Universities as drivers of AI research and innovation

Lav R. Varshney
Assistant Professor, University of Illinois at Urbana-Champaign
Principal Research Scientist, Salesforce Research
Chief Scientist, Ensaras, Inc.

@lrvarshney

December 3, 2019

salesforce

ENSARAS

Intelligence for Wastewater

ILLINOIS

National Science Foundation
• Foundations of Belief Sharing in Human-Machine Systems
• Towards an Information Theory of Attention
• Food and Data: Interoperability through the Food Pipeline
• Matching Non-Native Transcribers to the Distinctive Features of the Language Transcribed

DARPA and MARCO
• Systems on Nanoscale Information fabriCs (SONIC) Center
• Energy Efficient Image Storage and Compression with RRAM Arrays

IBM
• Center for Cognitive Computing Systems Research
• Toward Cognitive Systems for Scientific Discovery

Chan Zuckerberg Initiative
• Compression of Structural, Cartographic, and Multimodal Cell Data

FACE Foundation (Thomas Jefferson Fund)
• Reliable Artificial Intelligence on Energy-Efficient Hardware

Air Force Research Laboratory
• UMIMMI: Universal Multivariate Information Measures for Multisensor Inference [Phase I]
• UMIMMI: Universal Multivariate Information Measures for Multisensor Inference [Phase II]

USDA National Institute of Food and Agriculture
• Comparative Connectome of the Soybean Cyst Nematode and Establishment of an Online Anatomical Atlas

Siebel Energy Institute
• Incentives, Choices, and Analytics for Electric Vehicle Fleets in Jointly Managing Urban Traffic and the Smart Grid

Illinois OTM
• Machine Learning for Nanopore Bio-Detection

Los Alamos National Laboratory
• Computational Creativity for Materials

Illinois Learning Science Design Initiative
• Investigating the Neural Correlates of Learning in Cognitive Training

Living Analytics Research Centre (Singapore Management Univ.)
• Crowdsourcing, Choice Modelling, and Prediction for Healthy Eating

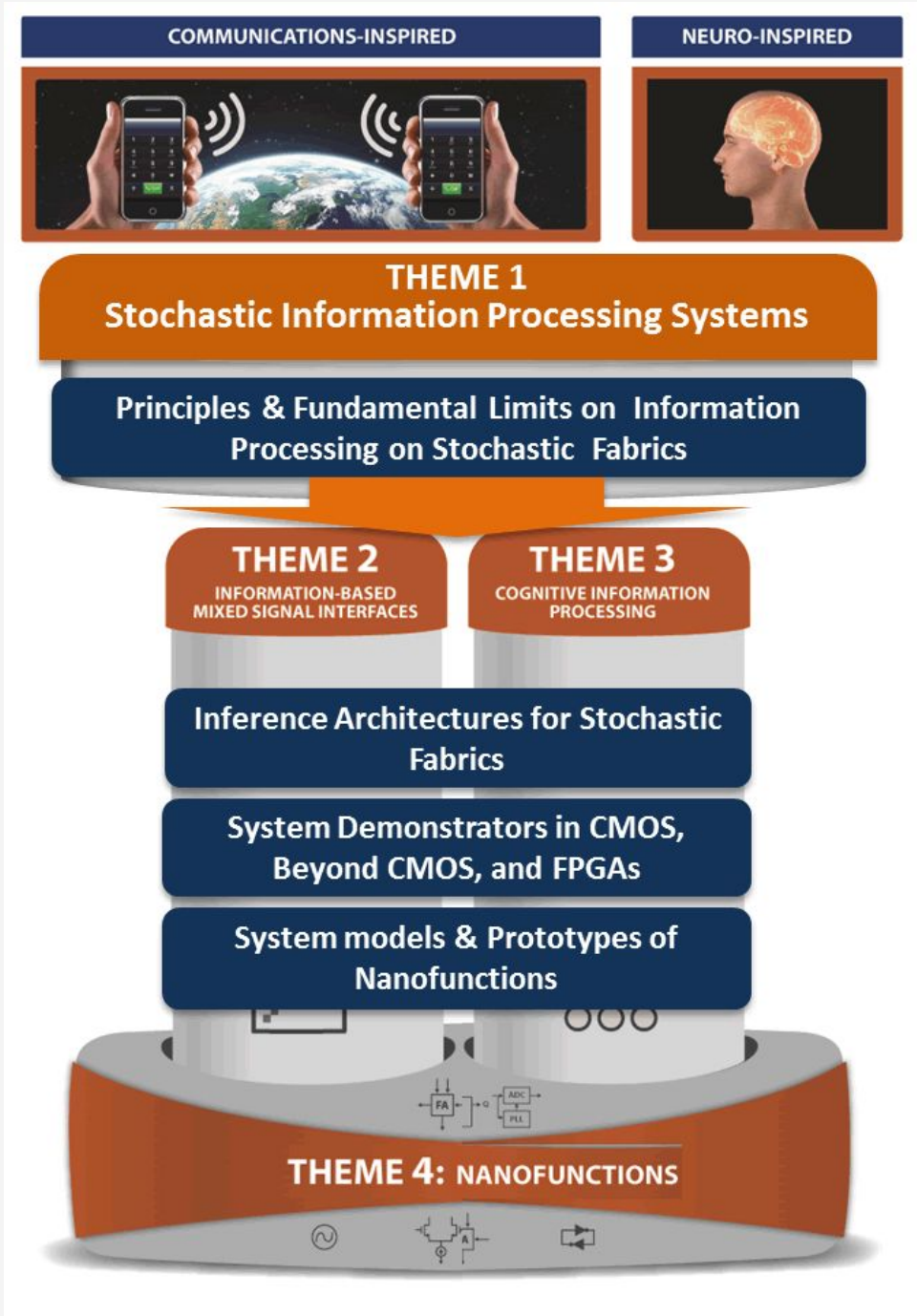
Zhejiang University
• Universal Compression of Deep Networks at the Information-Theoretic Limit

