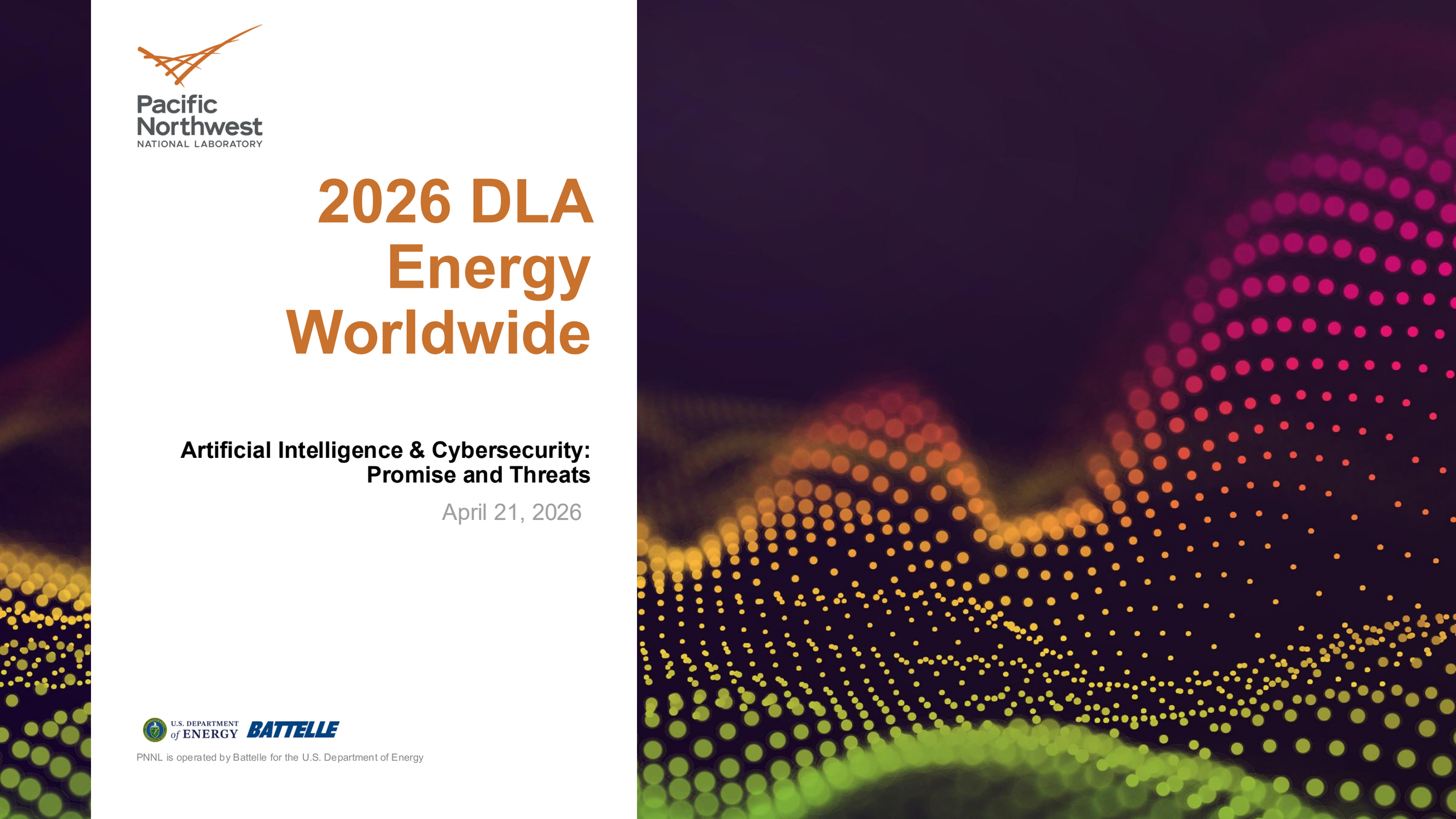
Pacific Northwest
NATIONAL LABORATORY

2026 DLA Energy Worldwide

Artificial Intelligence & Cybersecurity: Promise and Threats

April 21, 2026



U.S. DEPARTMENT of ENERGY BATTELLE

PNNL is operated by Battelle for the U.S. Department of Energy

Pacific Northwest
NATIONAL LABORATORY

Artificial Intelligence and Cybersecurity: Promise and Threats

- Heather Bevacqua, Pacific Northwest National Laboratory
- David Manz, Pacific Northwest National Laboratory
- William Hutton, Pacific Northwest National Laboratory
- Mark Hadley, Pacific Northwest National Laboratory

Implementing cybersecurity protections should be in place from the beginning of virtually all projects continue throughout their lifecycles. This session will address what those protections should look like at various stages and the consequences if they are not in place. AI is also an issue with major cyber security implications. our speakers will address the personnel, technological, and funding implications of AI for the military.

