



HAL
open science

The rise and fall of promises: a preliminary method to trace promises in science and technology

Stefan D M Gaillard, Sanda Hachana

► To cite this version:

Stefan D M Gaillard, Sanda Hachana. The rise and fall of promises: a preliminary method to trace promises in science and technology. *Journal of Responsible Innovation*, 2025, 12 (1), <10.1080/23299460.2024.2440965>. <hal-04982768>

HAL Id: hal-04982768

<https://hal.science/hal-04982768v1>

Submitted on 7 Mar 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



The rise and fall of promises: a preliminary method to trace promises in science and technology

Stefan D. M. Gaillard & Sanda Hachana

To cite this article: Stefan D. M. Gaillard & Sanda Hachana (2025) The rise and fall of promises: a preliminary method to trace promises in science and technology, Journal of Responsible Innovation, 12:1, 2440965, DOI: [10.1080/23299460.2024.2440965](https://doi.org/10.1080/23299460.2024.2440965)

To link to this article: <https://doi.org/10.1080/23299460.2024.2440965>



© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 07 Jan 2025.



Submit your article to this journal [↗](#)



Article views: 355



View related articles [↗](#)



View Crossmark data [↗](#)

RESEARCH ARTICLE



The rise and fall of promises: a preliminary method to trace promises in science and technology

Stefan D. M. Gaillard ^a and Sanda Hachana^b

^aInstitute for Science in Society, Radboud University, Nijmegen, the Netherlands; ^bUniv. Grenoble Alpes, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

ABSTRACT

Nanoscience has faced several scandals since its inception, with Theranos being the most recent. This scandal resulted from continuous overpromises and deceptive promises, practices not uncommon in science. Promises foster mutual confidence and cooperation, but to remain reliable, breaking them must have consequences. One such consequence is decreased reputation, for example through scandal. Although the Theranos scandal exposed the harms of overpromising and deceptive promises, methods to identify and monitor promises remain scarce. Such methods would allow analysis of the literature to detect broken promises over time. To address this gap, we develop a method to trace promises over time using a training corpus of nanobiology documents. This corpus was queried through the CQPweb system. We combined query development with the corpus metadata to enable the analysis when promises were made. Further integration of this method with other analytical tools could help ensure that consequences follow more quickly from over-promising.

ARTICLE HISTORY

Received 21 February 2024
Accepted 6 December 2024

KEYWORDS

Promises; Search query;
Scandals; Signalling; Science
communication

Introduction

The Theranos scandal, which shook academic and business communities alike, starkly illustrates the perils of broken promises, deceptive promises, and overpromises in science.¹ In 2003, a charismatic entrepreneur founded Theranos, promising a revolutionary technology that could conduct a wide range of blood tests using just a few drops of blood. Expectations were monumental: faster, cheaper, and more accessible diagnostic testing (Carreyrou 2023).

Throughout the years, the CEO of Theranos carefully cultivated these expectations. Her vision and promises, as well as the potential of her technology, attracted significant attention and investment, propelling the company to a valuation of \$9 billion at its peak. However, investigations revealed that behind the appearance of innovation and progress,

CONTACT Stefan D. M. Gaillard  stefan.gaillard@ru.nl  Institute for Science in Society, Faculty of Science, Radboud University, PO Box 9102, 77, 6500 HC Nijmegen, The Netherlands

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

the technology was deeply flawed and incapable of producing accurate and reliable results. Theranos, it turned out, was using standard equipment for most tests while misleading investors, customers, and the public about their capabilities. Theranos promised investors unrealistic timelines for the development of their product as well as unrealistic revenue projections (Carreyrou 2023). In 2006, Theranos assured potential investors that the company *would* generate revenues of \$120 million to \$300 million over the next eighteen months, that their products *would* be placed in patients' homes during clinical trials and that patients' use of their products *would* reduce pharmaceutical companies' research costs by as much as 30 per cent (Carreyrou 2023, 7–8).

As the Theranos story unfolded, the gap between the company's public promises and the reality of its technological capabilities grew increasingly wide. This disconnect not only fuelled scepticism among some observers but also placed enormous pressure on the company, which started to deceive its investors and partners. Theranos fostered a culture of secrecy, discouraging employees from discussing the company's technology and processes both internally and externally. This environment of fear and isolation was conducive to perpetuating the myth of a revolutionary breakthrough, even as evidence mounted that the technology did not work as claimed. The company's leadership, particularly its CEO, played a pivotal role in crafting this narrative, using her charisma and bold promises to maintain the illusion of progress and success. This strategy of obfuscation and deception was not sustainable in the long term, leading to the eventual exposure of the truth through investigative journalism and regulatory scrutiny.

The unravelling of Theranos is a textbook example of deceptive promising and overpromising in science, where the desire for prestige and power overshadowed the virtues of honesty and integrity. This extreme case underscores a broader issue in the scientific community: the pressure to deliver revolutionary findings can sometimes lead to overpromising or overstating the capabilities of a technology or research. Although the Theranos scandal might seem rather unique, it resulted from practices which are not uncommon elsewhere in science,² namely continuous overpromising and deceptive promising – both resulting in broken promises.

In this methodological article, we aim to develop a method to systematically trace promises over time, as currently existing methods are insufficiently effective, allowing sensationalist claims in science and science policy to go unchecked. Our method offers a significant advancement by making it easier to systematically trace predictions and promises over time. This in turn allows for the detection of actors who continuously make and break promises over time, indicating whether the actor is trustworthy or not. In addition, it can also be used to track whether a particular research line keeps making the same predictions and promises over time. This could help identify research programmes which fail to achieve success due to a variety of reasons, or are particularly successful.

Finally, it facilitates a deeper understanding of the evolving nature of promises in science. By tracking these promises, we can gain insights into the alignment or divergence of promised outcomes with actual research developments. Further research can then be conducted on how to better align promises with rational expectations.

To address the current lack of methods to identify continuously broken promises, we develop a method to identify repeated overpromising using a training corpus of nanobiology documents. In this methodological article, we first present the signalling

theory principles which justify our method. Subsequently, we describe how we selected and collected our corpus. After transforming the corpus PDFs into computer-readable and searchable VRT files, the corpus was analysed through the web-based CQPweb system. After optimizing for precision and recall, the query was then tested against the judgement of two independent coders, resulting in a differentiation between promises and other types of predictions. Combining this with the available metadata of the corpus enables an analysis of who broke which promises and when, providing a crucial component for the analysis of broken promises.

Whether a promise has been broken can of course sometimes be a matter of debate; the promiser and the promisee might have different viewpoints on the agreement that follows a promise or the meaning of certain definitions used in the promise. To circumvent potential discussions about whether promises were broken or not, we (1) focus on providing tools which anyone can use to identify promises made over time, so they can subsequently determine for themselves whether said promises were broken or not, and (2) focus on promises which are repeated over time by the same actor, providing clear benchmarks for moments when the promise remained unfulfilled – if the promise was fulfilled, there would be no need to repeat the promise. Continual repetition of a promise indicates that it has yet to be fulfilled, acting as a clear benchmark for (lack of) progress.

Identifying broken promises retroactively can serve five purposes:

- (1) Revealing behaviour which we might want to discourage, namely the continuous breaking of promises;
- (2) Indicating something about the trustworthiness of the promiser, namely whether they routinely keep their promises or not;
- (3) Alternately to (2), indicating something about the trustworthiness of the institution, namely whether they routinely keep their promises or not;
- (4) Identifying institutions which fail to correct untrustworthiness over time;
- (5) Highlighting stagnant or particularly successful research programs.

Our method opens up numerous other potential applications, such as research into the more positive aspects of promising. For instance, it can help identify the types of researchers who *do* fulfil their promises and the conditions under which promises are or are not made. This in turn can be used to inform scientific reform regarding for example research funding and peer review.

