BOWERBIRDS OF TECHNOLOGY

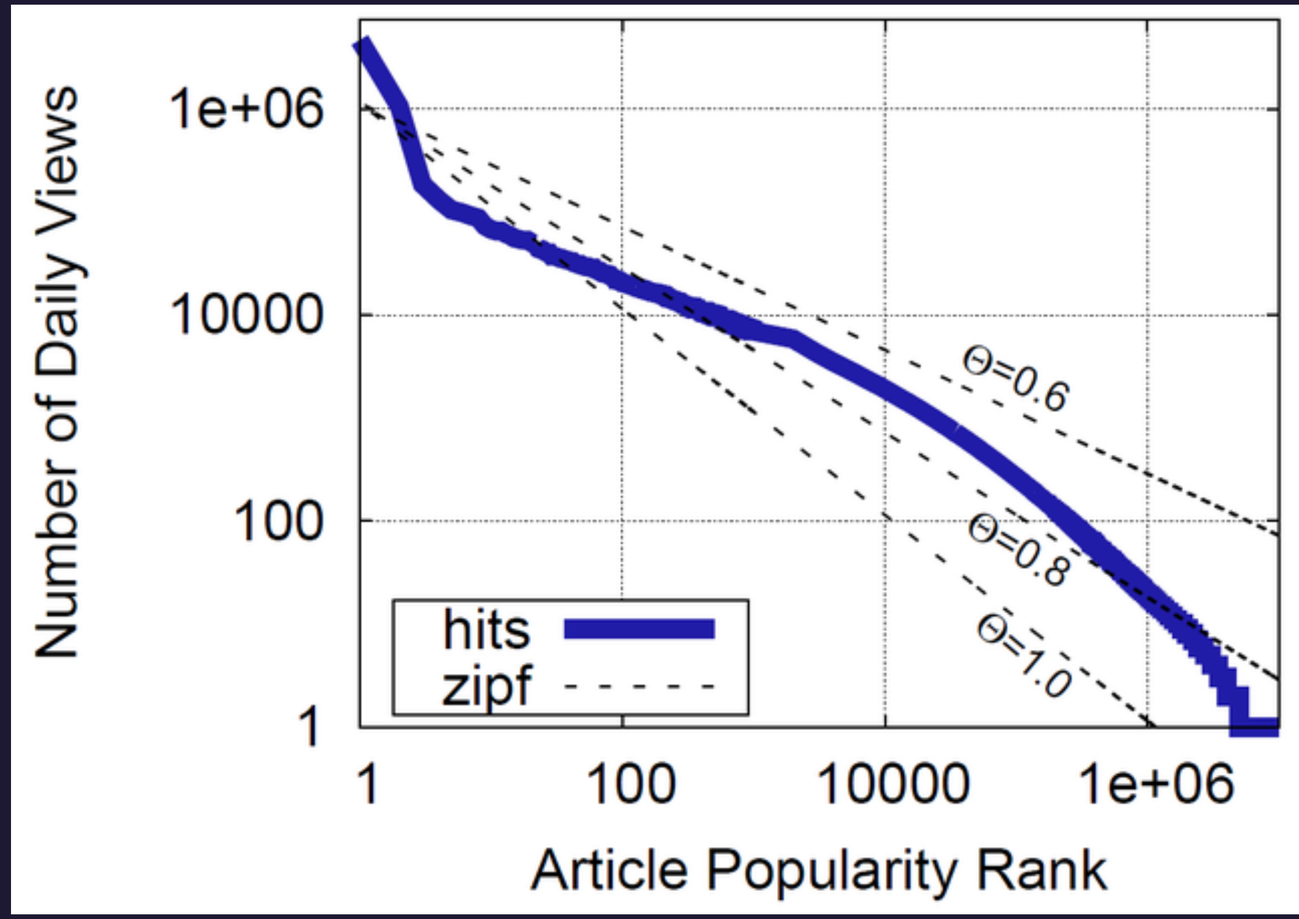
PYCON 2018

SAM KITAJIMA-KIMBREL @SKIMBREL

"Most websites aren't in the top 100 web sites."

CAL HENDERSON, DJANGOCON US 2008

ZIPF DISTRIBUTION



@skimbrel

YOU ARE NOT GOOGLE,
AND THAT'S OKAY!

