

# Aggregator Specification V1

## Audience

Aggregators that use GetFTR

## Use Case

An Aggregator wishes to share the DOIs that they hold with GetFTR, so that our internal API can send entitlement requests for those DOIs to both the Publisher and Aggregator.

## DOI Ingestion

### Proposal

Without a list of DOIs held by the aggregator, GetFTR is unable to determine which entitlement requests it should send to both Publisher and Aggregator APIs. We propose ingesting lists of DOIs per aggregator platform.

### Process

At a high level:

1. `.jsonl.gz` files are periodically submitted to GetFTR using an AWS S3 Bucket
2. GetFTR ingests these files and adds or removes DOI holdings for the aggregator

3. Our internal API receives an entitlement request for a DOI held by an aggregator and sends an entitlement request to both the Publisher and Aggregator

## Onboarding

Once an aggregator has agreed to join GetFTR, they will provide the name of their platform.

GetFTR then provides access to submit files for ingestion to an AWS S3 Bucket.

**Note:** A test/staging bucket is also available for testing purposes.

## Criteria

### Files

Each file submitted must conform to the following:

- A maximum of 10,000 DOIs
- [Contains newline delimited JSON](#)
- Is gzipped before submission
- Filename is unique, includes a UUID (eg.

```
f0e17914-9c70-4520-8196-4f8b47d40876.jsonl.gz)
```

### Schema

See the below [JSON Schema](#) for each line within the file:

```
{
  "$schema": "https://json-schema.org/draft/2020-12/schema",
  "title": "Aggregator Deposit Schema",
  "description": "Schema for associating a DOI with an aggregator",
  "type": "object",
  "required": ["doi"],
  "additionalProperties": false,
  "examples": [
    {
```

```

    "doi": "10.1234/5678",
    "accessType": "paid",
    "vor": [
      {
        "url": "https://example.com/content.pdf",
        "contentType": "application/pdf"
      }
    ]
  },
  {
    "doi": "10.9876/54321",
    "deleted": true
  },
  {
    "doi": "10.5678/1234",
    "accessType": "paid"
  }
],
"properties": {
  "deleted": {
    "type": "boolean",
    "description": "If true, the record will be disassociated"
  },
  "doi": {
    "type": "string",
    "description": "DOI to store information against"
  },
  "accessType": {
    "type": "string",
    "description": "Access Type for the DOI",
    "enum": ["paid", "open", "free", "permFree"]
  },
  "vor": {
    "type": "array",
    "minItems": 1,
    "description": "List of additional content links for the DOI,
typically used for providing a link directly to a PDF",
    "items": {
      "type": "object",
      "required": ["url"],
      "additionalProperties": false,
      "properties": {
        "url": {
          "type": "string",
          "description": "URL to the content",
          "pattern": "^https?://"
        },
        "contentType": {
          "type": "string",
          "description": "Content Type for the URL, eg.
application/pdf",

```

```
    "enum": [
      "application/pdf",
      "text/html",
      "application/epub+zip",
      "other"
    ]
  }
}
}
```

## Updates

Updates to records will overwrite existing data for the association of a DOI with an Aggregator. When `deleted` is set to `true`, the association of a DOI for an Aggregator is removed.

## Error Handling

The ingestion pipeline will be monitored by GetFTR. The pipeline will automatically handle recoverable failures. Failures that cannot be recovered from automatically will be manually handled by GetFTR. GetFTR will reach out to the producer if their input is required.

## Entitlement Resolution

For DOIs submitted by an aggregator with an `accessType` of `paid`, requests will be sent to the Aggregator Entitlement API. For DOIs with an `accessType` of `open`, `free` or `permFree`, the platform will respond on behalf of the aggregator, using the `vor` property submitted alongside the DOI.

# Entitlement API

An Aggregator providing an Entitlement API may follow the Publisher Entitlement API specification for implementation information, provided here:

<https://www.getfulltextresearch.com/guidelines-specifications>