

DH ・ 技術要素 ・ DTS

DTS入門

テキストに共通の「窓口」をつくる

DH入門 / 技術要素シリーズ

中村

※実験的な取り組みです（構成・図・AI音声合成を含む）。内容をご確認・ご注意のうえご利用ください

この動画について

- ✓ **クリエイティブ・コモンズ**のオープン教材を参照し、独自に構成した解説です
- ✓ スライド・図は新規作成、ナレーションは**本人声のAI音声合成**
- ✓ これは**実験的な取り組み**です。内容は**ご確認・ご注意のうえ**ご利用ください
- ✓ 誤りに気づいたら概要欄からご指摘ください。出典・ライセンスは末尾と概要欄に記載しています

この回のゴール

テキストのコレクションに、共通の作法で触れる仕組みをつかむ

- ✓ **DTS** が「テキストに共通のAPIで触れる」ための仕様だと説明できる
- ✓ **3つのエンドポイント**（コレクション/ナビゲーション/ドキュメント）の役割を区別できる
- ✓ **引用構造**と、JSON-LD・TEI/XML でデータが返る流れをイメージできる
- ✓ DTSに準拠すると**相互運用**につながることを説明できる

前提知識は特にありません。プログラミングの細部には立ち入らず、考え方を中心に見ていきます。

今日の流れ

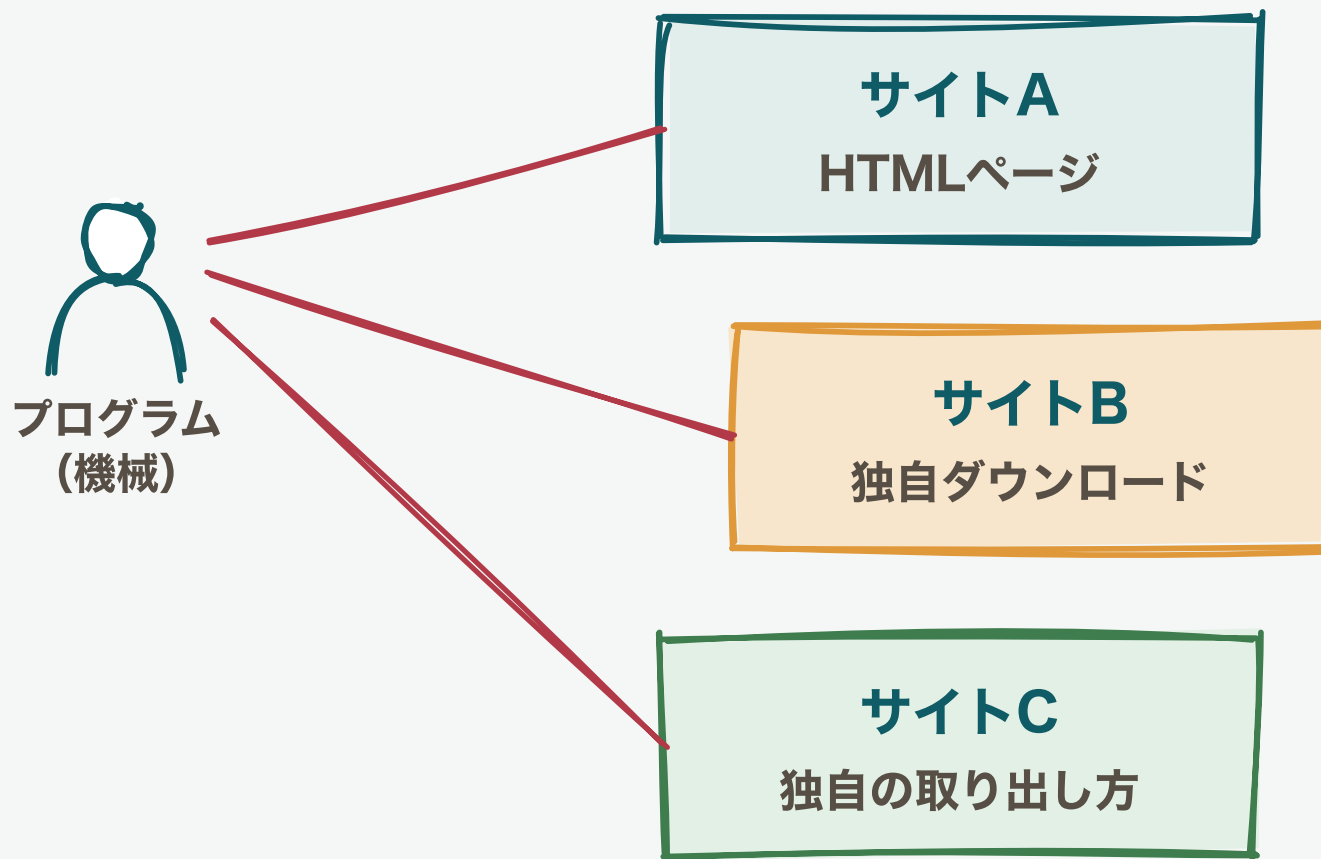
- ✓ なぜテキストに共通の「窓口」が要るのか
- ✓ DTSの仕組み — 3つの窓口（エンドポイント）
- ✓ どんなデータでつながるのか（JSON-LD・TEI/XML）
- ✓ どこを「一単位」として指すか — 解釈の話
- ✓ 始め方・学ぶには

