Cal Poly (DB/DS) after COVID

Alex Dekhtyar
Computer Science and Software Engineering
Cal Poly, San Luis Obispo
Faculty (Databases)

Alex Dekhtyar
- Databases
- Data Science
- Bioinformatics
- Req. Engineering

Lubomir Stanchev
- Databases
- Data Science
- Distributed Systems
- Knowledge management
Faculty (Databases+HPC)

Alex Dekhtyar
- Databases
- Data Science
- Bioinformatics
- Req. Engineering

Lubomir Stanchev
- Databases
- Data Science
- Distributed Systems
- Knowledge management

Maria Pantoja
- Parallel/GPU Comp.
- Distributed Systems
- HPC
- Deep Learning
Faculty (Databases+HPC+Data Science)

Alex Dekhtyar

Lubomir Stanchev

Foaad Khosmood

Jonathan Ventura

Paul Anderson

Maria Pantoja

Computer Vision

NLP

AI in Games

ML

Sumona Mukhopadhay

Mugizi Rwebangira

ML

Ranking in IR

Bioinformatics
Our News

❖ Noyce School of Applied Computing
❖ Quarters -> Semesters
❖ Central Coast Data Science Partnership
❖ Cross-Disciplinary Studies Minor In Data Science
❖ Faculty Projects
Cal Poly is pleased to announce the establishment of The Noyce School of Applied Computing, a new interdisciplinary school (the first of its kind at Cal Poly) combining three departments under one umbrella — Electrical Engineering, Computer Science and Software Engineering, and Computer Engineering — to create interdisciplinary collaboration opportunities for departments and faculty doing applied computing across the university in fields such as statistics.

**Computer Science and Software Engineering**

**Computer Engineering**  **Statistics**

**Electrical Engineering**

Chris Lupo Named Founding Director of The Noyce School of Applied Computing

College of Engineering Dean Amy S. Fleischer is pleased to announce Chris Lupo, Ph.D., has been named the founding director of The Noyce School of Applied Computing.
Quarters

3 + 1 DB courses

3 + 3 ML/NLP courses

6 Data Science courses

Data Engineering Concentration

Semesters

?
Central Coast Data Science Partnership

UCSB  Cal Poly

3-year NSF HDR grant

Data Science courses
● Capstone for UCSB
● Data Literacy for Cal Poly

Data Science Fellowship
Data Science Program Update

Cross-Disciplinary Studies Minor In Data Science

- 2016 - present (6th graduating class this year)
- 80 quarter units common for all students
- CS + Statistics + (now) Math
- Culminating experience (revised)
  - Data Science Process
  - Foundations of ML
  - Lab
  - Capstone (2 quarters)
- Converting to Semesters
Data Science Capstone

Classify species found in water samples in Florida estuaries /The Smithsonian/

Detect marine animal "songs" in underwater audio recordings /Marine Science/

Automated face blurring in videos of human physical activity /Cal Poly Kinesiology-Sentimetrux/ (see poster!)

Network analysis of Illicit Massage Businesses in Colorado to detect human trafficking /Global Emancipation Network/

Predicting market success of startups from OSINT data /OCOB - Chase/
Data Science Program Update

B.S. In Data Science

- MOU between Math, Statistics and CS
- 13 faculty
- Semester BS in Data Science Degree proposal
  - working to put in place in Fall 2026
Faculty Projects
Digital Democracy Resurrected

Foaad Khosmood
Alex Dekhtyar
Digital Democracy + AI4Reporters

- Original digital democracy is rebooted (single comprehensive normalized DB for all legislative data)
- But instead of end-user (citizen) experience, a new AI system designed to help reporters cover the state legislature is the focus (See Poster)
  - News organization partner (CalMatters) is helping to fund and develop
Lubomir Stanchev

- IEEE International Conference on Semantic Computing Paper on extracting genealogical relations from natural language text using BERT models. Accuracy of above 70%.
- Working a paper on semantic document clustering using BERT models+ structured data (e.g, WordNet, DBPedia).
- Presented SIGCSE poster on data science seminar covering history of data science.

Adam has a brother, Jeffrey. They grew up in California, raised by their married parents Eva and James.
Maria Pantoja

- Reliable Distributed Hyper Parameter Optimization for Deep Learning Training
- Adding Uncertainty to Reliable Distributed Deep Learning Model Training
Urban Tree Detector: Estimated 43.5 million trees in California’s urban forest
Jonathan Ventura + Alex Dekhtyar

Urban Forest Quality Assessment and Estimation

188 TREES