xBRL EUROPE

25th XBRL EUROPE DIGITAL WEEK

Online conference

15-18th June 2020

Integrated Reporting with xBRL-CSV

- Mark Goodhand
  - mrg@corefiling.com
- Head of Research, CoreFiling
- XBRL UK

What is “Integrated Reporting”?

- Deloitte’s definition?

  *Integrated Reporting brings together material information about an organisation's **strategy, governance, performance** and prospects in a way that reflects the commercial, social and environmental context within which it operates. It leads to a clear and **concise** articulation of **your value creation story** which is useful and relevant to all stakeholders.*

What is “Integrated Reporting”?

- integratedreporting.org definition?

  *Integrated reporting is a process founded on integrated thinking that results in a **periodic integrated report** by an organization about **value creation over time** and related communications regarding aspects of value creation.*

  *An integrated report is a **concise** communication about how an organization’s **strategy, governance, performance** and prospects, in the context of its external environment, lead to the creation of value in the short, medium and long term.*

What is “Integrated Reporting”?

- ECB definition ✔

  *The idea behind the ESCB **IReF** is to integrate, as far as possible, existing ESCB **statistical data** requirements for banks … into a **single framework**, thus ensuring **integration across countries and data domains**. In particular, the IReF would consist of:*

  *(i) an **integrated set of reports for banks**, aimed in the long run at **replacing national reporting templates** as far as possible; and*

  *(ii) a unique set of **transformation rules for compiling the derived statistics** required by authorities, possibly to be shared with the stakeholders involved.*

Scope of IReF

- Balance Sheet Items (BSI)
- Interest Rates (MIR)
- Securities Holdings Statistics (SHS)
- Credit and Credit Risk (AnaCredit)

- Also integrate with COREP & FINREP data from the EBA

Granularity & Consistency

- We want **one reporting format** that can cope with all needs
- The format must be suited to **dimensional data**
- The format must be able to handle **large volumes** of **granular data**
  - *The IReF is not meant to introduce new reporting requirements.*
  - *However, it may add some **additional granularity** compared to the datasets it is intended to integrate in order to ensure the integration of the existing reporting lines and avoid a duplication of the requirements.*
  - *users will be able to **drill down** from **aggregated data** to the **underlying granular information***

Why CSV?

- Simple, plain text, used since the 70s
- Supported by a wide range of popular tools & frameworks
  - R, Pandas, Spark, Beam
  - LibreOffice, Numbers, Microsoft Excel
  - sort, uniq, awk, grep
- Suitable for extremely large data volumes
  - Parallel & distributed processing of CSV lines
  - Can be processed in a streaming fashion, line-by-line

Problems with traditional CSV

- Columns are defined by out-of-band, unstructured documentation
- No standardised index for a collection of related documents
- Range of dialects
  - Character encoding, comments, header rows, quote characters, line terminators, trimming of whitespace
- No standardised semantics & validation
  - datatypes, co-constraints, arithmetic relationships
- Presentation considerations mixed with data modelling
- Cannot naturally represent hierarchical data

Previous CSV standardisation efforts

- 2005 – RFC 4180
  - Specifies line endings and whitespace handling
  - Requires MIME parameters for encoding and header row presence/absence
- 2012 – Frictionless Tabular Data Package
  - JSON metadata file, field types, primary & foreign keys, constraints
- 2015 – W3C Metadata Vocabulary for Tabular Data (“CSVW”)
  - Formalises and extends Frictionless spec
  - Adds support for RDF

