

Implementación de un buscador de documentos

Andrea Engdahl Yunge



Andrea Engdahl Yunge

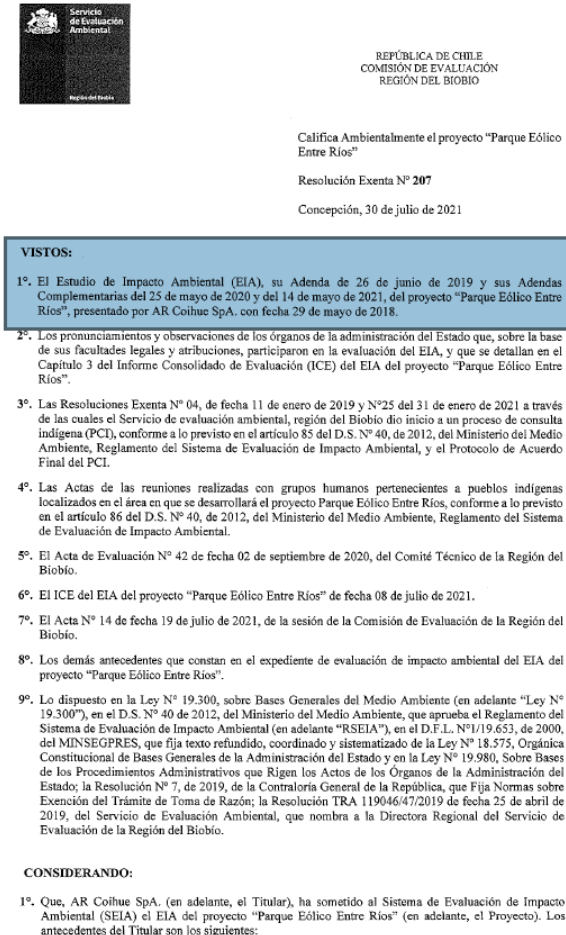
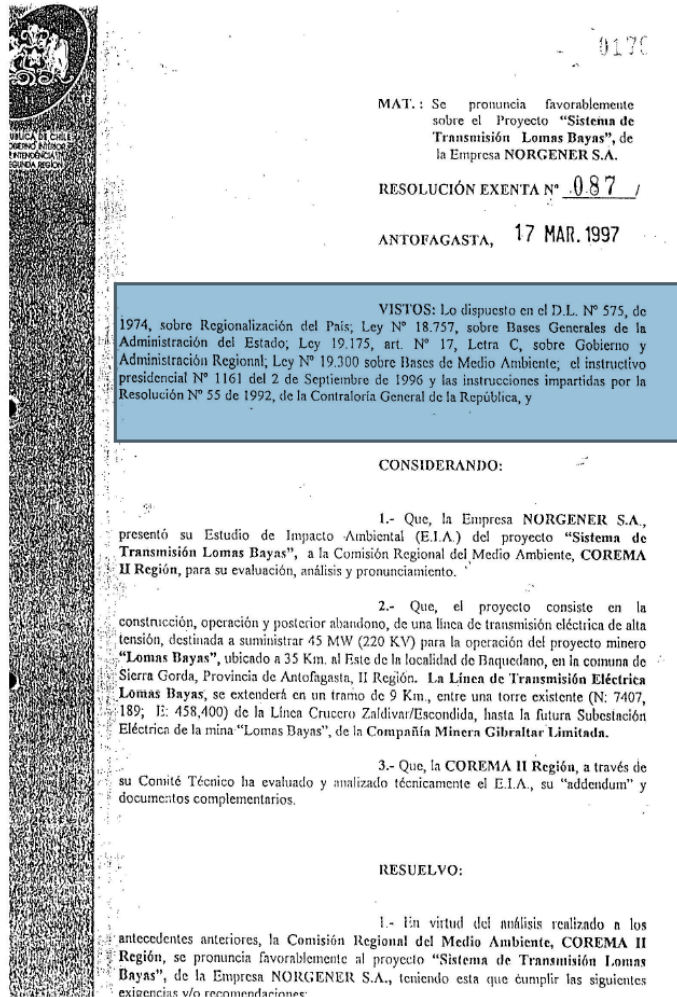
- ▶ Ing. Civil Matemática.
- ▶ Mg. en Cs. de la Ing. Mecánica.
- ▶ Diplomados en Big Data y Data Science.
- ▶ Mg(c) en Inteligencia Artificial.
- ▶ Funcionaria en la superintendencia del medio ambiente.

El problema...

Una gran cantidad de documentos en papel.
¿Quiénes tienen acceso?



Escanear o Digitalizar



¿Qué necesitamos?

