

Case Study L3-002

Semantic Integrity (Consistency)

The "Synonym Storm" (7 Variants for 1 Term)

Cédric Stéphany — Technical Translation & AI Alignment Specialist

Case Study Metadata

Dataset ID: L3-002

Category: Semantic Integrity — Component B3

Focus: Catastrophic Terminology Drift

Model: Generic NMT

Domain: Bio-Pharma Flow Systems

1 The Context: Naming the Core Invention

In this set of patent claims, the subject matter is a "**Flow process system**". To ensure the "Unity of Invention" and clear legal scope, this term must be translated identically every time it appears.

2 The Glitch: The "Goldfish Memory" Effect

The Generic NMT model treated each claim as an isolated event, creating a chaotic list of synonyms. It translated the **same source term** in **7 different ways** across 15 claims.

Critical Failure

The Drift Failure:

- **Claim 1:** "Système de traitement **en** écoulement"
- **Claim 2:** "Système de traitement **par** écoulement"
- **Claim 4:** "Système de traitement **d'**écoulement"
- **Claim 6:** "Système de traitement **de flux**"
- **Claim 7:** "Système de **processus** d'écoulement"
- **Claim 11:** "Système de **processus en** écoulement"
- **Claim 13:** "Système de **processus en** écoulement"

Legal Consequence: In patent litigation, a "Système de traitement" (Treatment System) is legally distinct from a "Système de processus" (Process System). This inconsistency fractures the patent's scope, potentially limiting protection to only the specific variations listed in the claims.

3 Alignment Methodology: Zero-Tolerance Locking

To cure this catastrophic drift, we implemented a **Term-Locking Protocol** (Checklist B3).

Alignment Methodology

The Protocol:

1. **Scan Phase:** The Agent scans the source text for high-frequency noun phrases before translation begins. "Flow process system" is identified as a Core Term ($N > 5$).
2. **Lock Phase:** The system assigns a Canonical Translation → "*Système de procédé d'écoulement*".
3. **Propagation Phase:** This string is forced into Claims 1-15, overriding the model's probabilistic desire for stylistic variety.

4 Key Insights

Key Concept

What This Case Study Demonstrates:

1. **NMT Has No Memory:** Without external constraints, generic models solve every sentence from scratch.
2. **Variety is a Vice:** The model's training to avoid repetition (good for literature) is fatal for patents (where repetition = precision).
3. **Unity of Invention:** Terminology consistency is not just about style; it is the primary mechanism for defining the legal boundaries of the invention.

Portfolio: Patent Translation AI Alignment Framework
Author: Cédric Stéphany
Specialization: Semantic Integrity in Intellectual Property
Last Updated: January 8, 2026