Pacific Northwest
NATIONAL LABORATORY

DOE's 17 National Laboratories

Pacific Northwest
NATIONAL LABORATORY

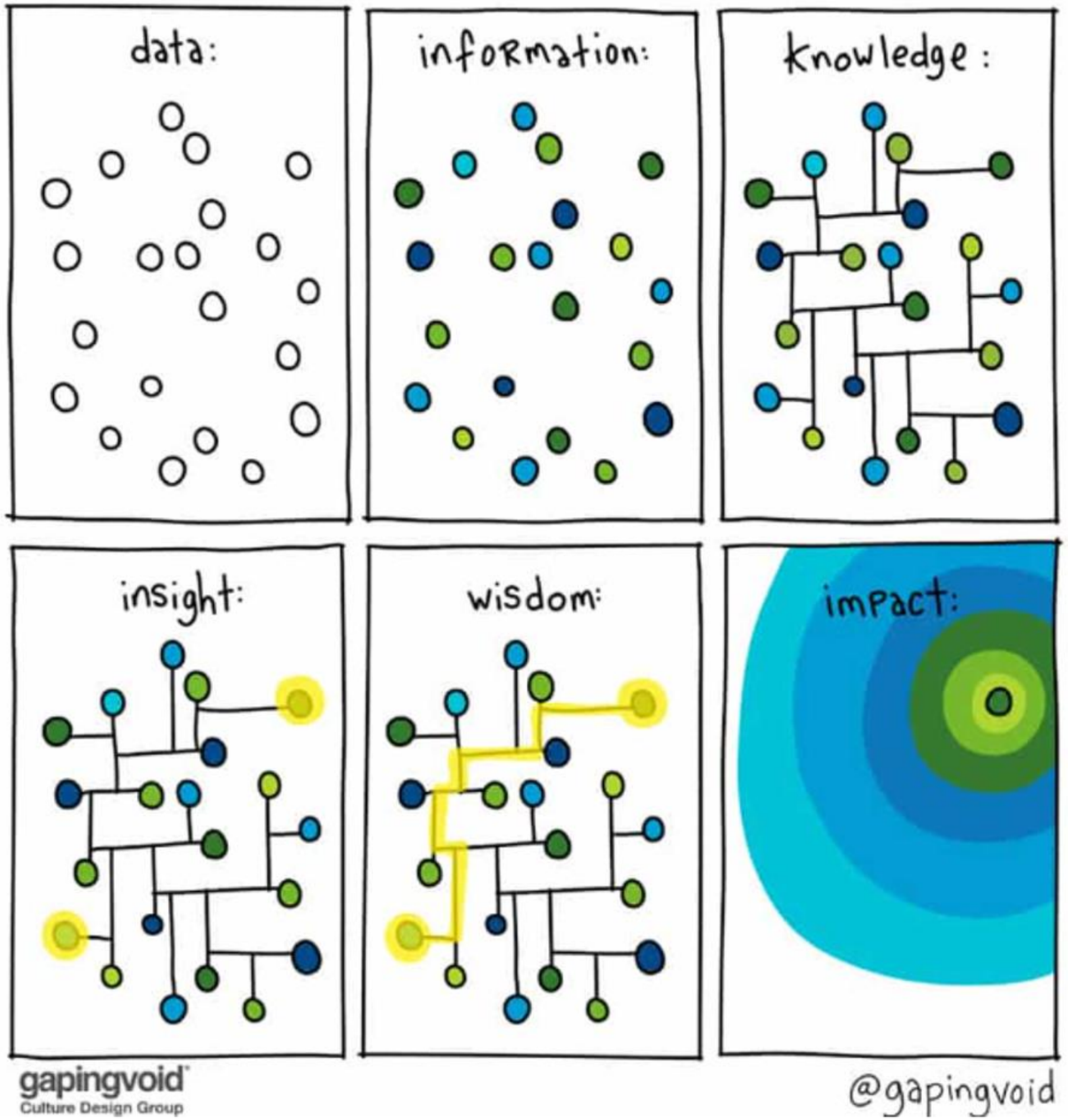
AI Winters

1. Machine translation (1960s)
2. LISP machine collapse (1980s)
3. LLMs?

Pacific Northwest
NATIONAL LABORATORY

What is an LLM, anyway?

- Just a graph
- So, what's a graph?
- Graphs with probability
- Ghosts of the first AI Winter

