Case Study L3-001
In-Context Consistency (Document-Wide)
The "Drifting Terminology" Trap (One-to-Many Mapping)

Cédric Stéphany — Technical Translation & AI Alignment Specialist

Case Study Metadata

Dataset ID:	L3-001
Category:	In-Context Consistency — Level 3
Focus:	Terminological Stability / Reference Locking
Model:	Generic NMT
Domain:	Medical Devices (Injector Systems)

1 The Context: NMT Has "No Memory"

Generic Neural Machine Translation (NMT) models process text sentence-by-sentence. When translating Sentence 50, the model has no "memory" of how it translated a key term in Sentence 1.

This leads to **Stochastic Variation** (random synonym selection). While variation is desirable in literature to avoid repetition, it is fatal in patent law.

Key Concept

The Legal Canon of Consistency:
"A word or phrase used consistently in a patent claim is presumed to have the same meaning throughout."
Conversely, **different words are presumed to mean different things**. If a translator refers to Component 1202 as a "*dispositif de commande*" in Claim 1 and a "*dispositif de contrôle*" in Claim 5, the Examiner assumes these are two separate physical devices.

2 The Glitch: The "Three-Headed" Device

In this medical injector patent, the source text consistently refers to a single component: **"control device (1202)."**

However, the generic NMT model, triggered by slight variations in sentence context, produced three different translations for this single term across the document.

- **Instance 1:** *dispositif de commande* (Command device)
- **Instance 2:** *dispositif de contrôle* (Monitoring device)
- **Instance 3:** *dispositif de régulation* (Regulation device)

Critical Issue

The Catastrophic Result (Indefiniteness):
By assigning three different names to Item 1202, the translation artificially creates a complex system with three distinct controllers. This triggers a rejection under **35 U.S.C. § 112(b)** for indefiniteness, as the scope of the claim becomes impossible to determine.

3 The Translation Failure Matrix

Source (English)	AI Hallucination (Failure)	Golden Rewrite (Correct)
Claim 10: "...a **control device** (1202) operatively associated..."	[cite$_s$*tart*]× "...un **dispositif de commande** (1202)..." [cite: 351]	✓ "...un **dispositif de commande** (1202)..."
Claim 20: "...the **control device** (1202) comprising..."	[cite$_s$*tart*]× "...le **dispositif de contrôle** (1202)..." [cite: 351]	✓ "...le **dispositif de commande** (1202)..."
Description: "...said **control device** configured to..."	[cite$_s$*tart*]× "...ledit **dispositif de régulation**..." [cite: 351]	✓ "...ledit **dispositif de commande**..."

Table 1: Terminological Drift: One Term → Three Translations

4 Alignment Methodology

To solve the "Memory Problem," we implement a **Global Term-Locking Protocol**.

Alignment Methodology

The "Freeze-and-Propagate" Workflow:

1. **Entity Extraction (First Pass):** The system scans the entire document to identify high-frequency technical nouns (e.g., "control device").
2. **Anchor Selection:** The first approved translation (e.g., *dispositif de commande*) is designated as the **Master Term**.
3. **Constraint Propagation:** This Master Term is loaded into a temporary "Session Dictionary."
4. **Forced Consistency:** For all subsequent sentences, the model is **forbidden** from generating synonyms for the locked term. It must retrieve the Master Term from the session cache.

This ensures that "control device" is translated identically in line 1, line 50, and line 500, simulating human "short-term memory."

5 Results & Impact

- **Consistency Score:** 100% terminology match across 45 claims.

- **Reference Integrity:** Reference numerals (1202) always align with the exact same text description.
- **Risk Mitigation:** Elimination of "Indefiniteness" rejections caused by synonym drift.