なぜ「窓口」が要るのか

まずは、テキストを「配る」ときの困りごとから

テキストはばらばらに公開されがち

場所も形式も触り方もバラバラ



共通の「窓口」があれば



F・A・I・R (フェア) にする

見つかる

たどり着ける

かみ合う

再利用できる

置き場が違っても**同じ作法**で触れれば、見つけ・たどり・つなぎ・再利用しやすくなる (**FAIR**)

DTS = テキストに触れるための共通API



中身の置き方が違っても、同じ作法で触れる

DTS (分散テキストサービス) は、テキストのコレクションに**標準化された作法**で触れるための仕様。コミュニティが策定し、2026年に 1.0 が公開されました

画像の「あれ」を、テキストにも



画像には **IIIF** という共通の作法が広まりました。DTSは、いわば**テキスト版の同じ発想**です

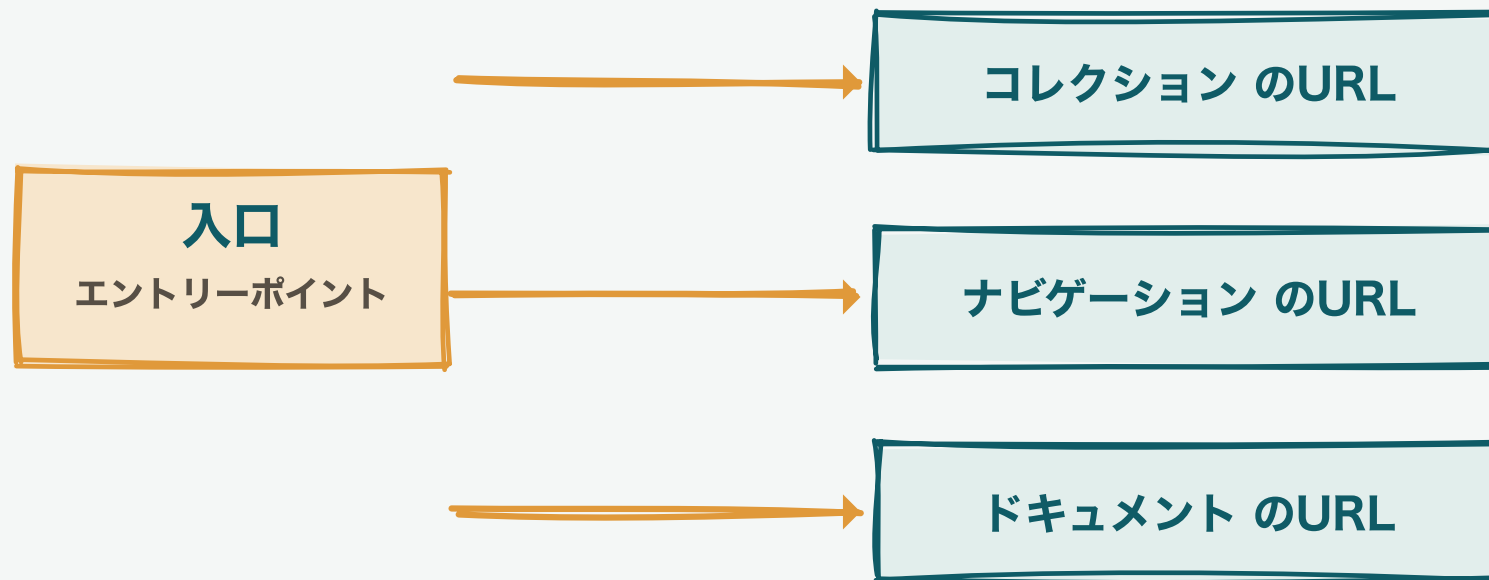
ここまでのポイント

- ✓ テキストは、置き場・形式・取り出し方が**ばらばら**になりがち
- ✓ **共通の窓口（標準API）**があれば、同じ作法で触れる=**FAIR** に近づく
- ✓ それを定めた仕様が **DTS**。画像の IIF と**同じ発想**のテキスト版
では、その「窓口」は具体的にどんな形をしているのでしょうか。 → 3つのエンドポイントへ

