Research Journal in Translation, Literature, Linguistics, and Education (RJTLLE)
https://scientific-board.com/rjtlle Volume 2, Issue 3, 2026
https://doi.org/10.64120/7hbtf208

CASE STUDY ARTICLE **Section:** *Translation Studies*

LSP translation challenges: The case of disruptive technologies discourse

Dieudonné Toukam[1]*
[1]Dalhousie University, Canada
https://orcid.org/0009-0007-7929-0512
***Correspondence:** *d.toukam@dal.ca*

Published by **Scientific-Board LLC** in the Research Journal in Translation, Literature, Linguistics, and Education (RJTLLE). **Volume 2, Issue 3, 2026**.

ISSN: *3067-6290 (Online)*
3068-8450 (Print)

Article Details
Received: 25 April 2026
Revised: 01 May 2026
Accepted: 17 June 2026
Published: 07 July 2026

Conflict of Interest: The author/s declared no conflict of interest.

Abstract
Against a backdrop of shortage of studies on the discourse of disruptive technologies (including artificial intelligence, data science, and blockchain), especially from an English-French contrastive perspective, this paper seeks to explore the main characteristics of such a specialized language and examine some of the many translation challenges it hauls. The database for our study is a bilingual comparable corpus made up of original texts by domain specialists, consisting of some five million words, and built following principles of creating corpora for terminology and translation purposes (Gamallo, 2019; Sharoff *et al.*, 2013). Our findings show that the disruptive technologies discourse is characterized by features such as the prevalence of lexical economy and synonymy, as well as the use of metaphor and metonymy in terms formation. Moreover, this study addresses some of the many translation challenges of the domain while suggesting tips for outstanding others.

Keywords: specialized language, disruptive technologies, translation challenges, terminology

How to Cite:
Toukam, D. (2026). LSP translation challenges: The case of disruptive technologies discourse. *Research Journal in Translation, Literature, Linguistics, and Education*, 2(3), 71-92. https://doi.org/10.64120/7hbtf208

1. Introduction

Language for Special Purposes (LSP) refers to the specialized aspects of the language concerned as it delineates a corpus of knowledge that is specific to a domain. Available literature on specialized languages encompasses very few studies of the specialized discourse of disruptive technologies such as artificial intelligence, data science, and blockchain. When it comes to such discourse in English and French, the existence of a few lexicons only appears as a priming work towards an in-depth study of the disruptive-technology terminology and translation challenges in the field. This paper seeks to outline the specialized language of disruptive technologies and address a number of English-French translation challenges raised in the domain. The data analysed come from a voluminous

comparable corpus of specialist texts dealing with disruptive technologies that were built using the key principles of creating corpora designed for contrastive, terminology and translation purposes (Gamallo, 2019; Sharoff *et al.*, 2013; Desagulier, 2017; Zanettin, 2014). Following this introduction, Section 3 gives a quick overview of the specialized discourse of disruptive technologies. Section 3 describes major translation challenges in the field, highlighting those addressed by our terminology project and proposing tips for many others that remain outstanding. Section 4 serves a conclusion.

2. Overview of the Specialized Discourse of Disruptive Technologies

The language for special purposes (LSP) of disruptive technologies is characterized by three main features. The first one is pervasive synonymy, which is primarily nurtured by diatopic and diastratic variations, extensive neologisms, and the need for reducing extensive borrowing (in the case of languages other than English). The second feature is the prevalence of lexical economy, the materialization of which accounts for abbreviated forms of terms, including truncated terms and shortened complex words. The last feature of disruptive technology discourse is the use of metonymy and metaphor for term formation.

2.1 Pervasive Synonymy

As part of a project aimed to building a bilingual glossary of terms in this specialized domain, we examined the impact synonymy in both English and French. We established a sample of 200 English terms that guided us in our analysis and provided interesting statistics. Findings of our analysis are as follows:

— Out of the 200 English-French terminology records selected, 62 (that is, 31%) have synonyms in both languages. When considering English entries alone, 93 out of 200 have synonyms. While such statistics show how terminologically rich the disruptive technology discourse may be, it also illustrates a lack of absolute unanimity in the use of terms among the community of experts worldwide. Despite the many scientific meetings on these technologies, researchers do not feel obliged to select the terms they use from a single basket. For example, for the same concept, specialists will show disparity in or preference for using (BLOCKCHAIN) CONSENSUS PROTOCOL, CONSENSUS ALGORITHM and VALIDATION MECHANISM.

— In the French language, synonymy is more prevalent in the LSP of transformative technologies: each of the 62 sampled terms in French displays a rate of 1.3 synonym, which means a total of 82 synonyms being shared by 62 records. This unprecedented proportion of synonymy represents a major constraint to specialized communication's clarity and fluidity and may orchestrate ambiguity in transferring technical knowledge. It is worth recalling that technical discourse burdened by ambiguity is a significant impediment to rationalized professional communication and adequate dissemination of specialized knowledge (Wüster, 1985: 89). In French, synonymy is intensified by English borrowings coexisting with neologisms, and the lack of terminology standardization efforts. Synonymy, which notably derives from divergence in naming concepts across the global expert community, is often cultivated by diatopic and diastratic disparities, particularly between Canada and Europe.

2.2 Prevalence of Lexical Economy

In both intralingual or interlingual communication processes, economy is defined as the act of transmitting information using sometimes extreme means of concision, but efficiently and without entropy. "Economy is not simply concerned with strategies for concise transmission, but with all aspects of the reduction of effort in the transmission of information" (Sager, 1990: 105). Following Kocourek (1982: 317–319), Sager also points out that in terminology, lexical economy illustrates itself[1] in various ways and leverages

1 Of course, we take it for granted that symbols, tables and graphs are also privileged means of economy shared by virtually all techno-scientific discourse.

several means, ranging from abbreviation and contraction of terms and truncation to the dropping out of logical links such as prepositions in multi-word terms. As evidenced by our corpus, the discourse of disruptive technologies largely relies on linguistic economy and mostly on lexical economy.

