

Watson Personality Insight API

Watson API 사용법 보고 와주세요

Personality Insight

사용법이 어렵지는 않지만,
결과를 이해하기 위해서는 **JSON** 프로파일과 **CSV** 프로파일을 이해해야 합니다.

입력은 **json** 형식 혹은 ,로 구분된 **csv** 파일 이어야만 합니다.

Personality Insight

또한 Big5로 나타나는 항목은
우측의 내용들을 의미합니다.

https://ko.wikipedia.org/wiki/5%EA%B0%80%EC%A7%80_%EC%84%B1%EA%B2%A9_%ED%8A%B9%EC%84%B1_%EC%9A%94%EC%86%8C

5가지 성격 특성 요소

위키백과, 우리 모두의 백과사전.

5가지 성격 특성 요소(Big Five personality traits)는 심리학에서 경험적인 조사와 연구를 통하여 정립한 성격 특성의 다섯 가지 주요한 요소 혹은 차원을 말한다. 신경성, 외향성, 친화성, 성실성, 경험에 대한 개방성의 다섯 가지 요소가 있으며, Costa & McCrae에 의해서 집대성된 모델로 다양한 나라들에서 그 유효성이 확인된 바 있다^[1]. 현대 심리학계에서 가장 널리 인정받고 있는 성격이론이다.^[2]

수많은 연구 결과 성격 5요인 이론이 개인의 행복, 신체적·정신적 건강, 종교성, 정체성뿐 아니라 가족, 친구, 연인 사이에서의 각종 관계적 결과들 및 직업 선택, 직무 만족도, 수행, 사회 참여, 범죄 행동, 정치적 입장 같은 요소들을 잘 예측한다는 것이 밝혀졌다.^[3]

이 이론을 토대로 한 검사로는 NEO - PI-R 성격 검사지가 있다. 이러한 Big 5 모델은 다양한 자료에서 신뢰성과 타당성을 가진다.

목차 [숨기기]

- 1 개괄
- 2 적용
- 3 같이 보기
- 4 각주
- 5 외부 링크

개괄 [편집]

성격이란 특정 상황에서 개인이 어떻게 행동할 것인가를 예측할 수 있게 해주는 것, 인간의 모든 행동과 관련되어 있는 것을 말한다. 5가지 성격 요소와 간단한 설명은 아래와 같다.

- 경험에 대한 개방성(Openness to experience) - 상상력, 호기심, 모험심, 예술적 감각 등으로 보수주의에 반대하는 성향
 - 개인의 심리 및 경험의 다양성과 관련된 것으로, 지능, 상상력, 고정관념의 타파, 심미적인 것에 대한 관심, 다양성에 대한 욕구, 꿈위 등과 관련된 특질을 포함
- 성실성(Conscientiousness) - 목표를 성취하기 위해 성실하게 노력하는 성향
 - 과제 및 목적 지향성을 촉진하는 속성과 관련된 것으로, 심사숙고, 규준이나 규칙의 준수, 계획 세우기, 조직화, 과제의 준비 등과 같은 특질을 포함
- 외향성(Extraversion) - 다른 사람과의 사교, 자극과 활력을 추구하는 성향
 - 사회와 현실 세계에 대해 의욕적으로 접근하는 속성과 관련된 것으로, 사회성, 활동성, 적극성과 같은 특질을 포함
- 친화성(Agreeableness) - 타인에게 반항적이지 않은 협조적인 태도를 보이는 성향
 - 사회적 적응성과 타인에 대한 공동체적 속성을 나타내는 것으로, 이타심, 애정, 신뢰, 배려, 겸손 등과 같은 특질을 포함
- 신경성(Neuroticism) - 분노, 우울함, 불안감 같은 불편한 정서를 쉽게 느끼는 성향
 - 걱정, 부정적 감정 등과 같은 바람직하지 못한 행동과 관계된 것으로, 걱정, 두려움, 슬픔, 긴장 등과 같은 특질을 포함(정서적 안정성은 정서적 불안정성과 반대되는 특징)