xBRL-CSV

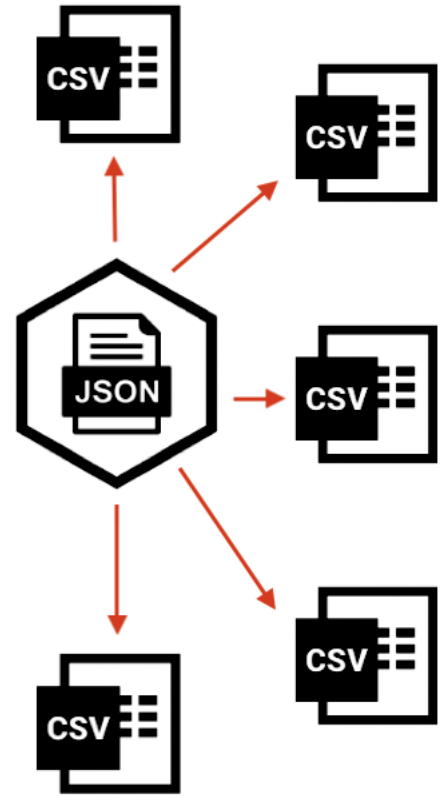
- Documentation in a standard, structured form (taxonomy)
  - Unlike CSVW, data model is decoupled from the report syntax
  - Benefits from taxonomy-side validation and rendering instructions
- Developed based on xBRL regulatory reporting experience
- JSON metadata
  - Points to a collection of CSV files
  - Defines dimensional bindings for tables, rows and columns
- Fixed dialect, based on RFC 4180
  - Additional constraints on encoding and column headers
- Special features for handling dimensional data

xBRL-CSV

- JSON metadata file + CSV files
- Share information
  - Common to all facts in a row
  - Common to all facts in a column
  - Common to all facts in a table
  - Common to all facts in a report

```
{
    "documentInfo": {
        "documentType": "http://xbrl.org/YYYY/xbrl-csv",
        "namespaces": {
            "ld": "http://xbrl.org/oim/conformance/firm-loa
            "lei": "http://standards.iso.org/iso/17442",
            "iso4217": "http://www.xbrl.org/2003/iso4217"
        },
        "taxonomy": [
            "https://xbrl.org/oim/conformance/firm-loans.xs
        ]
    },
    "dimensions": {
        "entity": "$entityLEI",
        "period": "$reportPeriod"

    },
    "tableTemplates": {
    "loan_data_template": {
        "columns": {
            "loan_id": {},
            "company_lei": {
                "dimensions": {
                    "concept": "ld:CompanyLEI"
                }
            },
            "country_inc": {
                "dimensions": {
                    "concept": "ld:CountryOfIncorporat
                }
            },
            "local_currency": {},
            "fixed_rate_period": {},
            "deposit_amount_hc": {
                "dimensions": {
                    "concept": "ld:DepositAmount",
                    "period": "$fixed_rate_period@start
                    "unit": "iso4217:EUR"
                }
            },
```



xBRL-CSV modelling approaches

- Datapoint-centric
- Table-centric
- Dimension-centric

Datapoint-centric modelling

- xBRL-CSV metadata defines datapoints
- Filers provide values for them (and some report-level parameters)

```
datapoint,value,unit
dp1234,42000,iso4217:EUR
dp1235,99000,iso4217:GBP
dp1236,r:FR
```

- Advantages:
  - Extremely compact format for data transfer
  - Trivial mapping to datapoint-centric data stores
- Disadvantages:
  - Can’t analyse data directly, as dimensional information is elsewhere
  - One-fact-per-line approach rules out some validation optimisations

Table-centric modelling

- xBRL-CSV metadata aligns with template row-column codes
- Filers provide values for them (and some report-level parameters)

```
_____,c0010,c0020,c0030
r0010,42000,41000,40000
r0020,99000,98000,97000
r0030,r:FR,r:GB,r:ES
```

- Advantages:
  - Reduced metadata size (don’t enumerate all combinations)
  - More potential for applying validation on a row-by-row basis
- Disadvantages:
  - Analysis requires understanding of row-column codes
  - Not suitable for all table arrangements

Dimension-centric modelling

- xBRL-CSV table for each allowed combination of dimensions
- Filers provide values for them (and some report-level parameters)

