

OPENEHR & FHIR DATA TYPE MAPPING

Gino Canessa

<http://aka.ms/ginoc/decks>

AGENDA

- About Me
- Context
- Work stream
- Progress!
- Details
- Next Steps
- Questions & Discussion

GINO CANESSA

- Principal Software Engineer @ Microsoft
 - MSR - Health Futures
 - Healthcare Standards and Interop
- Focused on FHIR since 2019
 - Infrastructure / Tooling
 - Code generation
 - Specification Development
 - FHIR-I, FMG, TSC, Argonaut Project
- Contact
 - Gino.Canessa@microsoft.com
 - [Zulip](https://chat.fhir.org) (chat.fhir.org)
 - [YouTube](#) (FHIR Educational Content)



CONTEXT

- Models are crystalized philosophy
- Structural choices based on ecosystem needs and goals
- Layers built by many hands over many years

ADDITIONAL CONTEXT

- "Data Type Mapping" is focused on the 'lowest' layers
- Reusable patterns and tools
- Guidance and dangers

EVEN MORE CONTEXT

OPENEHR

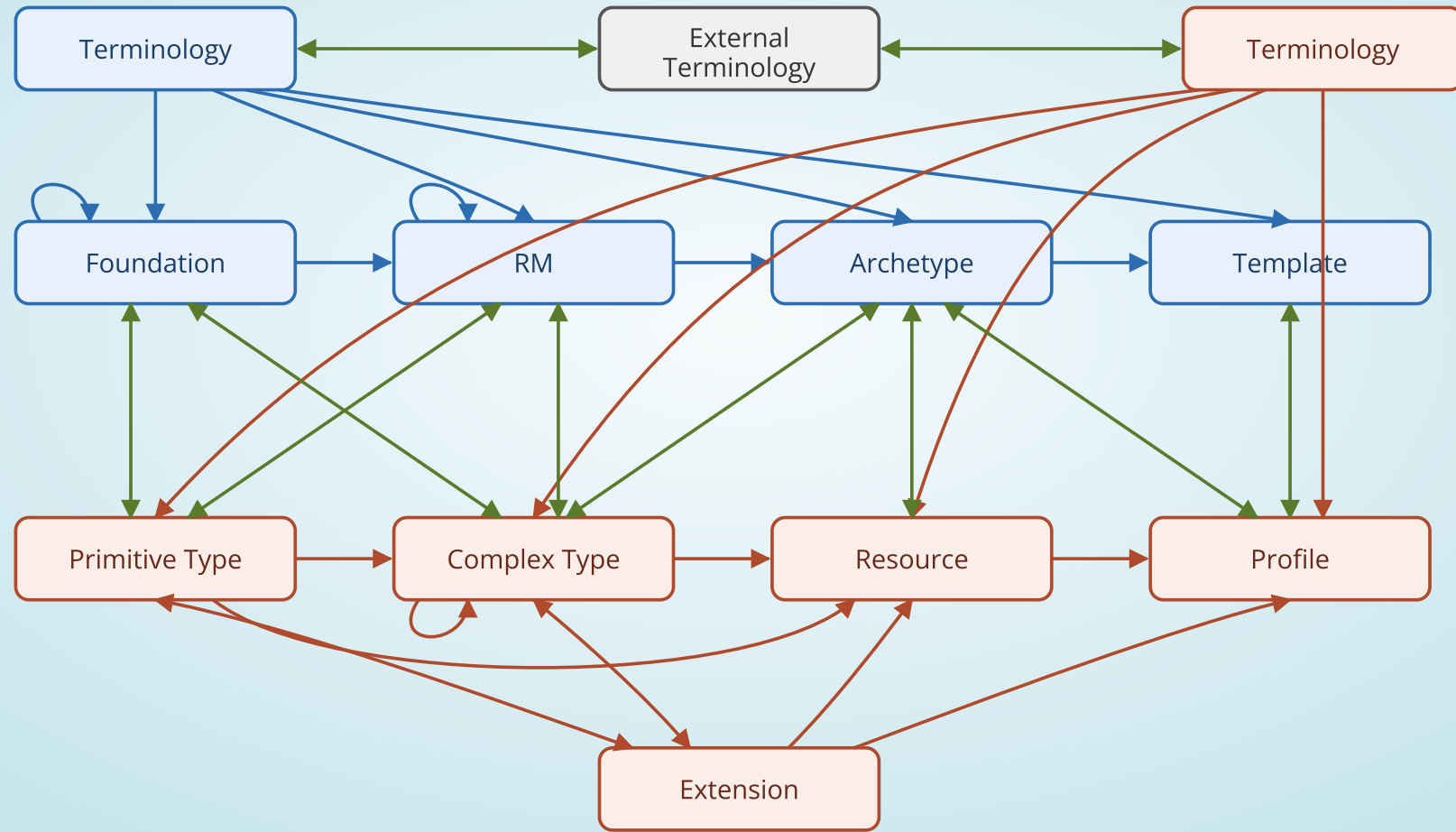


FHIR



LIES

REALITY



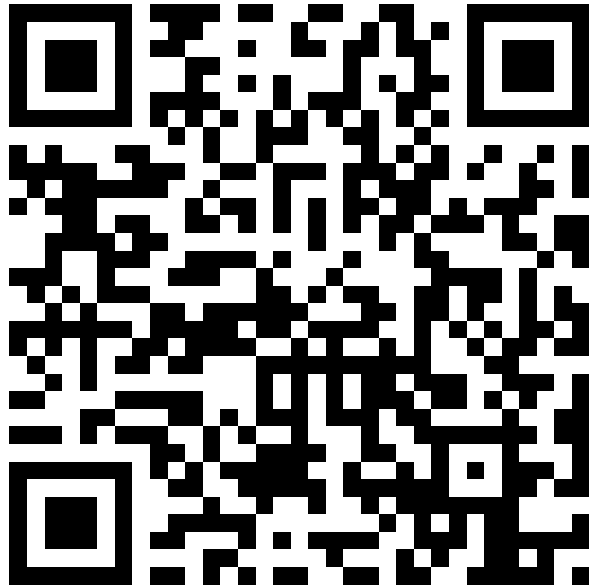
  HERE BE DRAGONS!  

- Disclaimer: This is work in progress and is subject to change
- Data types *cannot* always be mapped without context

GROUPS OF KINDS

- Boolean Data
- Numeric Primitives
- Resource-Locator Data and References
- Textual Data
- Coded Data
- Quantities
- Temporal Data
- Other Data (e.g, binary, etc.)

LINKS!



Draft Doc



Confluence

- [Draft Doc](#)
- [Intermediate Review Doc](#)
- [Long-term Working Doc](#)
- [WIP Confluence Overview](#)

BOOLEAN DATA