Theoretical framework: justification for the development of a new method

The significance of broken promises

Broken promises are often, but not necessarily, the consequence of overpromising or making deceptive promises. Overpromising means making a promise without taking into account the context of knowledge (Gaillard, Mody, and Halfman 2023), whereas deceptive promises are promises made without the aim of fulfilling them (Gaillard, Mody, and Halfman 2023; Markovits 2011). Unsurprisingly, the practice of promising

is a common one in the domains of science and technology (Rip 2006) – just like elsewhere, promises made in these domains signify the creation or willing of an obligation (Anscombe 1978) or signify the creation of an expectation (Gaillard, Mody, and Halffman 2023). Promises can therefore lead to mutual confidence and enable cooperation (Rand 1984). If promises are frequently broken or seen as unrealistic, they lose their signalling strength and effectiveness; subsequently, their value as signals diminishes (Przepiorka and Berger 2017; Gambetta 2009; Scanlon 1998, 312). In this context, signalling strength refers to the perceived commitment and likelihood that the promise will be fulfilled. In the context of promising, a strong signal is a promise which conveys credibility and reliability. On the other hand, signalling effectiveness pertains to how the promise influences the behaviour and decisions of those to whom it's made. An effective promise leads to actions based on the trust that it will be honoured. When the strength and effectiveness of these signals are compromised by a few irrational scientists with a pattern of unfulfilled promises, it undermines confidence and can lead to scepticism and caution in future interactions, stifling collaboration, innovation, and progress.

Promises are made by agents. In the case of nanobiology, these agents are often scientists, and when scientists make a promise, they either (1) assure to a promisee that they (will) do something, or (2) assure that a specific outcome will occur (Gaillard, Mody, and Halffman 2023). What reason do we have to potentially doubt these assurances? One way is to critically assess the promise itself, to investigate whether the promiser has properly taken into account the context of knowledge. If it seems that they did not, then we have reason to doubt the promiser. However, because at 'the beginning of every new science information is densest at the boundaries' (Coombes and Robinson 1997), it is sometimes very difficult to assess whether a promiser has properly taken into account the context of knowledge – the relevant context of knowledge is not self-evident.

What we can do in those cases is turn to another helpful resource to assess promises, namely looking at who is making the promise. If we have reason to distrust the *promiser*, we by extension have reason to distrust their promise. The past behaviour of the promiser plays a crucial role in determining whether we can justly distrust them, as it is indicative of their capability to fulfil promises. Is it an agent with a history of keeping their word, or is it an agent known for empty proclamations? After all, we have good reason to be extra sceptical when we encounter a new promise by someone who has a pattern of overpromising – when someone frequently overpromises, it means they habitually promise more than they can deliver. This discrepancy between what the promiser habitually promises and what they actually deliver can be due to a number of factors, namely a lack of realistic judgement, deliberate manipulation, or outside influences. The discrepancy between promises and results signals reliability issues, which are crucial in assessing the trustworthiness of an agent's promises. This is what makes it useful to look at their track record of promising and is precisely what makes our method useful.

Of course, individuals operate in an environment. The existence of outside influences means that even though a scientist habitually does not fulfil their promises, this does not necessarily mean that this was due to their lack of judgment or deliberate manipulation, 'much like a car crash doesn't automatically indicate reckless driving' (Gaillard, Mody, and Halffman 2023). This is especially the case in science, as the inherent uncertainty in science makes it likely for promises to be broken. In other words, the amount of

promises a scientist breaks is likely to be high to begin with, due to the uncertain nature of the scientific endeavour. Still, some scientists might have a strong record of breaking promises even if one accounts for this uncertainty. Indeed, ‘just as a car crash can lead to inquiries about reckless driving, an unfulfilled promise can instigate investigation into potential overpromising’ (Gaillard, Mody, and Halffman 2023), especially if there is a pattern of promises being broken. Further investigation might then reveal that there were no outside influences, revealing the promiser to be either a poor judge or a deliberate manipulator – for our purposes, it is not necessary to differentiate which of the two. After all, both characteristics indicate that we cannot trust the promiser to fulfil their promise.

Some might argue that it could very well be the case that a deceptive promiser suddenly decides to make a truthful promise, or that they, despite being a poor judge, make a realistic judgement and make a promise which *does* take into account the context of knowledge. However, unless we have good reason to believe so, it is a far safer assumption that their incapability to keep their promises has not changed. Just like a potential partner would do well to question the sudden fidelity of her notoriously philandering suitor, so too should a promisee be sceptical of any supposed sudden trustworthiness on the part of a notoriously untrustworthy promiser. That is not to say that the capability of the person in question cannot change, but rather that the onus is on the promiser to prove their sudden trustworthiness.

Others might even go as far as to argue that someone who promises each year that they will find a cure for skin cancer may very well realistically be able to fulfil that promise by the end of their career: their earlier promises might have brought in the necessary funding and staff, which combined with time could be enough to finally be able to fulfil their promise; because the circumstances have changed, the capability of the promiser might have changed with it. However, in principle, we have good reason to distrust a promiser who has a habit of breaking their promises. Although it might be true that circumstances change, it seems just as likely that this will not be sufficient to change the capability of the promiser to fulfil their promise. Cases such as Theranos have shown that changing circumstances alone may just as well lead to no concrete results at all down the line. In addition, such an attitude would boil down to condoning or even rewarding reckless promises and dishonesty. Ideally, the situation described above would never arise because success should be grounded in actual ability and proven results, not in the manipulation of hope or expectations. Rewarding overpromises or systematic failure to deliver on promises, even if this occasionally eventually leads to success, undermines the values of integrity, honesty, justice, and productiveness. A culture that favours genuine accomplishment will encourage more honest and effective approaches to problem-solving, leading to real and sustainable progress.

Consequences for broken promises

Therefore, continuously broken promises need to have consequences, such as the decreased reputation of the actor who breaks their promises, for example through a scandal. For this to occur, we need methods to identify such actors. On a more systemic level, identifying institutions which continuously break their promises can provide grounds for further inquiry into the mechanisms incentivizing scientists to make promises that they cannot fulfil. This can then lead to reform to prevent future promises from

being broken. If such continuously broken promises persist over time, this can also point us towards institutions which, for example, fail to correct overpromising.

Although the Theranos scandal has laid bare the adverse effects of overpromising and deceptive promises, methods to identify continuously broken promises in science are scarce. In the field of corpus linguistics, there have been various queries developed to find expressions of future action (e.g. Berglund 1997; Berglund 2000; Hansen 2016), but these are not tailored towards finding predictions about the future with regards to science or technology, let alone with the ultimate aim of identifying (broken) promises. In the field of vision assessment, attention has been given to *how* visions about the future are communicated (e.g. Roßmann 2019), but these often remain descriptive, pointing out things like the range of impact of the vision and which problem it would supposedly solve (e.g. Fiedeler 2008), but not assessing whether and which promises have been routinely broken over time. In both futurology and futures studies, systematic studies of possible, probable and preferable futures are conducted (Sardar 2010). However, no specific method seems to exist for systematically studying promises about those futures and whether they were broken or not. On the other side of the spectrum, we are not aware of any digital methods to track down broken promises in the field of history either.

Method

The need for a method to identify continuously broken promises

Broken promises can only be identified qualitatively, because the context in which the promise is made must be understood in order to identify agreements resulting from the promise and evaluate whether the promise was fulfilled. Such a qualitative analysis can be supported by digital tools which help track which promises were made, by whom, and when.

Developing a method to identify continuously broken promises

A method to identify continuously broken promises would then require a tool to identify promises over time. As such, we developed a search query to identify predictions about the future in a given collection of texts (corpus), enabling the tracking of predictions over time. Users of our query can then manually determine whether these predictions were promises (and if they were broken) or not.