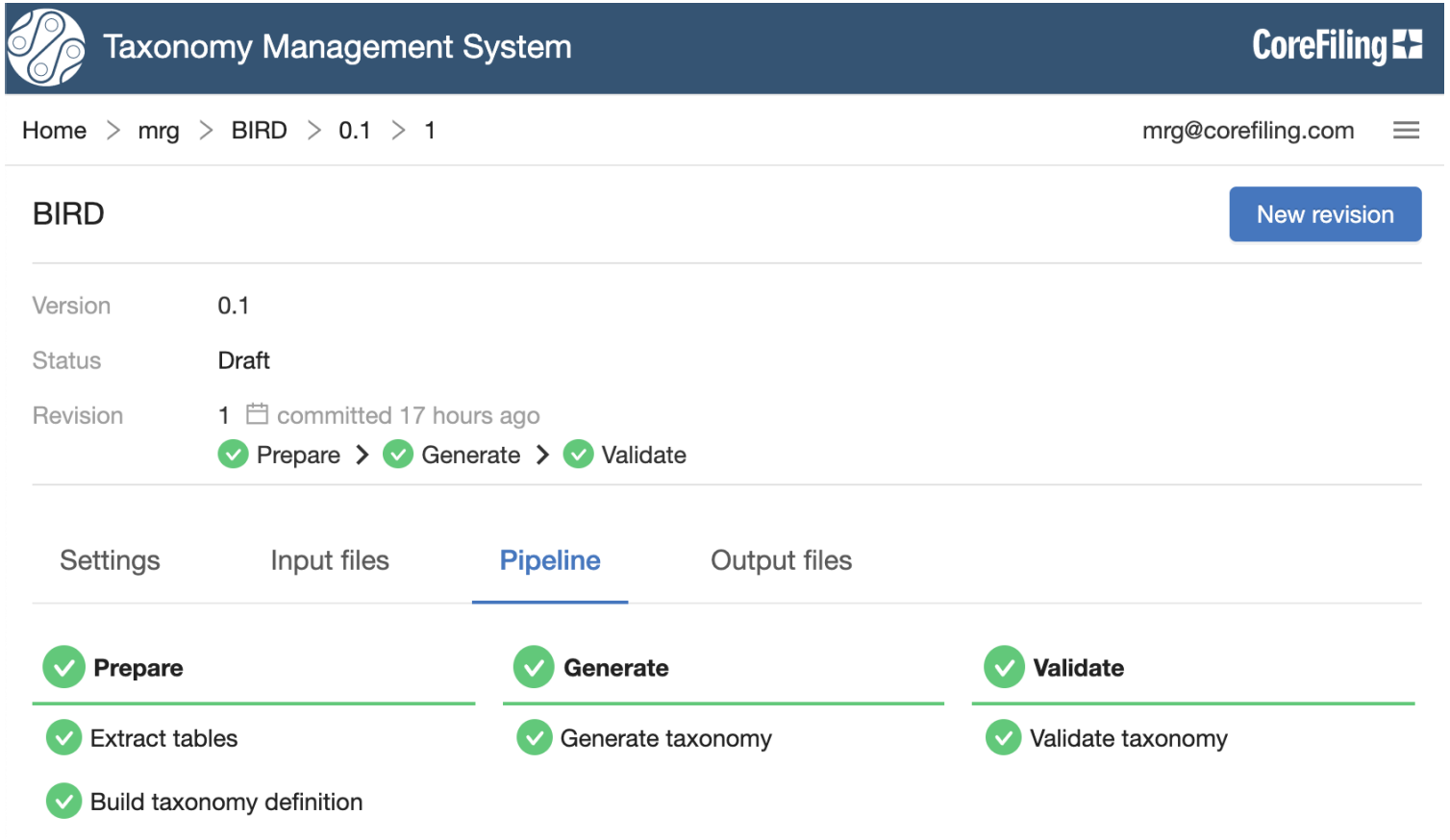
```
region,product,met1,met2,note1
R:X,P:Y,41000,40000,abc
R:X,P:Z,99000,98000,def
R:U,P:W,11000,12000,xyz
```

