



# Current & Near-Term AI for XBRL

Herm Fischer, Exbee Ltd

**DIGITAL REPORTING IN EUROPE**

3-5 JUNE 2025, FRANKFURT

Hosted by the European Central Bank



**XBRL | EUROPE**

# Does AI need XBRL?

- LLMs do ingest \*textual\* financial reports for
  - Financial News Analysis
  - Anomalies in transactions
  - Concise statements from financial reports
  - Virtual assistance and prediction
- Top LLMs for Financial Large Models
  - FinGPT: open source, analysis and prediction
  - FinRobot: market forecasting and trading analysis
  - BloombergGPT: sentiment analysis, questions answering
  - InvestLM: curated financial database
  - FinLlama: algorithmic trading

# What can AI get from non-structured data?

- Table extraction
  - Table extraction is used to process financial reports, balance sheets, and income statements. This enables quick compilation of financial metrics for analysis, forecasting, and regulatory reporting
- At best, summary and face statement data
  - Lower level may be BS

# Use Cases for AI with XBRL

| Application of AI              | When     | Example                     |
|--------------------------------|----------|-----------------------------|
| Taxonomy authoring             |          |                             |
| Model synthesis                | Now      | GPT (for classes)           |
| Model construction             | Now      | XENDER editor + CoPilot     |
| Validation construction        | Now      | XULE editor + CoPilot       |
| Instance authoring             |          |                             |
| Location of content            | ?        |                             |
| Validation                     | ?        |                             |
| Correction of tagging          | ?        |                             |
| Assistance in error correction | ?        |                             |
| Instance consumption           |          |                             |
| OIM Taxonomy Modeled instances | Next yr? | Models cubes, sects, tables |

# Taxonomy Authoring

- Model (taxonomy) development
  - Sections/cubes, concepts, primary relationships
  - Dimensions
    - Period, unit, lang, etc
    - Model-defined axes, domains, members
    - Constraints
- Validation rules development
  - Aggregation
  - Dimensionality
  - Model integrity

# Instance Authoring

- Location of taggable content
  - Identification of cubes, sections, tables and parentheticals
    - Mapping of datapoints to cubes, sections, tables, parentheticals
    - Mapping of human-readable to machine-readable (transformations)
- Validation
  - Cube, section, table model-based constraints
  - Interpretation of tagging errors and constraint misfits

# Taxonomy Model Consumption

- XBRL 2.1, dimensions, Xlink linkbases unlikely to be consumed
  - Requires XBRL processor
  - Semantics obfuscated with arcane syntax
- OIM Taxonomy purposefully designed for AI consumption
  - Cubes, tables and restrictions are well modeled
  - Processors can pre-associate facts to cubes before submitting to AI
  - AI has a semantic model for the reported data

# Asking AI to compare OIM Taxonomy to 2.1 version

- “please compare uploaded XBRL OIM Taxonomy of IFRS to published IFRS”
  - The uploaded XBRL OIM Taxonomy for IFRS is a faithful representation of the published IFRS standards for the specified version (2024-03-27). It provides the necessary structure for digital reporting, with detailed domains and relationships. However, it is limited to the standards and concepts included in the taxonomy and may not cover all IFRS standards or include all label and reference details. For full compliance, users should ensure the taxonomy version matches the reporting period's applicable IFRS standards.
  - **Uploaded Taxonomy:** The taxonomy is organized into:
    - **Domain Roots:** Lists of concepts grouped by domains (e.g., ifrs-full:BiologicalAssetsDomain, ifrs-full:ClassesOfFinancialInstrumentsDomain).
    - **Networks:** Relationships between concepts (e.g., parent-child, summation-item) for specific roles (e.g., cash flow statements, earnings per share).
    - **Labels and References:** Includes labels and references for concepts, but the uploaded file does not show detailed label types or reference types.