Center for Digital Agriculture
• Virtual Farming Networks for Smallholders through Digital Communities of Trust

Facebook
• AI Tools for Localized Concrete Formulation

SONIC

Members	
Analog Devices, Inc.	Arm Limited
EMD Performance Materials (a Merck KGaA affiliate)	IBM Corporation
Intel Corporation	Lockheed Martin Corporation
Micron Technology, Inc.	Northrop Grumman Corporation
Raytheon Company	Samsung Electronics Co., Ltd.
SK hynix Inc.	Taiwan Semiconductor Manufacturing Company Limited



C3SR

center for cognitive computing systems research

Blockchain

Comm|Scope

CUDA- and NUMA-Aware Multi-CPU / Multi-GPU communication benchmarks for C3SR Scope.

Near Memory Acceleration

MLModelScope (CarML)

A hardware/software agnostic, extensible and customizable platform for evaluating and profiling ML models across datasets/frameworks/hardware, and within AI application pipelines.

C3SR System for Review Assignment

Creative Experiential Learning Advisor

DISCvR

TensorCore

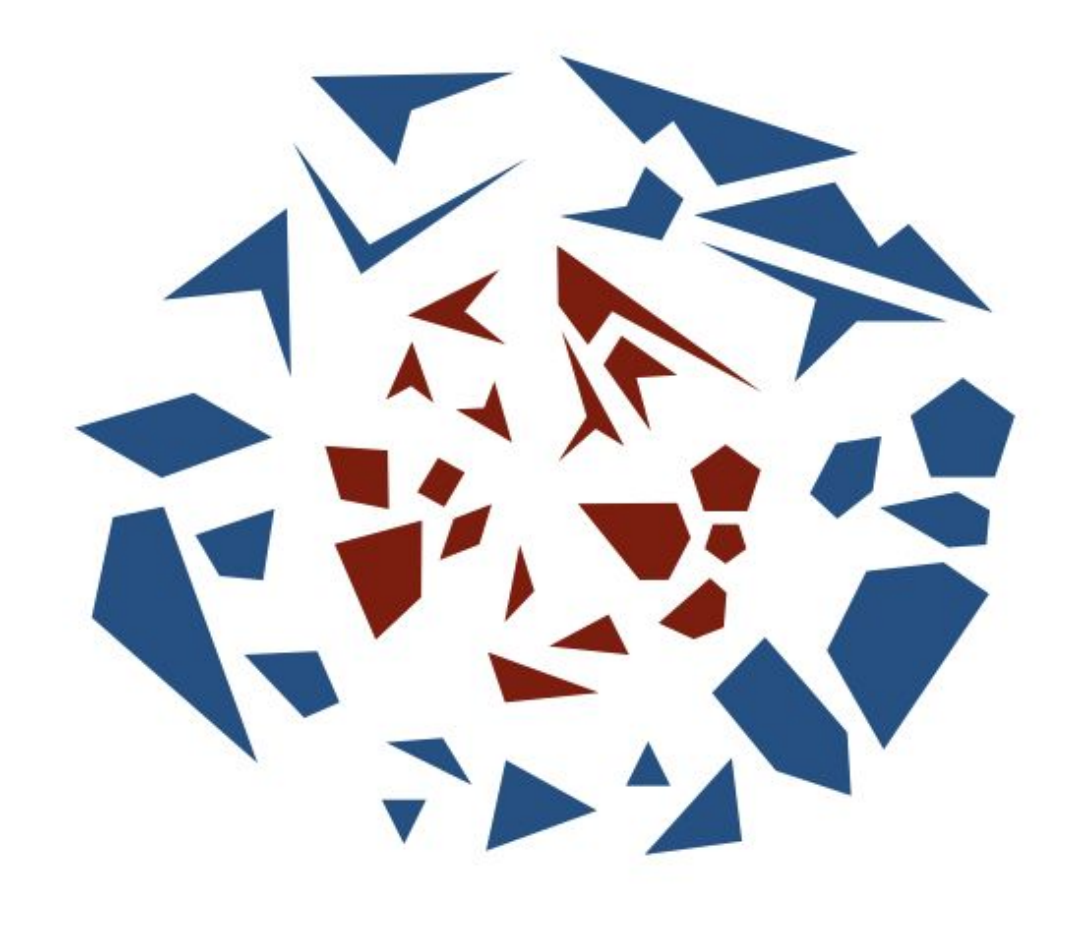
Leverage the TensorCore in modern GPUs and DNN accelerators to implement a high performance reduction and scan primitives.