- The simplest mapping in the entire effort
- DV_BOOLEAN ↔ FHIR boolean
- A single true/false value on each side
- No precision, accuracy, or auxiliary metadata to negotiate

ABSENCE $\neq$ FALSE

- In FHIR, a **missing element is not** equivalent to false
- A null openEHR DV_BOOLEAN should be **omitted** from FHIR output
- If absence itself is meaningful, use the [data-absent-reason](#) extension

NUMERIC PRIMITIVES

- openEHR Foundation primitives (Integer, Real, Double)
 - Generally not directly in archetypes
 - Typically in DV_*/DV_*<T>
- FHIR defines numeric primitives
 - integer, integer64, decimal, positiveInt, unsignedInt
 - Used directly in elements *and* complex types (e.g., Quantity)
- Target FHIR type: range and context

DECIMAL PRECISION

- JSON does *not* preserve trailing zeros
- When converting to FHIR, preserve precision via:
 - [quantity-precision](#)

INTEGER SIZING

- FHIR integer is 32-bit
- FHIR integer64 exists for values that exceed $2^{31} - 1$
- integer64 is uncommon — most elements use integer
- For openEHR 64-bit attributes targeting 32-bit elements:
 - Map values that fit into the integer element
 - Use an extension for overflow values



DANGERS!

- Mapping Real $\leftrightarrow$ decimal requires care to avoid precision loss
- Tracking precision metadata is the responsibility of the mapping engine
- Many openEHR numeric properties map as different FHIR elements
 - magnitude, precision, numerator/denominator, etc.