Corpus analysis tool

For our querying purposes, we use CQPWeb (Corpus Query Processor), a tool that allows us to easily find what we are looking for in a corpus (Hardie 2012). It is part of the collection of IMS Open Corpus Workbench (CWB) of open-source tools. The CWB collection of tools is used to work on big corpora to study the linguistics of texts. To do so, one has to use a query processor (CQP), collect a corpus, and annotate the linguistic information of this corpus. CQPWeb is an online tool that we will use in our project to do these studies. We used version v3.3.15 of CQPWeb.

To be able to work with CQPWeb, we have to upload a corpus. This requires collection of the corpus, then division into two files: the first one is the VRT file that corresponds to

the files' text and the second one is the META file that corresponds to the metadata of the files. We explain how to create these two files in Appendix 1.

CQPweb allows us to work efficiently with large corpora to find linguistic characteristics such as lexical (set of words in a text with a common theme), morpho-syntactic (set of structures that allow the grammatical construction of a statement), syntactic (grammatical rules for and between linguistic units) and semantic (the meaning, the significance). CQPWeb can be used for both qualitative and quantitative analysis. Quantitative analysis for example consists of studying the frequency of some words in the corpus, whereas qualitative analysis consists of analysing words or sentences in their context.

Corpus

To develop a method to trace the evolution of broken promises within the field of nanotechnology, it is necessary to construct a corpus of relevant documents to search through. Sometimes, relevant documents are not accessible due to proprietary constraints, such as in the cases of internal company documents and grant applications. We therefore limited our corpus to manuscripts accessible to third parties. This limitation complicates the identification of continuously made promises, as some relevant promises might be inaccessible.

Our corpus was assembled to facilitate a comprehensive analysis of promises made and their subsequent fulfilment or lack thereof over time. To achieve this, we selected a diverse array of dated documents, ensuring a rich and interconnected dataset. All documents consist of the whole manuscript rather than only a part, such as the abstract, which gives our corpus a distinct advantage over corpora which consist of only abstracts and executive summaries.

The chosen corpus comprises four distinct subcorpora, each serving a specific purpose in understanding the narrative of nanotechnology promises. The first subcorpus encompasses all twelve reports published by the National Nanotechnology Initiative (NNI) from its inception in 1999 until 2023. We selected these reports for their forward-looking nature. They are rich with scientists' aspirations, goals, and predictions, often intertwined with promises made to U.S. policymakers. These documents are pivotal in understanding how predictions and promises concerning nanotechnology have been made in the context of securing funding.

Subcorpora 2 and 3 were selected for their intrinsic connection to the first. Subcorpus 2 includes all scientific articles (43 in total) cited in the NNI reports. This linkage allows for the tracing of promises from their genesis in scientific discourse to their embodiment in policy-oriented documents. Subcorpus 3, consisting of patents (937 in total) that reference the scientific articles of subcorpus 2, extends this trajectory. It illuminates how these promises transition from scientific ideas to tangible technological applications, offering a unique perspective on the realization or modification of these promises in a practical domain.

The fourth subcorpus, although not directly linked to the preceding three, is crucial for its role in public discourse. It comprises a selection of popular science books on nanoscience and nanotechnology. The rationale for including these texts lies in their significant influence in disseminating knowledge of nanoscience and nanotechnology to a broader, non-specialist audience. This subcorpus enables us to capture a wider spectrum of narratives and promises about nanotechnology, including those that shape public

perception and understanding. By incorporating these texts, we aim to provide a more holistic view of the field's evolution, one that encompasses not only scientific advancements but also societal expectations and interpretations of nanotechnology.

In summary, the corpus consists of four subcorpora, namely (1) all twelve National Nanotechnology Initiative (NNI) reports from its founding to 2023, (2) all 43 scientific articles mentioned in the NNI reports, (3) 937 patents referencing the scientific articles of subcorpus 2,³ and (4) a selection of eight popular science books on nanoscience or – technology⁴ (see Table 1). The selection of these subcorpora was a strategic decision, designed to create a comprehensive and interconnected corpus. All subcorpora were processed via several steps to prepare them for use in the digital environment we use for the querying process.

A comprehensive process was undertaken beginning with the collection of PDF versions of the manuscripts. These PDFs were first transformed into machine-readable formats using an Optical Word Recognition (OCR) tool named Tesseract, and then converted into TEI XML files using GeneRation Of Bibliographic Data (GROBID) software. This conversion allowed for the extraction and systematic organization of both textual content and metadata (such as author names and publication dates) into a structured digital format. Subsequent analysis involved the use of Python scripts for advanced linguistic processing, including part-of-speech tagging. This processing classified words by their grammatical roles and identified specific entities, like chemicals or diseases, while also determining each word's base form (lemma). Although in the vast majority of cases, part-of-speech tagging occurs correctly, there are sometimes cases where words are mislabelled. This sometimes, although not often, can happen when a word (such as 'work') can be both a noun and a verb and the surrounding context does not provide sufficient clues for accurate disambiguation.

The final output was a Verticalized Text (VRT) file, which organized words and their linguistic attributes into a searchable table format, facilitating complex textual queries through the CQPweb interface. This meticulous process not only digitizes texts but also enriches them with detailed linguistic annotations, making them accessible for further scholarly analysis.

To facilitate the analysis of manuscripts, a process was developed to systematically compile the metadata. This was achieved using a Python script to extract metadata from TEI XML files, collecting various metadata attributes such as title, author(s), publication year, and abstract, with variations depending on the subcorpus. We manually checked whether the extracted metadata were accurate and complete – in some cases, some of the metadata were not available in the file itself and could therefore not be extracted automatically.

A unique identifier was also generated for each text, linking its content with its metadata to simplify retrieval on CQPWeb. This identifier allows researchers to efficiently

Table 1. Overview of the size and composition of the corpus.

Subcorpus	Type of manuscripts	Number of manuscripts	Total number of words	Average number of words per manuscript
1	NNI reports	12	208,775	17,398
2	Scientific articles	43	196,048	4,559
3	Patents	937	7,660,960	8,176
4	Popular Science Books	8	716,741	89,593

trace the source, date, and authorship of texts after conducting searches, aiding in the study of how promises are communicated and evolve over time. The META and VRT files are seamlessly integrated into the CQPWeb platform, enabling complex textual analyses without the need for manual file handling. See Appendix 1 for a more comprehensive description of how the corpus was prepared.

Query development

After the corpus was assembled, we developed a query to search through the corpus for predictions about the future. To develop that query, we first manually collected a list of sentences containing predictions from written sources – namely scientific articles, policy documents, books, editorials, and roadmaps – where claims about the future are made. These were sources that the first author consulted for other work. Every time the author came across a sentence containing a prediction⁵ in a written scientific source, he noted down the sentence in question. Sentences were purposefully collected from a variety of manuscript types to get a comprehensive overview of how predictions about the future are made in different media, so that we could tailor our query to find predictions about the future regardless of medium. For the same reason, sentences were collected from a variety of time periods and authors. These are sentences like

‘By 2010, nano-scale technologies will be the determining factor to profitability in virtually every sector of industrial economies’ (ETC Group 2003)

and

‘Within a decade, analysts foresee various nanotechnologies enabling things like complete medical diagnostic laboratories on one-inch computer chips; ubiquitous computers, some built into clothing; powerful, long-lasting batteries; efficient solar cells; new drugs and drug delivery systems; cheap visual displays; medical monitoring systems embedded in people’s bodies ready to sound an alert when a disease organism strikes or a cancer cell develops; and cheaper space travel’. (Bailey 2003)

From this list, we made another list with the common structures that these claims follow. Take the following example. In a 2006 article, the authors claim that

‘By 2010 a third generation of nanostructures will emerge, which feature nanosystems with thousands of interacting components’. (Nel et al. 2006)

Such a sentence was added to the list of sentences containing predictions. The claim in this sentence has the following, more general structure: By [year] a [Y] will emerge, which feature [feature].