I AM ALSO NOT GOOGLE

@skimbrel

@skimbrel

WHAT ARE GOOGLE'S PROBLEMS?

ABSURDLY HIGH THROUGHPUT AND STORAGE DEMANDS

TENS OF THOUSANDS OF SERVERS
HUNDREDS OF DATACENTERS

@skimbrel

THOUSANDS OF DEVS

@skimbrel

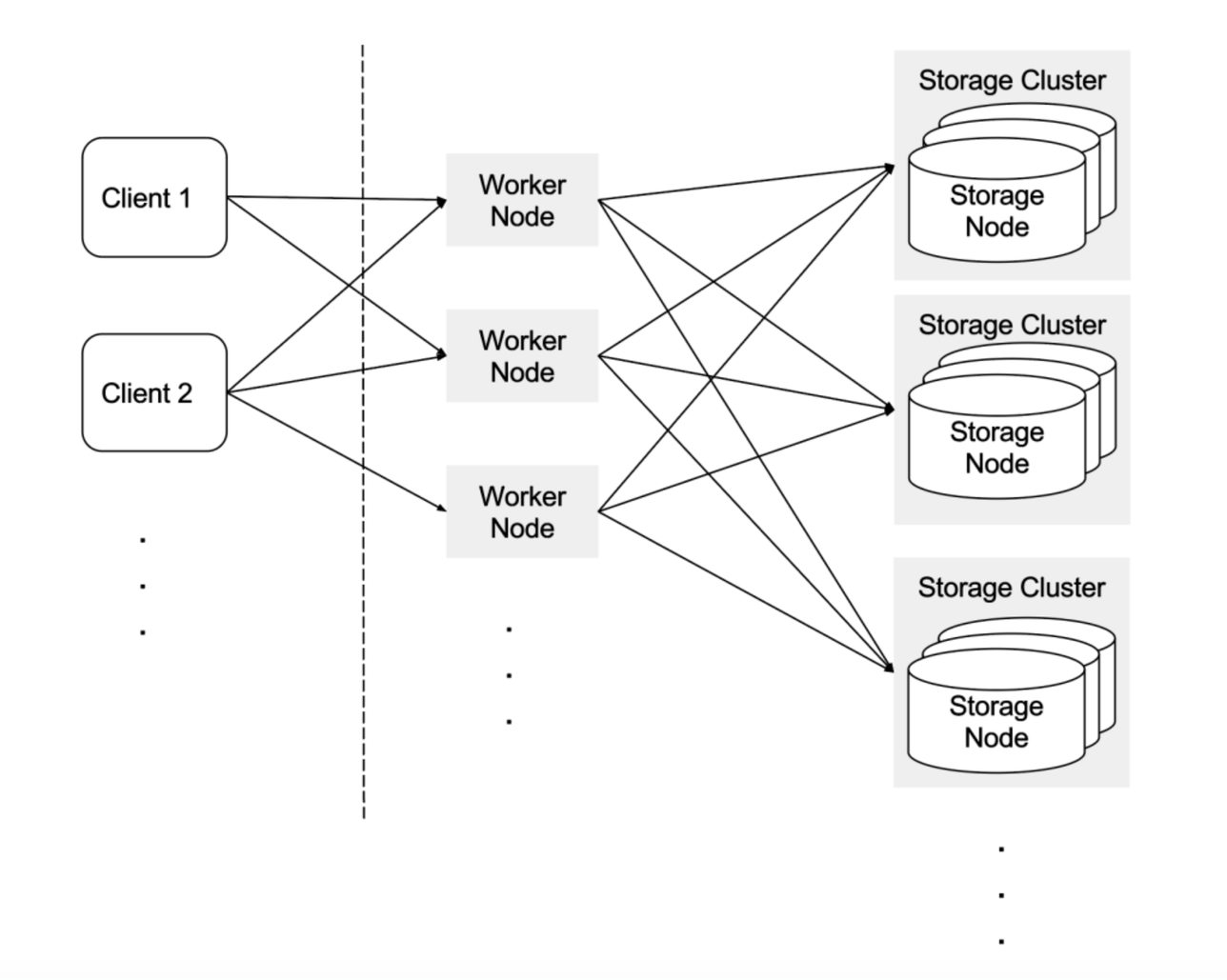
(NEAR-)UNLIMITED RESOURCES

@skimbrel

CASE STUDIES

UBER SCHEMALESS*

*INVOKING UBER AS A TECHNOLOGICAL EXAMPLE DOES NOT CONSTITUTE AN ENDORSEMENT OF UBER'S BUSINESS STRATEGY, ETHICS, OR CULTURE



"LINEARLY ADD CAPACITY
BY ADDING MORE SERVERS"

@skimbrel

"FAVOR WRITE AVAILABILITY OVER READ-YOUR-WRITE SEMANTICS"

@skimbrel

EVENT NOTIFICATIONS (TRIGGERS)

@skimbrel

WHAT IS IT, THEN?