# Asking AI, cont

- **4. Granularity**
  - **Uploaded Taxonomy:** Highly granular, with detailed line items and relationships (e.g., specific cash flow adjustments, reconciliation of changes in biological assets).
  - **Published IFRS:** Provides principles-based guidance but does not prescribe the same level of granularity as the taxonomy, which is designed for digital reporting.
- **5. Consistency with IFRS**
  - **Uploaded Taxonomy:** The concepts and relationships align with the published IFRS standards. For example:
    - The cash flow statement relationships in the taxonomy match the requirements of IAS 7.
    - The earnings per share disclosures align with IAS 33.
  - **Published IFRS:** The taxonomy is derived from the standards, so consistency is expected. However, the taxonomy may include additional technical details (e.g., XBRL-specific relationships) not present in the original standards.
- **7. Missing Elements**
  - The uploaded file does not include:
    - Detailed label types (e.g., terse, verbose labels).
    - Reference types (e.g., links to specific paragraphs in IFRS standards).
    - All IFRS standards (e.g., IFRS 9, IFRS 15, IFRS 16 are not explicitly referenced).

# Asking AI, cont

- “what cubes does Assets belong to?”
  - In the uploaded **IFRS XBRL Taxonomy (2024-03-27)**, the concept **ifrs-full:Assets** (or related asset concepts like **ifrs-full:PropertyPlantAndEquipment**, **ifrs-full:BiologicalAssets**, etc.) belong to several **cubes** (presentation networks) depending on the disclosure context. Below is a breakdown of the relevant cubes where **Assets** appear:
  - **1. General Asset Disclosure Cubes**
    - These cubes cover broad asset categories and reconciliations:
    - **ifrs\_entry\_point:ias\_1\_2024-03-27\_role-800000\_parent-child\_Network**
      - Used for the **Statement of Financial Position** (Balance Sheet), where Assets are presented hierarchically (e.g., Current Assets, Non-current Assets).
      - Example relationships:
      - {"source": "ifrs-full:StatementOfFinancialPositionAbstract", "target": "ifrs-full:AssetsAbstract", "order": 10.0}
    - **ifrs\_entry\_point:ias\_7\_2024-03-27\_role-800300\_parent-child\_Network**
      - Covers **Cash Flow Statement** adjustments related to assets (e.g., changes in working capital, depreciation).
      - Example:
      - {"source": "ifrs-full:CashFlowsFromUsedInOperatingActivitiesAbstract", "target": "ifrs-full:AdjustmentsForDecreaseInIncreaseInOtherAssets", "order": 160.0}

# Asking AI, cont

- “what cubes does Assets belong to?”
  - **2. Specific Asset-Type Cubes**
    - Assets are further detailed in standards-specific cubes:
      - **A. Property, Plant, and Equipment (IAS 16)**
        - **ifrs\_entry\_point:ias\_16\_2024-03-27\_role-822100\_parent-child\_Network**
          - Reconciliations for PPE (e.g., additions, disposals, depreciation).
      - **B. Biological Assets (IAS 41)**
        - **ifrs\_entry\_point:ias\_41\_2024-03-27\_role-824180\_parent-child\_Network**
          - Covers agricultural assets (e.g., by age, type).
      - **C. Intangible Assets (IAS 38)**
        - **ifrs\_entry\_point:ias\_38\_2024-03-27\_role-823180\_parent-child\_Network**
          - Reconciliations for intangible assets and goodwill.
      - **D. Investment Property (IAS 40)**
        - **ifrs\_entry\_point:ias\_40\_2024-03-27\_role-825100\_parent-child\_Network**
          - Fair value/cost model disclosures.
      - **E. Impairment (IAS 36)**
        - **ifrs\_entry\_point:ias\_36\_2024-03-27\_role-832410\_parent-child\_Network**
          - Impairment losses/reversals for assets.

# Asking AI, cont

- “what cubes does Assets belong to?”
  - **3. Domain-Specific Axes (Dimensions)**
    - Assets are also classified using **axes** (dimensional qualifiers) in table-style disclosures. For example:
      - **ClassesOfAssetsAxis**: Used to segment assets by type (e.g., PPE, intangibles).
      - **MeasurementAxis**: Distinguishes between cost/fair value models.
  - **Key Observations**
    - **Hierarchical Presentation**: Assets appear in parent-child networks for financial statements (e.g., Balance Sheet, Cash Flows).
    - **Standard-Specific Cubes**: Detailed reconciliations are in standards like IAS 16, IAS 38, IAS 40.
    - **Dimensional Reporting**: Axes like ClassesOfAssetsAxis enable granular tagging (e.g., by asset type, measurement basis).

