ContractShield

Andrew Choi
Mason Bulling

What already exists?

Expensive Software

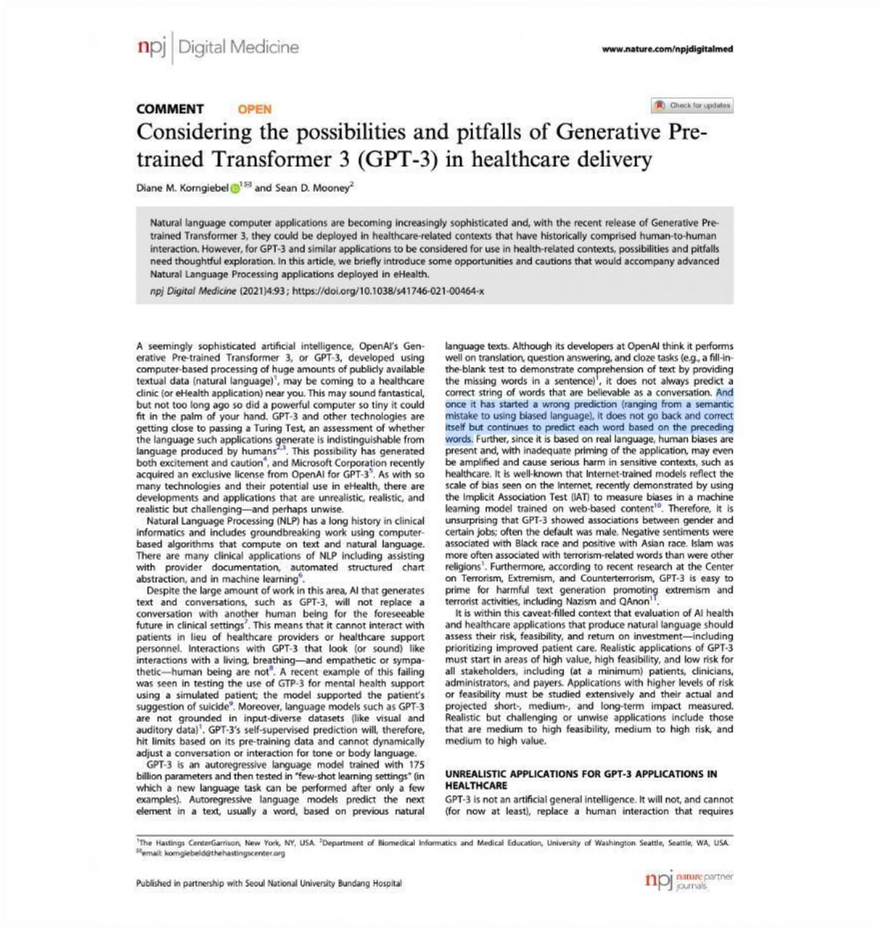
Lawyers

Simplify Legal Documents... For Everyone

Immigrants

Low-income groups

Small Business Owners



npj | Digital Medicine

www.nature.com/npjdigitalmed

COMMENT OPEN

Check for updates

Considering the possibilities and pitfalls of Generative Pre-trained Transformer 3 (GPT-3) in healthcare delivery

Diane M. Korngiebel[1] ✉ and Sean D. Mooney[2]

Natural language computer applications are becoming increasingly sophisticated and, with the recent release of Generative Pre-trained Transformer 3, they could be deployed in healthcare-related contexts that have historically comprised human-to-human interaction. However, for GPT-3 and similar applications to be considered for use in health-related contexts, possibilities and pitfalls need thoughtful exploration. In this article, we briefly introduce some opportunities and cautions that would accompany advanced Natural Language Processing applications deployed in eHealth.

npj Digital Medicine (2021)4:93 ; https://doi.org/10.1038/s41746-021-00464-x

A seemingly sophisticated artificial intelligence, OpenAI's Generative Pre-trained Transformer 3, or GPT-3, developed using computer-based processing of huge amounts of publicly available textual data (natural language)[1], may be coming to a healthcare clinic (or eHealth application) near you. This may sound fantastical, but not too long ago so did a powerful computer so tiny it could fit in the palm of your hand. GPT-3 and other technologies are getting close to passing a Turing Test, an assessment of whether the language such applications generate is indistinguishable from language produced by humans[2,3]. This possibility has generated both excitement and caution[4], and Microsoft Corporation recently acquired an exclusive license from OpenAI for GPT-3[5]. As with so many technologies and their potential use in eHealth, there are developments and applications that are unrealistic, realistic, and realistic but challenging—and perhaps unwise.