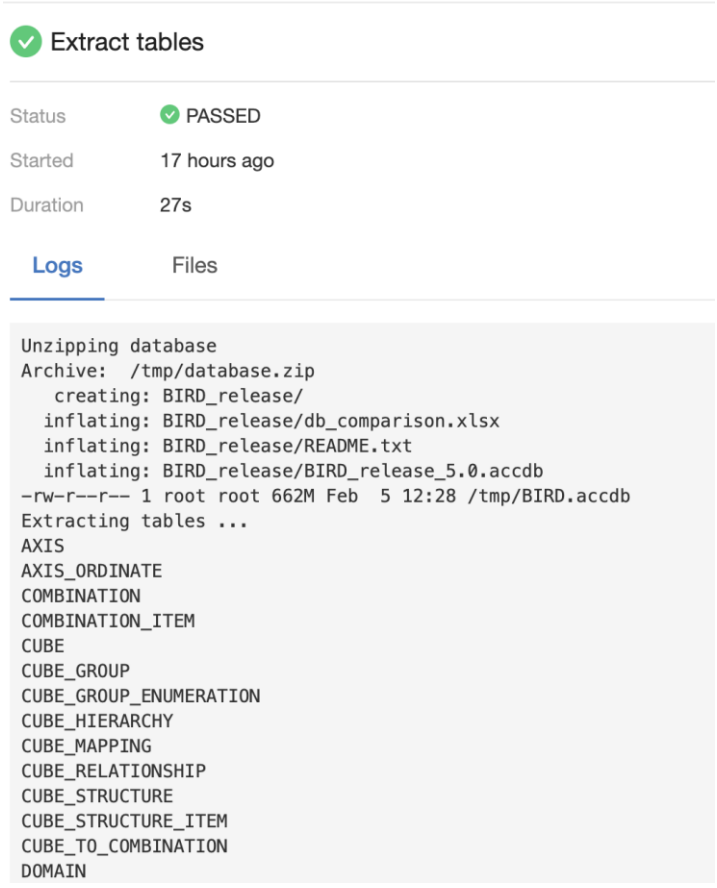
- Advantages:
  - All data in a given column has the same datatype
  - Can be analysed in-situ, using standard CSV tools
- Disadvantages:
  - More verbose, as dimension values are present on each row
  - Connection to business templates / datapoint model is not clear

Generating xBRL-CSV metadata

- Generate Taxonomy and xBRL-CSV metadata from DPM DB
  - Natural choice for Datapoint-centric modelling
  - Puts datapoints in an open, standard format
- Generate xBRL-CSV metadata from Table Linkbase
  - Good fit for Table-centric modelling
- Generate Taxonomy and xBRL-CSV metadata from cube model
  - Good fit for Dimension-centric models like BIRD

Taxonomy + xBRL-CSV metadata from BIRD





Benefits of xBRL-CSV over xBRL-XML

- Vastly reduced file size
- Simpler, more consistent format
- Potential for much faster validation