Cognitive Application Builder

Cognitive Application Builder

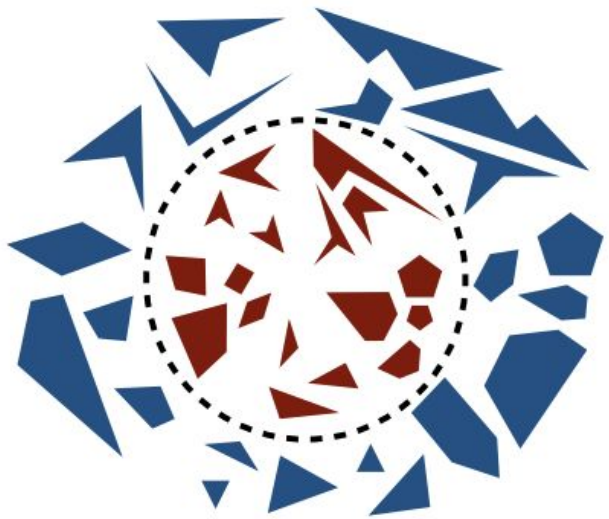
salesforce

A Group-Theoretic Approach to Computational Abstraction

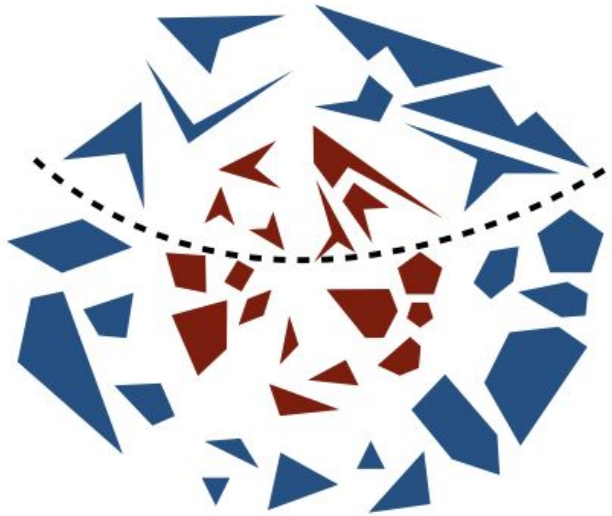


[H. Yu, I. Mineyev, and L. R. Varshney, “A Group-Theoretic Approach to Abstraction: Hierarchical, Interpretable, and Task-Free Clustering,” arXiv:1807.11167 [cs.LG].]

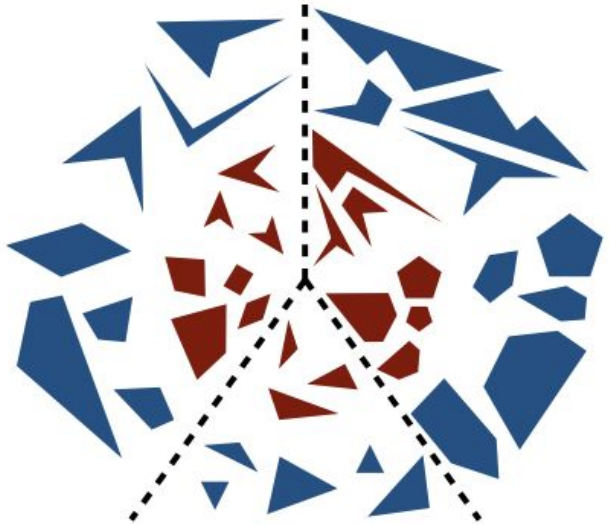
A Group-Theoretic Approach to Computational Abstraction