From this list of common claim structures, 26 common phrases were identified (see Table 2), i.e. combinations of words which indicate that a claim, and by extension the sentence containing it, features a prediction. For the general claim ‘By [year] a [Y] will emerge, which feature [feature]’, these phrases would be ‘By [year]’ and ‘[Y] will emerge’.

Subsequently, these phrases were expressed in query language using corpus query processor syntax. We used the web-based corpus analysis system CQPweb, which allows users to perform sophisticated queries that can involve specific word patterns, phrases, part-of-speech tags, and syntactic structures. These queries are much more advanced than simple keyword searches, thus facilitating the search for predictions in a given

Table 2. List of 26 phrases commonly occurring in sentences containing a prediction in texts about nanoscience and nanotechnology.

by [year with four numbers]
by about [year with four numbers]
by [year with four numbers – year with four numbers]
by about [year with four numbers – year with four numbers]
developments in the next [word]
developments in the next [X years]
developments in the next [XX years]
in [X years]
in [XX years]
in [X] to [X] years
in [X] to [XX] years
in [XX] to [XX] years
future
available *** in [year with four numbers – year with four numbers]
Nano[...] will
breakthroughs **** include
Nano[...] may
Nano[...] is going to be
Nano[...] might
Nano[...] is set to
will benefit from nano[...]
may benefit from nano[...]
is going to benefit from nano[...]
might benefit from nano[...]
[year with four numbers] may
revolutionize

corpus. For our 26 common phrases (see [Table 2](#)) we coded 22 queries, as some common phrases could be captured by the same query, which we then combined into a single query (see [Figure 1](#)).

Results

Precision and recall

While developing the query, we aimed to balance precision and recall. To understand these concepts, let us first look at the underlying concepts of true positives, true negatives, false positives, and false negatives. True positives refer to instances when the query gets a hit when it should get a hit, whereas false positives refer to instances when the query gets a hit when it should not get a hit. True negatives refer to instances when the query does not get a hit when it should not get a hit, whereas false negatives refer to instances when the query does not get a hit when it should get a hit. (Stefanowitsch 2020, 111)

So how do these concepts relate to precision and recall? Precision measures the accuracy of the results returned by the query. It is the ratio of relevant results to the total number of results returned by the query, or the ratio of true positives to the sum of true positives and false positives. (Stefanowitsch 2020, 112) Recall measures the



Figure 1. The process of distilling sentences containing a prediction into a single search query.

completeness of the results returned by the query. It is the ratio of relevant results returned by the query to the total number of relevant results that exist within the system, or the ratio of true positives to the combination of true positives and false negatives. (Stefanowitsch 2020, 113)

While developing our query aimed at finding predictions about the future in the corpus, we prioritized recall over precision; false positives can always be ignored during the manual determination of whether the hits contain promises or not, whereas false negatives are lost to the subsequent analysis. To find this balance between precision and recall we used an iterative process, where we first developed a query to test on a test corpus (Test Corpus 1) to determine which false positives it found.

This test corpus consisted of two manuscripts of each of the subcorpora. Manuscripts were selected manually to ensure diversity, with different first authors whenever possible. Additionally, one manuscript was chosen as the most recently published and the other as the oldest, adjusting slightly to avoid author overlap. This diversity was important because promises might be expressed differently in different time periods and by different authors, and we wanted to make sure the query was optimized for this variety of linguistic expressions.

The query was subsequently refined to find fewer false positives. For example, the query `'[word = "in"] [word = "\w*"] {, 4} [lemma = "year"]'`, intended to find phrases containing predictions such as

'in (about) 30 years'

and

'in one year',

also found phrases such as 'in this year' and 'in my eleventh year'. Furthermore, in some instances when the query found the intended phrasing, it did not concern a prediction at all, such as in the case of

'Rats mature in months, and then age and die in about two years'.

In fact, out of 32 hits, 18 were false positives, indicating a precision of around 44%.

To remedy this low precision rate, we subsequently replaced this query with the queries `'[tag = "MD"] [word = "\w*"] * [lemma = "year"]'` and `'[lemma = "year"] [word = "\w*"] * [tag = "MD"]'`. Combining these queries into a single query `'([tag = "MD"] [word = "\w*"] * [lemma = "year"] | [lemma = "year"] [word = "\w*"] * [tag = "MD"])'` and applying it to the test corpus resulted in 66 hits, 15 of which were false positives, or a precision of around 77%.

Subsequently, after initial refinements, the latest iteration of the query was used to test the corpus again, this time to determine the number and content of false negatives. To do so we manually coded the predictions in the test corpus. Then, we compared the hits of the query to the manually coded predictions to determine which predictions the query did not find. After determining the recall rate, we modified certain queries and added some queries to increase the recall. An unintended but unsurprising consequence of this was that the precision decreased again. However, because we prioritize recall over precision, this was considered an acceptable tradeoff.

Intercoder reliability

To determine whether the query finds all predictions in the corpus, two independent coders were asked to classify whether sentences from another test corpus (Test Corpus 2) contained a prediction and whether they contained a promise. The former was done to determine whether the query should prioritize recall or precision. If there is high intercoder agreement, this indicates that the linguistic cues are fairly straightforward and it should be relatively easy to find all predictions. The query can then be optimized for precision without a significant loss in recall. If there is low intercoder agreement, the linguistic cues are likely ambiguous or complex, suggesting that it might be more challenging to identify all predictions accurately. In this case, the query should prioritize recall to ensure that most, if not all, predictions are captured, even at the expense of precision.

The test corpus consisted of two randomly selected manuscripts from each subcorpus, excluding those used in Test Corpus 1, to ensure a representative sample. The coding instructions were pre-tested by two colleagues to ensure comprehension and feasibility, meaning that they coded a sample number of sentences based on the instructions given and subsequently provided feedback on how instructions could be made clearer. See Appendix 3 for the coding instructions.

The coders each coded the same 778 sentences, consisting of the first 100 sentences from each manuscript in the test corpus. Two manuscripts consisted of less than 100 sentences, but rather 84 and 94 sentences, respectively, leading to a total of 778 sentences. Coder 1 classified 198 sentences as containing a prediction and 69 sentences as containing a promise. For coder 2 these numbers were 156 and 42 respectively. The coders agreed on 624 sentences whether they contained a prediction or not and agreed on 701 sentences whether they contained a promise or not. Using Cohen's Kappa, we calculated the intercoder reliability to be 0.439 for predictions and 0.256 for promises, indicating a moderate and low level of agreement respectively. Cohen's Kappa was chosen because we had two coders and categorical data (yes or no) and wanted to adjust for chance agreement.

Most of the disagreement between the coders seems to have arisen for predictions and promises where the linguistic expression was rather ambiguous, such as when the temporal component is unclear. In addition, coder 1 classified some sentences which use the word 'can' instead of 'will' as containing promises whereas coder 2 did not. Furthermore, coder 1 coded almost every sentence containing the word 'will' as containing a promise, even if those sentences contained words like 'maybe' or 'probably', whereas coder 2 did not code these sentences as containing a promise.

Because of the significant disagreement between the two coders, we decided to prioritize recall over precision; the query was further tailored to try to find all sentences containing what a user might consider to be a prediction. This means that there will also be more false negatives, but as the number of results already factored into the tens of thousands, some filtering mechanism is needed anyway, so an extra few hundred or thousand of hits will not make a difference. The results of the query can for example be filtered in CQPWeb by looking at the most frequently occurring words in the hits or by looking at which words occur near each other more than might be expected. In addition, future research could look at filtering the results using artificial intelligence.