Openness to experience
Conscientiousness
Extraversion
Agreeableness
Neuroticism

OCEAN model

67

이 다섯 가지 요소로 이루어진 모델은 영문 스펠링의 첫 자를 따서 **OCEAN 모델**이라고 불리기도 한다. 이 5가지 요소들은 보통 50%를 기준으로 측정된다. 예를 들어 성실성 수치가 80%라면 책임감과 질서 정연함을 남보다 중시 여긴다는 뜻이고, 외향성 수치가 5%라면 타인과의 교류보다는 고독함과 평정심을 즐긴다는 것을 나타낸다.

JSON Profile

프로파일 오브젝트

Profile 오브젝트는 서비스에서 리턴하는 최상위 레벨 JSON 오브젝트입니다. 오브젝트에는 다음 필드가 있습니다.

- word_count** (정수)는 프로파일을 생성하는 데 사용된 입력 콘텐츠의 단어 수를 제공합니다. 요청이 대용량의 콘텐츠를 제출할 경우, 이 숫자는 입력의 단어 수보다 적을 수 있습니다. 단어 수가 최소 임계값을 충족하지 못하는 경우, 결과에는 추가 안내를 제공하는 **word_count_message** 필드가 포함됩니다.
- processed language** (문자열)는 서비스가 입력을 처리하는 데 사용한 언어 모델(**ar** (아랍어), **en** (영어), **es** (스페인어), **ja** (일본어) 또는 **ko** (한국어))를 기술합니다.
- personality**는 입력 텍스트에서 추론된 빅 파이브 특질과 성향을 기술하는 **Trait** 오브젝트의 반복 배열입니다.
- needs**는 입력 텍스트에서 추론된 욕구를 기술하는 **Trait** 오브젝트의 배열입니다.
- values**는 입력 텍스트에서 추론된 가치를 기술하는 **Trait** 오브젝트의 배열입니다.
- behavior**는 요일 및 일중 시간 중에 콘텐츠의 배포를 기술하는 **Behavior** 오브젝트의 배열입니다. 서비스는 시간소인이 있는 JSON 입력에 대해서만 필드를 리턴합니다.
- consumption_preferences**는 소비 성향의 각 범주에 대한 결과를 제공하는 **ConsumptionPreferencesCategory** 오브젝트의 배열입니다. 배열의 요소는 해당 범주의 개별 성향에 대한 정보를 제공합니다. 서비스는 요청의 **consumption_preferences** 조회 매개변수가 **true**로 설정된 경우에만 필드를 리턴합니다.
- warnings**는 입력 텍스트에 대한 메시지를 제공하는 **Warning** 오브젝트의 배열입니다. 입력이 경고를 생성하지 않으면 배열은 비어 있습니다.

<https://cloud.ibm.com/docs/personality-insights?topic=personality-insights-output>

JSON Profile

성격 특성 결과

`Profile` 오브젝트에는 항상 모든 유형의 입력에 대해 `personality`, `needs` 및 `values` 필드가 포함되어 있습니다. 이러한 각 필드에는 해당 특성 유형의 속성에 대한 성격 특성을 기술하는 `Trait` 오브젝트의 배열이 포함되어 있습니다. 욕구 및 가치 특성의 경우, 배열에는 특성을 기술하는 단일 레벨이 있습니다. 빅 파이브 특성의 경우, 최상위 레벨 배열은 특질을 기술하며 2차 레벨 배열은 각 특질의 성향을 기술합니다.

