

국립국어원 어휘 의미 분석 말뭉치

(버전 1.0)

- **자료명:** 국립국어원 어휘 의미 분석 말뭉치
- **공개일**
 - (버전 1.0) 2020. 8. 25.
- **자료 유형:** 텍스트
- **관련 사업:** 어휘 의미 분석 말뭉치 구축(2019)
- **자료 설명**
 - **내용**
 - 총 300만 어절 규모(문어 200만 어절, 구어 100만 어절)의 어휘 의미 분석 말뭉치
 - **분량**
 - 총 300만 어절(문어 200만 어절, 구어 100만 어절)
 - **파일 형식:** JSON(UTF-8 인코딩)
- **인용:** 국립국어원(2020). 국립국어원 어휘 의미 분석 말뭉치(버전 1.0). URL: <https://corpus.korean.go.kr>
- **예시**

```
{
  "id": "NXLS1902008050",
  "metadata": {},
  "document": [
    {
      "id": "NWRW1800000022.417",
      "metadata": {},
      "sentence": [
        {
          "id": "NWRW1800000022.417.1.1",
          "form": "[제주·서울] ₩세계환경수도 조성위해 10개년 실
천계획 만들겠다₩" 김태환 지사 밝혀",
          "word": {
            "id": 1,
            "form": "[제주·서울]",
            "begin": 0,
            "end": 7
          }
        }
      ]
    }
  ]
}
```

```

{
  "id": 2,
  "form": "₩"세계환경수도",
  "begin": 8,
  "end": 15
},
{
  "id": 3,
  "form": "조성위해",
  "begin": 16,
  "end": 20
},
{
  "id": 4,
  "form": "10개년",
  "begin": 21,
  "end": 25
},
{
  "id": 5,
  "form": "실천계획",
  "begin": 26,
  "end": 30
},
{
  "id": 6,
  "form": "만들겠다₩",
  "begin": 31,
  "end": 36
},
{
  "id": 7,
  "form": "김태환",
  "begin": 37,
  "end": 40
},
{
  "id": 8,
  "form": "지사",
  "begin": 41,
  "end": 43
},
{
  "id": 9,
  "form": "밝혀",
  "begin": 44,
  "end": 46
}
],
"morpheme": [
  {
    "id": 1,
    "form": "[",
    "label": "SS",
    "word_id": 1,
    "position": 1
  },
  {
    "id": 2,
    "form": "제주",
    "label": "NNP",
    "word_id": 1,
    "position": 2
  }
]