After determining the reliability and recall of our query with regard to the test corpus and updating the query again, we applied the new query to the full corpus. The new query was as follows:

```
([tag = "MD"]) | ([lemma = "future"][word = "s"]?([word = "[,;:\(\)\|\-\]" ]* | [word = "\w*"])* [tag = "MD"]) | ([tag = "MD"] ([word = "[,;:\(\)\|\-\]" ]* | [word = "\w*"])* [lemma = "future"][word = "s"]?) | ([word = "by" ] [word = "\w*" ] {, 6} [word = "20.{2}"]) | ([word = "in" ] [word = "\w*" ] {, 4} [lemma = "year"]) | ([word = "in" ] [word = "\w*" ] {, 4} [lemma = "decade"]) | ([word = "nano.*" ] [lemma = "be" ] [pos = "ADV"]? [word = "\w*" ] [word = "to"][pos = "VERB"]) | ([word = "20.{2}"] [tag = "MD" & word != "can"]) | ([word = "nano.*" ] ([word = "[,;:\]" ]* | [word = "\w*"])* [lemma = "potential"]) | ([lemma = "potential" ] ([word = "[,;:\]" ]* | [word = "\w*"])* [word = "nano.*" ]) | ([lemma = "promise" ]) | ([word = "will" ] [word = "revolutionize" ]) | ([word = "expected" ] [word = "to" ]) | ([word = "slated" ] [word = "to" ]) | ([word = "predicted" ] [word = "to" ])
```

This query was able to find almost all of the sentences coded by the coders as containing a prediction or promise. These were predictions such as

‘Future nanoscale computers may consist of extremely large-configuration memories that are programmed for specific tasks by a tutor that locates and tags the defects in the system.’ (Heath et al. 1998, 1716)

and

‘Ultimately, nanotechnology is expected to be embedded throughout our economy, its contributions ranging from barely detectable to wholly dominant.’ (President’s Council of Advisors on Science and Technology 2005, 12)

and promises such as

‘Numerous newly designed, advanced materials and manufacturing processes will be built by 2015 using control at the nanoscale level in at least one key component.’ (Roco 2006, 229)

and

‘The Force Amplified Biological Sensor (FABS) is a desktop or portable instrument currently under development at the Naval Research Laboratory. FABS will use a rapid, automated immunoassay to detect analytes such as proteins, viruses, and bacteria.’ (Baselt et al. 1997, 672)

From the hits, we manually selected all predictions which were also promises. Then we first grouped these promises by subject and subsequently by author or institution. Combined with the metadata about when these promises were made, this enabled us to see whether we could identify promises which were continuously broken over time, and whether these were made by the same actors. To test whether the query is able to find occurrences of predictions and promises concerning the same topic over time, we checked whether our query results did indeed find such occurrences for a topic where we already knew it should be able to find them: nanoscience as a trillion dollar industry by 2015. From the early 2000s onwards, the discourse surrounding nanoscience was filled with optimistic predictions and promises about its potential to revolutionize the industry.

The nanotechnology industry would supposedly become a trillion-dollar industry by 2015.

Whether the nanotechnology industry actually became a trillion-dollar industry depends partially on how one calculates the economic contribution of nanotechnology. This highlights a limitation of our method: there can be disagreement on whether a promise was fulfilled based on the definitions used. However, as argued previously, repeated promises indicate that a promise has not been met. After 2015, predictions and promises concerning the market value of nanotechnology continued but with radically lower numbers. For example, Binns (2021) mentions a prediction of a global nanotechnology market of 125 billion dollars by 2024.

The query found multiple instances of predictions and one of a promise concerning nanotechnology as a trillion dollar industry by 2015. For example, the query found an early instance of such a prediction in the 2003 popular science book *Nanotechnology: A Gentle Introduction to the Next Big Idea*: ‘The National Science Foundation predicts that nano-related goods and services could be a \$ 1 trillion market by 2015, making it not only one of the fastest-growing industries in history but also larger than the combined telecommunications and information technology industries at the beginning of the technology boom in 1998.’ (Ratner and Ratner 2003, 3) This prediction was then reiterated in two other popular science books from the early 2000s (Booker and Boysen 2005, 22; Jurvetson 2006, 42). By 2007, the prediction starts to appear as a promise, for example in the popular science book *Nanotechnology 101*. Although the first – ‘The National Science Foundation predicts that the global marketplace for goods and services using nanotechnologies will be worth a trillion dollars by 2015.’ (Mongillo 2007, p. xvii) – and second – ‘By 2015, nanotechnology could be a \$1 trillion industry’ (3) – references are predictions, the statement transforms into a promise by the time of the third reference – ‘As was stated earlier, the global marketplace for goods and services using nanotechnologies *will* grow to \$1 trillion by 2015’ [our emphasis] (4).

The previous example shows that the query is indeed capable of finding predictions and promises over time in a given corpus, which – together with the provided metadata – allows the user to track predictions over time. In addition, the frequency filtering option provided by CQPweb indicated the high prevalence of the phrase *by 2015* in the list of hits, indicating that CQPweb filtering options indeed provide useful support for searching through the hits.

Conclusions and discussion

Applications

Although promising occurs in every field of science, notably in grant applications, we hypothesize that promises about the future tend to occur more in some disciplines than others, as well as more so in some document types than others. Notably, disciplines which leverage their supposed future potential to secure funding and generate interest seem to us more likely to have more promises about the future than disciplines which focus primarily on either theoretical understanding or immediate practical applications (rather than future ones). For example, disciplines such as classical physics or organic chemistry, while fundamental, typically emphasize the importance of their current

findings and methodologies over speculative future applications. In contrast, fields like quantum computing and artificial intelligence, which are often touted as having transformative potential, are more likely to feature bold predictions and promises about their future impact. Because our framework is developed on the basis of data from such a discipline – namely from nanoscience and nanotechnology, with a particular focus on nanobiology – it will most likely work better for these types of disciplines. More research is needed to determine the prevalence of (broken) promises in different disciplines and, relatedly, how these promises are expressed. The latter is needed to be able to tailor the query to these disciplines.

To apply our method to other disciplines and/or media, we recommend that researchers go through some of the same steps that we did, but not necessarily all. Because we hypothesize that promises are expressed differently in different fields and media, our current list of phrases indicating that a sentence features a prediction is probably by no means complete. A different selection of texts might contain a different set of linguistic cues – for example, more metaphorical or colloquial ones (e.g. nanotechnology is on the cusp of rewriting reality) – signaling statements about the future. It would therefore be useful for potential users to first gather sample sentences from the field and/or media they want to study. Subsequently, they should reduce these sentences to common phrases and turn these into corresponding queries. Together with the more general queries for finding predictions we have provided, this should allow users to find predictions in the context of their interest.

Our goal for developing this method is not to develop a panopticon for scientists to constantly check each other's promising track record. Rather, our framework ideally should be used to identify the occasional rotten apple, help track down behaviour which we want to disincentivize, and, most importantly, as the starting point for further analysis of why phenomena such as broken promises and overpromising fail to be corrected over time.

Further care should be taken that the identification and analysis of (broken) promises do not inadvertently and unproductively lead to scandals, especially if the behaviour of the promiser was not sufficiently immoral to warrant such a response. In addition, singular, unique rotten apples may lead to scandals which unjustly tarnish the reputation of a whole field or, reversely, the damaged reputation of a whole field might unjustly tarnish the reputation of a single subdiscipline or technology. An example of the latter is when investors in the early 2000s were hesitant to invest in voice recognition due to its association with artificial intelligence, which had built a reputation for unwarranted hype and broken promises (Crevier 1993). Yet by that time, voice recognition had advanced far enough to be actually promising (Reddy 1976). Therefore, reducing funding for that specific technology because it would be unfruitful was unwarranted.