RESOURCE-LOCATORS & REFERENCES

- openEHR types that point at something else
 - DV_URI, DV_EHR_URI
 - LINK (clinical, internal)
 - PARTY_IDENTIFIED (operational, external)
- FHIR equivalents
 - URI primitives: `uri`, `url`, `oid`, `uuid`, `canonical`
 - Reference and CodeableReference

URI PRIMITIVES

- FHIR URI primitives have progressively **narrower** constraints
 - `uri / canonical` — general
 - `url` — addressable
 - `oid / uuid` — strict formats
 - `canonical` — adds versioning semantics
- Target choice depends on the **element definition**, not the openEHR source

THE ehr: SCHEME

- ⚠ Still under discussion ⚠
- DV_EHR_URI constrains the scheme to ehr:
- Points at items inside an openEHR server:
 - compositions, sections, entries, sub-elements
- FHIR typically uses resource-level links or are context-specific
 - Resource references can have an extension for element-level
 - Questionnaire.item references use linkId

SUB-ELEMENT ADDRESSING

- LINK.target can address a sub-element of a composition
- FHIR Reference.reference is resource-level
 - #fragment reserved for contained resources
- Two standard FHIR extensions:
 - `targetElement` — targets `Element.id`
 - `targetPath` — uses restricted `FHIRPath`
- ⚠ Uncommon to use — produce only when *truly* needed ⚠

CLINICAL VS. OPERATIONAL

- openEHR splits:
 - LINK — clinical, internal record links
 - PARTY_IDENTIFIED — operational, external party references
- FHIR uses Reference for both
- WG investigating consolidating LINK and PARTY_IDENTIFIED onto a common RM class (e.g., OBJECT_REF)

CodeableReference

- FHIR CodeableReference combines:
 - a coded concept, *and*
 - a reference
- openEHR has no direct equivalent
- Mapping into openEHR generally requires splitting:
 - A coded element (e.g., BodySite)
 - A LINK (or enclosing CLUSTER)

TEXTUAL DATA

- openEHR DV_TEXT — attribute-rich class
 - value, formatting, encoding, language, mappings, hyperlink (deprecated)
- FHIR plain-text primitives
 - string (≤ 1 MB UTF-8)
 - markdown (CommonMark + string)
 - xhtml (constrained XHTML + string)
- Rendering extensions:
 - [rendering-markdown](#), [rendering-xhtml](#)

CHARACTER ENCODING

- openEHR allows different character_sets
 - UTF-8, UTF-16, US-ASCII, etc.
- FHIR mandates UTF-8
- Non-UTF-8 content must be converted before mapping
- The original encoding is not preserved in FHIR

EMPTY VALUES & LENGTH

- FHIR strips leading/trailing whitespace
- FHIR does not allow empty `string` values
 - Empty optional openEHR text → drop
 - Empty mandatory openEHR text → upstream data error
 - Can use `data-absent-reason` extension if absence is meaningful
- FHIR `string` capped at **1 MB**
 - openEHR has no formal cap; oversize values are exceptional

LANGUAGES

- FHIR represents language at multiple levels
 - `Resource.language` (resource-wide)
 - Element-level extensions:
 - `language`
 - `translation`
 - `additional-language`
- openEHR carries `DV_TEXT.language` directly on the value
- Mapping must choose the appropriate FHIR target by context

FHIR NARRATIVE

- `DomainResource.text` — human-interpretable summary
- Mapping *highly* context-dependent:
 - Fully generated from structured content → may be dropped
 - Authored / contains additional content → must be preserved
- openEHR targets considered:
 - Description / story / clinical_synopsis archetypes
 - `INSTRUCTION.narrative` (safety-simplified — use with care)
 - Draft `CLUSTER.fhir_narrative` for general import

CODED DATA

- ⚠ The most structurally divergent area ⚠
- openEHR
 - CODE_PHRASE (terminology_id + code_string + preferred_term)
 - Wrapped by DV_CODED_TEXT, DV_ORDINAL.symbol, DV_SCALE.symbol, DV_STATE.value, null_flavour
- FHIR
 - Coding (system, version, code, display, userSelected)
 - CodeableConcept (text + 0..* coding)
 - code primitive for closed value sets

TERMINOLOGY

- Common external sources
 - SNOMED CT, LOINC, SNOMED CT, UCUM, SNOMED CT, etc.
- Using SNOMED CT with HL7 Standards
 - Formatting system to identify edition/version
- Some CodeSystems embed versions in URI
`http://snomed.info/sct/90...008/version/20240101`