A note on XPath in Formula

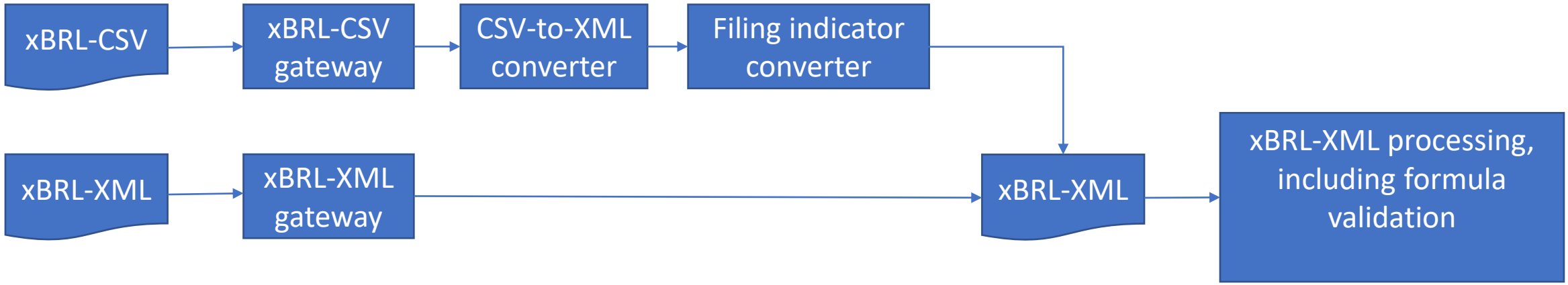
- XPath node navigation in Formula is inherently problematic
  - Prevents application of rules in non-XML representation
  - Forces processors to retain an XML model in memory (e.g. DOM)
  - Hampers performance even with xBRL-XML syntax
  - Processing XBRL as XML is inherently tricky and bug-prone
- XBRL International has a long-standing plan to restrict XPath expressions in Formula to remove node-navigation features
  - In the short term, this will be a restricted XPath mode to allow processors to optimise

Phased introduction of xBRL-CSV

- For existing mandates (e.g. CRDIV, EIOPA), regulators may choose to support xBRL-CSV as a filing option in parallel with xBRL-XML.
- For as long as the xBRL-XML route is provided, it is desirable to keep the xBRL-XML representation as close to the current xBRL-XML representation as possible.
- Filing indicators as used in Eurofiling require special handling
  - Tuples and custom attributes are not supported by OIM
- The closer you get to a pure OIM filing system, the more benefits you realise

Stage 1: bolt-on xBRL-CSV filing route

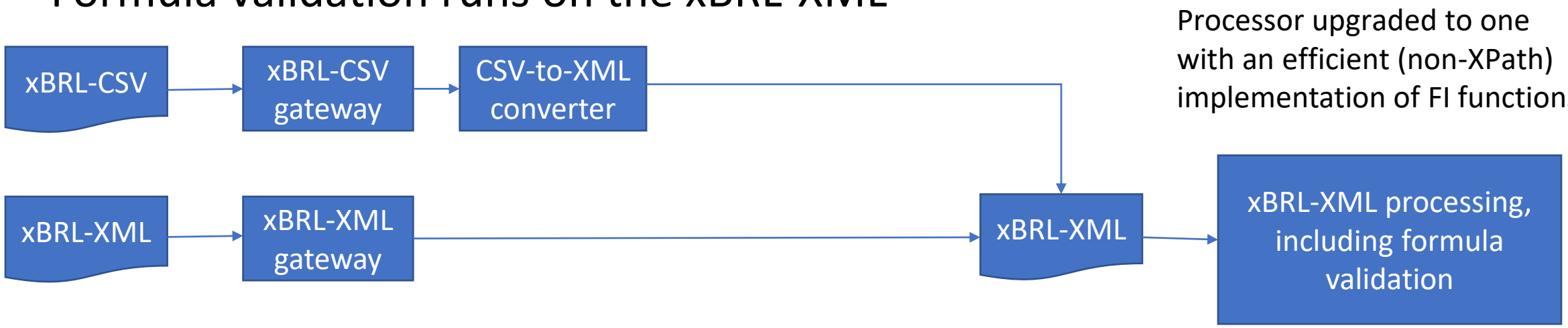
- Filers may submit in xBRL-CSV or xBRL-XML
- xBRL-CSV submissions are converted to xBRL-XML
  - with tuple-based filing indicators
- Formula validation runs on the xBRL-XML