Limitations

One major limitation of the method we have developed concerns the timeline for assessing broken promises. Scientific advancements often require decades to come to fruition. This extended timeline complicates the evaluation of whether a field is suffering from continuously broken promises or if individual agents within the field have repeatedly failed to deliver on their commitments. The method is therefore probably most useful

to assess agents and disciplines over a longer period of time, in the range of multiple decades. However, the method can also be fruitfully applied to smaller timescales for other types of research, such as the impact of policy changes on the prevalence of promises.

In addition, whether a promise was broken or not can be contested. As we saw with the example of nanotechnology as a trillion-dollar market, whether the promise was broken or not depended on the definition of the nanotechnology market that one uses. One therefore needs to be able to clearly define the terms of the original promise, or show that the promise was reiterated; thereby indicating that the original promise remained unfulfilled.

Future research

Future research should focus on further developing complementary tools for the query, such as artificial intelligence applications to more easily group related promises together by both the subject of the promise as well as by who made the promise. These AI tools could employ natural language processing techniques to identify and categorize promises, providing users with a more intuitive and comprehensive understanding of the data. Machine learning algorithms could also predict the likelihood of fulfilment based on historical patterns, offering a predictive dimension to the analysis of promises.

In addition, network analysis could serve as a useful supplement to the query. By mapping out the connections between different individuals, organizations, and topics, network analysis can reveal patterns and relationships that are not immediately apparent. This could help identify key influencers or recurrent themes in the promises made, and track how certain promises propagate through networks. Furthermore, network analysis can uncover clusters of related promises and highlight how their interconnectedness impacts the fulfilment of promises, thereby offering deeper insights into the dynamics of promises over time.

In conclusion, the method proposed in this article allows users to track predictions and promises over time. The integration of our method with other analytical tools represents a significant step forward in addressing the challenges of overpromising. It opens up new avenues for ensuring that the promises driving cooperation and funding in science are grounded in reality, thereby fostering a culture of trust and integrity within the scientific community. Future research should focus on refining these methods and exploring their application across different scientific disciplines.

Notes

1. A promise becomes broken when it remains unfulfilled at the point in time it was supposed to become fulfilled.
2. We take science to be systematic knowledge production, covering both the activities pursued in academia as well as in research and development divisions and other institutes.
3. These patents were found by searching for the related articles on Dimensions AI in the period of June 2023 until December 2023 and downloading the related patents available at the time. Of the downloaded files, 16 were discarded: 12 because they were empty and 4 because they were too big to process.
4. See Appendix 2 for details on the selection and content of this subcorpus.

5. A sentence was considered to contain a prediction if it implicitly or explicitly contained a statement about the future, indicated by linguistic cues such as modal verbs (e.g. 'will', 'might', 'could'), adverbs related to the future (e.g. 'soon', 'eventually', 'later'), and specific time references (e.g. 'by 2030', 'in the coming two decades').
6. The earliest filing of a variant of the patent.
7. The filing date of the patent in a specific country or territory.
8. The date on which a patent application is first published.

Acknowledgements

We would like to thank Willem Halfman, Cyrus Mody, Caspar Safarlou, and two anonymous peer reviewers for their extensive and valuable comments on earlier versions of this text. In addition, we would like to thank Cyril Labbé for his invaluable help with the processing of the corpus and development of the query. Finally, we are grateful to Aly Rogers for her meticulous copy editing, which greatly improved the clarity and readability of the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was funded by ERC NanoBubbles [grant number 951393].

ORCID

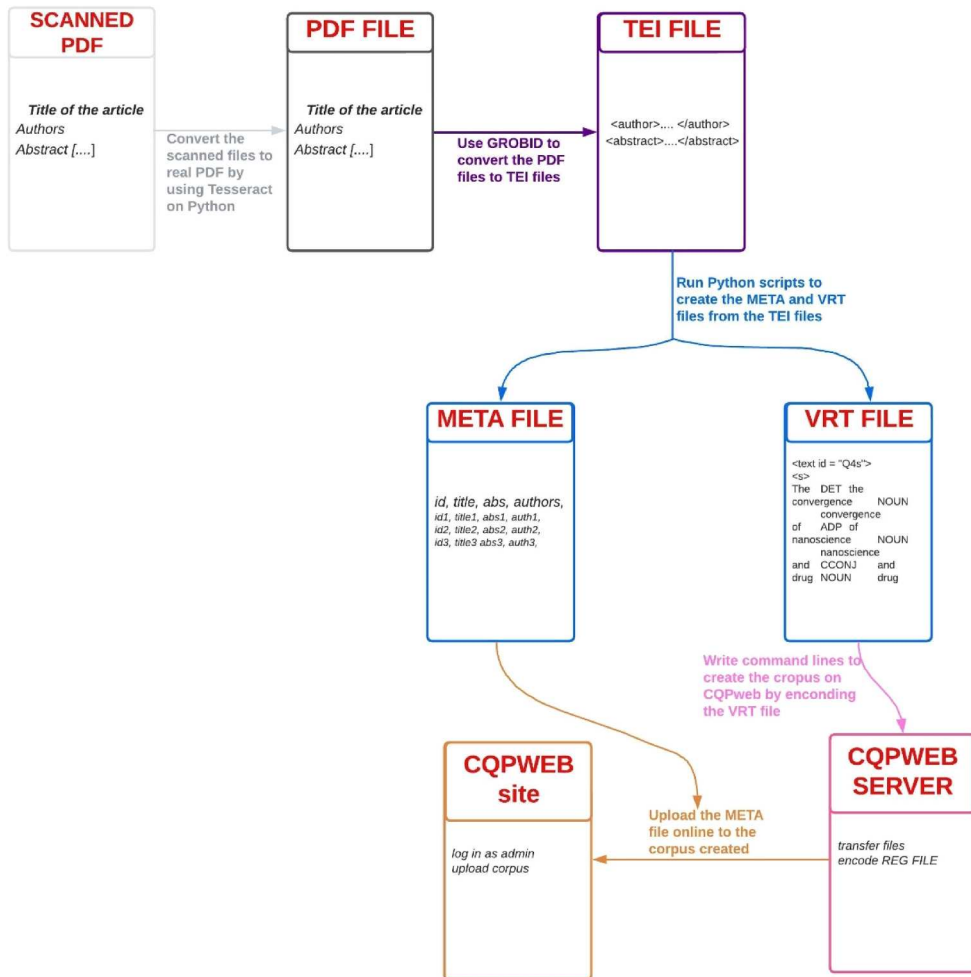
Stefan D. M. Gaillard  <http://orcid.org/0000-0003-1956-7325>

References

- Anscombe, G. E. M. 1978. "Rules, Rights, and Promises." *Midwest Studies in Philosophy* 3 (1): 318–323. <https://doi.org/10.1111/j.1475-4975.1978.tb00364.x>.
- Bailey, Ronald. 2003. "The Smaller the Better." Reason. Accessed February 18, 2024. <https://reason.com/2003/12/01/the-smaller-the-better-2/>.
- Baselt, D. R., G. U. Lee, K. M. Hansen, L. A. Chrisey, and R. J. Colton. 1997. "A High-Sensitivity Micromachined Biosensor." *Proceedings of the IEEE* 85 (4): 672–680. <https://doi.org/10.1109/5.573755>.
- Berglund, Y. 1997. "Future in Present-day English: Corpus-Based Evidence on the Rivalry of Expressions." *ICAME Journal* 21:7–20.
- Berglund, Y. 2000. "Utilising Present-day English Corpora: A Case Study Concerning Expressions of Future." *ICAME Journal* 24:25–64.
- Binns, C. 2021. *Introduction to Nanoscience and Nanotechnology*. Hoboken: John Wiley & Sons.
- Booker, R. D., and E. Boysen. 2005. *Nanotechnology for Dummies*. Hoboken: John Wiley & Sons.
- Carreyrou, J. 2023. *Bad Blood: Secrets and Lies in a Silicon Valley Startup*. New York: Vintage Books.
- Coombes, Steve, and Dave Robinson. 1997. "The Science of Sex." *New Yorker*, September 29.
- Crevier, D. 1993. *AI: The Tumultuous History of the Search for Artificial Intelligence*. New York: Basic Books, Inc.
- ETC Group. 2003. "The Big Down: Technologies Converging at the Nano-scale." Accessed February 18, 2024. <https://repstein.faculty.drbu.edu/Nanotech/bigdown.htm>.