Below are some terms of the domain whose formation leverages lexical economy through a mash-up and blend of two or more words to create a new one:

Term	Economy-based technique used	Note
datapulation	Conservation of 1st word and aphaeresis of the 2nd	data+(mani)pulation
dataveillance		data+(sur)veillance
deepfake$_{(N)}$	Apocope of the 1st word and full retention of the 2nd.	Deep (learning)+fake
info-obesity		info(rmation)+obesity
mempool		me(mory)+pool
Metaverse$_{(N)}$	Apocope of the 1st word and aphaeresis of the 2nd	meta+(uni)verse
sextortion		sex+(ex)torsion
to voxelize		vo(lume)+(pi)xelize ["a voxel is a unit of graphical information that defines a point in three-dimensional space" (Basdevant et al. 2022)

Table 1. Terms formed with lexical economy.

Lexical economy is also expressed through ellipsis of portions, which sometimes leads to terms or expressions that are rarely accepted in general language. Some of these terms are

- TO LIQUID STAKE [instead of 'to stake for liquidity']
- TO HARD-FORK [instead 'to do create a hard fork'], and
- TURING-COMPLETENESS [instead of 'completeness as for Turing principles'].

The last materialized forms of lexical economy in the specialized language of disruptive technologies are abbreviations that are systematically used in exchange for the long forms of corresponding terms. Here are some examples:

BERT (for 'Bidirectional Encoder Representations from Transformers')
ECDSA (for 'Elliptic Curve Digital Signature Algorithm')
GAN (for 'Generative Adversarial Network')
GPT (for 'pre-trained transformer)
KNN (for 'k Nearest Neighbours' algorithm)
LSTM (for 'Long Short-Term Memory' network)
mempool (for 'Memory Pool') – lexicalized acronym
MEV (for 'maximum extractable value')
nonce (for 'Number Only Used Once') – lexicalized acronym
PBS (for 'proposer-builder separation')
PoW (for 'Proof-of-Work')
Relu (for 'Rectified Linear Unit) – lexicalized acronym

2.3 Metonymy and Metaphor in Term Formation

In the specialist discourse of transformative technologies, many terms are formed on a metonymic or metaphorical pattern in English, and in most cases, such pattern is replicated in other languages. While metonymy and metaphor may lead to difficulties in figuring out terms meaning in some countries (Schmitz, 2007: 52), it is not the case with the specialized language of disruptive technologies. Furthermore, we can argue that contrary to Edem Antia (2007: xvi), who disclosed that the 21st century is witnessing a "back-off from the extremes of nominalization and […] metaphor" in "the language of science (and technology)," the discourse of AI, blockchain and data science continues to massively build on metaphorical and metonymic features. Here are some of the multiple terms formed through leveraging such rhetorical images.

Table 2 and Table 3 describe a sample of terms formed by metonymy and by metaphor, respectively.

English term	Metonymy
Bourgain adversarial generative network	Contiguity between the deep learning algorithm/ method and its creator.
Fléchet inception distance	
Wasserstein adversarial generative network	
Merkle tree	Contiguity between the crypto chart and the place where this type of stock market representation was first observed.
Kolmogorov-Arnold network	Contiguity between the algorithm and the creators of its main foundational theory (K-A representations).
satoshi	Contiguity between the unit of measurement and the inventor of the architecture using this indicator (Satoshi Nakamoto).

Table 2. Metonymic formation of terms.

English term	Metaphor
(artificial) neural network	Resemblance in function: a machine learning algorithm operates like a biological neural network.
candle; candlestick	Similarity in both shape and function: a candle is a crypto chart in the shape of a candle, and it informs (illuminates like a traditional candle) about the price of a crypto asset.
cloud	Resemblance between the virtual environment and the cloud (elusive, with no apparent limits, distant, etc.).
data lake	The similarity in volume between a lake and a gigantic reservoir of data
gas	Functional similarity: gas (here) is the 'fuel' for Ethereum's computing power, which is supplied to it by the network user, whereas gas in ordinary language is a source of energy needed to operate a device or machine.
hammer candlestick pattern	The shape of the crypto chart in question resembles that of a hammer.
smart object	Anthropomorphizing non-human entities: devices are rendered "intelligent", making them look like humans. As a result, there is an ontological shift of paradigm (an artificial intelligence-endowed entity operates like a naturally intelligent entity).
smart contract	Function resemblance: this is a self-executing algorithm that irrevocably seals a transaction, just like a signed contract.
snowification	The image transformation process adds and removes the blur, which looks like snow.
whale; crypto whale	Resemblance in form: the holder of huge quantities of crypto assets can move the share price on the market, just like a whale in the sea (it is huge in volume and moves the tide).
wormhole attack	Resemblance in form: the attack takes place through small network cracks similar to wormholes in a piece of fruit.

Table 3. Metaphorical formation of terms.

Such findings corroborate many other studies (Kocourek, 1982; Sager, 1990; Temmerman, 2002; Diki-Kadiri, 2022, *inter alia*), for which metonymy (relation of contiguity) and metaphor (relation of resemblance…) are natural tools for creating both general language and specialized language items, no matter the speciality. On the one hand, neo-lexicalizations, more common in science and technology, reflects metaphorical (and metonymic) reasoning. It "amounts to the understanding of a new fact, situation, process, or whatever type of category based on the imagined *analogy* between what one is trying to come to grips with, to understand, and something one knows and understands already." (Temmerman, 2000: 70). On the other hand, "The use of metonymy and metaphor [both of which are referred to as metasemia] as a naming process in advanced or highly specialized disciplines has been commonly observed (Diki-Kidiri, 2022: 202)." The author further discloses that "Metasemia consists in comparing the unknown entity with something that is already known, in order to better understand, explain, and know about it." (*Ibid*.).

3. Major Translation Challenges in the Area of Disruptive Technologies

Translation challenges in the field of disruptive technologies, to our opinion, encompass all difficulties that arise from the use of the related terminology, phraseology, and other syntactic features such as metaphorical and metonymic expressions, syntactic and lexical economy, and neologisms. While some translation challenges may be shared by several specialized domains (e.g. VIRTUALIZATION, in computer science; POLYNOMIAL ALGORITHM, in mathematics; and DECENTRALIZED EXCHANGE, in finance), many others may concern only terms known as belonging to the specific field of disruptive technologies (e.g. CONVOLUTIONAL NEURAL NETWORK).