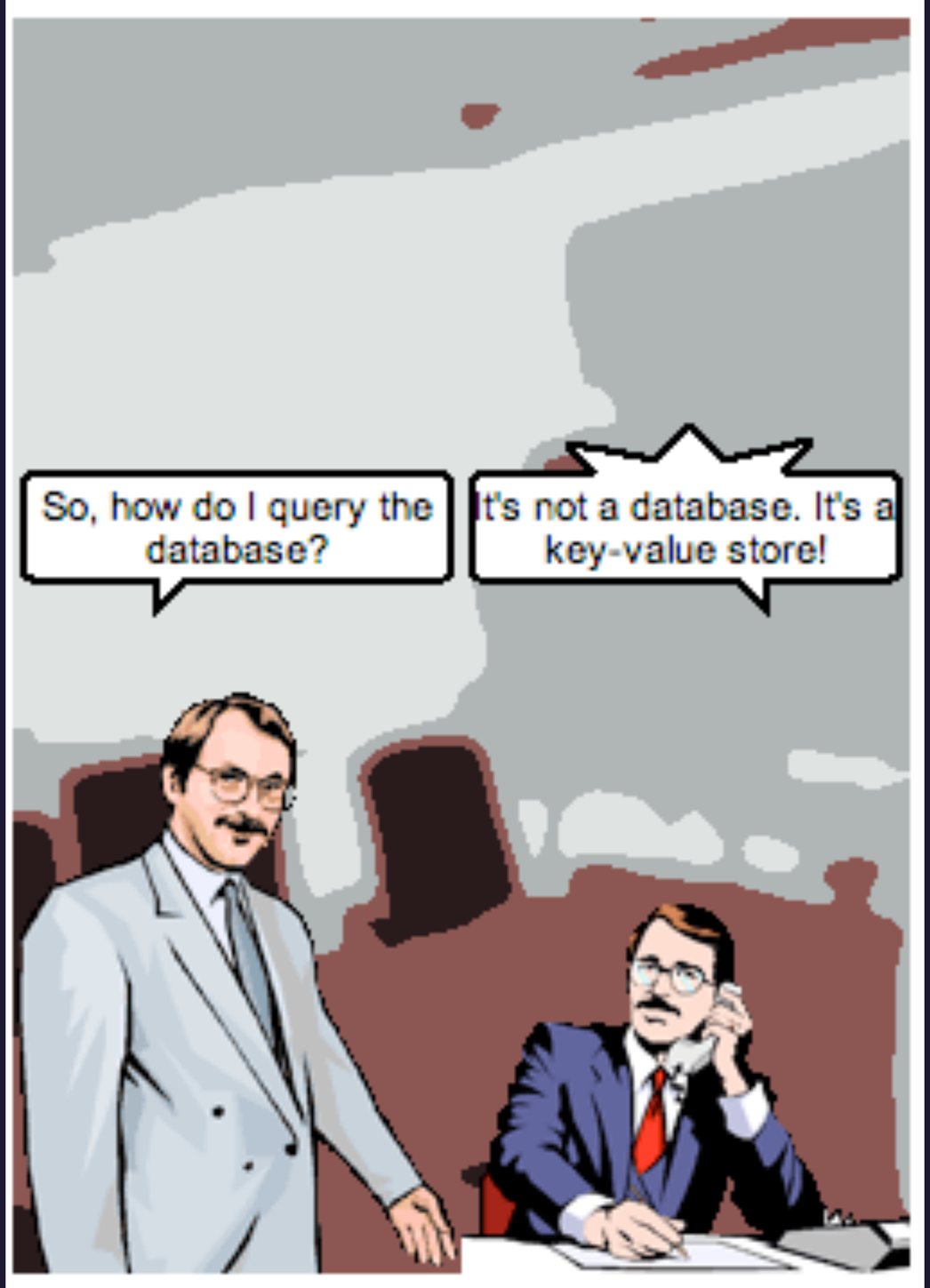
"APPEND-ONLY SPARSE THREE-DIMENSIONAL PERSISTENT HASH MAP, VERY SIMILAR TO GOOGLE'S BIGTABLE"

@skimbrel

Fault-tolerance

"So, how do I query the database?"