- Fiedeler, U. 2008. "Technology Assessment of Nanotechnology: Problems and Methods in Assessing Emerging Technologies." In *Presenting Futures. The Yearbook of Nanotechnology in Society, vol 1*, edited by E. Fischer, C. Selin, and J. M. Wetmore, 241–263. Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-8416-4_21
- Gaillard, S., C. Mody, and W. Halfman. 2023. "Overpromising in Science and Technology: An Evaluative Conceptualization." *TATuP-Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis* 32 (3): 60–65. <https://doi.org/10.14512/tatup.32.3.60>.
- Gambetta, D. 2009. *Codes of the Underworld: How Criminals Communicate*. Princeton: Princeton University Press.
- Hansen, K. R. 2016. "News from the Future: A Corpus Linguistic Analysis of Future-Oriented, Unreal and Counterfactual News Discourse." *Discourse & Communication* 10 (2): 115–136. <https://doi.org/10.1177/1750481315611240>.
- Hardie, A. 2012. "CQPweb — Combining Power, Flexibility and Usability in a Corpus Analysis Tool." *International Journal of Corpus Linguistics* 17 (3): 380–409. <https://doi.org/10.1075/ijcl.17.3.04har>.
- Heath, J. R., P. J. Kuekes, G. S. Snider, and R. S. Williams. 1998. "A Defect-Tolerant Computer Architecture: Opportunities for Nanotechnology." *Science* 280 (5370): 1716–1721. <https://doi.org/10.1126/science.280.5370.1716>.
- Jurvetson, S. 2006. "Nanotechnology Commercialization: Transcending Moore's Law with Molecular Electronics and Nanotechnology." In *Nanotechnology: Science, Innovation, and Opportunity*, edited by L. E. Foster, 33–56. Upper Saddle River, NJ: Prentice Hall PTR.
- Markovits, D. 2011. "Promise as an Arm's-Length Relation." In *Promises and Agreements: Philosophical Essays*, edited by H. Sheinman, 295–326. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195377958.003.0013>
- Mongillo, John. 2007. *Nanotechnology*. Vol. 101. Westport: Greenwood.
- Nel, A., T. Xia, L. Madler, and N. Li. 2006. "Toxic Potential of Materials at the Nanolevel." *Science* 311 (5761): 622–627. <https://doi.org/10.1126/science.1114397>.
- President's Council of Advisors on Science and Technology. 2005. *The National Nanotechnology Initiative at Five Years: Assessment and Recommendations of the National Nanotechnology Advisory Panel*. Washington, DC: National Science and Technology Council.
- Przepiorka, W., and J. Berger. 2017. "Signaling Theory Evolving: Signals and Signs of Trustworthiness in Social Exchange." In *Social Dilemmas, Institutions and the Evolution of Cooperation*, 373–392. <https://doi.org/10.1515/9783110472974-018>
- Rand, A. 1984. "Causality Versus Duty." In *Philosophy: Who Needs It*, 128–136. New York: Signet.
- Ratner, M. A., and D. Ratner. 2003. *Nanotechnology: A Gentle Introduction to the Next Big Idea*. Upper Saddle River, NJ: Prentice Hall Professional.
- Reddy, D. R. 1976. "Speech Recognition by Machine: A Review." *Proceedings of the IEEE* 64 (4): 501–531. <https://doi.org/10.1109/PROC.1976.10158>.
- Rip, A. 2006. "Folk Theories of Nanotechnologists." *Science as Culture* 15 (4): 349–365. <https://doi.org/10.1080/09505430601022676>.
- Roßmann, M. 2021. "Vision as Make-Believe: How Narratives and Models Represent Sociotechnical Futures." *Journal of Responsible Innovation* 8 (1): 70–93. <https://doi.org/10.1080/23299460.2020.1853395>.
- Roco, M. C. 2006. "Convergence and Integration." In *Nanotechnology: Science, Innovation, and Opportunity*, edited by L. E. Foster, 225–232. Upper Saddle River, NJ: Prentice Hall PTR.
- Sardar, Z. 2010. "The Namesake: Futures; Futures Studies; Futurology; Futuristic; Foresight—What's in a Name?" *Futures* 42 (3): 177–184. <https://doi.org/10.1016/j.futures.2009.11.001>.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge: Harvard University Press.
- Stefanowitsch, A. 2020. *Corpus Linguistics: A Guide to the Methodology*. Berlin: Language Science Press.

Appendix 1. Corpus Preparation



We first collected PDF files of all the manuscripts that the corpus consists of. Scanned PDFs were converted to real PDFs using the optical word recognition (OCR) tool Tesseract in Python. Subsequently, the PDF files were converted to TEI XML files using the open-source software tool Generation Of Bibliographic Data (GROBID). GROBID was used to extract and parse textual information from the PDF documents to create TEI XML files, which are indexed with tags separating the information (the metadata as authors, publication date etc.) and sentences of the PDF file. We then used multiple Python scripts to create both a Verticalized Text (VRT) file containing the sentences from the manuscript and a META file containing the metadata.

VRT files

Subsequently for the VRT file, multiple Python scripts were used on the PDFs for part-of-speech tagging and named entity recognition. Part-of-speech tagging (POS) is a technique where each

word in a text is labeled with its grammatical type, like for example verb or adjective. The tool spaCy was used for this, which enabled us to label each word using two different categorization lists, namely the spaCy list and the Penn Treebank list. Each list has a different way of naming and categorizing grammatical types – for example, a verb in spaCy will be tagged as ‘VERB’ while the Penn Treebank uses the tag ‘VB’. We chose to work with both of them, because spaCy’s list is more accurate for our corpus due to being tailored to (bio)medical texts, whereas the Penn Treebank list gives more specified information about the part of speech (for example for a verb, it can specify whether it’s the base form (‘VB’), a verb conjugated to a past tense (‘VBD’), and more). Although in the vast majority of cases, part-of-speech tagging occurs correctly, there are sometimes cases where words are mislabelled. This sometimes, although not often, can happen when a word (such as ‘work’) can be both a noun and a verb and the surrounding context does not provide sufficient clues for accurate disambiguation.

Another technique we get by using spaCy’s model ScispaCy is named entity recognition. For each word, named entity recognition gives us its referential linguistic expression, i.e. something specific in the world that the word refers to (if there is one). We mostly found results for proper nouns. In our corpus, only two named entities exist, both proper nouns, namely ‘CHEMICAL’ (chemical and biological words) and ‘DISEASE’ (proper nouns of diseases). Finally, to get the lemma of a word we also used spaCy. A lemma is the base of the word, without any prefix and/or suffix. For example, break is the lemma of, among others, break, breaks, broke, broken, and breaking.

