

[Skip to main content](#)> [cs](#) > arXiv:1902.06006

quick links

- [Login](#)
- [Help Pages](#)
- [About](#)

Computer Science > Computation and Language

arXiv:1902.06006 (cs)


[Submitted on 15 Feb 2019 ([v1](#)), last revised 17 Apr 2020 (this version, v3)]

Contextual Word Representations: A Contextual Introduction

[Noah A. Smith](#)[Download PDF](#)

This introduction aims to tell the story of how we put words into computers. It is part of the story of the field of natural language processing (NLP), a branch of artificial intelligence. It targets a wide audience with a basic understanding of computer programming, but avoids a detailed mathematical treatment, and it does not present any algorithms. It also does not focus on any particular application of NLP such as translation, question answering, or information extraction. The ideas presented here were developed by many researchers over many decades, so the citations are not exhaustive but rather direct the reader to a handful of papers that are, in the author's view, seminal. After reading this document, you should have a general understanding of word vectors (also known as word embeddings): why they exist, what problems they solve, where they come from, how they have changed over time, and what some of the open questions about them are. Readers already familiar with word vectors are advised to skip to Section 5 for the discussion of the most recent advance, contextual word vectors.

Subjects: **Computation and Language (cs.CL)**Cite as: [arXiv:1902.06006](#) [cs.CL](or [arXiv:1902.06006v3](#) [cs.CL] for this version)

<https://doi.org/10.48550/arXiv.1902.06006> Focus to learn more

arXiv-issued DOI via DataCite

Submission history

From: Noah Smith [[view email](#)][\[v1\]](#) Fri, 15 Feb 2019 23:28:36 UTC (14 KB)[\[v2\]](#) Tue, 19 Feb 2019 05:25:19 UTC (14 KB)[\[v3\]](#) Fri, 17 Apr 2020 17:16:08 UTC (78 KB)☐ Bibliographic Tools

Bibliographic and Citation Tools

☐ Bibliographic Explorer ToggleBibliographic Explorer ([What is the Explorer?](#))☐ Litmaps ToggleLitmaps ([What is Litmaps?](#))☐ scite.ai Togglescite Smart Citations ([What are Smart Citations?](#))☒ Code, Data, Media

Code, Data and Media Associated with this Article

☐ Links to Code ToggleCatalyzeX Code Finder for Papers ([What is CatalyzeX?](#))☐ DagsHub ToggleDagsHub ([What is DagsHub?](#))☐ Links to Code TogglePapers with Code ([What is Papers with Code?](#))☐ ScienceCast ToggleScienceCast ([What is ScienceCast?](#))☐ Demos

Demos

☐ Replicate ToggleReplicate ([What is Replicate?](#))☐ Spaces ToggleHugging Face Spaces ([What is Spaces?](#))☐ Related Papers

Recommenders and Search Tools

☐ Link to Influence FlowerInfluence Flower ([What are Influence Flowers?](#))☐ Connected Papers ToggleConnected Papers ([What is Connected Papers?](#))

☐ Core recommender toggle

CORE Recommender ([What is CORE?](#))

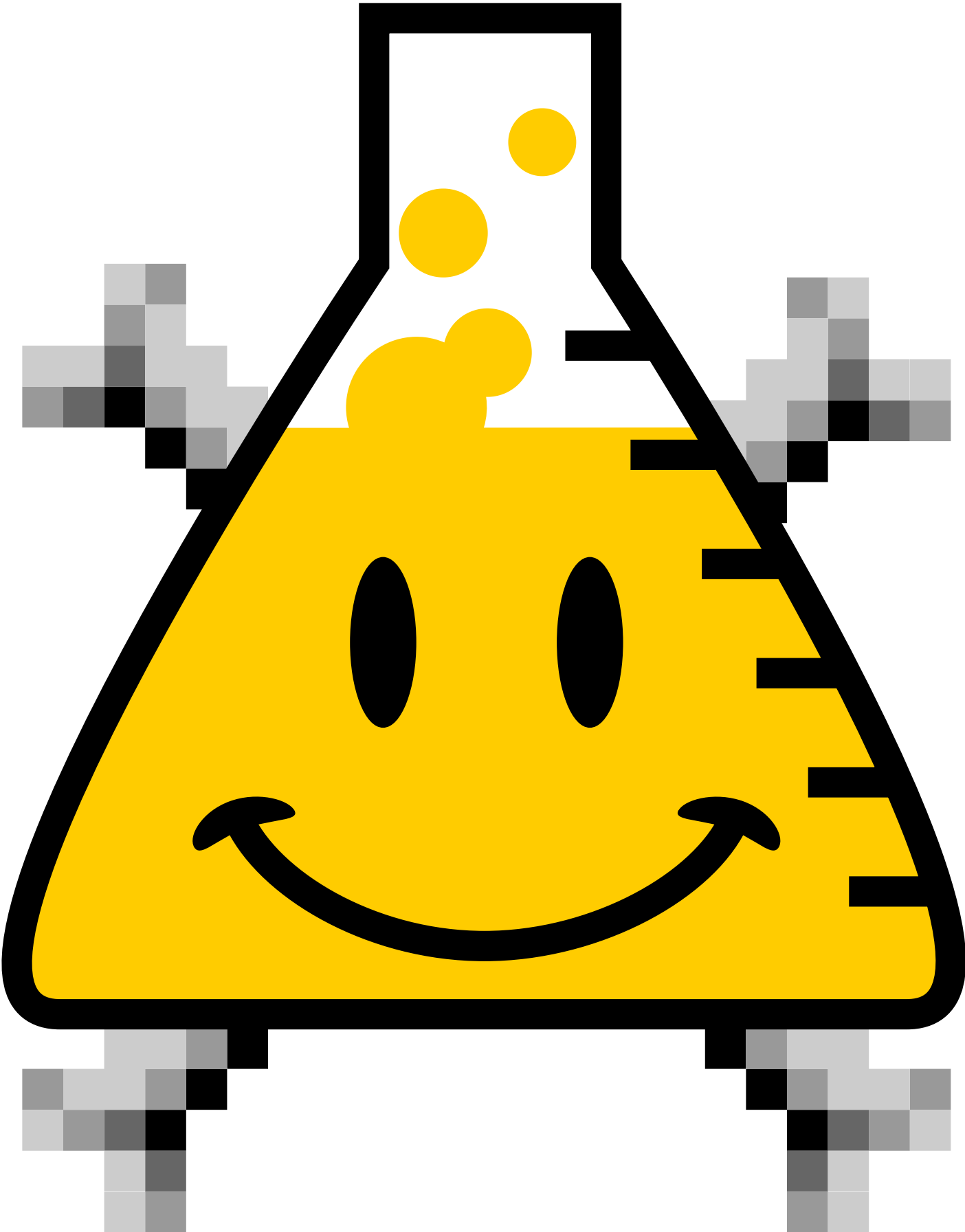
☐ About arXivLabs

arXivLabs: experimental projects with community collaborators

arXivLabs is a framework that allows collaborators to develop and share new arXiv features directly on our website.

Both individuals and organizations that work with arXivLabs have embraced and accepted our values of openness, community, excellence, and user data privacy. arXiv is committed to these values and only works with partners that adhere to them.

Have an idea for a project that will add value for arXiv's community? [Learn more about arXivLabs](#).



[Which authors of this paper are endorsers?](#) | [Disable MathJax](#) ([What is MathJax?](#))