"It's not a database. It's a key-value store!"



@skimbrel

"Ok, it's not a database. How do I query it?"

"You write a distributed map reduce function in Erlang!"

@skimbrel

"Did you just tell me to go **** myself?"

"I believe I did, Bob."

@skimbrel

THIS HAS A COST

NEW ABSTRACTIONS

@skimbrel

EVENTUAL CONSISTENCY

@skimbrel

FLEXIBLE QUERIES

@skimbrel

DEVELOPER FAMILIARITY

@skimbrel

AMAZON AND SERVICE ARCHITECTURE

"All teams will henceforth expose their data and functionality through service interfaces."

@skimbrel

"Teams must communicate with each other through these interfaces."

@skimbrel

"There will be no other form of interprocess communication allowed [...]the only communication allowed is via service interface calls over the network."

@skimbrel

"It doesn't matter what technology they use. HTTP, Corba, Pubsub, custom protocols – doesn't matter."

@skimbrel

"All service interfaces, without exception, must be designed from the ground up to be externalizable."

@skimbrel

"Anyone who doesn't do this will be fired."

@skimbrel

AMAZON LEARNED SOME THINGS

@skimbrel

"pager escalation gets way harder, because a ticket might bounce through 20 service calls before the real owner is identified"

@skimbrel

"every single one of your peer teams suddenly becomes a potential DOS attacker"

@skimbrel

"monitoring and QA are the same thing…"

@skimbrel

"…the only thing still functioning in the server is the little component that knows how to say 'I'm fine, roger roger, over and out' in a cheery droid voice"

@skimbrel

"you won't be able to find any of them without a service-discovery mechanism [...] which is itself another service"

@skimbrel

MASSIVELY SCALABLE
INFRASTRUCTURE
COSTS DEVELOPER TIME

BUT I WANT TO BE
GOOGLE!

GOOGLE WASN'T GOOGLE OVERNIGHT...

@skimbrel

"When I joined Google, it would take us about a month to crawl and build an index of about 50 million pages."

BEN GOMES

(HTTP://READWRITE.COM/2012/02/29/INTERVIEW_CHANGING_ENGINES_MID-FLIGHT_QA_WITH_GOOG/)

AND SOMETIMES
STILL IS NOT GOOGLE.

@skimbrel

"BORING" TECH CAN GO REALLY FAR

@skimbrel

EXPONENTIAL GROWTH FEELS SLOW AT FIRST

@skimbrel

ITERATE
ITERATE
ITERATE

@skimbrel

OK, I'M NOT GOOGLE (YET)
.... SO WHAT?

USER TRUST ABOVE ALL ELSE

@skimbrel

FAST, SAFE ITERATION

@skimbrel

HEALTHY TEAMS

@skimbrel

INCLUSIVE TEAMS

@skimbrel

LET'S BE BOWERBIRDS!

@skimbrel

BE A PICKY BIRD

PROJECT MATURITY

@skimbrel

MAINTAINERSHIP

@skimbrel

SECURITY

@skimbrel

STABILITY

@skimbrel

PROJECT ECOSYSTEM

@skimbrel

"OUT OF THE BOX"

@skimbrel

DOCUMENTATION

@skimbrel

SUPPORT AND CONSULTANTS

@skimbrel

LICENSING LANDMINES

@skimbrel

@skimbrel

BUY VS BUILD

WHAT WOULD IT COST TO BUILD IT?

@skimbrel

HOW LONG WOULD IT TAKE?

@skimbrel

HOW HARD IS IT TO REPLACE?