Natural Language Processing (NLP) has a long history in clinical informatics and includes groundbreaking work using computer-based algorithms that compute on text and natural language. There are many clinical applications of NLP including assisting with provider documentation, automated structured chart abstraction, and in machine learning[6].

Despite the large amount of work in this area, AI that generates text and conversations, such as GPT-3, will not replace a conversation with another human being for the foreseeable future in clinical settings[7]. This means that it cannot interact with patients in lieu of healthcare providers or healthcare support personnel. Interactions with GPT-3 that look (or sound) like interactions with a living, breathing—and empathetic or sympathetic—human being are not[8]. A recent example of this failing was seen in testing the use of GTP-3 for mental health support using a simulated patient; the model supported the patient's suggestion of suicide[9]. Moreover, language models such as GPT-3 are not grounded in input-diverse datasets (like visual and auditory data)[1]. GPT-3's self-supervised prediction will, therefore, hit limits based on its pre-training data and cannot dynamically adjust a conversation or interaction for tone or body language.

GPT-3 is an autoregressive language model trained with 175 billion parameters and then tested in "few-shot learning settings" (in which a new language task can be performed after only a few examples). Autoregressive language models predict the next element in a text, usually a word, based on previous natural language texts. Although its developers at OpenAI think it performs well on translation, question answering, and cloze tasks (e.g., a fill-in-the-blank test to demonstrate comprehension of text by providing the missing words in a sentence)[1], it does not always predict a correct string of words that are believable as a conversation. And once it has started a wrong prediction (ranging from a semantic mistake to using biased language), it does not go back and correct itself but continues to predict each word based on the preceding words. Further, since it is based on real language, human biases are present and, with inadequate priming of the application, may even be amplified and cause serious harm in sensitive contexts, such as healthcare. It is well-known that Internet-trained models reflect the scale of bias seen on the Internet, recently demonstrated by using the Implicit Association Test (IAT) to measure biases in a machine learning model trained on web-based content[10]. Therefore, it is unsurprising that GPT-3 showed associations between gender and certain jobs; often the default was male. Negative sentiments were associated with Black race and positive with Asian race. Islam was more often associated with terrorism-related words than were other religions[1]. Furthermore, according to recent research at the Center on Terrorism, Extremism, and Counterterrorism, GPT-3 is easy to prime for harmful text generation promoting extremism and terrorist activities, including Nazism and QAnon[11].

It is within this caveat-filled context that evaluation of AI health and healthcare applications that produce natural language should assess their risk, feasibility, and return on investment—including prioritizing improved patient care. Realistic applications of GPT-3 must start in areas of high value, high feasibility, and low risk for all stakeholders, including (at a minimum) patients, clinicians, administrators, and payers. Applications with higher levels of risk or feasibility must be studied extensively and their actual and projected short-, medium-, and long-term impact measured. Realistic but challenging or unwise applications include those that are medium to high feasibility, medium to high risk, and medium to high value.

UNREALISTIC APPLICATIONS FOR GPT-3 APPLICATIONS IN HEALTHCARE

GPT-3 is not an artificial general intelligence. It will not, and cannot (for now at least), replace a human interaction that requires

[1]The Hastings CenterGarrison, New York, NY, USA. [2]Department of Biomedical Informatics and Medical Education, University of Washington Seattle, Seattle, WA, USA. ✉email: korngiebeld@thehastingscenter.org