SPECIFICITY & UNIQUENESS

- When adding a system, preserve *conceptual uniqueness*
- If a local code PROC123 has no system, the engine:
 - **SHOULD** add a system if scope is understood and unique
 - **SHOULD NOT** include system otherwise
- The chosen system must be unique for:
 - the **concept** within the facility
 - the **meaning** across facilities

OPEN QUESTIONS

- `system + version ↔ terminology_id` still under discussion
 - `system|version` (FHIR canonical-style)
 - `system(version)` (parenthetical)
 - `system#version` (FHIR package-style)
- `defining_code` semantics in FHIR
 - `coding.userSelected` (working consensus)
 - `coding-purpose` `#original`
 - New `#defining_code` value ([HTA-170](#))



MISSING SYSTEM OR CODE



- When system AND code are both absent
 - Map FHIR display → openEHR DV_TEXT.value (not DV_CODED_TEXT)
- When only one is absent
 - Inherently **unsafe** — resolution requires local informatics advice
 - Options: degrade to DV_TEXT, inject a default system, treat as exception

NULL FLAVOURS

- openEHR NULL_FLAVOUR codes
 - 253 unknown, 271 no information, 272 masked, 273 not applicable
- FHIR targets (preferred → legacy)
 - [data-absent-reason](#) extension
 - `Observation.dataAbsentReason`
 - [iso21090-nullFlavor](#) extension
- ⚠ "no information" (271) has no exact DAR match ⚠
 - Maps to unknown with logged loss; proposal pending

QUANTITIES

- openEHR DV_AMOUNT branch
 - DV_QUANTITY, DV_COUNT, DV_PROPORTION, DV_DURATION
 - Share common props: accuracy, magnitude_status, etc.
- FHIR quantity family
 - Quantity, SimpleQuantity, Age, etc.
 - Plus Ratio, Range, Period, RatioRange, Timing, RelativeTime

DV_AMOUNT PATTERN

Inherited Property	FHIR Equivalent
magnitude_status	Quantity.comparator
accuracy	quantity-accuracy extension
normal_range	Observation.referenceRange (type = normal)
other_reference_ranges	Observation.referenceRange (type ≠ normal)
normal_status	Observation.interpretation

- Reference-range / interpretation are resource-level (typically Observation)

COMPARATOR MAPPING

openEHR magnitude_status	FHIR comparator
<, <=, >, >=	<, <=, >, >=
~	~ (<i>added in R6</i>)
=	(<i>implicit</i>)
(<i>none</i>)	ad (<i>no openEHR equivalent</i>)

- ~ resolved per [FHIR-56000](#), default tolerance $\pm 10\%$
- ad = "sufficient as part of a sum"

PROFILED VARIANTS

- `SimpleQuantity`: no comparator
- `Age`, `Distance`, `Duration`: UCUM-restricted units
- `Count` — `system = http://unitsofmeasure.org`, `code = 1` (32-bit value)
- `Money / MoneyQuantity` — `system = urn:iso:std:iso:4217`, ISO 4217 code
- Mappings may need to do unit conversions or refuse incompatible data
 - e.g., a comparator on a `SimpleQuantity` target = modelling error

UNITS & PRECISION

- FHIR Quantity invariant qty-3: system required when code present
 - When openEHR units_system is absent → add <http://unitsofmeasure.org>
- Precision metadata sources:
 - openEHR precision and magnitude lexical form
 - FHIR decimal lexical form and [quantity-precision](#) extension
- Use the most granular precision available across both
- precision = -1 (unlimited) ≡ absence of quantity-precision



DV_PROPORTION GAPS



- DV_PROPORTION has no unit fields — units come from the archetype
- FHIR Ratio carries units on numerator / denominator
 - e.g., 5 mg / 100 mL
- openEHR typically models unit-bearing ratios as two DV_QUANTITY values
- Fraction rendering hint (pk_integer_fraction) WIP:
 - [FHIR-56001](#)

TEMPORAL DATA

- openEHR DV_TEMPORAL family
 - DV_DATE, DV_TIME, DV_DATE_TIME, DV_DURATION
- FHIR temporal primitives & types
 - date, time, dateTime, instant, Duration, Period, Timing
- Date/time intervals → Period
- Numeric intervals → Range (covered in Quantities)