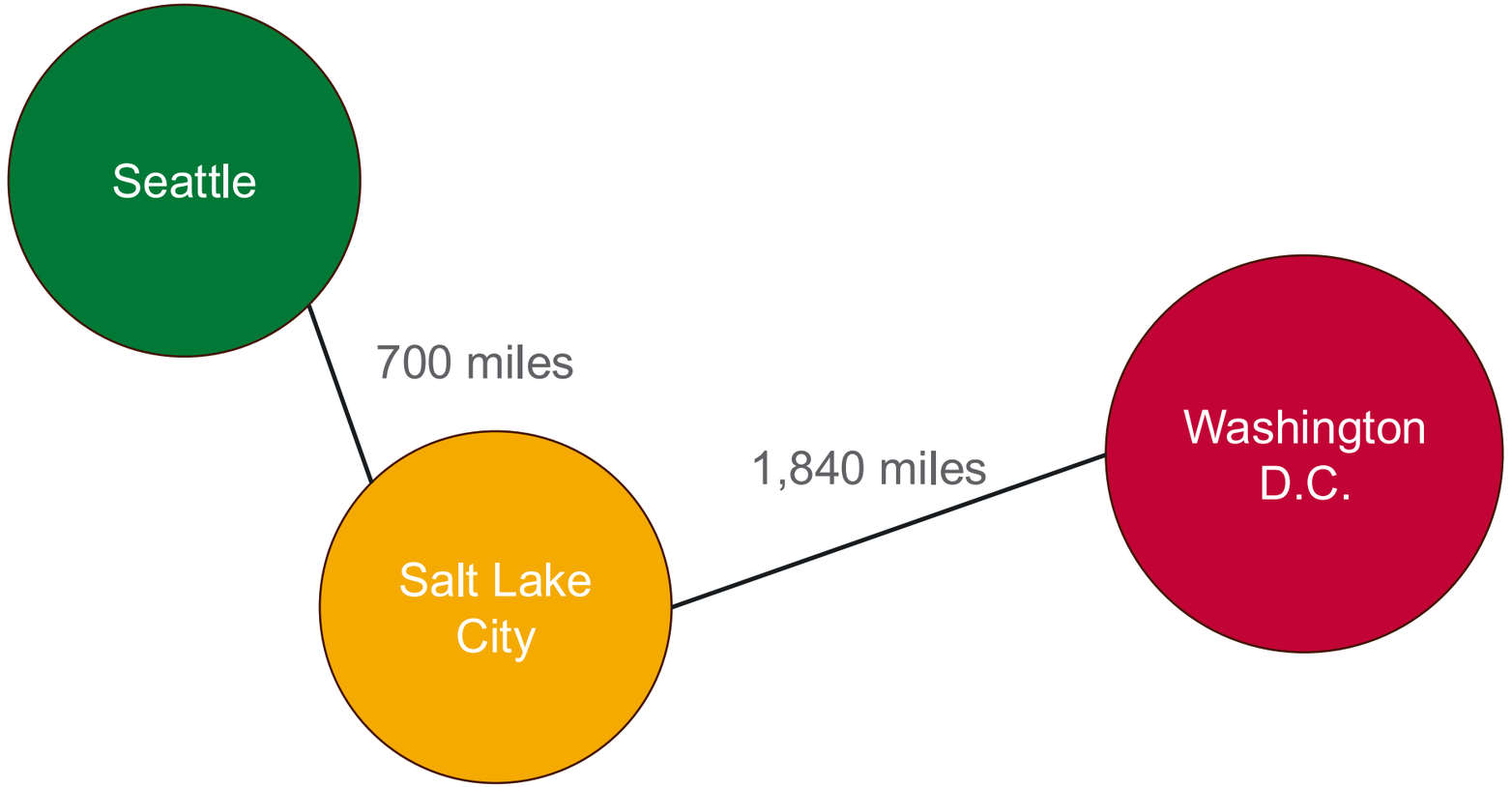


Pacific Northwest
NATIONAL LABORATORY

Graph Theory

- Graph *G*
- Set of *V* vertices (i.e. nodes) (circles)
- Set of *E* edges (lines)