```



```
    },  
    {  
      "id": 3,  
      "form": ".",  
      "label": "SP",  
      "word_id": 1,  
      "position": 3  
    },  
    {  
      "id": 4,  
      "form": "서울",  
      "label": "NNP",  
      "word_id": 1,  
      "position": 4  
    },  
    {  
      "id": 5,  
      "form": "]",  
      "label": "SS",  
      "word_id": 1,  
      "position": 5  
    },  
    {  
      "id": 6,  
      "form": "₩",  
      "label": "SS",  
      "word_id": 2,  
      "position": 1  
    },  
    {  
      "id": 7,  
      "form": "세계",  
      "label": "NNG",  
      "word_id": 2,  
      "position": 2  
    },  
    {  
      "id": 8,  
      "form": "환경",  
      "label": "NNG",  
      "word_id": 2,  
      "position": 3  
    },  
    {  
      "id": 9,  
      "form": "수도",  
      "label": "NNG",  
      "word_id": 2,  
      "position": 4  
    },  
    {  
      "id": 10,  
      "form": "조성",  
      "label": "NNG",  
      "word_id": 3,  
      "position": 1  
    },  
    {  
      "id": 11,  
      "form": "위하",  
      "label": "VV",  
      "word_id": 3,  
      "position": 2  
    }  
  ],  
  "tokens": "  
    .  
    서울  
    ]  
    ₩  
    세계  
    환경  
    수도  
    조성  
    위하  
  "
```



```
    },  
    {  
      "id": 12,  
      "form": "아",  
      "label": "EC",  
      "word_id": 3,  
      "position": 3  
    },  
    {  
      "id": 13,  
      "form": "10",  
      "label": "SN",  
      "word_id": 4,  
      "position": 1  
    },  
    {  
      "id": 14,  
      "form": "개년",  
      "label": "NNB",  
      "word_id": 4,  
      "position": 2  
    },  
    {  
      "id": 15,  
      "form": "실천",  
      "label": "NNG",  
      "word_id": 5,  
      "position": 1  
    },  
    {  
      "id": 16,  
      "form": "계획",  
      "label": "NNG",  
      "word_id": 5,  
      "position": 2  
    },  
    {  
      "id": 17,  
      "form": "만들",  
      "label": "VV",  
      "word_id": 6,  
      "position": 1  
    },  
    {  
      "id": 18,  
      "form": "갯",  
      "label": "EP",  
      "word_id": 6,  
      "position": 2  
    },  
    {  
      "id": 19,  
      "form": "다",  
      "label": "EF",  
      "word_id": 6,  
      "position": 3  
    },  
    {  
      "id": 20,  
      "form": "₩",  
      "label": "SS",  
      "word_id": 6,  
      "position": 4  
    }  
  ]  
}
```



```
    },  
    {  
      "id": 21,  
      "form": "김태환",  
      "label": "NNP",  
      "word_id": 7,  
      "position": 1  
    },  
    {  
      "id": 22,  
      "form": "지사",  
      "label": "NNG",  
      "word_id": 8,  
      "position": 1  
    },  
    {  
      "id": 23,  
      "form": "발히",  
      "label": "VV",  
      "word_id": 9,  
      "position": 1  
    },  
    {  
      "id": 24,  
      "form": "어",  
      "label": "EF",  
      "word_id": 9,  
      "position": 2  
    }  
  ],  
  "WSD": [  
    {  
      "word": "제주",  
      "sense_id": 8,  
      "pos": "NNP",  
      "begin": 1,  
      "end": 3,  
      "word_id": 1  
    },  
    {  
      "word": "서울",  
      "sense_id": 2,  
      "pos": "NNP",  
      "begin": 4,  
      "end": 6,  
      "word_id": 1  
    },  
    {  
      "word": "세계",  
      "sense_id": 2,  
      "pos": "NNG",  
      "begin": 9,  
      "end": 11,  
      "word_id": 2  
    },  
    {  
      "word": "환경",  
      "sense_id": 1,  
      "pos": "NNG",  
      "begin": 11,  
      "end": 13,  
      "word_id": 2  
    }  
  ],  
}
```



```
{
  {
    "word": "수도",
    "sense_id": 12,
    "pos": "NNG",
    "begin": 13,
    "end": 15,
    "word_id": 2
  },
  {
    "word": "조성",
    "sense_id": 5,
    "pos": "NNG",
    "begin": 16,
    "end": 18,
    "word_id": 3
  },
  {
    "word": "개년",
    "sense_id": 3,
    "pos": "NNB",
    "begin": 23,
    "end": 25,
    "word_id": 4
  },
  {
    "word": "실천",
    "sense_id": 1,
    "pos": "NNG",
    "begin": 26,
    "end": 28,
    "word_id": 5
  },
  {
    "word": "계획",
    "sense_id": 1,
    "pos": "NNG",
    "begin": 28,
    "end": 30,
    "word_id": 5
  },
  {
    "word": "김태환",
    "sense_id": 777,
    "pos": "NNP",
    "begin": 37,
    "end": 40,
    "word_id": 7
  },
  {
    "word": "지사",
    "sense_id": 10,
    "pos": "NNG",
    "begin": 41,
    "end": 43,
    "word_id": 8
  }
}
]
```