 13,
  "sentence_form": "성년이 된 히데키는 아내(쿠로키 하루)와 예쁜 딸 치사와 가정을 꾸렸습니다."
},
{
  "sentence_id": 14,
  "sentence_form": "히데키는 결혼해서 남부럽지 않은 가정을 꾸리고 싶은 마음에 들떴고, 딸이아
태어나자 열심히 블로그에 소식을 업로드하며, 육아 모임도 나가며 정성을 쏟죠."
},
{
  "sentence_id": 15,
  "sentence_form": "그러나 그것이 다시 단란한 가정을 찾아오며 일상을 잠식해 갑니다."
},
{
  "sentence_id": 16,
  "sentence_form": "보기왕은 당신의 두려움을 먹고 산다"
},
{
  "sentence_id": 17,
  "sentence_form": "당연하게 생각했던 일도 누군기에게는 불편한 감정, 상처주는 언행일 수 있으며
편향된 시각과 이기심은 사람 사이의 관계를 멀어지게 하고 두려움을 양생하는지 알 수 있죠."
},
{
  "sentence_id": 18,
  "sentence_form": "보기왕은 이런 관계의 빈틈을 찾아 찾아온답니다."
},
{
  "sentence_id": 19,
  "sentence_form": "자신을 찾는 손님 방문을 전한 후배의 원인불명 병환, 전화나 계속되는 괴이한 일
의 반복."
},
{
  "sentence_id": 20,
  "sentence_form": "시달리다 못한 히데키는 민속학자인 친구에게 도움을 요청하고, 언론인마 영매사
를 소개받습니다."
},
{
  "sentence_id": 21,

```

	"sentence_form": "과연 이 가족은 보기왕의 저주로부터 살아남을 수 있을까요?"
	},
	{
	"sentence_id": 22,
	"sentence_form": "캐릭터는 다들 말 못할 상처를 갖고 있다"
	},
	{
	"sentence_id": 23,
	"sentence_form": "영화와 소설은 세 사람의 시점으로 진행됩니다."
	},
	{
	"sentence_id": 24,
	"sentence_form": "소설은 나눠있지만 영화는 정확히 자르지 않습니다."
	},
	{
	"sentence_id": 25,
	"sentence_form": "먼저 저주가 쓰인 히데키, 그의 아내, 영매의 남자친구의 이야기로 구성됩니다."
	},
	{
	"sentence_id": 26,
고 강렬한 색채, 빠른 화면 전환이 자칫 불친절하고, 과도하게 느껴질 수 있었습니다."	"sentence_form": "비주얼 리스트인 감독의 영화 답게 마치 뮤직비디오, CF를 보는 듯하며, 감각적이
	},
	{
	"sentence_id": 27,
	"sentence_form": "고마츠 나나 비주얼 변신, 후반부에 알아 본. "
	},
	{
	"sentence_id": 28,
는 인간의 나아함과 불안함을 영화 가득 채워 넣었습니다."	"sentence_form": "실체가 명확하지 않는 그것을 구현한다는 것 자체가 어불성설이며, 그것에게 당하
	},
	{
	"sentence_id": 29,
	"sentence_form": "원작《보기왕이 온다》는 잊고 싶었던 기억, 내면 밑바닥의 공포를 건드립니다."
	},
	{
소설인데요 "	"sentence_id": 30,
	"sentence_form": "그 공포를 좀 먹고 자라난 마음은 어떤 수법으로 사람을 현혹시키기에 주목하는
	},
	{
다."	"sentence_id": 31,
	"sentence_form": "일본의 구절 설화와 서양의 부가맨을 결합해 그럴듯한 이야기보따리를 풀어 냈니
	},
	{
소가 큼니다."	"sentence_id": 32,
	"sentence_form": "원작 《보기왕이 온다》 원작은 호러, 심령의 느낌이 강했다면 영화는 오컬트적 요
	},
	{
로감을 선사할 것 입니다."	"sentence_id": 33,
	"sentence_form": "후반부 거대한 굿판은 비주얼의 극치를 달리며 누군가에겐 황홀을 누군가에겐 피
	},
	{

```

    "sentence_id": 34,
    "sentence_form": "이 부분은 <곡성>이 떠오르기도 했으며, 한국 무속인도 언뜻 보입니다."
  },
  {
    "sentence_id": 35,
    "sentence_form": " [제23회 부천판타스틱영화제] 온다"
  },
  {
    "sentence_id": 36,
    "sentence_form": "감독 나카시마 테츠야"
  },
  {
    "sentence_id": 37,
    "sentence_form": "출연 오카다 준이치, 쿠로키 하루, 고마츠 나나, 마츠 다카코, 츠부키 시토시"
  },
  {
    "sentence_id": 38,
    "sentence_form": "개봉 미개봉"
  },
  {
    "sentence_id": 39,
    "sentence_form": "평점: ★★★"
  },
  {
    "sentence_id": 40,
    "sentence_form": "한 줄 평: 원작의 지나친 변주"
  }
],
"opinions": [
  {
    "opinion_polarity": "conflict",
    "category": "영상/시각 요소#연출력"
  },
  {
    "opinion_polarity": "negative",
    "category": "영화#연출력"
  },
  {
    "opinion_polarity": "neutral",
    "category": "영화#일반"
  }
]
},

```

· 자료 내용 문의: 02-2669-9638