Published in partnership with Seoul National University Bundang Hospital

npj nature partner journals

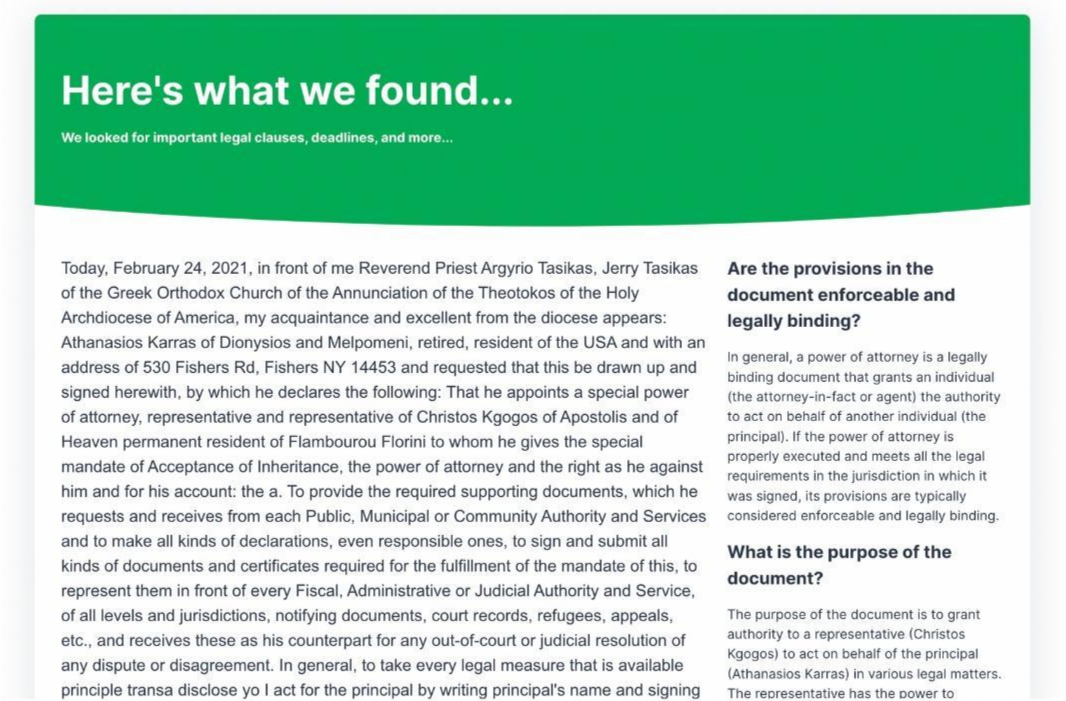
"And once it has started a wrong prediction, it does not go back and correct itself but continues to predict each word based on the preceding words"

KORNGIEBEL & MOONEY

The Problem...

Scalability & Performance

Narrow Training Sets



Here's what we found...

We looked for important legal clauses, deadlines, and more...

Today, February 24, 2021, in front of me Reverend Priest Argyrio Tasikas, Jerry Tasikas of the Greek Orthodox Church of the Annunciation of the Theotokos of the Holy Archdiocese of America, my acquaintance and excellent from the diocese appears: Athanasios Karras of Dionysios and Melpomeni, retired, resident of the USA and with an address of 530 Fishers Rd, Fishers NY 14453 and requested that this be drawn up and signed herewith, by which he declares the following: That he appoints a special power of attorney, representative and representative of Christos Kgogos of Apostolis and of Heaven permanent resident of Flambourou Florini to whom he gives the special mandate of Acceptance of Inheritance, the power of attorney and the right as he against him and for his account: the a. To provide the required supporting documents, which he requests and receives from each Public, Municipal or Community Authority and Services and to make all kinds of declarations, even responsible ones, to sign and submit all kinds of documents and certificates required for the fulfillment of the mandate of this, to represent them in front of every Fiscal, Administrative or Judicial Authority and Service, of all levels and jurisdictions, notifying documents, court records, refugees, appeals, etc., and receives these as his counterpart for any out-of-court or judicial resolution of any dispute or disagreement. In general, to take every legal measure that is available principle transa disclose yo I act for the principal by writing principal's name and signing

Are the provisions in the document enforceable and legally binding?

