Case Study C2-002

Structural Compliance in Patent NMT

Morpho-Syntactic Alignment of French Method Claims

Cédric Stéphany — Technical Translation & AI Alignment Specialist

Case Study Metadata

Dataset ID:	C2-002
Category:	Structural Compliance — Constraint 1
Focus:	Verb Nominalization
Model:	Generic NMT
Domain:	Photonics / Optics

1 The Context: Prepositions Define Relationship

In patent claims, prepositions are not filler words; they define the geometric or logical relationship between elements. A critical failure point for generic NMT engines is prioritizing "Fluency" over "Accuracy," often attempting to "smooth" awkward legal phrasing into more conversational language.

Key Concept

The Legal Distinction:

- **"Of" (De):** Implies direct possession, composition, or integral definition. It signals a definitive calculation using the specific items mentioned.
- **"Among" (Parmi):** Implies a selection within a group (subset logic).

Confusing these alters the claim's scope from a mandatory calculation to a selection-based option.

2 The Glitch: Hallucinating "Fluency"

The NMT model analyzes the complex source syntax "*...of the at least...*" and identifies it as clumsy or repetitive. To "fix" the sentence structure for the French reader, it hallucinates a smoother preposition: *parmi* (among).

2.1 Why This Matters

Critical Issue

Litigation Risk: Scope Drift
This is not merely a stylistic error. Using "Among" implies a subset selection that wasn't in the original invention. In litigation, this allows opposing counsel to argue that the claim covers only a **Selection-Based** logic rather than the intended **Definitive Calculation**, potentially invalidating the patent for lack of clarity or narrowing its protective scope.

3 The Alignment Challenge

3.1 The Translation Failure

Source (English)	AI Hallucination (Failure)	Golden Rewrite (Correct)
"...a difference between any two presentation times **of the at least** two unequal..."	× **Scope Drift / Fluency:** "...une différence... **parmi les** au moins deux temps..." (Implies Selection / Subset)	**Literal Compliance:** "...une différence... **des au moins**..." (Mandatory Contraction / Integral)

Table 1: Prepositional Scope Failure

4 Alignment Methodology

4.1 The "Prepositional Lock"

To prevent this "Scope Drift," we implement a **Prepositional Lock** constraint via Relation Extraction in Label Studio.

Alignment Methodology

Annotation Process:

1. **Trigger Identification:** Annotators tag the complex determiner "*of the at least*" as a `COMPLEX_DETERMINER` trigger.
2. **Scope Enforcement:** We enforce a literal translation rule: "*Of + The + At Least*" must map to "*Des au moins*", regardless of phonetic harshness.
3. **Syntactic Unit Locking:** By treating this prepositional chain as a fixed syntactic unit, we prioritize legal accuracy over stylistic smoothness.

This explicitly forces the model to suppress the "Fluency" vector and activate the "Literal" logic required for Antecedent Basis compliance.

5 Results & Impact

5.1 Practical Implications

- **Scope Preservation:** Ensures the French filing maintains the exact logical scope of the English original, preventing "Scope Drift."
- **Defensibility:** The translation remains defensible in litigation by adhering to the strict definition of the source text.
- **Consistency:** Eliminates variability in how complex determiners are handled across large patent portfolios.