3.1 Cross-disciplinary Translation Challenges Found in the Disruptive-Technology Discourse

In their fast-evolving track, emerging technologies, including those with a particularly disruptive impact, have been bringing in their wake a few concepts whose French designations, for those that exist, have yet to be validated through reliable technical writing or translations. To translate most of such terms into French, we have noticed that 'calque' and borrowing account for common and efficient translation approaches. While 'calque' illustrates the transparency of terms in technology and the proven propensity of specialists in these fields to create terms based on the English model, the *lingua franca* of techno-scientific production, borrowing basically shows translators' and terminologists' fear of denaturing English meaning, seeing newly created terms eluded in use, among others. Cross-disciplinary terms feature among such items whose translation in the field of disruptive technologies keep posing problems. We have compiled a non-exhaustive list of such English terms used in emerging technologies whose French translation is either non-existent or remains questionable.

— ENIGMATIC SUPERCOMPUTER (Haenlein & Kaplan, 2019).

The French literal translation, *SUPERORDINATEUR ÉNIGMATIQUE*, is valid, but even with the help of specialists, language specialists face challenges in understanding and properly defining the term.

— Translation of three theories used by technoscientific fields:

1. THEORY OF PLANNED BEHAVIOUR [TPB] (fr. THÉORIE DE LA PRÉVISION DES COMPORTEMENTS; THÉORIE PRÉVISIONNELLE DES COMPORTEMENTS; THÉORIE DE LA PRÉVISION COMPORTEMENTALE; THÉORIE DE COMPORTEMENTS PLANIFIÉS*)
2. TECHNOLOGY ACCEPTANCE MODEL [TAM] (fr. MODÈLE D'ACCEPTATION DE TECHNOLOGIE; MODÈLE D'ACCEPTATION TECHNOLOGIQUE…)
3. TRUST THEORY (fr. THÉORIE DE LA CONFIANCE; THÉORIE DE LA FIABILITÉ (TECHNOLOGIQUE); THÉORIE RELATIVE À LA CONFIANCE*; etc.)

As can be seen, none of the above terms has a standardized French translation, either through frequent use or by an approved body. The above translations come from a few sources or are our own, but a large number of texts found opt for the full English borrowings, especially the abbreviations. It should be that the abbreviations TPB and TAM are adopted almost unanimously in French. Based on our own experience, we have discarded some of the translations of theories proposed above. Those deemed inappropriate, in terms of accuracy in renderings, bear an asterisk.

— Translation of other terms designating various cross-disciplinary concepts:

Surprisingly, the following terms remain without standardized French equivalents, despite the age of some of the concepts – e.g. the most cited definitions of the terms TECHNOLOGY PERCEIVED BEHAVIOURAL CONTROL and TECHNOLOGY PERCEIVED USEFULNESS date back to Ajzen (1991). In the absence of a commonly agreed translation of these terms, we propose one or more translations for each of them, according to gathered data, after setting aside any renderings deemed inaccurate. Preference in the use of these accepted equivalents will therefore depend on each user.

anglais	**français**
average variance extracted (AVE)	*variance moyenne extraite ; extraction de la variance moyenne ; variance moyenne retenue*
composite reliability (CR)	*fiabilité composite*
Cronbach's alpha	*alpha de Cronbach ; coefficient alpha de Cronbach*
discriminant validity	*validité discriminante*
outer loading	*charge extérieure*
technology adoption and use	*adoption et utilisation de [une] technologie ; adoption et usage de [une] technologie*
technology adopter	*adopteur de technologie ; utilisateur primaire de [une] technologie*
technology adoption re-intention behaviour	*comportement relatif à l'intention d'adopter une/la technologie ; attitude-intention d'adoption de [la/une] technologie*
technology enabler	*catalyseur de [une] technologie ; facilitateur de [une] technologie*
technology mainstream user	*utilisateur primaire de [la/une] technologie*
technology perceived anthropomorphism (PATM)	*anthropomorphisme perçu de [la/une] technologie*
technology perceived behavioural control (PBC)	*maîtrise comportementale perçue face à de [une] technologie ; perception de la maîtrise comportementale à l'égard de [la/une] technologie*
technology perceived ease of use (PEU)	*facilité d'utilisation perçue de [la/une] technologie ; perception de la facilité à utiliser la/une technologie*
technology's perceived intelligence	*intelligence perçue de [la/une] technologie*
technology perceived usefulness (PU)	*utilité perçue de [une] technologie*
technology's utilitarian value	*valeur utilitaire de [la/une] technologie*
technology use and re-use intention	*intention d'utiliser et de réutiliser une technologie*
technology use continuance	*utilisation continue de [la/une] technologie ; usage permanent de [la/une] technologie ; utilisation ininterrompue de [une] technologie*

Table 4. Translation of a number of cross-disciplinary terms used in technology areas.

3.2 Translation Challenges resulting from the Specific Terminology of Disruptive Technologies

Apart from the case in point in section 1 above, we examined translation issues raised by the contextual use of disruptive-technology terminology. Some of the most recurrent described below.

3.2.1 Translation problems stemming from the multiplicity of network/blockchain attacks

Such translation challenges are underpinned not only by the large number of cyber-attacks targeting intelligent networks and their connected objects – which may lead to confusions –, but also by the technicality of related terms. In our terminology project, we attempted to list a good number of them, distinguishing between attacks affecting typical networks (AI-based networks, neural networks, IoT, etc.) and blockchain attacks. The more networks are multiplied and intertwined, the more complex the task of understanding and translating terms appears to be. For example, there are multiple denial-of-service attacks (known as DoS attacks), hence an increased complexity of their designations:

flooding (attack)	*attaque par saturation*
jamming (attack)	*attaque par brouillage*
ping of death (attack)	*attaque par « pings mortels »*
smurf (attack)	*attaque par couche de réseau*
teardrop (attack)	*attaque par injection de données fragmentées*
distributed DoS attack; DDoS attack	*attaque par déni de service distribuée (= attaque DoS distribuée)*

Subdivision of DoS attacks can be scaled, as demonstrated by the types of jamming attacks, often with unusual names, including the following:

- constant flooding [attack] (fr. attaque par brouillage constant; attaque par brouillage ininterrompu)
- deceptive flooding [attack] (fr. attaque par brouillage à intervalles fixes)
- granular flooding [attack] (fr. *attaque par brouillage granulaire*).