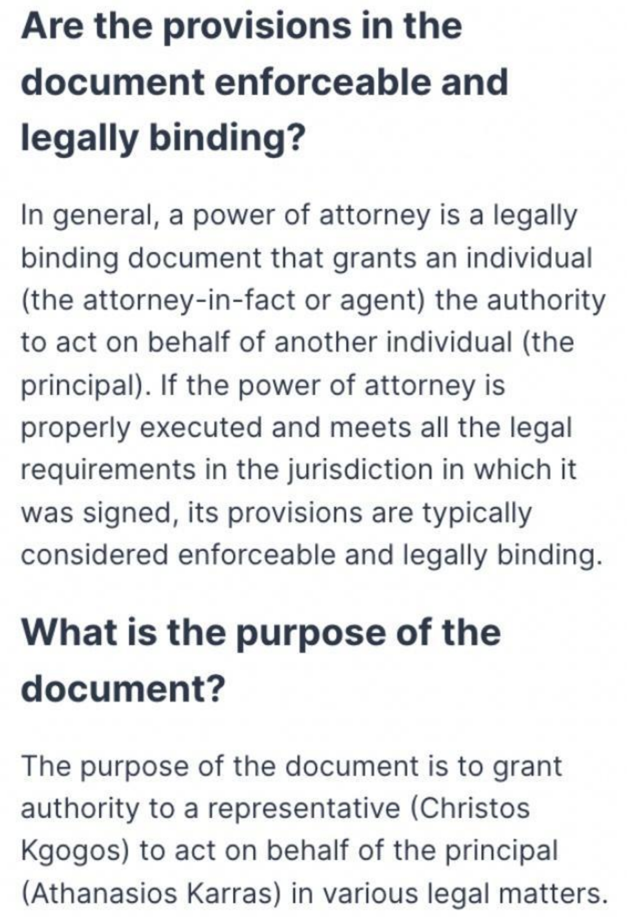
In general, a power of attorney is a legally binding document that grants an individual (the attorney-in-fact or agent) the authority to act on behalf of another individual (the principal). If the power of attorney is properly executed and meets all the legal requirements in the jurisdiction in which it was signed, its provisions are typically considered enforceable and legally binding.

What is the purpose of the document?

The purpose of the document is to grant authority to a representative (Christos Kgogos) to act on behalf of the principal (Athanasios Karras) in various legal matters. The representative has the power to

An Easy Platform

Simplifying Documents, One Clause at a Time.



Are the provisions in the document enforceable and legally binding?

In general, a power of attorney is a legally binding document that grants an individual (the attorney-in-fact or agent) the authority to act on behalf of another individual (the principal). If the power of attorney is properly executed and meets all the legal requirements in the jurisdiction in which it was signed, its provisions are typically considered enforceable and legally binding.

What is the purpose of the document?

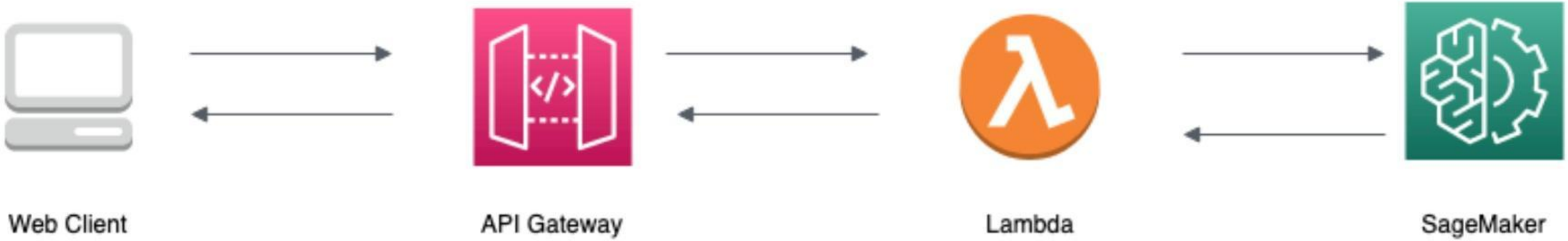
The purpose of the document is to grant authority to a representative (Christos Kgogos) to act on behalf of the principal (Athanasios Karras) in various legal matters.