# Asking AI, cont

- “what additional semantics does the OIM taxonomy convey compared to the published taxonomy?”
  - (jumping ahead...) **Conclusion**
    - The OIM taxonomy adds **semantic precision**, **modularity**, and **dynamic dimensionality** to traditional IFRS XBRL taxonomies. It aligns closer to the logical structure of IFRS standards while enabling more efficient digital reporting.
  - (back to start...):
    - The **XBRL Open Information Model (OIM) Taxonomy** for IFRS introduces several **additional semantics** compared to the traditional published IFRS taxonomy (e.g., the IFRS Foundation's XBRL Taxonomy). These enhancements aim to improve interoperability, reduce redundancy, and enable richer data relationships. Below is a detailed comparison:**Conclusion**
    - The OIM taxonomy adds **semantic precision**, **modularity**, and **dynamic dimensionality** to traditional IFRS XBRL taxonomies. It aligns closer to the logical structure of IFRS standards while enabling more efficient digital reporting.

# Asking AI, cont

- “what additional semantics does the OIM taxonomy convey compared to the published taxonomy?”
  - **1. Enhanced Relationship Semantics**
    - **Published IFRS Taxonomy (Traditional XBRL)**
      - Uses basic XBRL relationships:
        - parent-child (presentation).
        - calculation (summation).
        - reference (links to IFRS standards).
      - Relationships are flat and lack contextual depth (e.g., no distinction between "operating" vs. "investing" cash flows beyond labels).
    - **OIM Taxonomy**
      - Introduces **explicit semantic roles** for relationships:
        - **xbml:parent-child**: Retained but enriched with order and preferredLabel (e.g., xbrl:totalLabel for subtotals).
        - **Custom relationship types**:
          - E.g., adjustmentFor (links cash flow adjustments to specific line items).
          - reconciliationOf (ties reconciliations to primary statements).
        - **Contextual hierarchies**:
          - Example: In cash flows, ProceedsFromSalesOfPropertyPlantAndEquipment is explicitly tagged as investing activity via its network role (ias\_7\_2024-03-27\_role-800300).

# Asking AI, cont

- “what additional semantics does the OIM taxonomy convey compared to the published taxonomy?”
  - **2. Structured Domain-Root Taxonomies**
    - **Published Taxonomy**
      - Domains (e.g., BiologicalAssetsDomain) are static and predefined.
      - Limited ability to dynamically extend or link domains to external taxonomies.
    - **OIM Taxonomy**
      - **Dynamic domain roots:**
        - Domains like ifrs-full:BiologicalAssetsDomain are linked to **axes** (e.g., BiologicalAssetsByAgeAxis) with explicit membership rules.
        - Enables multidimensional reporting without redundant concepts.
      - **Cross-domain relationships:**
        - Example: BiologicalAssets can be linked to FairValueHierarchyDomain to show valuation techniques.

# Asking AI, cont

- “what additional semantics does the OIM taxonomy convey compared to the published taxonomy?”
  - **3. Modular and Reusable Components**
    - **Published Taxonomy**
      - Concepts are often duplicated across disclosures (e.g., Depreciation appears separately in PPE, intangibles, and cash flows).
    - **OIM Taxonomy**
      - **Reusable primitives:**
        - Common concepts (e.g., ifrs-full:Depreciation) are defined once and referenced across networks.
        - Example: The same Depreciation concept is used in:
          - IAS 16 (PPE).
          - IAS 38 (Intangibles).
          - IAS 7 (Cash Flows).
      - **Parameterized relationships:**
        - Relationships can carry attributes (e.g., weight=1.0 for additive calculations, order=60.0 for sequencing).
    - ...