There are other types of network attacks for which the French translation is far from easily coping with, as noticed in our comparable corpus. These include SIDE-CHANNEL ATTACK (fr. *ATTAQUE PAR CANAUX AUXILIAIRES*); SINKHOLE ATTACK (fr. *ATTAQUE PAR CONTRÔLE DE NŒUDS, ATTAQUE TROU D'ABÎME*); ZERO-DAY ATTACK (fr. *ATTAQUE PAR FAILLE LOGICIELLE*). This last translation has been criticized by several specialists, who are more inclined to the English borrowing on grounds that the expression "*par faille logicielle*" would be simplistic. Moreover, there are several types of attacks that build on software deficiency across a network.

Zero-day attacks include the following:

software supply-chain attack	*attaque de la chaîne d'approvisionnement logicielle*
backdoor attack	*attaque par portes dérobées*
Byzantine attack	*attaque byzantine*
malware attack	*attaque par logiciel malveillant*

In addition, network attacks target a variety of networks and can involve a wide range of methods. All this adds to the difficulty of translating the terms concerned into French, since understanding the

concepts is a prerequisite for a good translation – let's remind that translation problems arise both upstream (where French-speaking specialists forge French equivalents of English terms) and downstream (where language professionals do in terminology and translation activities). Despite their increasing collaboration, language professionals and domain experts tend to refrain from mechanically using the terminology disseminated by domain specialists in their work. Once again, typical systems affected by such attacks range from distributed/federated systems and platforms to intelligent networks (IoT), neural networks, and databases/data centres.

With regard to attacks on blockchain-type networks, they are also numerous and bring about designations that are difficult to translate. Even if the recurring trend is to use literal or calque translations, we should regularly question the existence and level of acceptability and implementation (Sager, 1990; Delavigne & Gaudin, 1996) of most terms, which in the field of disruptive technologies, are generally as young as the concepts they designate.

Here are a few examples of difficult translations, which require a perfect understanding of the underlying concepts:

1. FRONTRUNNING ATTACK (fr. ATTAQUE PAR ANTÉTRANSACTION)

While this translation is inspired by the term "*antétransaction*" – which is mainly used in the field of finance –, then the translator must have been lucky enough to be familiar with the concept and, above all, the term, as it is recent (in 2017, according to the French Government's Official Bulletin). In addition, it is important to place it in the context of the blockchain. This allows us to accept "*ATTAQUE PAR NŒUD EMBUSQUÉ*" as another translation of FRONTRUNNING ATTACK. In this subdomain, *frontrunning* and *backrunning* go hand in hand, even if the former is a genuine and feared attack. In case of doubt, either term should be translated by the generic "*(attaque par) manipulation de l'ordre des transactions*."

2. GAS GOLFING ATTACK; GAS GOLFING (fr. *ATTAQUE PAR SIPHONNAGE*)

This term is difficult to translate if you want to Frenchify its equivalent in the other language, as full borrowing is extremely frequent in French communications. Moreover, the number of occurrences of the translation "*attaque par siphonnage*," which is semantically very good but scarcely used in the relevant literature (only a couple occurrences in our corpus and a few examples on the Internet).

3. REENTRANCY ATTACK (fr. *ATTAQUE DE RÉENTRANCE*)

Though the word 'reentrancy' is well known in computing, the term REENTRANCY ATTACK is thought to have originated with the advent of blockchain technology, and particularly Ethereum and smart contracts. The French translation seems literal, but it may hide a pitfall in understanding.

4. MAXIMAL EXTRACTABLE VALUE ATTACK; MEV ATTACK; MINER EXTRACTABLE VALUE ATTACK (fr. *ATTAQUE MEV*)

Translating the English term incorporating an abbreviated form is a challenge for language professionals. Is there a known translation of "maximal extractable value"? The answer is yes, because the multilexemic unit "*valeur maximale extractible*" does exist and refers to a 'maximum profit that the miner/validator makes from producing blocks or manipulating the order of pending transactions' (according to Ethereum, 2023 and Larsson, 2023). However, nowhere have we seen a longer form of "*attaque MEV*" (e.g. "*attaque basée sur/visant/en vue de la valeur maximale extractible*"). The popularity of the shorter form "*attaque MEV*" can be explained in two ways: (i) domain specialists do not hesitate to take advantage of lexical economy, thus getting rid of both the connectors *'basé sur'* or *'visant'* or *'en vue de'* and the long phrase *'la valeur maximale extractible'*; and (ii) the fact that the longer form of the term refers to the purpose of the attack, rather than to the type of attacker or the process used, may have somehow motivated the

specialist who coined the French equivalent. In fact, even the English, there are far more occurrences of 'MEV attack' than its longer form. Another interesting translation fact concerns the last synonym of the English term (*miner extractable value attack*). It is gradually falling into disuse because the merger area in the blockchain world, concretized by transition from the PoW [proof of work] consensus mechanism to the PoS [proof of stake] resulted in the progressive disappearance of mining processes, and thus that of the "miner" concept. Therefore, the appropriate term now is 'maximal extractable value attack.'

5. SANDWICH ATTACK (fr. ATTAQUE EN SANDWICH; ATTAQUE SANDWICH)
Should we say "*en sandwich*" or "*sandwich*"? We have encountered both types of spelling, and it seems adequate to use any of both variants.

6. SYBIL ATTACK; ATTACK BY THE 51%; ATTACK BY MAJORITY (fr. *ATTAQUE SYBIL; ATTAQUE DES 51 %*)
Note that "Sybil" as a proper noun remains capitalized, and that the figure "51%" is written as a number instead of being fully spelled out.

7. TIME-BANDIT ATTACK (fr. ATTAQUE DE BANDIT TEMPOREL)
Some subject specialists argue that the phrase "time-bandit" doesn't technically say much, which can bother a translator eager to use only terms accepted by experts.
For their linguistic competence, the translator or interpreter should bear in mind that the above-mentioned lexical units whose central component is "attack" (fr. '*attaque*') have a single government in English [attack on/against Y by X] but two in French [*attaque de X contre Y,* and *attaque de Y par X*]. For instance,

English:	French :
This is an adversarial attack against/on the training model [***against<on>* ART N**] by infiltrators [***by* ART N**]	1. Il s'agit d'une attaque antagoniste des agents infiltrés [***de* ART N**] contre nos modèles d'entraînement [***contre* ART N**]. 2. Il s'agit d'une attaque antagoniste de nos modèles d'entraînement [***de* ART N**] par des agents infiltrés [***contre* ART N**].