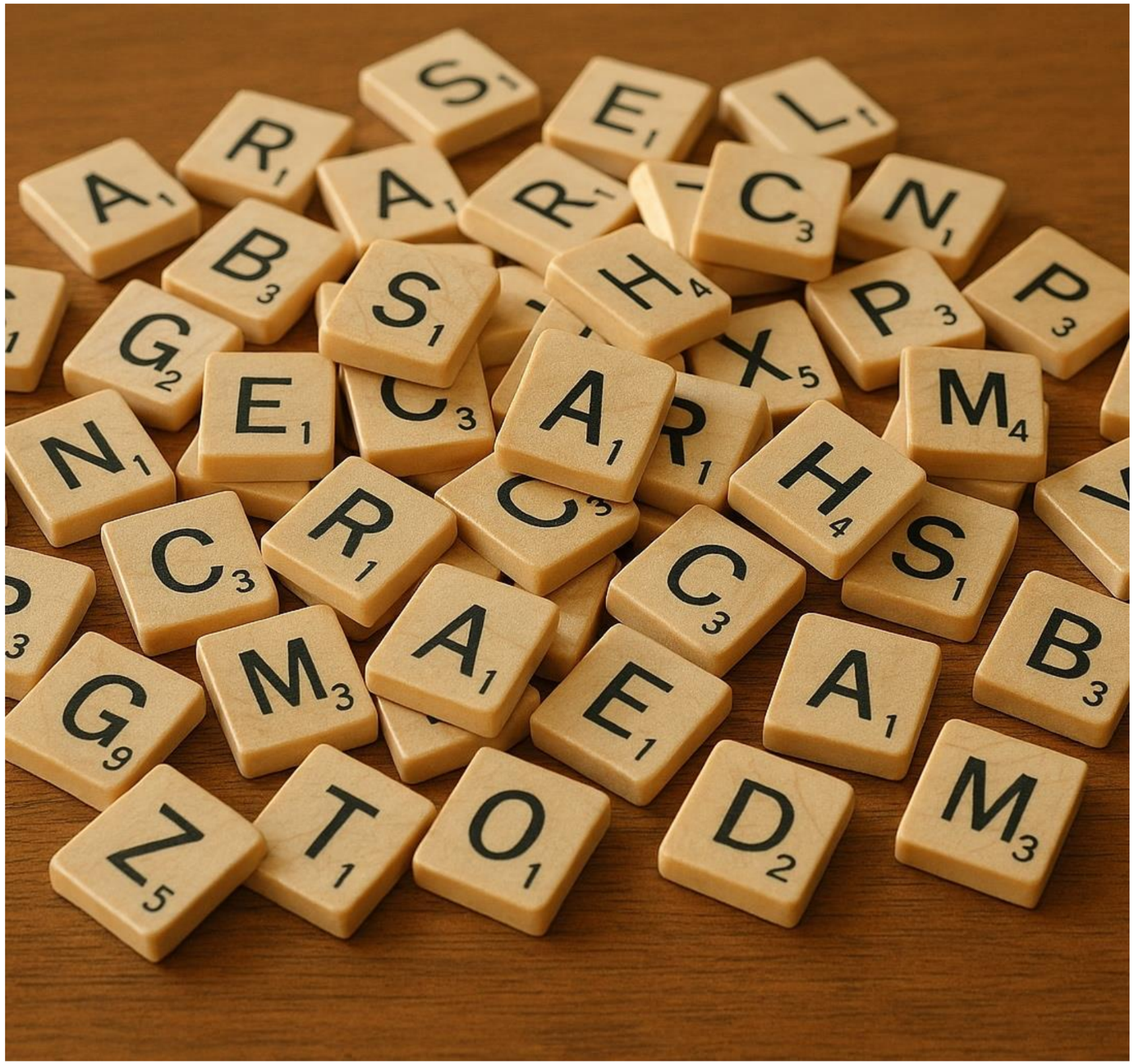
$G = \{V, E\}$



Pacific Northwest
NATIONAL LABORATORY

Claude Shannon's work on Information Theory and Entropy

- Zero-order approximation
- First-order approximations
- Second-order approximations
- Third-order approximations



Pacific Northwest
NATIONAL LABORATORY

Zero-order approximation

- All symbols (i.e. letters) are independent and equi-probable
- Too many Z's and W's
- Not enough E's or spaces

