FAST, RELIABLE, CATASTROPHICALLY FAILING?

SAFELY AVOIDING BEHAVIORAL INCIDENTS IN COMPLEX PRODUCTION ENVIRONMENTS
WHO?

- Software Engineer
- Working on Data Science teams as the engineer
- Exposed to “proper science”
- Put this model/data product into prod
WHAT ARE WE TALKING ABOUT

AND WHAT AREN’T WE TALKING ABOUT
COMPLEXITY
EMERGENT BEHAVIOR
UNKNOWN
DISCOVERED AT SCALE

SUBTLETY
DORMANT BEHAVIOR
FORENSIC INCIDENT DATA
DATA AS A NEW THREAT VECTOR
IS "ACTING WEIRD" AN INCIDENT?
LETS TALK ABOUT STUFF
THIS TALK IS NOT ABOUT
PREMISE

COMPLEXITY AND FAILURE
GO HAND IN HAND

RELIABILITY AND ROBUSTNESS
COME FROM DIRECT EXPERIENCE
WITH FAILURE
COMPLEXITY

SIDE EFFECT OF SUCCESS
COMPLEXITY

ESSENTIAL

ACCIDENTAL
COMPLEXITY
ACCIDENTAL
COMPLEXITY

ESSENTIAL
ACCEPT

SOLUTION TO COMPLEXITY IS NOT SIMPLICITY
ACCEPT

COMPLEXITY HAS TO BE EMBRACED AND MANAGED
FAILURE
HAZARDOUS LAYERED DEFENSES BUILT OVER TIME CATASTROPHE INVOLVES MULTIPLE FAILURES ERROR DETECTION IS HARD
COMPLEX SYSTEMS

NO ROOT CAUSE
OPERATORS HAVE DUAL ROLE
2 THREATS TO AVAILABILITY

THE SOFTWARE CHANGES THE ENVIRONMENT CHANGES

@rmn
THE ENVIRONMENT CHANGES

NETWORK LATENCY
RESOURCE CONTENTION / NOISY NEIGHBOR
DISK IS FULL
TIME IS WRONG
IN TERMS OF DUAL ROLE

THE SOFTWARE CHANGES

YOU ARE A PRODUCER
2 THREATS TO AVAILABILITY

THE ENVIRONMENT CHANGES

DEFENDER
IT IS DOING SOMETHING YOU DIDN’T ANTICIPATE
ALLOWS YOU TO BE A PARTICIPANT IN SYSTEM DURING AN INCIDENT

AAAAND WE’LL TAKE THIS AWAY LATER TOO 😎
MOSTLY WHAT AREN’T WE TALKING ABOUT
WHY TELL IS NOT THE STUFF RMIN?!?

MEANS OF DEALING WITH THIS ARE BECOMING WELL UNDERSTOOD

BECOMING A SOLVED PROBLEM
LETS REVISIT COMPLEXITY
THE AMAZON/NETFLIX KIND
FOREGO CORRECTNESS ADOPT SAFETY
RECOGNIZE HAZARDOUS SHARP EDGE
SHARP EDGE

SAFETY APPARATUS IS BUILT INTO SYSTEM

HUMANS AND CULTURE
3 CONCEPTS

TEST IN PROD
PROGRESSIVE DELIVERY
ERROR BUDGETS
I DON'T ALWAYS TEST MY CODE.

BUT WHEN I DO, I TEST IN PROD. STAY ON CALL MY FRIENDS.
TEST IN PROD DOESN'T MEAN RELEASE WITHOUT TESTING
TESTING IN PROD MEANS EXTENDING THE SOFTWARE DEVELOPMENT LIFECYCLE BEYOND RELEASE
“REAL USERS, REAL TRAFFIC, REAL SCALE, REAL UNPREDICTABILITIES”

@MIPSYTIPSY
PROGRESSIVE DELIVERY
“PROGRESSIVE DELIVERY IS CONTINUOUS DELIVERY WITH FINE-GRAINED CONTROL OVER THE BLAST RADIUS.”

James Governor, RedMonk (@monkchips)
SEPARATE DEPLOY FROM RELEASE

WHY?