REPÚBLICA DE CHILE
COMISIÓN DE EVALUACIÓN
Región Metropolitana de Santiago

Califica Ambientalmente el proyecto "Línea 7 Metro de Santiago"

Resolución Exenta N°541/2021

Santiago, 26 de julio de 2021

VISTOS:

1°. El Estudio de Impacto Ambiental (EIA) admitido a trámite con fecha 27 de junio de 2019 mediante Resolución Exenta N°343/2019 de la Comisión de Evaluación de la Región Metropolitana de Santiago, su Adenda de fecha 01 de abril de 2020, su Adenda Complementaria de fecha 15 de enero de 2021 y su Adenda Excepcional de fecha 13 de mayo de 2021 del Proyecto "Línea 7 Metro de Santiago" presentado por Empresa de Transporte de Pasajeros Metro S.A.

2°. Los pronunciamientos y observaciones de los órganos de la administración del Estado que, sobre la base de sus facultades legales y atribuciones, participaron en la evaluación del EIA, y que se detallan en el Capítulo 3 del Informe Consolidado de Evaluación (ICE) del EIA del proyecto "Línea 7 Metro de Santiago".

3°. El Acta de Evaluación N°01/2019 de fecha 23 de julio de 2019 y el Acta de Evaluación N°01/2021 de fecha 05 de julio de 2021, ambas del Comité Técnico de la Región Metropolitana.

4°. El ICE del EIA del proyecto "Línea 7 Metro de Santiago" de fecha 05 de julio de 2021.

5°. El acuerdo de la Comisión de Evaluación de la Región Metropolitana de fecha 13 de julio de 2021.

6°. La Resolución Exenta N°414/2012, de fecha 14 de septiembre de 2012, la Resolución Exenta N°469/2012, de fecha 25 de octubre de 2012, la Resolución Exenta N°589/2013, de fecha 09 de diciembre de 2013, la Resolución Exenta N°243/2014, de fecha 22 de abril de 2014, y la Resolución Exenta N°558/2015, de fecha 07 de diciembre de 2015, todas de la Comisión de Evaluación, Región Metropolitana de Santiago, que se modifican a través de la presente Resolución.

7°. Los demás antecedentes que constan en el expediente de evaluación de impacto ambiental del EIA del proyecto "Línea 7 Metro de Santiago".

8°. Lo dispuesto en la Ley N°19.300, Sobre Bases Generales del Medio Ambiente; el D.S. N°40, de 2012, del Ministerio del Medio Ambiente, Reglamento del Sistema de Evaluación de Impacto Ambiental ("Reglamento del SEIA"); la Ley N°19.880, que Establece Bases de los Procedimientos Administrativos que Rigen los Actos de los Órganos de la Administración del Estado; la Ley N°18.575, Ley Orgánica Constitucional de Bases Generales de la Administración del Estado; la Resolución TRA N°119046/163/2018 de fecha 25 de octubre de 2018, de la Dirección Ejecutiva del Servicio de Evaluación Ambiental; la Resolución Exenta RA N° 119046/83/2021, de 29 de enero de 2021, de la Dirección Ejecutiva del Servicio de Evaluación Ambiental; y en la Resolución N°7/2019 de fecha 26 de marzo de 2019, de la Contraloría General de la República, que Fija Normas sobre Exención del Trámite de Toma de Razón.

1°. El Estudio de Impacto Ambiental (EIA) admitido a trámite con fecha 27 de junio de 2019 mediante Resolución Exenta N° 343/2019 de la Comisión de Evaluación de la Región Metropolitana de Santiago, su Adenda de fecha 01 de abril de 2020, su Adenda Complementaria de fecha 15 de enero de 2021 y su Adenda Excepcional de fecha 13 de mayo de 2021 del Proyecto "Línea 7 **Metro de Santiago**" presentado por Empresa de Transporte de Pasajeros Metro S.A.

¿Qué herramienta podemos usar?