Pacific Northwest
NATIONAL LABORATORY

First-order approximation

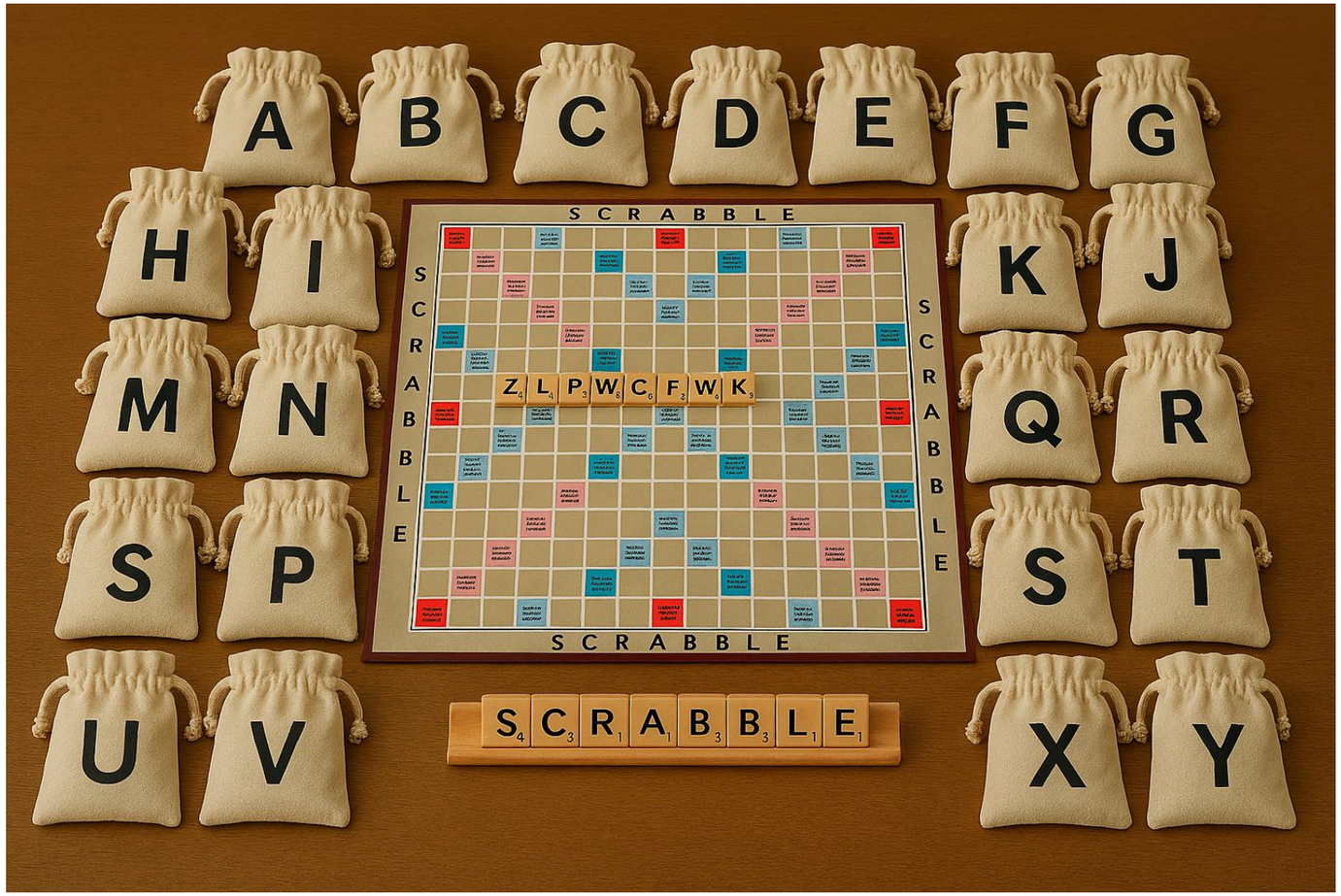
- Symbols (letters) are independent, but with frequencies of English
- More E's
- Few W's and Z's

TABLE 1. English letter frequency distribution

a 8%	h 6%	n 7%	t 9%
b 1.5%	i 6.5%	o 8%	u 3%
c 3%	j < 1%	p < 1%	v 1%
d 4%	K < 1%	q < 1%	w 1.5%
e 13%	l 3.5%	r 6.5%	x < 1%
f 2%	m 3%	s 6%	y 2%
g 1.5%			z < 1%

Second-order approximation

- Digram probability
- 9% of drawing a T from the regular Scrabble bag
- Then, draw our next tile from the “T” bag.
- 37% of “TH” given “T”