THINGS ARE DEFINITELY GOING TO GO WRONG IN WAYS YOU DIDN’T ANTICIPATE. EMBRACE IT

@rmn
FEATURE FLAGS

TARGET SPECIFIC USERS FOR NEW “FEATURES”
ABILITY TO TOGGLE EXPOSURE ON/OFF
CANDY

EXPOSE SOME % OF LIVE TRAFFIC TO A NEW SERVICE
MONITOR KEY BUSINESS METRICS FOR THAT POPULATION
A/B TEST OUTCOME OF NEW DEPLOYMENT
WIDER RELEASE WHEN YOU ARE COMFORTABLE
ERROR
BUDGETS
ERROR BUDGETS
OPPORTUNITY FOR LEARNING
IT IS NOT YOUR JOB TO CREATE INFINITELY RELIABLE SOFTWARE

WHAT IF YOU COULD CREATE MORE BUSINESS VALUE BY LETTING SOFTWARE BE MORE BROKEN
MAYBE THE NATURAL DISTRIBUTION OF FAILURE HAS SPARED YOU
YOU MIGHT HAVE SOME 9S TO PLAY WITH
PERMIT AUDACITY
WHEN AUDACITY CAPITAL AVAILABLE
RECOGNIZE SHARP EDGE

VOCABULARY FOR MANAGING COMPLEXITY SAFELY

TEST IN PROD

PROGRESSIVE DELIVERY

ERROR BUDGETS

@rmn
EXPERIMENT

DELIBERATELY EXPLORE WEIRD BEHAVIOR (CHAOSENG)

TRY NEW THINGS INSIDE YOUR BUDGET

PERMIT AUDACITY WHEN AUDACITY CAPITAL AVAILABLE

ALLOW AN ACCIDENTAL “OVERAGE” OF SLA TO BE YOUR PLAYGROUND

YOU HAVE HEADROOM TO TAKE RISKY CHANGES
WHAT WE ARE TALKING ABOUT
SUBTLE

OBLIQUE
MODELS IN PRODUCTION

LATENT CATASTROPHOIC BEHAVIOR

WHAT ARE THE FAILURE COMPONENTS?

ARE THESE OUTAGES?
MODEL ARTIFACT IS COMPLEX

BUT NOT COMPLEX DEPENDENCIES

COMPLEX RESPONSE TO INPUT
DATA NEW VECTOR OF FAILURE
2. 3 THREATS TO AVAILABILITY?

THE SOFTWARE CHANGES THE ENVIRONMENT CHANGES THE DATA WAS UNANTICIPATED
THE DATA IS HAZARDOUS

DATA ISN'T A TRADITIONAL COMPONENT IN A COMPLEX SYSTEM
AN INCIDENT
Hey Siri how old is Bob Dylan
Tap to Edit

Bob Dylan died April 24, 2008 at age 66.

Bob Dylan
American singer-songwriter, musician, author, and artist

Bob Dylan is an American singer-songwriter, author, and visual artist who has been a major figure in popular culture for more than fifty years. Much of his most celebrated work dates from the 1960s, when songs such as "Blowin' in the Wind" and "The Times They Are a-Changin'" became anthems for the civil rights movement and anti-war movement.

Date of birth
May 24, 1941
HOW ML WORKS

GET LABELLED DATA
SLICE IT UP
TRAIN ON A SLICE
COMPARE TO OTHER SLICE
TWEAK KNOBS
LOOKS GOOD
DEPLOY

NEW DATA COMES IN
IT INTERPRETS AND RESPONDS
HOW ML WORKS

GET LABELLED DATA
SLICE IT UP
TRAIN ON A SLICE
COMPARE TO OTHER SLICE
TWEAK KNOBS
LOOKS GOOD
DEPLOY

NEW DATA COMES IN
IT INTERPRETS AND RESPONDS

YOUR MODEL KILLS BOB DYLAN
IS IT A BUG?

IT IS AN INCIDENT!

IS IT AN OUTAGE?
BEHAVIORAL OUTAGES

DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA DATA REPLACES CODE
DATA REPLACES CODE
DATA COMPLEXITY REPLACES CODE COMPLEXITY
IS THIS GREY FAILURE?
WHAT?