{red, blue}



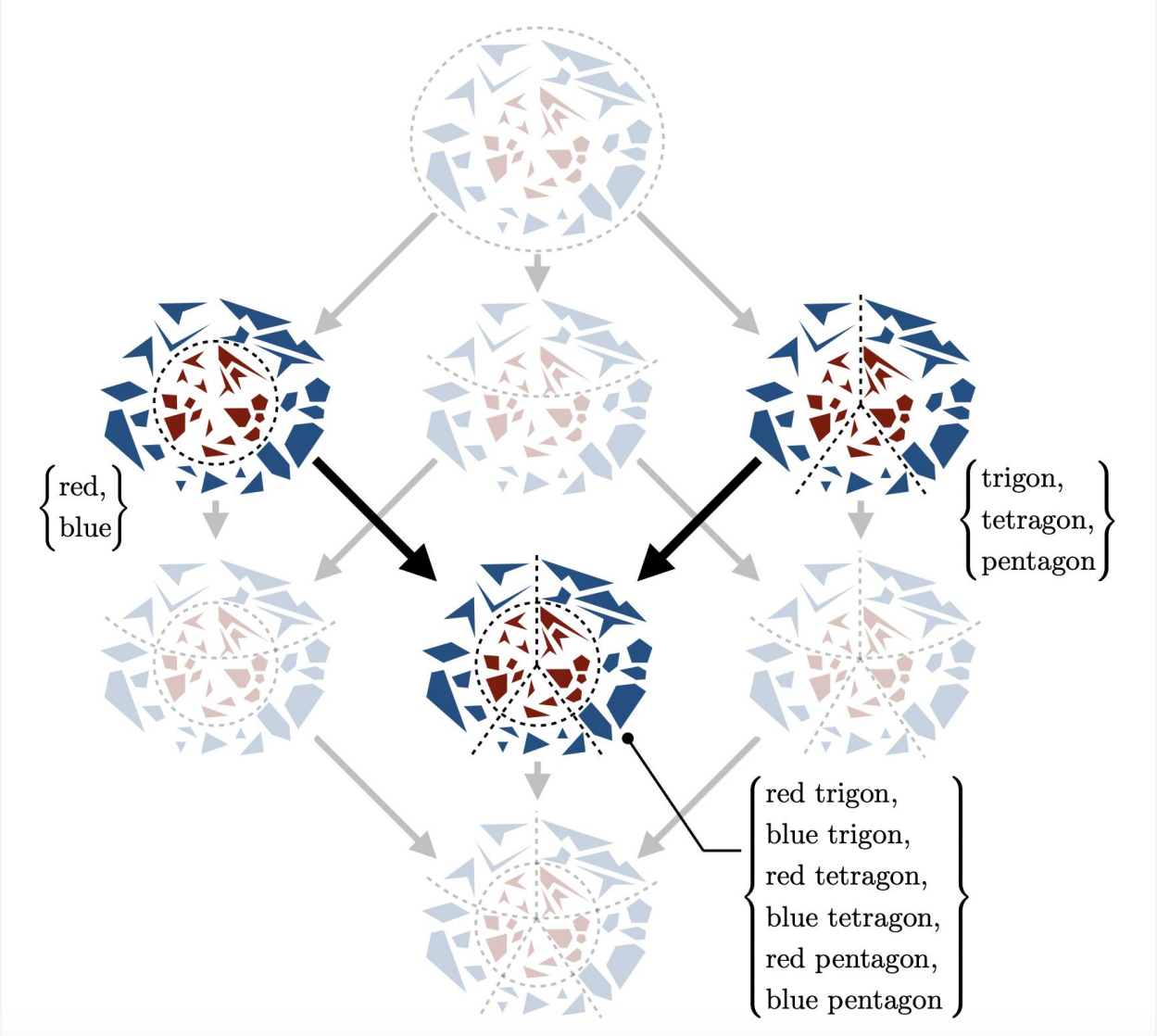
{convex, concave}



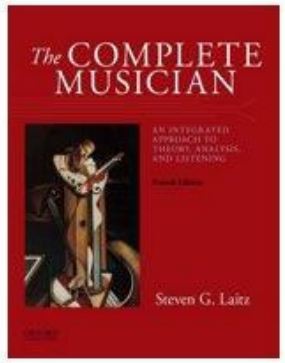
{trigon, tetragon, pentagon}

[H. Yu, I. Mineyev, and L. R. Varshney, “A Group-Theoretic Approach to Abstraction: Hierarchical, Interpretable, and Task-Free Clustering,” arXiv:1807.11167 [cs.LG].]