DTSの仕組み

3つの窓口（エンドポイント）を見ていきます

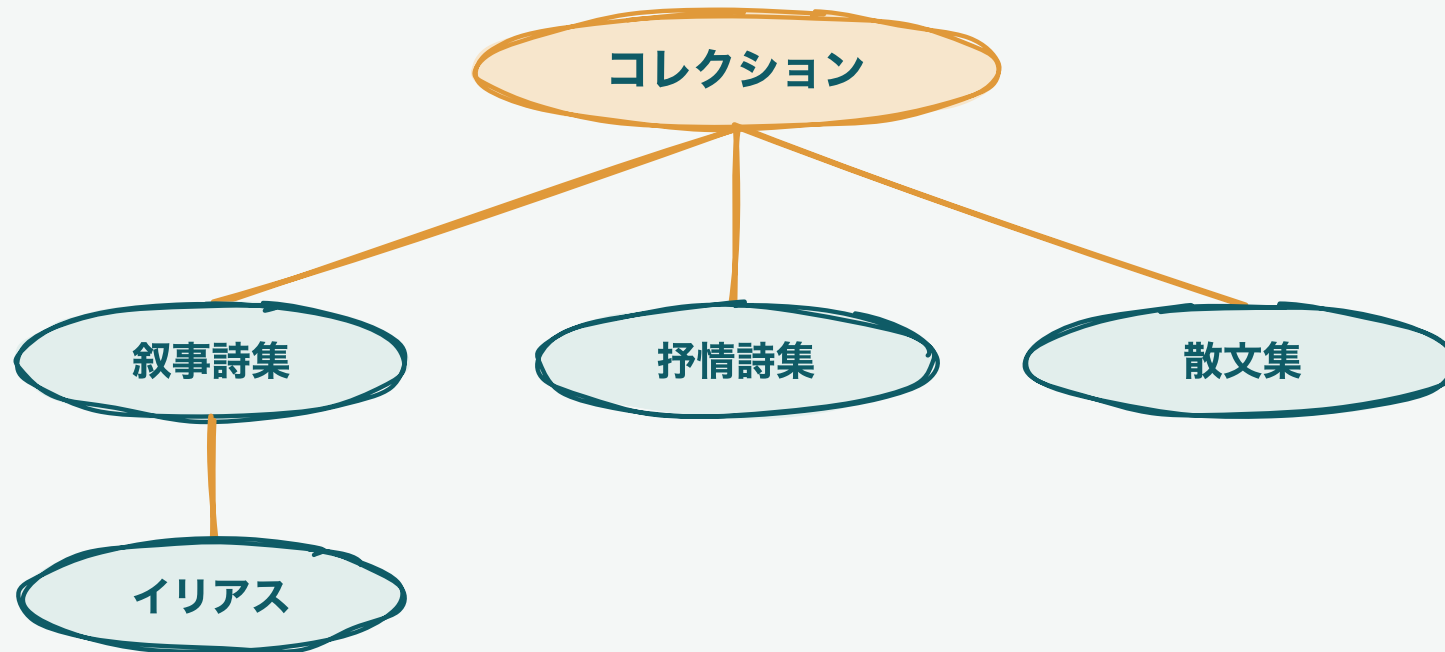
まずは入口（エントリーポイント）



まず入口を叩くと、3つの窓口のURLが返ってくる

入口をひとつ叩けば、**3つの窓口の場所 (URL)** が分かります

① コレクション — テキストの「間」

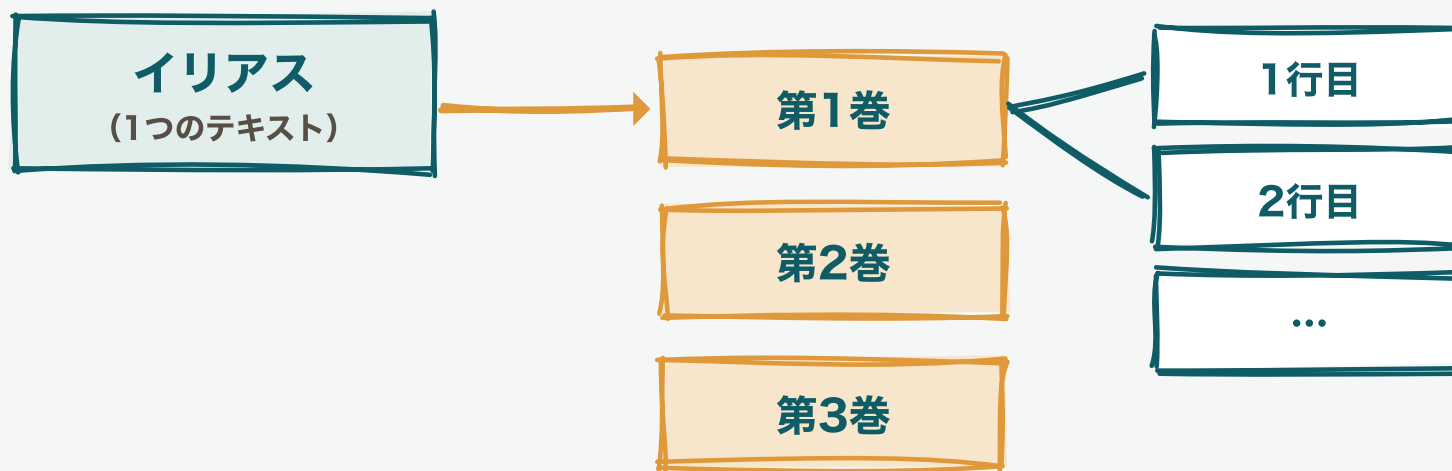