3.2.2 Translation problems due to abundant synonymy
According to data from our corpus, a great deal of terms has one or more synonyms in the discourse of transformative technologies. Out of 200 sampled terms in English, it appeared that 62 had synonyms, accounting for a proportion of 31%. This prevalence of synonymy obviously leads to a number of translation difficulties. In French, the rate is even higher, with many coined terms coexisting with English borrowings, coupled with disparities in term acceptability among both domain specialists and terminologists/translators. Here are a few examples to illustrate the point:

English	French	Comments
BIG DATA $_1$	*MÉGADONNÉES; DONNÉES MASSIVES; DONNÉES GIGANTESQUES; DONNÉES DÉLUGE; BIG DATA* $_1$	While the French translation of the term offers several alternatives, which can make it easier to find equivalents, the multiplicity of synonyms can also lead the language professional astray if he or she is unaware of all these semantic equivalents. This is even truer when it comes to French-English translation: the translator needs to know that '*mégadonnées*' and the other three terms above refer to one and the same concept, and that they are interchangeable, including with several collective nouns for 'big data$_1$' such as 'mountains of data' (*montagne de données*), 'cart of data' (*tombereau de données*), 'avalanche of data' (*avalanche de données*), 'tsunami of data' (*tsunami de données*), 'crowd of data' (*foule de données*), 'mass of data' (*masse de données*), etc. The translator's practical knowledge is therefore called into play, especially when these synonyms are scattered throughout a text or when the translator wishes to diversify his/her terminological usage. As we explain below (subsection 3.2.3), polysemy is another major source of translation difficulty. The polysemic nature of the term big data creates an additional embarrassment – note that big data [that is, BIG DATA$_2$] also refers to the technology aimed to collect, process and store massive data, while the last semanteme [BIG DATA$_3$] refers to the corresponding (sub) domain.
EDGE ARTIFICIAL INTELLIGENCE; ARTIFICIAL INTELLIGENCE OF THINGS	*INTELLIGENCE ARTIFICIELLE EN PÉRIPHÉRIE DE RÉSEAU; INTELLIGENCE ARTIFICIELLE EMBARQUÉE; INTELLIGENCE ARTIFICIELLE PROXIMALE*	For language professionals, the example in (b) also highlights the challenge of recognizing, from the outset, the synonymous relationship between 'edge artificial intelligence' and 'artificial intelligence of things (AIoT)'. The two synonymous terms can appear in the same text to be translated and thus lead to confusion. Better still, if the translation is done in English, the large number of synonyms for "*intelligence artificielle en périphérie*" (up to four) is also challenging. Furthermore, the presence in the French text of a synonym like "*intelligence artificielle proximale*", which is not yet widely used, would add to the difficulty of understanding and translating this terminological unit.

HODLING	*HODLAGE; DÉTENTION PRUDENTE DE CRYPTOACTIFS*	This term belongs to the subdomain of blockchain, and its French translation (apart from borrowing) is a neologism that has not yet taken root. The major difficulty in translating the English term lies not only in its novelty, but also in its non-transparency, impeding easy inference of the term's meaning. And even if a translator is familiar with the French equivalent "*hodlage*", the need for diversifying one's style may require them they often use the synonym "*détention prudente de cryptoactifs*", which is semantically more accessible to non-specialists due to its propositionality.

Table 5. Examples of translation problems due to synonymy.

3.2.3 Translation problems stemming from polysemy

As technical translation, interpretation and writing are largely dependent on understanding and delimiting the semantic scope of concepts and the semantemes of corresponding terms, polysemy is obviously a source of translation challenges. Admittedly, the context can help to determine the meaning of a specialized lexical unit – particularly through its collocates, even though the real polysemous terms are those that survive the discursive context (Kleiber (2008: 90). However, our approach to polysemy strikes a balance between the proponents of polysemy as a lexical relationship that is intrinsic to a lexical unit and those who believe that the multiplicity of meanings for a term is also dependent on the linguistic community concerned and the discourse used. Cases of polysemy as a paradigmatic relationship inherent in a term therefore deserve to be recorded in terminology resources, in view of the derived problems of understanding and translating.

In the disruptive-technology discourse, we note that the main translation challenges attributed to polysemy arise when the copolysemes of a term in the source language (English in our case) are graphically different from their equivalents in the target language. French terms that cannot be easily inferred from their English equivalents, owing to their morphological similarities, imply that literal or calque translation would be risky. A few examples from the domain under study can illustrate our say.

English copolyseme in context	French translation	Difficulty/Remark
hash$_1$ "Each block has its own hash, with a unique ID, and it also brings the **hash** of the previous block and generates more transaction security." [Queiroz & Fosso-Wamba, 2019]	empreinte cryptographique (= hache$_{(N)}$; hash; haché$_{(N)}$)	In contrastive terms, the polysemy is only present in the source language, since the copolysemes are translated by two different terms in French: the action noun, "*hachage*", is different from the result noun ("*empreinte cryptographique*"), even if rare occurrences grant a second meaning to "*hachage*" as a synonym of "*empreinte cryptographique*".
hash$_2$ (= hashing; hash coding) "For secure user authentication of Data Owners and Data Users, we proposed a SHA3 hashing algorithm, which performs a **hash** of user information and store it in the Trust Center. [Narayanan *et al.*, 2020]	*hachage*	
frontrunner$_1$	*extracteur embusqué*	Note that the term "*extracteur embusqué*" (literally, a blockchain searcher acting as a malicious extractor of transactions) also applies to the rare English term "backrunner1" (= backrunning node), while "*favori*" is the translation of "backrunner$_2$" (= backrunning bot).
frontrunner$_2$	*favori*	

distributed ledger technology$_1$ (DLT) "These same cryptocurrencies created the need for solutions surrounding blockchain; and this marked the beginning of new networks like Holochain, DAG and Hashgraph, that are based on distributed ledger technology (DLT). " [Iredale 2021]	technologie de registre distribué (DLT)	To distinguish between the technology itself and the device based on it, France recommends DLT, on the one hand, and DEEP on the other. It should be noted that while this polysemic case is not really contested, the second copolyseme (DEEP) is slow to take hold, especially across the Atlantic. As a reminder, a DEEP is a network or system based on DLT.
distributed ledger technology$_2$ (DLT) [= distributed ledger technology network/ system]	dispositif d'enregistrement électronique partagé (DEEP)	

Table 6. Examples of translation problems driven by polysemy.