- `trait_id` (문자열)는 결과가 관련된 특성의 고유 ID입니다.
 - `big5_{characteristic}`: 빅 파이브 성격 특질의 경우
 - `facet_{characteristic}`: 빅 파이브 성격 성향의 경우
 - `need_{characteristic}`: 욕구의 경우
 - `value_{characteristic}`: 가치의 경우
- `name` (문자열)은 사용자에게 보이는 특성의 이름입니다.
- `category` (문자열)는 특성의 범주입니다.
 - `personality`: 빅 파이브 성격 특성의 경우
 - `needs`: 욕구의 경우
 - `values`: 가치의 경우
- `percentile` (double)은 특성에 대한 정규화된 백분위수 점수입니다. 자세한 정보는 [성격 특성에 대한 백분위수](#)를 참조하십시오.
- `raw_score` (double)은 특성에 대한 원 점수입니다. 이 필드는 `raw_scores` 조회 매개변수를 `true`로 설정하여 원 점수를 요청하는 경우에만 리턴됩니다. 자세한 정보는 [성격 특성에 대한 원 점수](#)를 참조하십시오.
- `significant` (부울)은 특성이 입력 언어에 대해 의미가 있는지 여부를 표시합니다. 영어, 스페인어 및 일본어 입력의 모든 특성에 대해 이 필드는 항상 `true`입니다. 이 필드는 서비스의 모델이 의미 있는 결과를 생성할 수 없는 아랍어와 한국어 입력의 특성의 서브세트에 대해 `false`입니다. 자세한 정보는 [아랍어와 한국어 입력에 대한 제한사항](#)을 참조하십시오.
- `children`은 입력 텍스트에서 추론된 각각의 빅 파이브 특질의 성향에 대한 보다 자세한 결과를 제공하는 `Trait` 오브젝트의 배열입니다. 배열은 빅 파이브 특질에 대해서만 리턴됩니다.

JSON Profile

소비 성향 결과

`consumption_preferences` 조회 매개변수가 `true`로 설정된 경우, `Profile` 오브젝트에는 `consumption_preferences` 필드가 포함되어 있습니다. 이 필드에는 성향의 각 범주에 대한 `ConsumptionPreferencesCategory` 오브젝트가 포함되어 있습니다.

- `consumption_preference_category_id` (문자열)는 `consumption_preferences_{category}` 양식으로 결과가 관련되는 소비 성향 범주의 고유 ID입니다.
- `name` (문자열)은 사용자에게 보이는 소비 성향 범주의 이름입니다.
- `consumption_preferences`는 범주의 개별 성향에 대한 결과를 제공하는 `ConsumptionPreferences` 오브젝트의 배열입니다.

범주의 각 개별 성향은 `ConsumptionPreferences` 오브젝트를 통해 기술됩니다. 일부 범주에는 단일 성향만 있으며, 기타 범주에는 다수의 성향이 있습니다.

- `consumption_preference_id` (문자열)는 `consumption_preferences_{preference}` 양식으로 결과가 관련되는 소비 성향의 고유 ID입니다.
- `name` (문자열)은 사용자에게 보이는 소비 성향의 이름입니다.
- `score` (double)는 작성자가 항목을 선호할 가능성을 표시하는 점수입니다. 자세한 정보는 [소비 성향에 대한 점수](#)를 참조하십시오.

행동 결과

서비스에 대한 입력이 개별 콘텐츠 항목에 대한 시간소인이 있는 JSON인 경우, `Profile` 오브젝트에는 `behavior` 필드가 포함되어 있습니다. 이 필드에는 각 요일 및 일중 시간에 대한 `Behavior` 오브젝트가 포함되어 있습니다.

- `trait_id` (문자열)는 결과가 관련된 특성의 고유 ID입니다.
 - `behavior_{day}`: 요일의 경우(예: `behavior_sunday`).
 - `behavior_{hour}`: 일중 시간의 경우(예: `behavior_0000`).
- `name` (문자열)은 사용자에게 보이는 특성의 이름입니다.
- `category` (문자열)는 특성의 범주이며, 이는 항상 `behavior`입니다.
- `percentage` (double)는 해당 요일 또는 일중 시간 중에 발생한 콘텐츠 항목의 백분율입니다. 자세한 정보는 [행동 특성에 대한 백분율](#)을 참조하십시오.