コレクションの中の「どの作品か」をたどる

コレクションの中を渡り歩いて、目的のテキストを見つける窓口です

② ナビゲーション — テキストの「内」

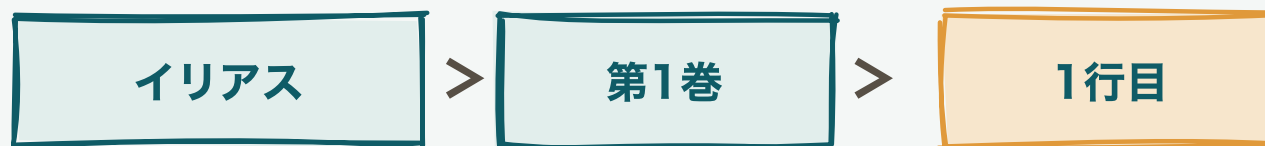
巻 → 行… と、目次のようにたどる



1つのテキストの中の**引用構造 (目次のような階層)** をたどる窓口。どこに何があるかが分かります

引用構造 — 「ここ」を指す住所

ひとつの引用単位を、住所で正確に指す

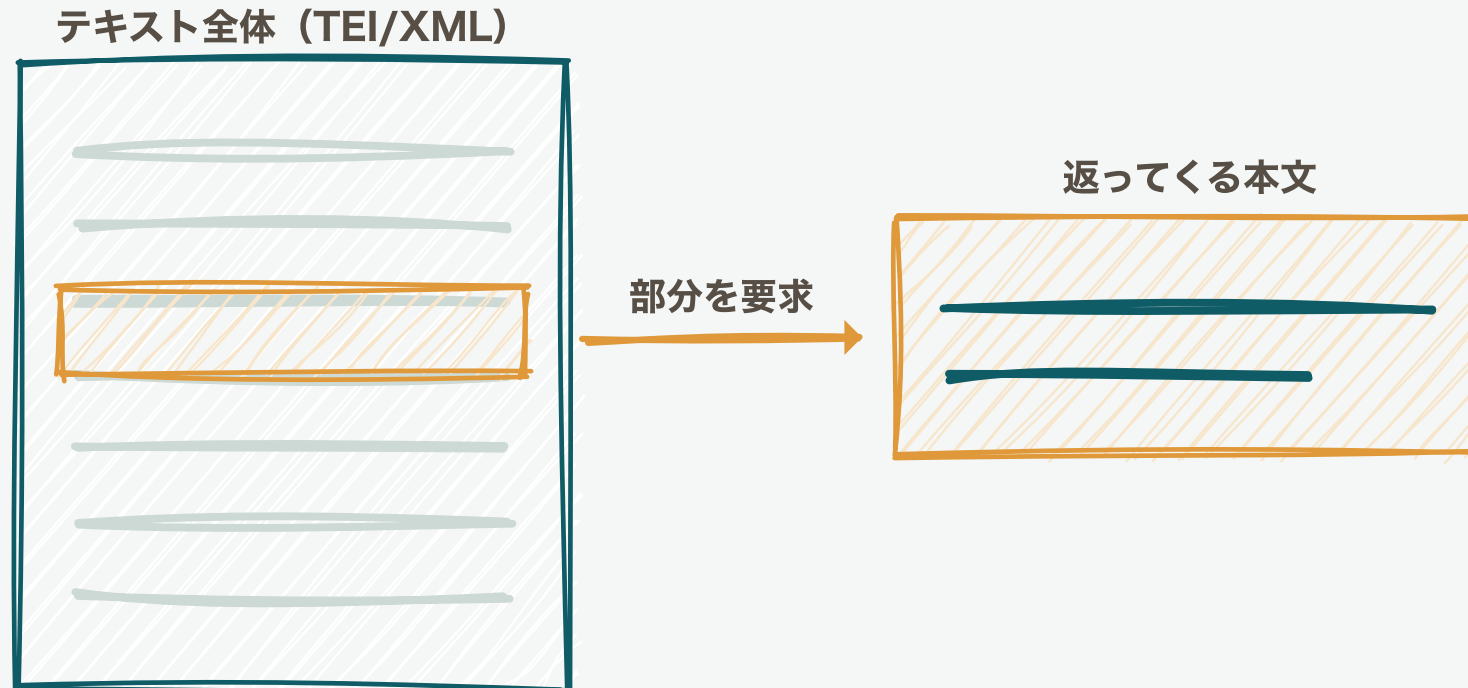


```
ref = イリアス . 第1巻 . 1
```

この住所 (ref) で、本文のどこを呼び出すかが決まる

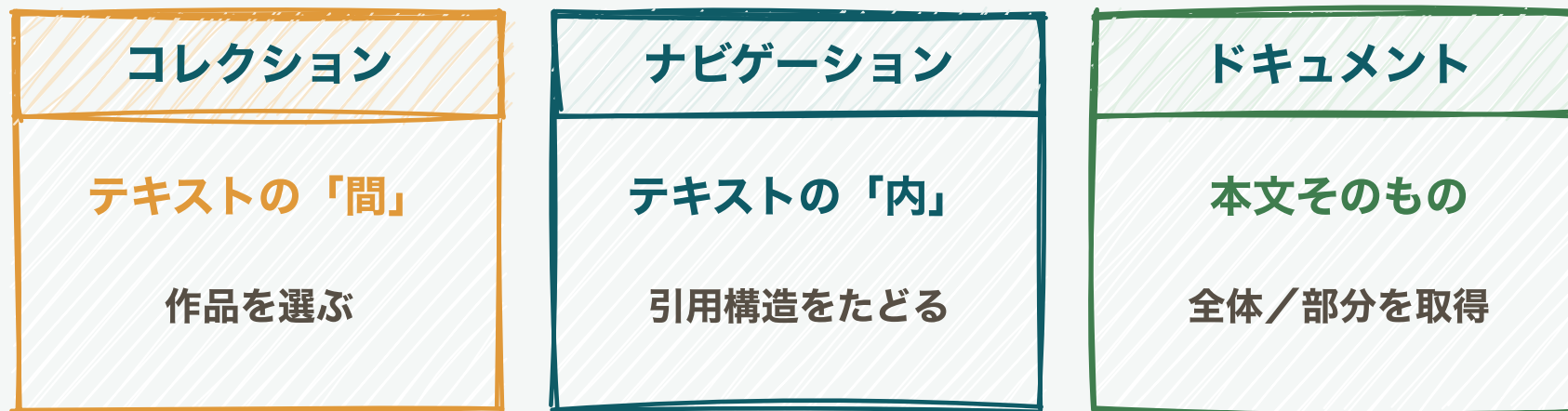
「第1巻の1行目」のように、**引用できる一単位**へ住所 (**ref**) を与えます

③ ドキュメント — 本文そのもの