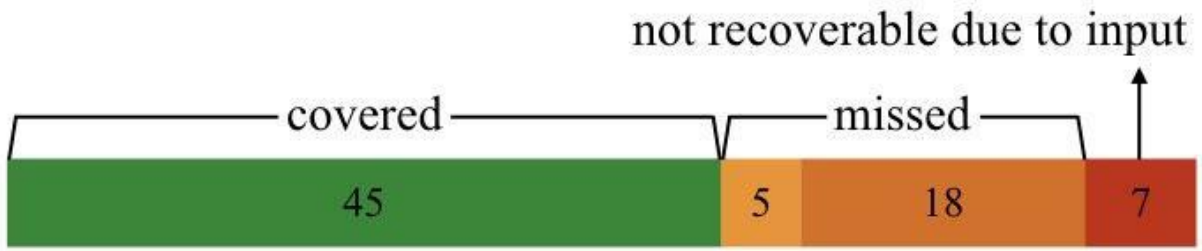
A Group-Theoretic Approach to Computational Abstraction



MUS-ROVER recovers most known music theory



MUS 101, 102, 201 (75 topics in total):



[H. Yu and L. R. Varshney, "Towards Deep Interpretability (MUS-ROVER II): Learning Hierarchical Representations of Tonal Music," in *Proc. 5th International Conference on Learning Representations (ICLR)*, April 2017.]

MUS-ROVER discovers new music theory

Interesting probabilistic pattern

Unresolved tritone (TT):

TT ⟶ m7

"harmonic" escape tone or changing tone?