ISO 8601 SUBSETS

- Both systems use ISO 8601 — but with different allowed forms
- openEHR permits **compact** forms
 - 20250301, T143000
- FHIR requires **extended** form
 - 2025-03-01, 14:30:00
- Compact values must be expanded when mapping to FHIR
- Formal subset comparison is ~90% aligned (undocumented)

PRECISION PRESERVATION

- Partial values must be preserved by truncation, not padding
 - 202604 → 2026-04 
 - 202604 → 2026-04-01 
- Fractional seconds
 - FHIR allows up to **9** decimal digits
 - openEHR restricts to **3**
 - Excess precision is truncated when FHIR → openEHR
- openEHR accuracy / magnitude_status are dropped

TIME ZONES

- FHIR `time` cannot carry a timezone
- FHIR `dateTime` and `instant` can/must (context-dependent)
- openEHR DV_TIME with timezone → FHIR `time` + `timezone` extension
- In practice, openEHR data without timezone is rare

SPLIT FIELDS

- Some FHIR resources split date and time across elements
- Example: `Patient.birthDate`
 - Element typed date
 - Time conveyed via `patient-birthTime` extension
- Mapping engines must be aware of **element-specific** patterns

DURATION CONVERSION

- openEHR DV_DURATION uses ISO 8601 strings
- FHIR Duration uses UCUM code + value
- Conversion needed (per [duration-units](#))
 - $PT\{n\}S \leftrightarrow s$, $PT\{n\}M \leftrightarrow min$, $PT\{n\}H \leftrightarrow h$
 - $P\{n\}D \leftrightarrow d$, $P\{n\}W \leftrightarrow wk$, $P\{n\}M \leftrightarrow mo$, $P\{n\}Y \leftrightarrow a$
-  Calendar units (months, years) are not fixed UCUM durations ([UCUM §31](#))
- Sub-millisecond ISO 8601 values can lose precision

OTHER DATA

- Data types that do not fit cleanly elsewhere
 - DV_MULTIMEDIA ↔ FHIR Attachment
 - DV_IDENTIFIER ↔ FHIR Identifier
- Both areas have **active gaps** under discussion
- Both benefit from **extended openEHR CLUSTER archetypes** for full fidelity

MULTIMEDIA — FIELD HIGHLIGHTS

DV_MULTIMEDIA	FHIR Attachment
data	data
uri	url
media_type	contentType
size	size
alternate_text	title
integrity_check	hash
compression_algorithm	<i>(no equivalent)</i>

DV_MULTIMEDIA

thumbnail

FHIR Attachment

(no equivalent)



KNOWN GAPS



- `compression_algorithm` — no FHIR equivalent
 - Tracked in [FHIR-56003](#)
 - Decompress on conversion to avoid loss
- `thumbnail` — no FHIR equivalent
 - Tracked in [FHIR-56002](#)
- `integrity_check_algorithm` — pending core extension
`alternate-hash`
 - [FHIR-55422](#) — resolved, awaiting publication
- FHIR-only fields (`creation`, `height`, `width`, `frames`, `duration`, `pages`)
 - Map to `CLUSTER.mediafile` extension archetype, or drop

IDENTIFIER — FIELD MAPPING

DV_IDENTIFIER	FHIR Identifier
id	value
issuer	system
assigner	assigner (Reference)
type	type (CodeableConcept)
(none)	use
(none)	period

SYSTEM VS ISSUER

- FHIR system is a **URI**; openEHR issuer is a **free-text string**
- When issuer is **not a valid URI**:
 - Placeholder `http://openehr.org/identifier/{issuer}`
 - Pending TSMG/ITI guidance on URN vs. URI
- When importing FHIR → openEHR:
 - **Skip** the placeholder URLs (they exist only for the reverse direction)
 - If `::` notation is provided, use the embedded system



ROUND-TRIP LOSS



- FHIR use and period have no native DV_IDENTIFIER equivalent
 - Rarely significant in routine clinical use
 - Matters for demographics and temporary IDs
- FHIR Identifier.assigner.display is not preserved by a bare round-trip
- Multiple Identifier.type.coding[] entries cannot all be preserved
- For full fidelity, draft
CLUSTER.extended_demographics_identifier archetype

QUESTIONS & DISCUSSION