# Instance Consumption

- Pre-requisite: cube semantics
  - Must model cubes, constraints, sections and tables
  - Publish association of facts to cubes(?)
- Well-modeled instances can be consumed today
  - Poorly modeled instances have no semantics
    - XBRL 2.1 not useful
    - xBRL-JSON useful for fact identification (not semantics)
    - xBRL-JSON + OIM Taxonomy suitable for semantic consumption

# Instance Consumption, a new goal, a new format

- Single Json structure, ingestible into current LLMs, without XBRL tools
- Start with the OIM Taxonomy Model
  - Resolved into single model of all imported Taxonomy references
  - OIM Reports, fact objects and footnotes are reformatted
    - Fact map/dictionary becomes OIM-style arrays of named fact objects
    - Footnote relationships become OIM Taxonomy
      - Networks, relationships and labels
  - Unused resolved taxonomy model components are removed
- Ensure consistent behavior across multiple LLMs, cloud and local

# Privacy

- AI must not compromise privacy of
  - Pre-filed data, pre-issued taxonomies, non-public rules and validation
- Expect prohibition of sending AI operations off-site
  - Do queries leak non-public information? (probably! Redit users say “definitely”!)
  - How are silos of private information added to local AI models
- Run your own LLM?
  - Yes, it is now reasonable: To run your own LLM, you'll need the model, an inference engine, sufficient hardware resources (CPU and RAM, with a GPU recommended), and the necessary software environment for the chosen engine.
  - Mac M3 ultra or M4 pro/max (or go buy an Nvidia)
- DIY: <https://thegeekinsights.com/running-deepseek-locally/>

# Using locally loaded 4GB open-source AI

- **Loaded OIM-Taxonomy of IFRS:**

Okay, so the user provided three JSON strings with some text inside. They asked about the XBRL IFRS taxonomy and how it compares to XBRL FUS (U.S. GAAP).

First, I need to figure out if these JSON strings are related to XBRL taxonomies or something else. Looking at the content inside each string, they seem to be long texts with various identifiers like citation numbers and some random characters. There's a mix of symbols, letters, and what appears to be code snippets, possibly related to XBRL since it uses XML schemas.

Since none of these strings clearly contain taxonomies or detailed information about IFRS or U.S. GAAP, I can't directly extract the relevant information from them. That means I don't have specific data on the user's OIM Taxonomy version of IFRS they mentioned.

In that case, I should respond by explaining what a taxonomy is and then provide a general comparison between IFRS and U.S. GAAP taxonomies since I can't get specifics from their examples.

I'll start by defining XBRL taxonomies for both standards, highlighting their key features like depth, focus areas, language support, application, interoperability, and maybe some typical use cases or differences in handling specific accounting rules.

This way, even without the specific data from the user's OIM Taxonomy, I can offer a comprehensive comparison based on general knowledge.

# Using locally loaded 4GB open source AI

“In the just-loaded OIM Taxonomy of IFRS what cubes does Assets belong to”

- In the OIM Taxonomy of IFRS (International Financial Reporting Standard), the concept of "Assets" is typically mapped under the **Cube: Governance, Asset Management, and Intangibles**. Within this framework, Assets are further subdivided into more specific dimensions such as:
- **Tangible Fixed Assets**
- **Intangible Assets**
- These sub-dimensions help structure accounting and reporting concepts related to assets, their measurement, classification, and presentation in financial statements. Each asset type is assigned a cube or dimension within the OIM Taxonomy to facilitate easier navigation and understanding of IFRS standards.

(Not happy: when home from Frankfurt I will try larger locally loaded models.)

# Provenance

- Attribute each AI result to source of trained inference and pattern sources
  - May be public
  - May be within private offline silo of trained data
  - Identify generated fictitious results (faked up examples)
- Nobody does this 🤖
- But wait:
  - R1 seems to track citations. (Investigating whether they are useful.)

# Conclusion

- This is exciting
  - What can be gleaned from better modeled Taxonomies is amazing
  - Open-source *\*is\** available
  - Running locally and offline *\*is possible\** and should protect filers
  - How do we integrate open-source AI into our tools
  - How do we train users

Thanks for your attention

# DIGITAL REPORTING IN EUROPE

3-5 JUNE 2025, FRANKFURT

Hosted by the European Central Bank