@skimbrel

@skimbrel

RELATIONSHIPS

TEAMS

SUSTAINABLE ON-CALL

@skimbrel

168 HOURS ÷
40 HOUR WORK WEEK =
4.2 PEOPLE

@skimbrel

168 HOURS ÷
40 HOUR WORK WEEK =
~~4.2~~ 5 PEOPLE

@skimbrel

168 HOURS ÷
40 HOUR WORK WEEK =
~~4.2~~ ~~5~~ 6 PEOPLE

@skimbrel

EMPOWERMENT AND AUTOMATION

@skimbrel

APPROPRIATE AVAILABILITY AND SCALABILITY

@skimbrel

99% UPTIME?

3.65 days/year

7.2 hours/month

1.68 hours/week

14.4 minutes/day

THREE NINES

8.76 hours/year

43.8 minutes/month

10.1 minutes/week

1.44 minutes/day

@skimbrel

FOUR NINES?

52.56 minutes/year

1.01 minutes/week

8.64 seconds/day

FIVE?!

5.26 minutes/year

DO YOU NEED THAT SLA?

@skimbrel

HEALTHY TEAMS

PSYCHOLOGICAL SAFETY

@skimbrel

INCLUSIVITY,
NOT JUST DIVERSITY

@skimbrel

SET GROUND RULES

@skimbrel

ADDRESS BIAS

@skimbrel

RETAIN AND PROMOTE

@skimbrel

GET PROFESSIONAL ASSISTANCE

@skimbrel

SUMMARY: HUMANS FIRST

HAPPY USERS

CUSTOMER EMPATHY

@skimbrel

KNOW YOUR USER TECH BASE

@skimbrel

KNOW YOUR IMPACT

@skimbrel

SET EXPECTATIONS

@skimbrel

DEGRADE GRACEFULLY

@skimbrel

(OVER)COMMUNICATE

@skimbrel

"we were unable to update the individual services' status on the AWS Service Health Dashboard (SHD) because of a dependency the SHD administration console has on Amazon S3"

AMAZON

(HTTPS://AWS.AMAZON.COM/MESSAGE/41926/)