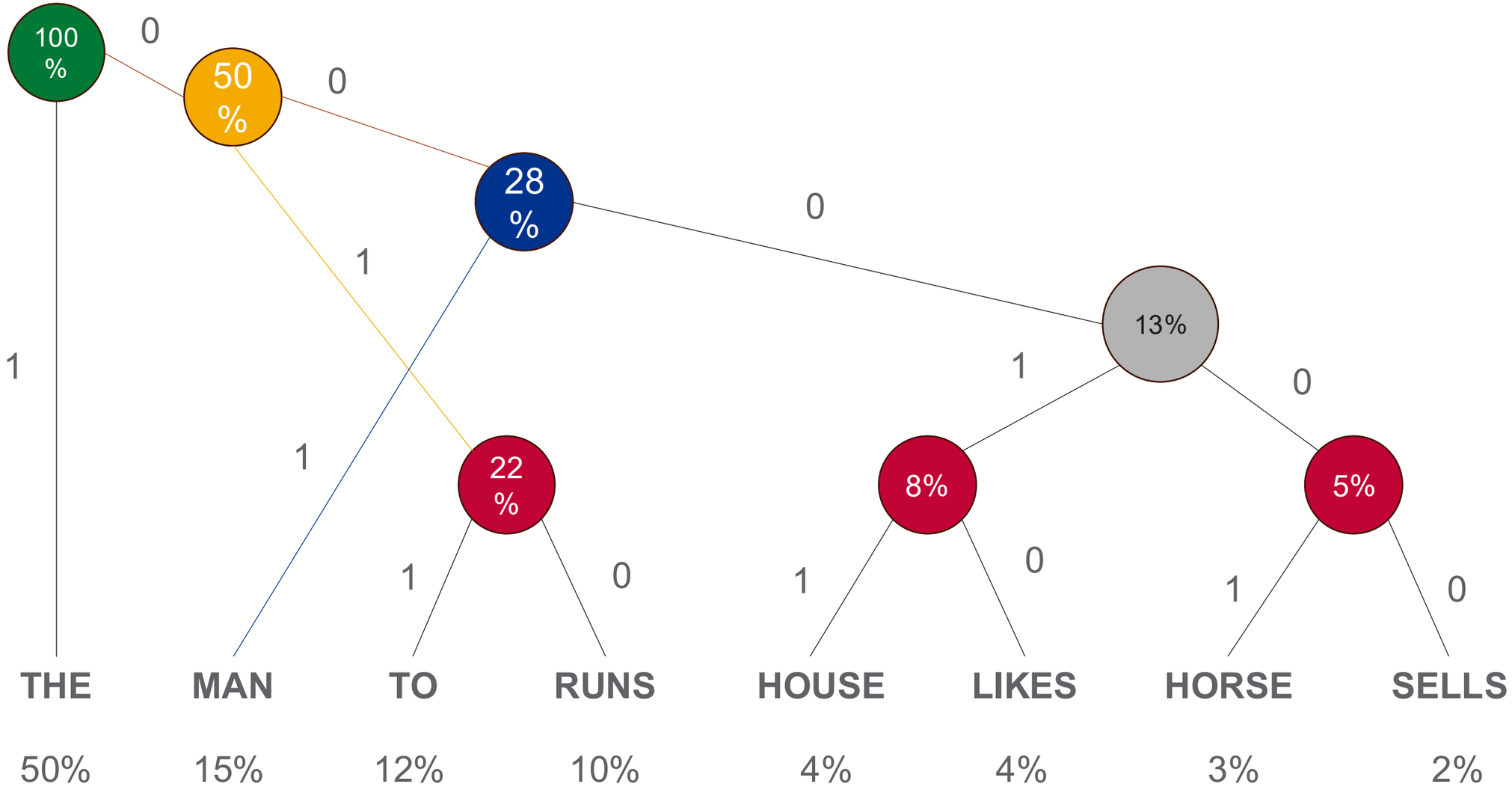


Third-order approximation

- Like a digram, but the third tile we draw is based on the probability of the last TWO tiles.
- Bears an increasing resemblance to English

IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PONDENOME OF DEMONSTRURES OF THE REPTAGIN IS RECOGACTIONA OF CRE

A “Small Language Model” (SLM)



Large Language Models (LLMs)

- Computational model (graph)
- Designed for *natural language processing* tasks
  - Language generation
  - Image creation
- Generate, summarize, translate, and reason over text
- Context is crucial!

Goals of Security & Trust

- Security Goals
  - Prevention
  - Detection
  - Recovery
- Policies
  (What is allowed or forbidden?)
- Controls
  (Enforce policies)
- **Assumptions about trust?**
- How can we trust unpredictable outcomes?

Study Finds 45% of AI Queries Produce Errors

- Based on a mathematical model that correlates statistical relationships of tokens (words) to other tokens.
- Dangerously confident
- Biased to give an answer (any answer!) over saying, “I don’t know.”

Pacific Northwest
NATIONAL LABORATORY

Thank you