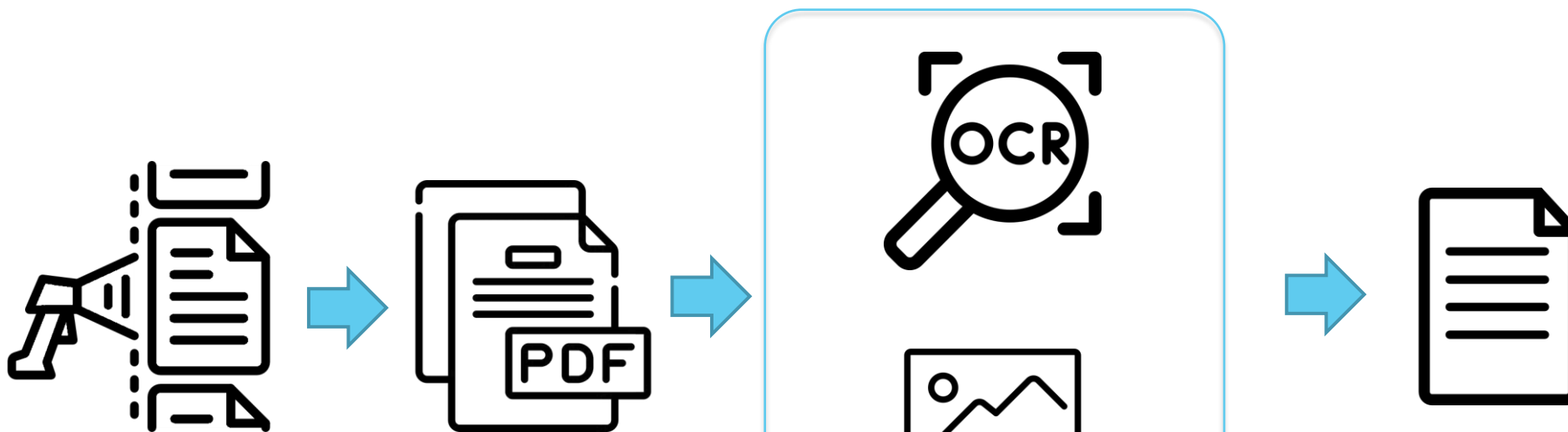
Python

+



ElasticSearch

Digitalización de Documentos.



Tesseract OCR



<https://smashicons.com>
<https://www.freepik.com>

Elasticsearch



- Es un motor de búsqueda, una herramienta de indexación que almacena información y luego nos permite localizarla.
- Se utiliza para monitorizar aplicaciones y ver los logs.

Elasticsearch	Base de datos relacional
Index	Base de datos
Type (*)	Tabla
Document	Fila
Field	Columna

¿Cómo funciona?

Índice Invertido

Documents 1 & 2



ID	Term	Document
1	butterfly	1
2	blue	1,2
3	bright	1,2
4	retire	2
5	wind	2

¿Cómo funciona?

Ejemplo: Metro de Santiago

Tokenización	Análisis
metro	metr
de	de
santiago	santiag

Creación y Conexión.

ElasticSearch & Python

```
!pip install elasticsearch
```

Conexión a ElasticSearch.

```
from elasticsearch import Elasticsearch  
es=Elasticsearch([{'host':'localhost', 'port':9200}])
```

Creación de un índice.

```
request_body = {  
    "settings" : {  
        "number_of_shards": 1,  
        "number_of_replicas": 0  
    },  
  
    'mappings': {  
        'properties': {  
            'author': {'type': 'keyword'},  
            'texto': {'type': 'text', 'analyzer':'spanish'}  
        }  
    }  
}  
  
es.indices.create(index = 'test-index', body = request_body)
```

Añadir un Documento

Añadir información.

```
doc = {
    'author': 'SEA',
    'texto': '1°. El Estudio de Impacto Ambiental (EIA) admitido a trámite con fecha 27 de junio de 2019 mediante Resolu'
}
res = es.index(index="test-index", body=doc)

print(res['result'])
```

Búsqueda de una Frase

Query.

```
query = {
    'query': {
        'match_phrase': {
            'texto': 'metro de santiago' #va a buscar: metr ? santiag
        }
    }
}
```

```
result=es.search(index='test-index', body=query)
print(result)
```

```
{'took': 2, 'timed_out': False, '_shards': {'total': 1, 'successful': 1, 'skipped': 0, 'failed': 0}, 'hits': {'total': {'value': 1, 'relation': 'eq'}, 'max_score': 0.5753642, 'hits': [{'_index': 'test-index', '_type': '_doc', '_id': 'U0S_nnsBqAakVMBYnE8H', '_score': 0.5753642, '_source': {'author': 'SEA', 'texto': '1°. El Estudio de Impacto Ambiental (EIA) admitido a trámite con fecha 27 de junio de 2019 mediante Resolución Exenta N°343/2019 de la Comisión de Evaluación de la Región Metropolitana de Santiago, su Adenda de fecha 01 de abril de 2020, su Adenda Complementaria de fecha 15 de enero de 2021 y su Adenda Excepcional de fecha 13 de mayo de 2021 del Proyecto "Línea 7 Metro de Santiago" presentado por Empresa de Transporte de Pasajeros Metro S.A.'}]}}}
```

Reflexiones Finales.

- ▶ Digitalización e implementación de un buscador de documentos
 - ▶ Optimizar la búsqueda y respuesta ante solicitudes.
 - ▶ Eficiencia organizacional.

Agradecimiento



Gianina Salomón

Programadora Web Full Stack



SMA

Superintendencia
del Medio Ambiente
Gobierno de Chile

Gracias!