本文を取り出す窓口。**テキスト全体**でも、住所で指した**一部分 (チャンク)**でも受け取れます

3つの窓口、それぞれの役割

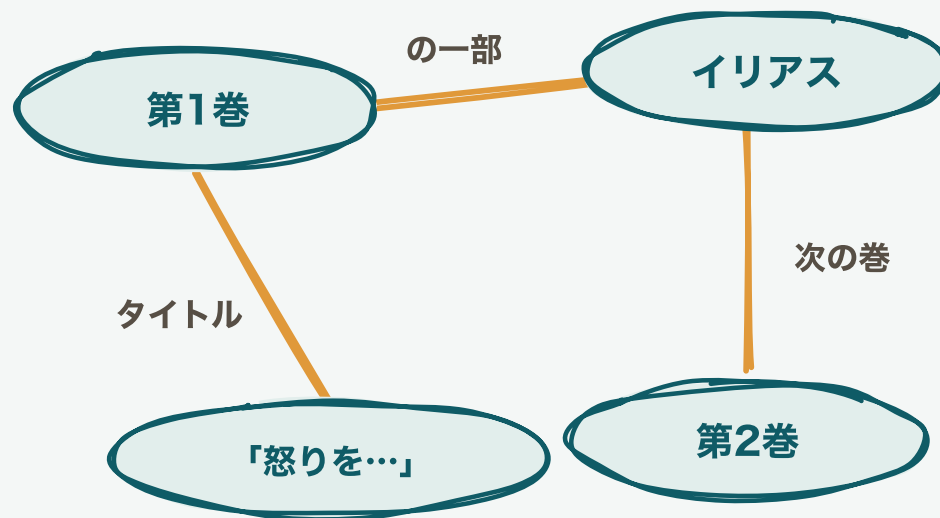


入口で場所を知り、**間**→**内**→**本文**と降りていく。これがDTSの基本の流れです

どんなデータでつながるのか

返ってくるデータの「形」を見ます

一覧や目次は JSON-LD で返る



機械が意味の「つながり」をたどれる形 (リンクトデータ)