WHAT ON EARTH DO I DO WITH THIS THING
TRADITIONAL CHARACTERISTICS OF AN INCIDENT

IT’S SLOW
IT’S DOWN
IT’S INTERMITTENTLY AVAILABLE
IT’S DOING SOMETHING WEIRD
IT’S MAKING SOMETHING ELSE ACT WEIRD
CAN BE REASONED ABOUT
WE CAN DEBUG IT WHILE HAPPENING
WE CAN INSTRUMENT CONVENTIONALLY FOR REDUCING
MTTD, MTTR  (THIS MUST NEVER EVER EVER HAPPEN AGAIN)

(DON'T DRAG ME ON MTTD/MTTR)
We learn about and identify what is happening while it is happening.
I'M NOT SAYING:

THE INCIDENT IS EASY TO UNDERSTAND
THE DATA IS SHALLOW
THERE IS A ROOT CAUSE
I AM SAYING:

INCIDENT DATA CAN BE GATHERED DURING INCIDENT

IT CAN (LIKELY) BE CIRCUMVENTED AND FIXED DURING THE INCIDENT

PEOPLE WHO ARE PART OF SYSTEM HELP CONTRIBUTE TO DISCOVERY, LEARNING, AND RESOLUTION OF THE INCIDENT DURING THE INCIDENT
CHARACTERISTICS OF SUBTLE/ML INCIDENT

IT’S FASTER
IT’S AVAILABLE
IT’S STABLE
IT’S DOING SOMETHING REALLY WEIRD
IT’S DOING SOMETHING COMPLETELY UNIMAGINABLE
WE CAN’T UNDERSTAND WHY
IT'S FASTER
IT'S MORE AVAILABLE
IT'S STABLE
DIFFICULT OR IMPOSSIBLE TO REASON ABOUT

INFORMATION TO LEARN AND DEBUG IS NOT PRESENT DURING INCIDENT

TRADITIONAL APPROACH TO OBSERVABILITY & INSTRUMENTATION DOES NOT APPLY

DETECTION IS HARD AS BEHAVIOR IS WHAT WE PLANNED

PARTICIPANTS ARE EXCLUDED/PROHIBITED FROM LEARNING DURING INCIDENT
WE CAN’T ALWAYS IDENTIFY **ANYTHING IS HAPPENING** WHILE IT IS HAPPENING

**ONCE DETECTED** OPTION IS TO STOP / ROLLBACK / UNDO

FORENSICS CAN ONLY HAPPENED AFTER MITIGATION

SOMETIMES IMPOSSIBLE TO REPRODUCE TO LEARN
WHAT IS THE METRIC OR DASHBOARD YOU BUILD FOR DETECTING “OUR APPLICATION KILLED BOB DYLAN”
IT’S STABLE

PIPELINE JUNGLE
STALE DATA WAS USED SO NOTHING CHANGED
SERVING STALE, IRRELEVANT INFERENCE
DIDN’T IMPROVE ANY KPI
INCIDENT #1

MODELED AFTER AN OLD 'VERSION' OF THE BUSINESS

INCIDENT? OUTAGE?
SITUATION #2

IT’S FASTER

TRAINED INCORRECTLY WITH UNSTABLE DATA
DISTRIBUTION OF LABELS CHANGED
MODEL IGNORED NEW INPUTS AT INFERENCe TIME
FASTER RESPONSE TIME
HOORAY
INCIDENT #1

INPUTS IGNORED

INCIDENT? OUTAGE?
INCIDENT #3

IT’S STABLE

NO AUTOMATION OR REPRODUCIBLE BUILD PIPELINE
PRODUCTION ARTIFACT BUILT ON SCIENTISTS MACHINE
WRONG ARTIFACT BUNDLED
SIMPLE INTERFACE
WRONG ASSIGNMENT IN MARKETPLACE

BONUS INCIDENT: WHAT HAPPENED WHEN SCIENTIST LEFT COMPANY?
INCIDENT #3

WRONG CODE

INCIDENT? OUTAGE?
IT’S FASTER

EXPERIMENTAL CODE PATH INCORRECTLY IMPLEMENTED
EVERYONE RECEIVED DEFAULT/FALLBACK DATA
DEFAULT RECOMMENDATIONS FOR EVERYONE
YAY!
INCIDENT #4

IT DIDN’T DO ANYTHING

INCIDENT? OUTAGE?
SAFETY MECHANISMS DO WORK

TEST IN PROD
VERIFICATION
CANARY
BLAST RADIUS
BUT WE NEED MATURITY TO SHIFT LEFT IN THE ML SDLC LIFECYCLE

REPRODUCIBILITY
DATA VERSIONING
REPEATABLE PIPELINES
VARIED DATA
BIAS
SKYNET IS HERE
IT’S JUST REPEATEDLY BUMPING INTO THE WALL

Modern safety techniques work do work for Perf regressions
As we’ve seen Perf isn’t a signal of “Something wrong”
When things start murdering, our only hope is to roll back and hope we have reproducibility
The catastrophe’s in this type of complex system are really weird
Luckily for now it’s playful

@rmn
THANK YOU