CSV Profile

빅 파이브 친화성 백분위수 (7개 열)	big5_agreeableness / facet_altruism / facet_cooperation / facet_modesty / facet_morality / facet_sympathy / facet_trust	이름 지정된 특질 또는 성향에 대해 텍스트 작성자의 정규화된 백분위수 점수.
빅 파이브 성실성 백분위수 (7개 열)	big5_conscientiousness / facet_achievement_striving / facet_cautiousness / facet_dutifulness / facet_orderliness / facet_self_discipline / facet_self_efficacy	이름 지정된 특질 또는 성향에 대해 텍스트 작성자의 정규화된 백분위수 점수.
빅 파이브 외향성 백분위수 (7개 열)	big5_extraversion / facet_activity_level / facet_assertiveness / facet_cheerfulness / facet_excitement_seeking / facet_friendliness / facet_gregariousness	이름 지정된 특질 또는 성향에 대해 텍스트 작성자의 정규화된 백분위수 점수.
빅 파이브 감정 범위 백분위수 (7개 열)	big5_neuroticism / facet_anger / facet_anxiety / facet_depression / facet_immoderation / facet_self_consciousness / facet_vulnerability	이름 지정된 특질 또는 성향에 대해 텍스트 작성자의 정규화된 백분위수 점수.
빅 파이브 개방성 백분위수 (7개 열)	big5_openness / facet_adventurousness / facet_artistic_interests / facet_emotionality / facet_imagination / facet_intellect / facet_liberalism	이름 지정된 특질 또는 성향에 대해 텍스트 작성자의 정규화된 백분위수 점수.

CSV Profile

욕구 백분위수 (12개 열)	need_liberty / need_ideal / need_love / need_practicality / need_self_expression need_stability / need_structure / need_challenge / need_closeness need_curiosity / need_excitement / need_harmony	이름 지정된 욕구에 대해 텍스트 작성자의 정규화된 백분위수 점수.
가치 백분위수 (5개 열)	value_conservation / value_hedonism / value_openness_to_change value_self_enhancement / value_self_transcendence	이름 지정된 가치에 대해 텍스트 작성자의 정규화된 백분위수 점수.
요일 백분율 (7개 열)	behavior_sunday / behavior_monday / behavior_tuesday / behavior_wednesday / behavior_thursday / behavior_friday / behavior_saturday	입력 텍스트에 시간소인이 있는 경우, 각 요일과 연관된 입력의 백분율. 그렇지 않으면 모든 백분율은 0.0입니다.
일중 시간 백분율 (24개 열)	behavior_0000 through behavior_2300	입력 텍스트에 시간소인이 있는 경우, 각 일중 시간과 연관된 입력의 백분율. 그렇지 않으면 모든 백분율은 0.0입니다.
단어 개수 및 언어 (2개 열)	word_count processed_language	입력 텍스트에 존재하는 단어의 수를 표시하는 정수이며, 텍스트 분석을 위해 서비스가 사용한 언어 모델의 두 문자 ID입니다.

Personality Insight API 활용

```
from ibm_watson import PersonalityInsightsV3
from ibm_cloud_sdk_core.authenticators import IAMAuthenticator
from os.path import join, dirname
import json
```

```
authenticator = IAMAuthenticator('발급받은 API키 입력')
personality_insights = PersonalityInsightsV3(
    version='2017-10-13',
    authenticator=authenticator
)
personality_insights.set_service_url('발급받은 URL 입력')
```

```
with open(join(dirname('profile'), './profile.json')) as profile_json:
```

```
    profile = personality_insights.profile(
        profile_json.read(),
        accept = 'application/json',
        content_language = 'en',
        content_type='application/json',
        consumption_preferences=True,
        accept_language='en',
        raw_scores=True
```

accept 항목입니다. 리턴 형식을 어떻게 받을것인가 입니다. text/csv도 가능합니다.