コレクションとナビゲーションの応答は **JSON-LD**。前回の**リンクトデータ**と同じく、機械が関係をたどれます

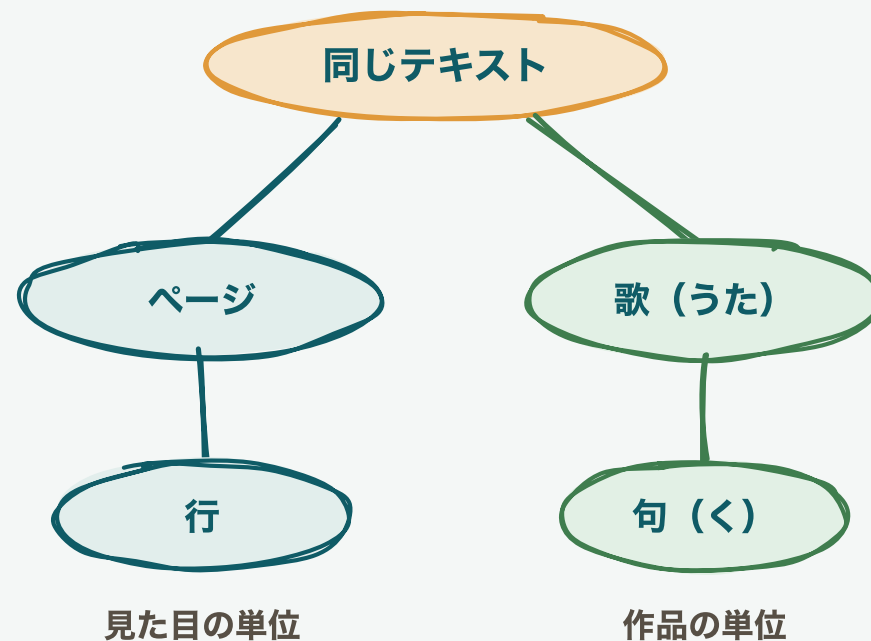
本文は TEI/XML で返る

本文は、人文テキストの標準 TEI/XML で返る

```
<TEI>  
<text><body>  
<l n="1">怒りを歌え、女神よ……</l>  
<l n="2">ペレウスの子アキレウスの……</l>  
</body></text>  
</TEI>
```

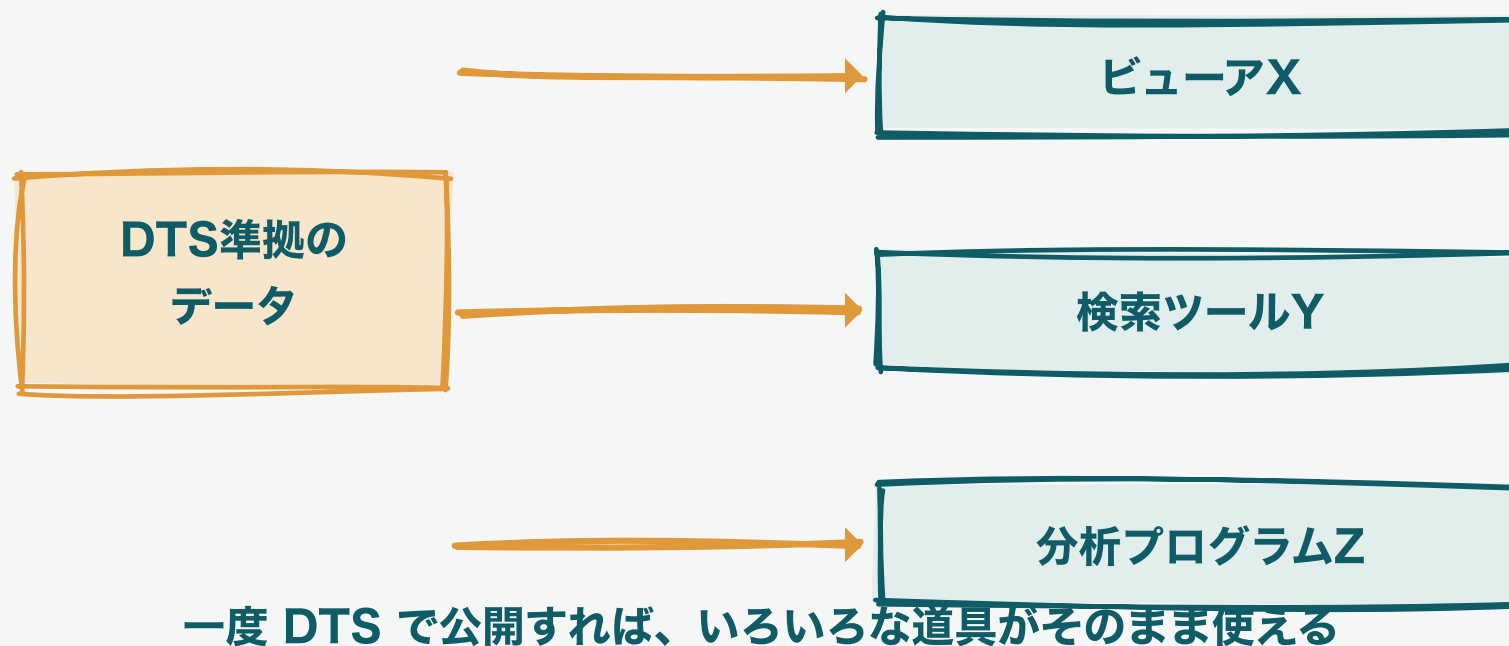
ドキュメントが返す本文は、前回見た**TEI/XML**。構造のタグごと受け取れるので、そのまま研究に
使えます