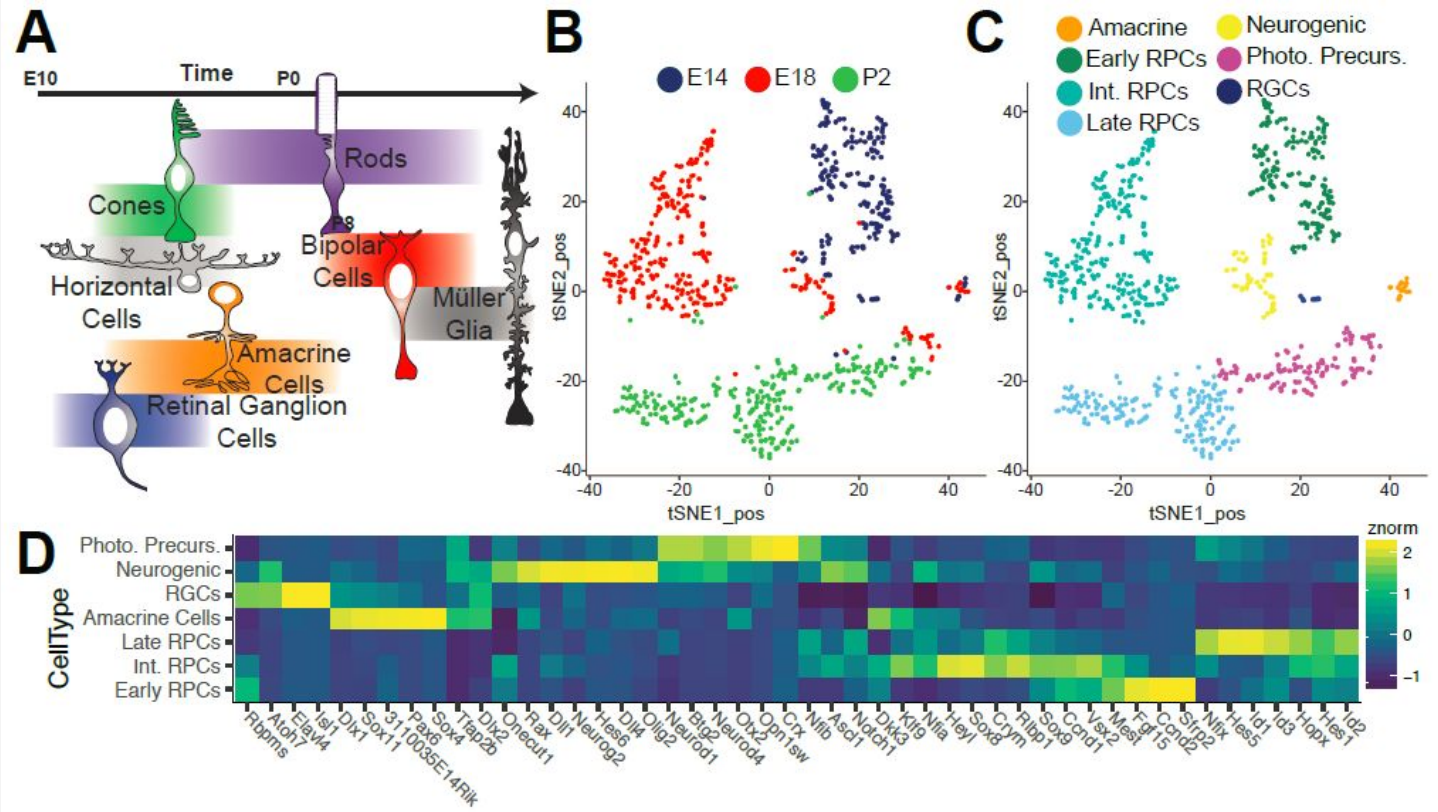
Interesting abstraction

Rule Trace

1	$\texttt{order} \circ w_{\{1,2,3,4\}}$
2	$\texttt{order} \circ \texttt{diff} \circ \texttt{sort} \circ w_{\{1,2,4\}}$
3	$\texttt{order} \circ \texttt{diff} \circ \texttt{mod}_{12} \circ w_{\{1,2,3\}}$
4	$\texttt{order} \circ \texttt{diff} \circ \texttt{diff} \circ w_{\{1,2,3,4\}}$
5	$\texttt{order} \circ \texttt{sort} \circ \texttt{mod}_{12} \circ w_{\{2,3,4\}}$
6	$\texttt{order} \circ \texttt{sort} \circ \texttt{mod}_{12} \circ w_{\{1,3,4\}}$
7	$\texttt{order} \circ \texttt{sort} \circ \texttt{mod}_{12} \circ w_{\{1,2,3,4\}}$
8	$\texttt{mod}_{12} \circ w_{\{1\}}$
9	$\texttt{mod}_{12} \circ \texttt{diff} \circ w_{\{2,3\}}$
10	$\texttt{mod}_{12} \circ \texttt{diff} \circ w_{\{3,4\}}$

[H. Yu and L. R. Varshney, "Towards Deep Interpretability (MUS-ROVER II): Learning Hierarchical Representations of Tonal Music," in *Proc. 5th International Conference on Learning Representations (ICLR)*, April 2017.]

Generalizing to other topical domains



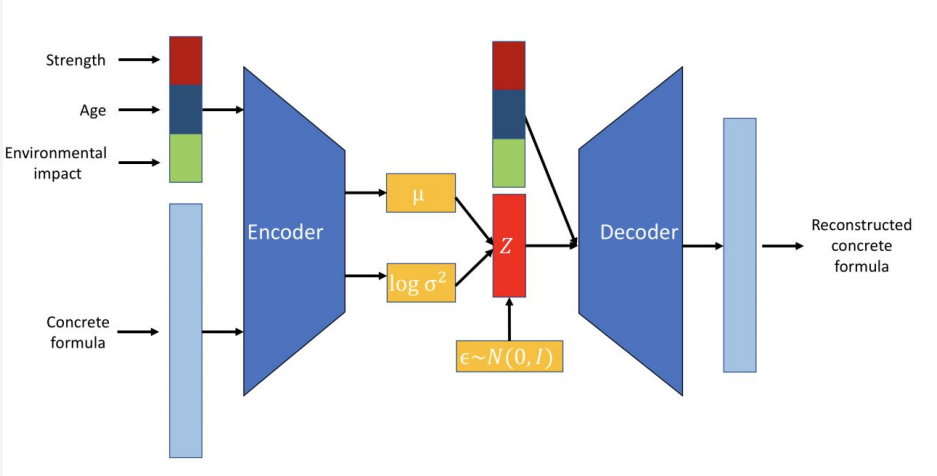
[B. Clark, et al., "Comprehensive analysis of retinal development at single cell resolution identifies NFI factors as essential for mitotic exit and specification of late-born cells," bioRxiv, 2018.]

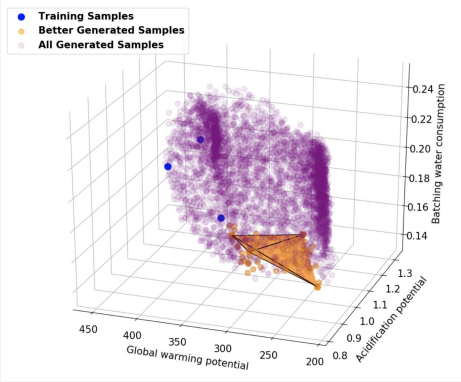
Single-cell RNA sequence data analysis for understanding the rules that govern pattern formation in neurodevelopment

[H. Yu, L. R. Varshney, and G. Stein-O'Brien, "Towards Learning Human-Interpretable Laws of Neurogenesis from Single-Cell RNA-Seq Data via Information Lattices," to be presented at *Learning Meaningful Representations of Life Workshop at NeurIPS 2019*, December 2019.]