Auto-Scaling

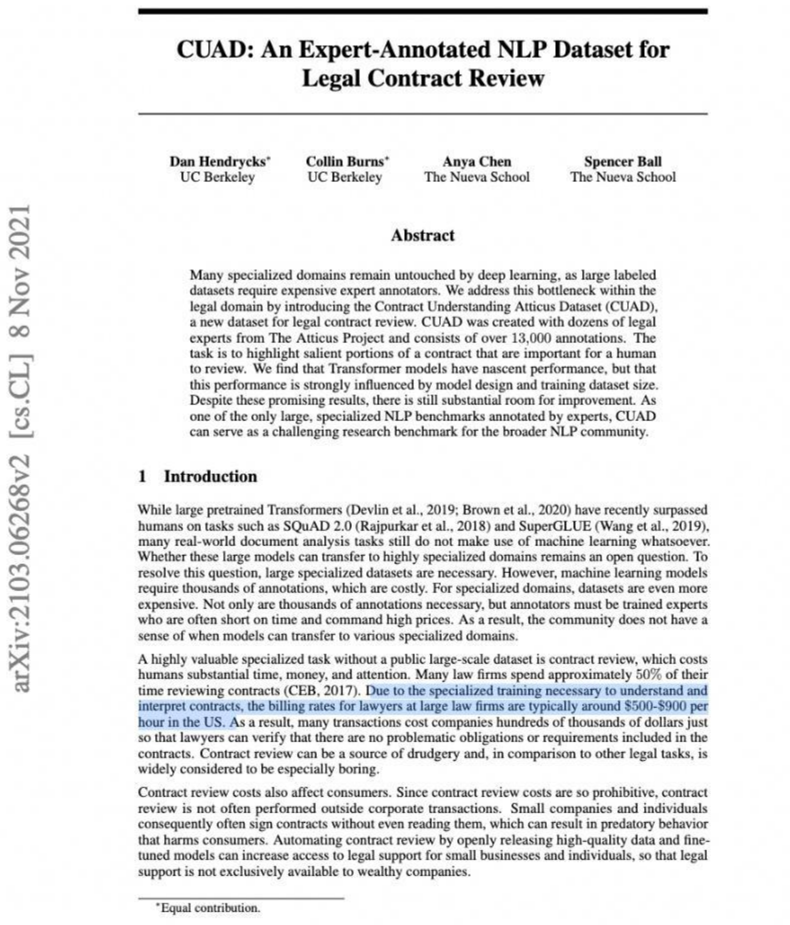
Pseudo-serverless architecture allows for easy, fast horizontal scaling

Spot Instances

We use 10 ml.c6g.xlarge spot instances, to run our jobs at the fraction of the original cost



MADE WITH
beautiful.ai



CUAD: An Expert-Annotated NLP Dataset for Legal Contract Review

Dan Hendrycks* UC Berkeley
Collin Burns* UC Berkeley
Anya Chen The Nueva School
Spencer Ball The Nueva School

Abstract

Many specialized domains remain untouched by deep learning, as large labeled datasets require expensive expert annotators. We address this bottleneck within the legal domain by introducing the Contract Understanding Atticus Dataset (CUAD), a new dataset for legal contract review. CUAD was created with dozens of legal experts from The Atticus Project and consists of over 13,000 annotations. The task is to highlight salient portions of a contract that are important for a human to review. We find that Transformer models have nascent performance, but that this performance is strongly influenced by model design and training dataset size. Despite these promising results, there is still substantial room for improvement. As one of the only large, specialized NLP benchmarks annotated by experts, CUAD can serve as a challenging research benchmark for the broader NLP community.

1 Introduction

While large pretrained Transformers (Devlin et al., 2019; Brown et al., 2020) have recently surpassed humans on tasks such as SQuAD 2.0 (Rajpurkar et al., 2018) and SuperGLUE (Wang et al., 2019), many real-world document analysis tasks still do not make use of machine learning whatsoever. Whether these large models can transfer to highly specialized domains remains an open question. To resolve this question, large specialized datasets are necessary. However, machine learning models require thousands of annotations, which are costly. For specialized domains, datasets are even more expensive. Not only are thousands of annotations necessary, but annotators must be trained experts who are often short on time and command high prices. As a result, the community does not have a sense of when models can transfer to various specialized domains.

A highly valuable specialized task without a public large-scale dataset is contract review, which costs humans substantial time, money, and attention. Many law firms spend approximately 50% of their time reviewing contracts (CEB, 2017). Due to the specialized training necessary to understand and interpret contracts, the billing rates for lawyers at large law firms are typically around $500-$900 per hour in the US. As a result, many transactions cost companies hundreds of thousands of dollars just so that lawyers can verify that there are no problematic obligations or requirements included in the contracts. Contract review can be a source of drudgery and, in comparison to other legal tasks, is widely considered to be especially boring.

Contract review costs also affect consumers. Since contract review costs are so prohibitive, contract review is not often performed outside corporate transactions. Small companies and individuals consequently often sign contracts without even reading them, which can result in predatory behavior that harms consumers. Automating contract review by openly releasing high-quality data and fine-tuned models can increase access to legal support for small businesses and individuals, so that legal support is not exclusively available to wealthy companies.

*Equal contribution.

"Due to the specialized training necessary to understand and interpret contracts, the billing rates for lawyers at large law firms are typically around $500-$900 per hour in the US"

Documents Shape Our Lives

Cars

Data Plans

Liability Waivers

Insurance

ContractShield