同じ本文を、別々のたどり方で



1つのテキストに**引用構造を複数**持たせ、「ページと行」でも「歌と句」でも引けます

一度そろえれば、道具を選べる



同じ作法でそろえてあるので、**DTSに対応した道具なら**、作り直さずにそのまま開ける = **相互運用**

ここまでのポイント

- ✓ 一覧・目次（コレクション／ナビゲーション）は **JSON-LD** で返る
- ✓ 本文（ドキュメント）は **TEI/XML** で返る
- ✓ 1つの本文に**引用構造を複数**持たせて、別々のたどり方で引ける
- ✓ 作法をそろえるから、**いろいろな道具で開ける（相互運用）**

仕組みが見えたところで、少し立ち止まって考えてみましょう。

考えてみよう

あなたの扱うテキストを、**コレクション** → **テキスト** → **部分**に分けると？

どこを「引用できる一単位」にしますか。ここで少し、動画を止めて考えてみてください。

「一単位」の決め方は、解釈をともなう

- ✓ どこを**引用できる一単位**とするかは、機械的に決まるとは限りません
- ✓ 章・段落・行・歌… **どの単位で引けるようにするか**に、編集や研究の判断が入ります
- ✓ 同じ本文でも、引用構造の選び方しだいで、**たどれる見え方**が変わります
テキストを「どう構造化するか」という TEI の問いと地続きの、**人文学的な営み**でもあります

始め方・学ぶには

- ✓ まず触れる：公開されている **DTS対応ビューア** で、サンプルのコレクションを開いてみる
- ✓ 確かめる：応答の **JSON-LD** を眺め、コレクション→ナビゲーション→ドキュメントをたどってみる
- ✓ 体系的に：**DTS公式仕様 (dtsapi.org)** と FAQ・Cookbook
- ✓ つくる：まずは静的ファイルからでも始められる **Level 0** から

TEI/XML を用意していれば、それを「配る窓口」を足す、というのが出発点になります

まとめ

- ✓ **DTS**は、テキストのコレクションに**共通の作法で触れる**ための仕様
- ✓ 入口の先に **コレクション・ナビゲーション・ドキュメント**の3つの窓口
- ✓ 一覧・目次は **JSON-LD**、本文は **TEI/XML**。引用構造で「ここ」を指せる
- ✓ そろえれば**相互運用**。ただし「一単位」の決め方には**解釈**が入る

テキストに「タグを付ける」TEI、データを「線でつなぐ」RDF。それらを「配る」のが DTS です

出典・ライセンス

本動画は、以下のオープンライセンス文献を参照して作成しました。

- ✓ Distributed Text Services (DTS): A Community-Built API to Publish and Consume Text Collections as Linked Data / B. Almas, H. Cayless, T. Clérice 他 (Journal of the TEI, 2023) — CC BY 4.0
- ✓ 事実確認：DTS 公式仕様 1.0 (dtsapi.org, 2026)

スライド・図は中村による新規作成（概念を参照し、表現は新たに構成）。

ご清聴ありがとうございました