Sustainable Building Materials

- Concrete is the most widely used engineered material in the world with more than 10 billion tons produced annually
- Significant burden in terms of energy, water, and release of greenhouse gases and other pollutants
- We develop generative models to create novel concrete formulations that minimize environmental burden, while satisfying engineering performance requirements
  - Low carbon footprint concrete impacts Indicator 9.4.1 of SDG





New formulation: 44.86 MPa at 7 days, 56.60 MPa at 14 days (as experimentally tested)

- 51.45% reduction in global warming potential
- 33.50% reduction in acidification potential
- 6.29% reduction in batch water consumption

[X. Ge, R. T. Goodwin, J. R. Gregory, R. E. Kirchain, J. Maria, and L. R. Varshney," Accelerated Discovery of Sustainable Building Materials," in *Proceedings of the AAAI Spring Symposium on Towards AI for Collaborative Open Science*, Palo Alto, California, 25-27 March 2019.]

Nina Kshetry, President

Lav Varshney, Chief Scientist

water
ENVIRONMENT RESEARCH

Water Environment Federation
the water quality people

RESEARCH ARTICLE

Prediction of odor complaints at a large composite reservoir in a highly urbanized area: A machine learning approach

John Mulrow, Nina Kshetry, Dominic A. Brose, Kuldip Kumar, Darshan Jain, Mohil Shah, Thomas E. Kunetz, Lav R. Varshney