인풋 파일의 언어로, ar, en(디폴트), es, ja, ko가 분석 가능합니다.

입력 파일은 json 파일입니다. / text/csv도 가능합니다.

소비선호 항목 결과를 받습니다.

결과 언어를 영어로 받습니다. ko도 가능합니다.

분석 score 결과도 받습니다

만약 accpet가 text/csv일 경우 csv_header = True도 추가 할 수 있습니다.

항목에 대한 정보를 HEADER로 받겠다는 뜻입니다.

```
    ).get_result()
print(json.dumps(profile, indent=2))
```

Personality Insight API 실행 결과 (json 프로파일)

```
{
  "word_count": 15223,
  "processed_language": "en",
  "personality": [
    {
      "trait_id": "big5_openness",
      "name": "Openness",
      "category": "personality",
      "percentile": 0.8011555009552995,
      "raw_score": 0.7756540425503807,
      "significant": true,
      "children": [
        {
          "trait_id": "facet_adventurousness",
          "name": "Adventurousness",
          "category": "personality",
          "percentile": 0.8975586904731845,
          "raw_score": 0.5499070403121895,
          "significant": true
        },
        {
          "trait_id": "facet_artistic_interests",
          "name": "Artistic interests",
          "category": "personality",
          "percentile": 0.9770309419531928,
          "raw_score": 0.7663670485959848,
          "significant": true
        }
      ]
    }
  ]
}
```

```
{
  "trait_id": "facet_emotionality",
  "name": "Emotionality",
  "category": "personality",
  "percentile": 0.9947058875647474,
  "raw_score": 0.7524002152027129,
  "significant": true
},
{
  "trait_id": "facet_imagination",
  "name": "Imagination",
  "category": "personality",
  "percentile": 0.8733065387317427,
  "raw_score": 0.7915903144017664,
  "significant": true
},
{
  "trait_id": "facet_intellect",
  "name": "Intellect",
  "category": "personality",
  "percentile": 0.8717194796402048,
  "raw_score": 0.6597622585300698,
  "significant": true
},
{
  "trait_id": "facet_liberalism",
  "name": "Liberalism",
  "category": "personality",
  "percentile": 0.6405414845731188,
  "raw_score": 0.5343564751353818,
  "significant": true
}
]
```

결과 해석 :

(입력 예제) 여러 트윗들을 분석한
결과입니다

전반적으로 개방성이 높습니다.

children 항목으로

각각 모험심, 심미적 흥미, 감정,
상상력, 지능, 자율성(고정관념에서
어느정도 벗어나 있는가) 를
의미합니다.

Personality Insight API 실행 결과 (json 프로파일)

```
"behavior": [
  {
    "trait_id": "behavior_sunday",
    "name": "Sunday",
    "category": "behavior",
    "percentage": 0.21392532795156408
  },
  {
    "trait_id": "behavior_monday",
    "name": "Monday",
    "category": "behavior",
    "percentage": 0.425832492431887
  },
  {
    "trait_id": "behavior_tuesday",
    "name": "Tuesday",
    "category": "behavior",
    "percentage": 0.07164480322906155
  },
  {
    "trait_id": "behavior_wednesday",
    "name": "Wednesday",
    "category": "behavior",
    "percentage": 0.011099899091826439
  },
  {
    "trait_id": "behavior_thursday",
    "name": "Thursday",
    "category": "behavior",
    "percentage": 0.12209889001009082
  },
  {
    "trait_id": "behavior_friday",
    "name": "Friday",
    "category": "behavior",
    "percentage": 0.07769929364278506
  },
  {
    "trait_id": "behavior_saturday",
    "name": "Saturday",
    "category": "behavior",
    "percentage": 0.07769929364278506
  },
]
```

결과 해석 :

행동 소비 - 트윗 중 시간 요소에 대해
분류 한 것 입니다.

퍼센테이지가 높을 수록
트윗에 많이 언급된 것입니다.