GITLAB DATABASE FAILURE

HTTPS://ABOUT.GITLAB.COM/2017/02/10/POSTMORTEM-OF-DATABASE-OUTAGE-OF-JANUARY-31/

@skimbrel

(OVER)COMMUNICATE

@skimbrel

MEASURE SUPPORT PERFORMANCE

@skimbrel

DISASTER RECOVERY

IDENTIFY FAULT DOMAINS

@skimbrel

FAULT TOLERANCE HAS COSTS

@skimbrel

PRACTICE

@skimbrel

SECURITY

OWASP

@skimbrel

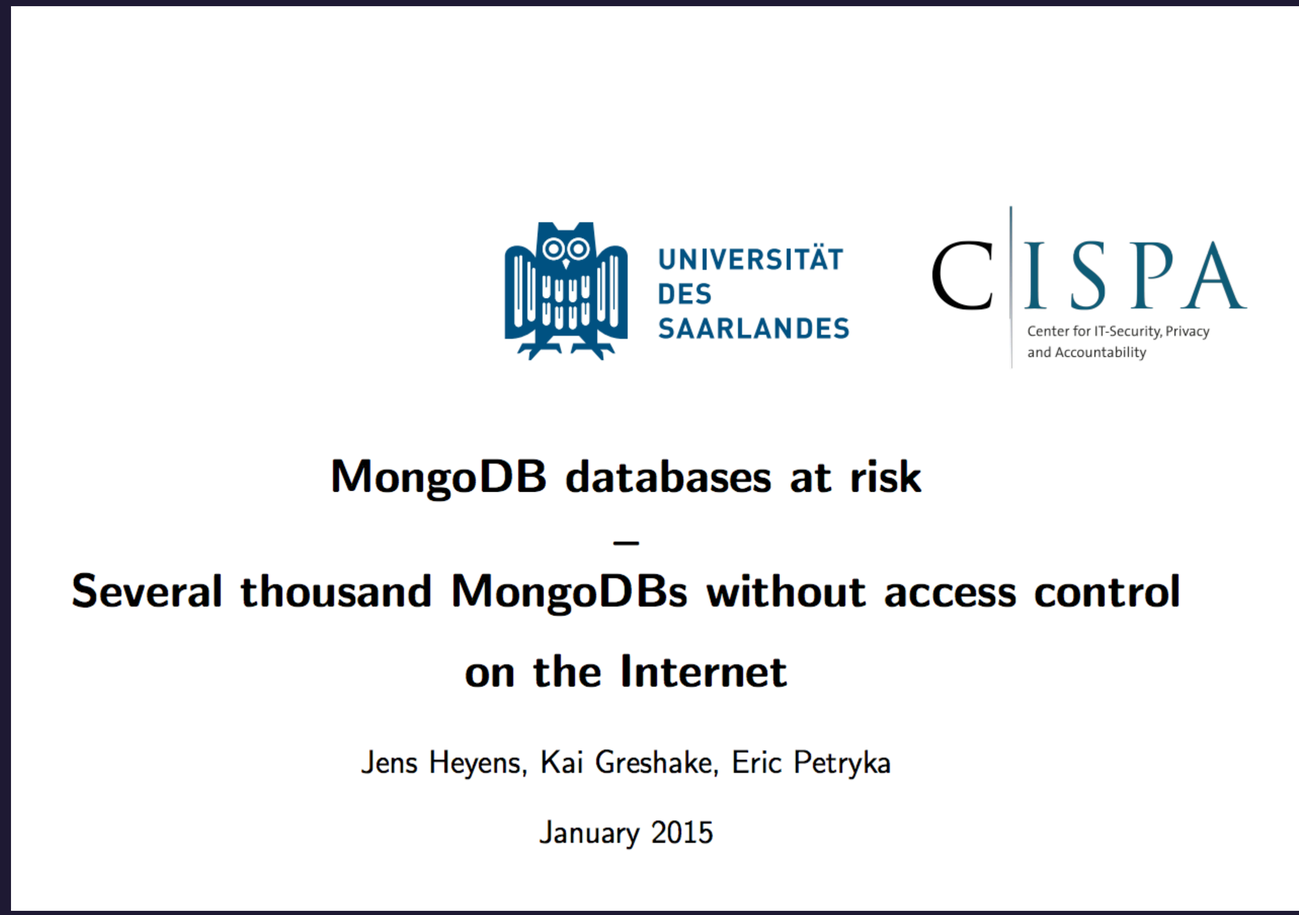
KNOW YOUR THREAT MODEL

@skimbrel

DON'T CHECK CREDENTIALS INTO GIT

@skimbrel

AND DON'T DO THIS



@skimbrel

COMMUNICATE (AGAIN)

@skimbrel

HAPPY BOWER-BUILDING!

FURTHER READING

https://samkimbrel.com/posts/bowerbirds.html

SOURCES

CC BY-NC-ND 2.0 ccdoh1 https://www.flickr.com/photos/ccdoh1/5282484075/
CC BY 2.0 https://www.flickr.com/photos/rileyfive/25506971724
CC BY-SA 3.0 Andrew West https://commons.wikimedia.org/wiki/File:Wikipedia_view_distribution_by_article_rank.png
CC BY-ND 2.0 Melanie Underwood https://www.flickr.com/photos/warblerlady/7664022750/
CC BY-NC 2.0 Nick Morieson https://www.flickr.com/photos/ngmorieson/8056110806/
CC BY-NC-ND 2.0 Julie Bergher https://www.flickr.com/photos/sunphlo/11578609646/
CC BY-NC-ND 2.0 Nathan Rupert https://www.flickr.com/photos/nathaninsandiego/20291539244/
CC BY-NC-ND 2.0 Nathan Rupert https://www.flickr.com/photos/nathaninsandiego/20726159958/
CC BY-NC-ND 2.0 Julie Burgher https://www.flickr.com/photos/sunphlo/11522540164/
CC BY-NC-ND 2.0 Neil Saunders https://www.flickr.com/photos/nsaunders/22748694318/
CC BY 2.0 thinboyfatter https://www.flickr.com/photos/1234abcd/4717190370/
CC BY-SA 2.0 Jim Bendon https://www.flickr.com/photos/jim_bendon_1957/11722386055
https://cispa.saarland/wp-content/uploads/2015/02/MongoDB_documentation.pdf
CC BY-NC-ND 2.0 Julie Bergher https://www.flickr.com/photos/sunphlo/11578609646/
CC BY-SA 3.0 Kay-africa https://commons.wikimedia.org/wiki/File:Flightless_Dung_Beetle_Circellium_Bachuss,_Addo_Elephant_National_Park,_South_Africa.JPG
CC BY-SA 2.0 Robyn Jay https://www.flickr.com/photos/learnscope/14602494872

@skimbrel