First published: 06 August 2019 | https://doi.org/10.1002/wer.1191

Salesforce Customer 360

Trusted · Smart · Flexible · Sustainable

Platform

Mobile · Blockchain · AI · Vision · Voice

Security · Privacy & Identity · Communities · Builder · Flow

Salesforce

Customer 360 Analytics

World's #1 trusted analytics solution

Einstein: Analytics Embedded in CRM

Actionable insights

AI-powered analytics

Enterprise-scale data platform

Datorama: Marketing Intelligence

Centralize your data

Get deeper insights more quickly from your data

Tableau: Enterprise-Wide Analytics for All Your Data

An integrated platform with powerful analytics

Flexibility to connect to all kinds of data on prem or in the cloud

Provides strong security and governance model

Creating a Better World, Together

SUSTAINABLE DEVELOPMENT GOALS

Signatory to the United Nations Global Compact

Business is the Greatest Platform for Change

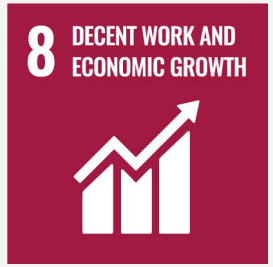
$90M donated for public education

100% renewable energy by FY2022

LEED Platinum offices

Supply chain engagement

$10M+ for equal pay

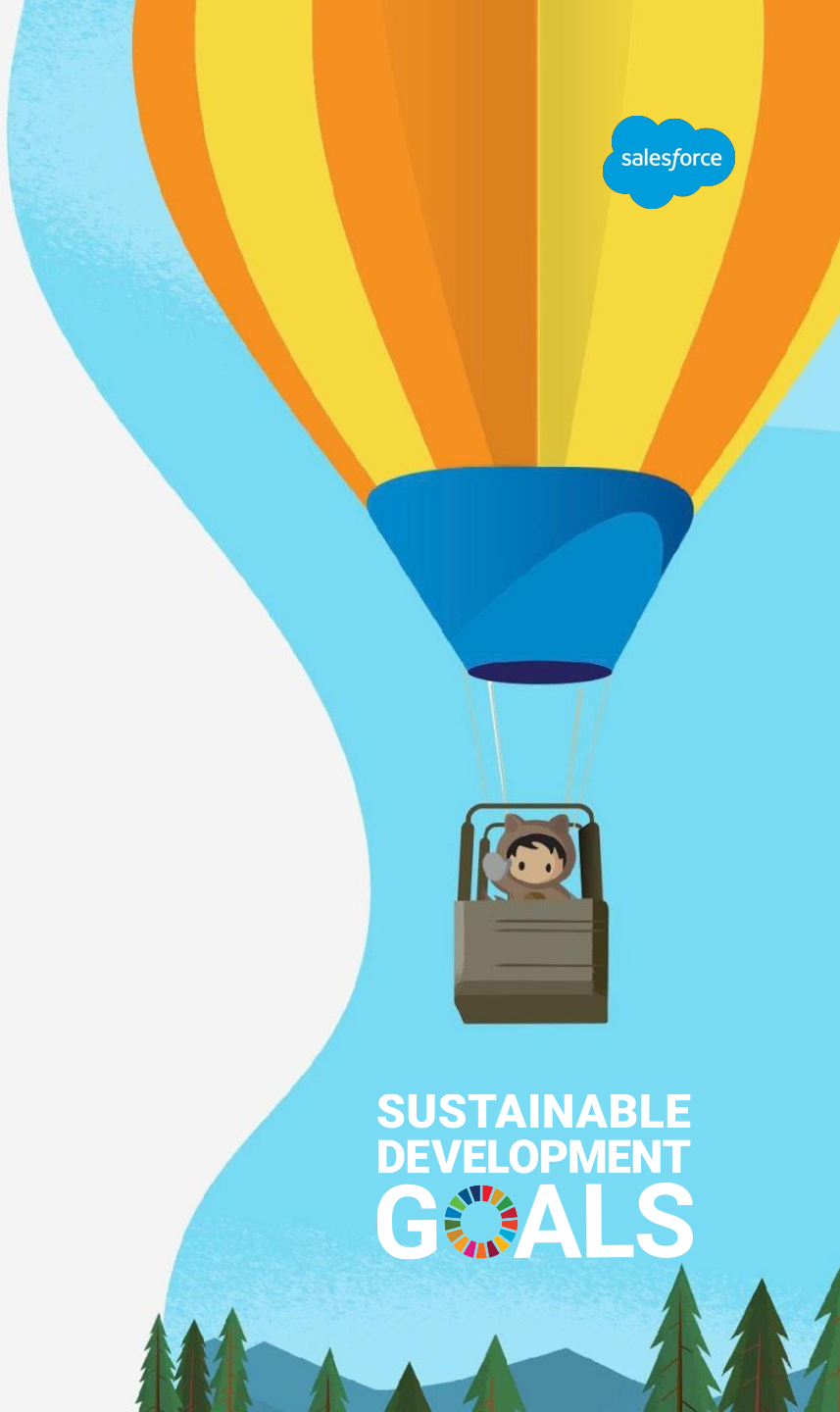


Pledge to America's Workers 1M jobs

Net-zero emissions

Carbon neutral cloud

Step Up Coalition



The Salesforce Economy

1 **Canada**
Jobs: 33K
Revenue: $20B

2 **United States**
Jobs: 869K
Revenue: $514B

3 **Mexico**
Jobs: 223K
Revenue: $38B

4 **Brazil**
Jobs: 491K
Revenue: $62B

5 **Spain**
Jobs: 49K
Revenue: $17B

6 **United Kingdom**
Jobs: 144K
Revenue: $71B

7 **Netherlands**
Jobs: 35K
Revenue: $23B

8 **Germany**
Jobs: 62K
Revenue: $30B

9 **France**
Jobs: 154K
Revenue: $79B

10 **Australia**
Jobs: 42K
Revenue: $26B

11 **Japan**
Jobs: 200K
Revenue: $110B

12 **India**
Jobs: 548K
Revenue: $67B

13 **Singapore**
Jobs: 9K
Revenue: $6B

Global Jobs

4.2M

New Salesforce economy jobs by 2025

Global Economic Impact

$1.2T

New business revenue by 2025

Source: IDC white paper sponsored by Salesforce, The Salesforce Economic Impact: 4.2 Million New Jobs, $1.2 Trillion of New Business Revenues from 2019 to 2024, October 2019. The statements are based on the data from 2019 through 2024.

Pretrained AI Models: Performativity, Mobility, and Change

Lav R. Varshney
Salesforce Research
Palo Alto, CA, USA
lvarshney@salesforce.com

Nitish Shirish Keskar
Salesforce Research
Palo Alto, CA, USA
nkeskar@salesforce.com

Richard Socher
Salesforce Research
Palo Alto, CA, USA
rsocher@salesforce.com

CTRL: A Conditional Transformer Language Model for Controllable Generation

Nitish Shirish Keskar[*], **Bryan McCann**[*], **Lav R. Varshney**, **Caiming Xiong**, **Richard Socher**
Salesforce Research[†]

Summary

Universities as key drivers of AI research and innovation

- Transfers of people, knowledge, entrepreneurship
- Functionally new applications in numerous domains, including fine-tuning of large-scale models
- New approaches to AI that are more robust (generalization, safety), resource-efficient (data, energy, volume, etc.), and bound by holistic view of ethics

Successful university AI labs

- Able to drive the development of new algorithmic approaches, fundamental limit theorems, and novel applications (often outside the core interest of large computing companies and startups)
- Examples throughout the world, including in places like Urbana-Champaign, Illinois that are not as close to industrial corridors like Silicon Valley

Academia-industry partnerships in AI

- Funded and joint research, leaves/sabbaticals/dual appointments, internships


THANK YOU
salesforce
Trail Blazer
EXPLORER
ADVENTURER
HIKER