Benefits: simpler for filers, smaller submissions, no taxonomy changes required

Stage 2: taxonomy & tool support for new FIs

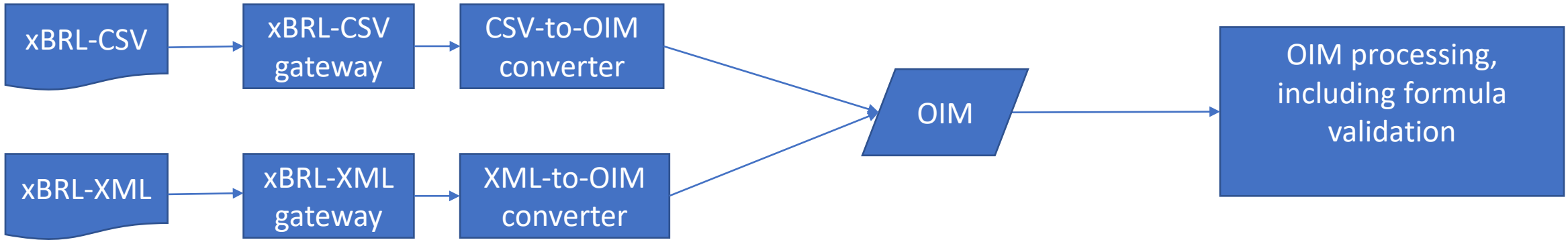
- Filers may submit in xBRL-CSV or xBRL-XML
- xBRL-CSV submissions are converted to xBRL-XML
- Formula functions support both old and new FIs
- Formula validation runs on the xBRL-XML



Benefits: simpler for filers, smaller submissions, no FI conversion step

Stage 3: OIM-centric filing system

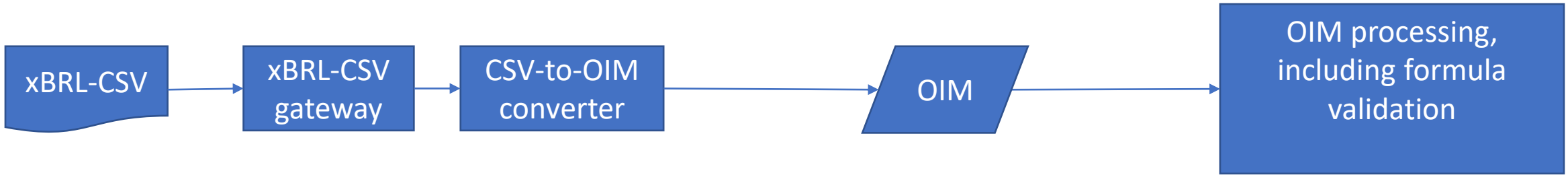
- Filers may submit in xBRL-CSV or xBRL-XML
- Submissions are loaded into OIM model (in-memory or DB)
- Formula functions replace all XPath XML navigation
- Formula validation runs on the OIM model



Benefits: simpler for filers, smaller submissions, more efficient formula input model (no XML baggage)

Stage 4: xBRL-CSV as the only filing format

- xBRL-XML is no longer a filing option
- Submissions are loaded into OIM model (in-memory or DB)
- Formula functions replace all XPath XML navigation
- Formula validation runs on the OIM model
  - Some rules moved to table/row level and applied early



Benefits: simpler for filers, smaller submissions, maximum potential for assertion optimisation

Progress of the xBRL-CSV standard

2020-05-06	Candidate Recommendation
2019-10-09	Candidate Recommendation
2019-08-07	Public Working Draft
2017-05-02	Public Working Draft

- At least one more *Candidate Recommendation* expected this year
- … then *Proposed Recommendation*
- … then *Recommendation*

- Strong interest from regulators
- EBA plans to accept both XBRL-XML format and XBRL-CSV format from DPM 3.1

Changes since 2020-05-06 CR

- Support for specifying decimals information within a cell (#403)
  - 4300d-2

- Support for properties associated with rows (#398)
  - Dimensional bindings in JSON metadata can apply to rows as well as columns
  - Replaces current Transposed tables feature

Get involved

- XEU Bank & Insurance Working Group
  - BIRD PoC
  - EBA TFERF updates

- XII OIM Working Group
  - Request membership: info@xbrl.org
  - Provide specification feedback: oim@xbrl.org