All of these previous techniques are used to create the VRT file so the text is ready to be searched through via CQPweb. The VRT file consists of a table with five columns, where each column represents a different type of information about each word in the text. These columns are:

1. **Word:** The actual word from the text.
2. **POS (Part of Speech):** This shows the grammatical category of the word, like noun, verb, adjective, etc.
3. **Penn TAG:** These are specific codes from the Penn Treebank system, used to give more detailed grammatical information about the word.
4. **Lemma:** The base or dictionary form of the word. For example, for ‘running,’ the lemma would be ‘run.’
5. **Named Entity:** If the word is a specific name of a person, place, organization, etc., this column will have that information.

The file has as many lines as there are words in the text it’s analyzing. Each line in the file corresponds to a single word from the text, along with its related information in the five columns. Special tags are included in the file to indicate the start of a new text or a new sentence. These help in organizing and differentiating between different sections or sentences in the text.

This file is then added to the CQPweb environment, allowing the user to look for a word, a sequence of words or even a sentence by its grammatical information in the query. For example, with the query ‘([word = ‘nano.*’] [lemma = ‘be’] [pos = “ADV”]? [word = “\w*”] [word = ‘to’][pos = “VERB”])’, it is possible to search for any word containing nano, followed by any conjugation of the word be, followed by any adverb, followed by any word, followed by the word to, followed by any verb.

META files

To create the META files, the metadata of each of the TEI XML files was added automatically to this new file using a Python script. The META file also has the format of a table where each line corresponds to one text. The columns correspond to the metadata we previously chose to get automatically from the TEI files on our script. It is advisable to manually check whether the extracted metadata is accurate and complete. This manual verification ensures that the automated process does not miss any critical information or introduce errors.

For subcorpora 1, 2, and 4, the metadata consists of title, author(s), authors' affiliation(s), publisher, publication year and the abstract. For subcorpus 3, the metadata consists of title, priority date,⁶ filing date,⁷ publication date,⁸ inventors, and patent ID.

In addition, the first column corresponds to a specific code identificatory created randomly and automatically. We used a code identificatory for each file to link its text (from the VRT file) to its metadata (from the META file) so it will be found easily on CQPWeb. The META file helps us, once the results are found after executing the query, to get the information about the text for each result, especially the outlet, the date and the author(s). This is crucial for our purposes, as this enables us to easily identify where, when and by whom promises are made, making it possible to chart the making and breaking of promises over time.

The VRT file and the META file need not be opened manually (except if there's a need to check something). They are directly uploaded on CQPWeb by typing command lines on the server and uploading some parameters on the website of CQPWeb. For the scripts used, please see <https://gricad-gitlab.univ-grenoble-alpes.fr/nanobubbles/cqpweb/-/tree/main>.

Appendix 2. Selection of Subcorpus 4

To compile subcorpus 4, we went to Amazon.com, a popular website for ordering books, and searched for 'nanotechnology' 'popular science'. We then selected all the books which were solely on nanoscience or nanotechnology, which were not academic (such as collections of articles), which cost \$35,00 or less, and which appeared in the first ten hits. We did the same after a search for 'nanotechnology' 'innovation', 'nano' 'popular science', and 'nano' 'innovation'. This turned up the following books:

'Nanotechnology' 'popular science'

- Size Really Does Matter: The Nanotechnology Revolution;
- Nanofuture: What's Next For Nanotechnology;
- Nanotechnology 101 (Science 101);
- No Small Matter: Science on the Nanoscale;
- Understanding Nanotechnology (Science Made Accessible);
- Radical Abundance: How a Revolution in Nanotechnology Will Change Civilization;

'Nanotechnology' 'innovation'

- Radical Abundance: How a Revolution in Nanotechnology Will Change Civilization;
- Nanofuture: What's Next For Nanotechnology;
- Nano : The Emerging Science of Nanotechnology;
- Nanotechnology: A Gentle Introduction to the Next Big Idea;
- Nanotechnology For Dummies;
- Nanotechnology Revolution;
- Nanotechnology: Science, Innovation, And Opportunity;

'Nano' 'popular science'

- Life's Ratchet: How Molecular Machines Extract Order from Chaos;
- Size Really Does Matter: The Nanotechnology Revolution;
- Nanotechnology 101 (Science 101);
- Engines of Creation: The Coming Era of Nanotechnology;
- Radical Abundance: How a Revolution in Nanotechnology Will Change Civilization;
- Understanding Nanotechnology (Science Made Accessible);
- Nanofuture: What's Next For Nanotechnology;

'Nano' 'innovation'

- Nano: The Emerging Science of Nanotechnology;
- Nanofuture: What's Next For Nanotechnology;

The list of unique books was as follows:

- Size Really Does Matter: The Nanotechnology Revolution;
- Nanofuture: What's Next For Nanotechnology;
- Nanotechnology 101 (Science 101);
- No Small Matter: Science on the Nanoscale;
- Understanding Nanotechnology (Science Made Accessible);
- Radical Abundance: How a Revolution in Nanotechnology Will Change Civilization;
- Nano: The Emerging Science of Nanotechnology;
- Nanotechnology: A Gentle Introduction to the Next Big Idea;
- Nanotechnology For Dummies;
- Nanotechnology Revolution;
- Nanotechnology: Science, Innovation, And Opportunity;
- Life's Ratchet: How Molecular Machines Extract Order from Chaos;
- Engines of Creation: The Coming Era of Nanotechnology;

Of these, *No Small Matter: Science on the Nanoscale*, *Understanding Nanotechnology (Science Made Accessible)*, *Nano: the Emerging Science of Nanotechnology*, *Nanotechnology Revolution* and *Life's Ratchet: How Molecular Machines Extract Order from Chaos* did not have PDF versions available, leading to the following list of books included in the subcorpus:

- Size Really Does Matter: The Nanotechnology Revolution;
- Nanofuture: What's Next For Nanotechnology;
- Nanotechnology 101 (Science 101);
- Radical Abundance: How a Revolution in Nanotechnology Will Change Civilization;
- Nanotechnology: A Gentle Introduction to the Next Big Idea;
- Nanotechnology For Dummies;
- Nanotechnology: Science, Innovation, And Opportunity;
- Engines of Creation: The Coming Era of Nanotechnology;

Appendix 3: Coding Instructions

Please read through the sentences in the Excel file you received and for each sentence indicate whether the sentence contains a prediction (with a 1) or not (with a 0), as well as whether the sentence contains a promise (with a 1) or not (with a 0). In case of doubt, please indicate that the sentence indeed does contain a prediction (in case of doubt about a prediction) or promise (in case of doubt about a promise). Please code any prediction and promise you come across, regardless of who makes it (the author or someone else). The first tab in the Excel file contains some example codings.

A sentence is considered to contain a prediction if it implicitly or explicitly contains a statement about the future, including but not limited to possible and (im)plausible futures.

A sentence is considered to contain a promise if it contains an assurance that a particular thing will happen (e.g. "Trust me, the police will definitely catch the perpetrator") or that one will do something (e.g. "If you give me the money I will buy you a new phone"). Although sentences containing these types of assurances will often contain the word 'will', they need not (e.g. the promise 'I give you my word that I shall bring you groceries tomorrow'). Similarly, sentences containing the

word 'will' need not necessarily be promises, such as in cases where the sentence does not contain an assurance (e.g. 'I think I will perhaps do groceries tomorrow').

A promise always contains an implicit or explicit reference to a time scale ('this research will cure cancer'/'this research will cure cancer in X years'/'this research will eventually cure cancer'). It is not necessarily linked to a person ('I will cure cancer' and 'this research will cure cancer' are both valid instances).

Each sentence coded as containing a promise (1) should also be coded as containing a prediction (1) – see the conceptualization above. However, it is possible for sentences with a prediction to include (1) a promise or not (0). An example of a sentence containing a prediction but not a promise would be '2030 could be the year we will start seeing self-replicating nanobots'.