3.2.4 Difficulties in translating full names of GANs

Generative adversarial networks (GANs), a type of deep learning architectures, have a particularly diverse typology. There are around thirty varieties, the full names of which are often rare in literature and are difficult to translate. In most cases, properly translating them into French requires a full grasp of their definitions and/or a regular use of techniques such as modulation or adaptation.

English	French	Observations
cycle-consistent generative adversarial network; CycleGAN (CCGAN)	*réseau antagoniste génératif de constance cyclique (CycleGAN)*	To put the term properly, there are two constants at the heart of this type of architecture: front cyclic constancy and rear cyclic constancy.
deep convolutional generative adversarial network (DCGAN)	*réseau antagoniste génératif profond à convolutions (DCGAN)*	A problem avoided by the translation is the juxtaposition of four adjectives in a row - "*réseau antagoniste génératif convolutionnel profond**". The use of pleonasm could also be avoided – a generative network is automatically a deep learning algorithm. Moreover, using the expression "*à convolutions profondes*" would suggest that there are superficial ones (all convolutions are inherent in deep learning).

dynamically grown generative adversarial network; dynamically grown GAN (DGGAN)	*réseau antagoniste génératif à croissance dynamique ; GAN à croissance dynamique (DGGAN)*	Translation by modulation of "dynamically grown".
information-maximizing generative adversarial network (InfoGAN)	*réseau antagoniste génératif d'optimisation d'informations (InfoGAN)*	The term '*maximisation*' is also used, but the most frequent is '*optimisation*'.
least squares generative adversarial network; least squares GAN (LSGAN)	*réseau antagoniste génératif par moindres carrés ; GAN par moindres carrés (LSGAN)*	Note the use of the preposition "*par*" rather than another.
progressive growing generative adversarial network (PGGAN, progressive GAN)	*réseau antagoniste génératif progressif (PGGAN) ; GAN progressif)*	The translation seems to suggest that the network is progressive, but also extensive. According to the term's definition, this is a GAN that improves by gradually increasing the number of layers in its generator and discriminator.
quantum generative adversarial network (QuGAN; quantum GAN)	*réseau antagoniste génératif à couches quantiques (QuGAN; GAN à couches quantiques)*	The translation reveals that the layers that are quantum-based and not the architecture itself.
sphere generative adversarial network (sphere GAN)	*réseau antagoniste génératif à métrique de probabilité intégrale (GAN-MPI, sphere GAN)]*	Typical notional translation: only the definition can effectively help achieve this rendering.
style-based generative adversarial network (StyleGAN)	*réseau antagoniste génératif de transfert de style (StyleGAN)]*	Translation aided by the definition: it's a GAN that reproduces image effects perfectly, hence its reputation in deepfake.

Loss-Sensitive GAN (LSGAN)	*réseau antagoniste sensible à la fonction de perte (LSGAN)*	By the way, all GANs are sensitive to the loss function. To be clearer, the following synonymous translation is more convincing (according to the definition): "*réseau antagoniste génératif à nouvelle fonction de perte*".

Table 7. Some translation problems inherent in GANs names.

3.2.5 Translation problems related to francization efforts

Disruptive technologies constitute a fast-evolving field, with a tremendous pace of English knowledge products. As a result, there is a huge gap between the proportion of English publications and those in French. Moreover, French-language communications, whether edited or translated, are still burdened with too many English borrowings, anglicisms and neologisms. French-speaking specialists in disruptive technologies, like translators and journalists, are making efforts to francize many terms to enrich this specialized language and boost dissemination of knowledge in French. As francization efforts are still rarely coupled with regular terminological standardization, the translation of a single English term into French gives rise to the multiplication of neologisms, which in turn fuels synonymy on the target language. This often results in challenging acceptance and rooting of French neologisms within the community of Francophone experts. In some cases, translation accuracy may be undermined, as is typically the case of lexical economy – francized terms tend to be longer and impact the size of texts (which is a real problem in subtitling).

Table 8 describes selected neologisms (francized terms) of the domain. We have created some of them and added others from both specialists and non-specialists (e.g., journalists), but while waiting for their widespread use, such francized terms represent a challenge for translators. We also suggest a few tips, based on our examples, to help those language specialists to keep francizing English terms in the field.

Useful tips	English	French
Systematically use the definition of the designated concept: the tendency towards terminological transparency (ensure the name of a concept helps to understand the semanteme of the said concept) is a major feature of techno-scientific fields.	event data	*données de contrats intelligents* [subdomain: blockchain]
	sphere generative adversarial network (sphere GAN)	*réseau antagoniste génératif à métrique de probabilité intégrale (sphereGAN)*
	brain organoid crosstalk	*transmission synergique entre cérébroïdes et organoïdes sensoriels*
	danksharding	*partitionnement en mode blob*
	brain-based biocomputing (= brain-directed organoid biological computing; biocomputing)	*bioinformatique organoïde [Qsyn : informatique neuromorphique]*
	Qsyn: neuromorphic computing	
	cyberflashing	*envoi non consenti d'images dénudées*
	biofeedback	*retour d'information de source biologique*
	location-based experience	*expérience vivante augmentée*
Make extensive use of layer translation where appropriate	native debt financing	
	blockchain trilemma	*trilemme de blockchain*
	synthetic biological intelligence (SBI)	*intelligence biologique synthétique; biointelligence synthétique*
	orthogenetic actuator	*actuateur orthogénétique*
	neural spheroid	*spéroïde neuronal*
	self-attention guidance	*guidage par autoattention*
	to tokenize; to crypto tokenize	*tokeniser; cryptotokeniser*

Reproduce the neologisation of existing terms. Make the most of the short forms, integrating abbreviations where necessary.	informer [inspired model : transformer] deepnude deepnuding app cheap nude pseudorandom rumber generator (PRNG) vanilla GAN time-based one-time password (TOTP) algorithm one-for-all (OFA) Microsoft common objects in context (MS-COCO)	*modèle informateur; réseau génératif informateur* *déshabillage truqué; déshabillage numérique* [*inspired model*: deepfake]. – another meaning: *image dénudée par (hyper)trucage.* *application de déshabillage (d'images)* *image dénudée non photoréaliste* *générateur PRNG* *GAN vanille; GAN original* *algorithme TOTP* *OFA* *[ex.: réseau OFA]* *jeu de données MS-COCO*
Dare to translate several English ending with "-ing" and "er" by "*-age*" and "*eur*", respectively, in French.	hodling doxing [derived from "dropping dox"]. proverb	*hodlage* *doxage* *prouveur* [Syn. of "*fournisseur de preuve*"]
Use one's imagination to draw on a range of French morphological resources, including affixation, prefixation, and compounding (Ahmad, 2015)	immersive (environment) user groomer [A paedophile who leverages online AI tools to deceive potential victims on their sex in order to trap them] grooming exaFLOP$_{[adj]}$ to convolute	*immersant*$_{(N)}$ [like many other French nouns ending with -ant] *groumeur* *pédopiégeage en ligne* *exaflopique* *convolutionner*

Table 8. A few tips and examples of francization in the specialized discourse of disruptive technologies.

