CS644: Big Data

Week 1

Introduction to the Course

History of Big Data

(presented on whiteboard)

Current State of Big Data

(presented on whiteboard)

Programming Paradigms

- **Procedural/Imperative Programming**
  Examples: C/C++, Java, Python, Fortran

- **Functional Programming**
  Examples: Lisp, ML, Haskell, Scala (can be), Idris

- **Declarative/Logic Programming**
  Examples: Datalog, Prolog, SQL

Why Functional Programming? Why Scala? Why Spark?

Slides courtesy of Heather Miller:

https://heather.miller.am/teaching/cs4240/spring2018/pdf/spark-1-1.pdf

Data-Parallel to Distributed Data-Parallel

https://heather.miller.am/teaching/cs4240/spring2018/pdf/spark-1-2.pdf

(slides courtesy of Heather Miller)

Latency

https://heather.miller.am/teaching/cs4240/spring2018/pdf/spark-1-3.pdf

(slides courtesy of Heather Miller)