3.2.6 Translation problems linked to a pleonastic or redundant use of terms

The pleonastic or redundant use of several terms, although considered acceptable in English, leads to translation difficulties in French. Most of such repetitions are not idiomatic and require deletions or reformulations. The table below gives some examples.

English	French		Note
	Raw translation	*Revised translation*	
Data mining of data$_{(Y)}$ from the hospital's database$_{(Z)}$ by the engineer$_{(X)}$.	Exploration des données des données contenues dans la base de données de l'hôpital par l'ingénieur.	Exploration des données contenues dans la base de données de l'hôpital par l'ingénieur.	The term "data mining" should be rendered as "*exploration*" or "*fouille*" if its 2nd actant [Y= what is explored] is 'data'.
"Text mining is data mining of textual data." (Ratner, 2017)	L'exploration de textes est l'exploration des données des données textuelles.	L'exploration de textes est la fouille de données textuelles.	
"However, it focuses on data mining of very large amounts of data…"; "data mining of commercial data"	Cependant, il [l'ouvrage] se concentre sur l'exploration des données de très grandes quantités de données ; l'exploration des données des données commerciales	Cependant, il [l'ouvrage] se concentre sur l'exploration de mégadonnées commerciales	
Data mining of inputs$_{(Y)}$ from the hospital's database$_{(Z)}$ by the_engineer$_{(X)}$.	Exploration des données des entrées/informations issues de la base de données de l'hôpital par l'ingénieur	Exploration des entrées/ informations issues de la base de données de l'hôpital par l'ingénieur	If Y (= what is explored) is a synonym for 'data' (input, information, etc.), then this synonym should be used and 'data' deleted.

"Involuntary dataveillance of limited data sets as a condition of airplane boarding, …" "dataveillance" of our online habits as well as the complexity of scientific data curation.	La dataveillance involontaire de petits ensembles de données, une condition de l'embarquement aéroportuaire, … = La surveillance par les données de petits ensembles de données, … La dataveillance de nos habitudes et la complexité de la curation des données scientifiques.	La dataveillance involontaire basée sur de petits ensembles de données, une condition de l'embarquement aéroportuaire, … La dataveillance de nos habitudes et la complexité de la curation des données scientifiques. [No change]	Above all, use the neologism "*dataveillance*" and avoid its synonym "*surveillance par les données*" in this context. Correct use
"The notion that new business models can be developed through the data monetization of AI and IoT data is incredibly compelling." (Ed Maguire, 2019) While the data monetization of **vehicle data continues to be a hyped buzzword, yet not everyone is aware of this possibility...**	L'idée du développement de nouveaux modèles économiques par le biais de la monétisation des données des données de l'IA et de l'IdO est incroyablement captivante. Alors que la monétisation des données des données de véhicules continue d'être un terme à la mode, personne n'a encore conscience de cette possibilité…	L'idée du développement de nouveaux modèles économiques par le biais de la monétisation des données de l'IA et de l'IdO est incroyablement captivante. Alors que la monétisation des données de véhicules continue d'être le terme en vogue dans les médias, personne n'a encore conscience de cette possibilité…	Delete redundant words

Table 10. Examples of redundancies causing translation problems.

3.2.7 Translation problems related to the contextual use of terms

Like in most specialized languages, the context imposes on the translator whether he/she can mechanically insert a term in a sentence or not. For the sake of idiomaticity and appropriate communication, terms are often modulated, glossed over, or replaced by an anaphoric word when the specific context of use requires to do so. This represents a challenge for language specialists and shows that the availability of a terminology resource is not sufficient to produce impeccable translations. The translator's competence in properly using each term according to the context is key. Translation issues related to terms use relate

not only to terminology, but also to the need for achieving accuracy of rendering and high-quality style and language as a holistic package. There is obviously an interplay between terminology usage, grammar, and stylistics.

Let's examine two translation problems resulting from how terms are inserted into specific contexts. The two issues were identified in an excerpt from EntryPoint (2023):

> 'While user experience is expected, user retention in Cosmos will require deeper liquidity and smooth UX for new users. User experience is full of friction and unfriendly to new and veteran users alike.'

The first challenge resides in the absolute need to modify terms some picked from a terminology database according to specific contexts. In the excerpt above (just two small sentences), the term 'user' appears five times, which would look clumsy if replicated in the target language. Reducing this number entails avoiding a mechanical insertion of the French equivalents of the terms 'user acquisition', 'user retention', and 'user experience,' for this specific context. Instead of the French translation in (a),

(a) '*Alors que l'acquisition utilisateur est attendue, la rétention utilisateur dans Cosmos exigera un volume de liquidités plus important et une expérience utilisateur (UX) souple de la part de nouveaux utilisateurs. L'expérience utilisateur est difficile et peu conviviale tant pour les nouveaux que pour les anciens utilisateurs.*'

the one in (b) is far better and recommended:

(b) '*Alors qu'il espère acquérir de nouveaux utilisateurs, le réseau Cosmos les retiendra en leur exigeant un volume de liquidités plus important et une expérience utilisateur (UX) souple. Il est difficile et peu convivial, tant pour les nouveaux que pour les anciens utilisateurs, d'acquérir de l'expérience.*'

The second translation challenge raised by the above-mentioned excerpt from EntryPoint is related to the translator's ability to identify the appropriate term and use it in the target language even when the original is partially expressed in context.

In the said excerpt, the term "fragmentation" is partial even though its aggregate meaning can be inferred. In fact, the whole is "fragmentation of (crypto) assets", the French equivalent of which is '*fragmentation d'actifs*' or '*fragmentation de cryptoactifs*' – which is the term to be used in the translation. The same kind of problem may arise from the verbal expression of a term that is well known in nominal form. For instance, let's examine the translation of this short excerpt from Ballav *et al.*, (2024):

> AI, particularly machine learning and deep learning, has the potential to transform the way one analyzes and interprets complex biological data. In the context of brain organoids, AI can be used to analyze large-scale data sets, identify patterns, and make predictions that would be impossible for humans to do manually.

The translator should be able to note that 'identify patterns' in the artificial-intelligence discourse – which is a verbal collocation – is known in French as '*génération des formes*', a true multi-word term that can be found in a relevant terminology resource. Therefore, the excerpt from Ballav *et al.*, (2024) above can be translated as follows:

> '*Au niveau des organoïdes cérébraux, l'intelligence artificielle (IA) intervient dans l'analyse de grands ensembles de données, la reconnaissance des formes, ainsi que dans des prédictions qui seraient impossibles à réaliser manuellement par les humains.*' (we underline identifiable terms)

Of course, other alterations of multi-word terms imposed by the context and idiomaticity include those

deriving from the impact of "universals of translation" (Baker, 1993: 242–243), linguistic features which typically characterize translated rather than original texts and include the tendency to be more explicit (hence, the notion of 'explicitation' by the author), to use more conventional grammar and lexis, and to sometimes simplify, alter or gloss over terminological units across a context. For example, the sentence

> "Predictions are part of supervised learning or unsupervised learning, where data are labelled or not."

can be translated into French as follows (thus altering the underlined recognized terms):

> 'Les prédictions s'effectuent de manière supervisée ou non supervisée, avec données étiquetées ou non.'

4. Conclusion

This paper aimed to provide the main characteristics of the specialized language of disruptive technologies and examine some of the translation challenges faced in the domain. Following our investigation based on an English-French comparable corpus of specialist texts totalling five million words, we have found that the disruptive technologies discourse abundantly leverages methods such as lexical economy, synonymy, metonymy and metaphor to form its terms and phraseological elements. The paper also addresses some translation challenges posed by the domain's terminology and suggests tips to solve others. In parallel, we vow for sustained research on the specialized discourse of disruptive technologies and the creation of multilingual terminology databases for this complex and highly evolving field.

References

Ahmad, Khurshid (2015). The role of specialist terminology in artificial intelligence and knowledge acquisition. In Wright; Sue Ellen *et al.* (Eds), *Terminology Handbook*, volume 2. Amsterdam/New York: Johns Benjamins, xxx.

Ajzen, I. (1991). The theory of planned behaviour. In Ajzen *et al.* (Eds.), *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T.

Baker, Mona (1993). Corpus Linguistics and Translation Studies. Implications and Applications. In Mona Baker, Gill Francis, & Elena Tognini Bonelli (Eds.), *Text and Technology: In Honor of John Sinclair*, Amsterdam/Philadelphia: John Benjamins, 233–250.

Ballav, Sangeeta, Amit Ranjan, Subhayan Sur, & Soumya Basu (2024). Organoid Intelligence: Bridging Artificial Intelligence for Biological Computing and Neurological Insights. In Ballav, S. *et al.* (Eds.), *Technologies in Cell Culture – A Journey from Basics to Advanced Applications*. Book chapter. InTechOpen, 27 p.

Desagulier, Guillaume (2017). *Corpus Linguistics and Statistics with R. – Introduction to Quantitative Methods in Linguistics*. Gewerbestrasse/New York, 359 p.

Diki-Kidiri, Marcel (2022). Cultural Terminology: An introduction to theory and method. In Faber, Pamela, & Marie-Claude L'Homme (Eds). *Theoretical Perspectives on Terminology – Explaining, Terms, Concepts and Specialized Knowledge*. Amsterdam/Philadelphia: Johns Benjamins, 197–218.

Edem Antia, Bassey (2007). Introduction – LSP studies: Factoring in indeterminacy. In Edem Antia, Bassey (Ed.). *Indeterminacy in Terminology and LSP: Studies in honour of Heribert Picht*. Volume 8. Amsterdam/Philadelphia: John Benjamins, xi–xxii.

EntryPoint (2023). Interchain Liquidity: An Analysis of Defi Markets in Cosmos. https://entrypoint.zone.

Gamallo, Pablo (2019). Strategies for building high quality bilingual lexicons from comparable corpora. In Doval, Irene, & Teresa Sánchez Nieto (Eds.), *Parallel Corpora for Constructive and Translation Studies*. Amsterdam/Philadelphia: John Benjamins, 251–266.

Kleiber, Georges (1988). Prototypes, stéréotypes : un air de famille. *DRLAV*, 38, 1–61.

Kocourek, Rotislav (1982). *La langue française de la technique et de la science*. Paris/Zurich, Brandstetter Verlag, 259 p.

Larsson, Richard (2023). MEV Bots 101: What are MEV Bots and How are They Making Millions? https://medium.com/coinmonks/mev-bots-101-what-are-mev-bots-and-how-are-they-making-millions-c91406d9334a.

Ratner, Bruce (2017). *Statistical and Machine-Learning Data Mining*. 3rd Edition. London/New York: CRC Press, 285 p.

Sager, Juan C. (1990). *A Practical Course in Terminology Processing*. Amsterdam/Philadephia: John Benjamins.

Sharoff, Serges, Reinhard Rapp, Pierre Zweigenbaum, & Pascale Fung (Eds.) (2013). *Building and Using Comparable Corpora*. New York/London, Springer, 333 p.

Schmitz, Klaus-Dirk (2007). Indeterminacy of terms and icons in software localization. In Edem Antia, Bassey (Ed.), *Indeterminacy in Terminology and LSP: Studies in honour of Heribert Picht*. Volume 8. Amsterdam/Philadelphia: John Benjamins, 49–58.

Temmerman, Rita (2002). Metaphorical Models and the Translation of Scientific Texts. *Linguistica Antverpiensia* 1, 211–226.

Wüster, Eugen (1985). *Einführung in die allgemeine Terminologielehre und terminologische Lexikographie*. 2nd Ed. (posthumous) [1st in 1979]. The LSP Centre, UNESCO Alsed Lsp Network. Copenhagen: The Copenhagen School of Economics.

Zanettin, Federico (2014). *Translation-Driven Corpora: Corpus Resources for Descriptive and Applied*