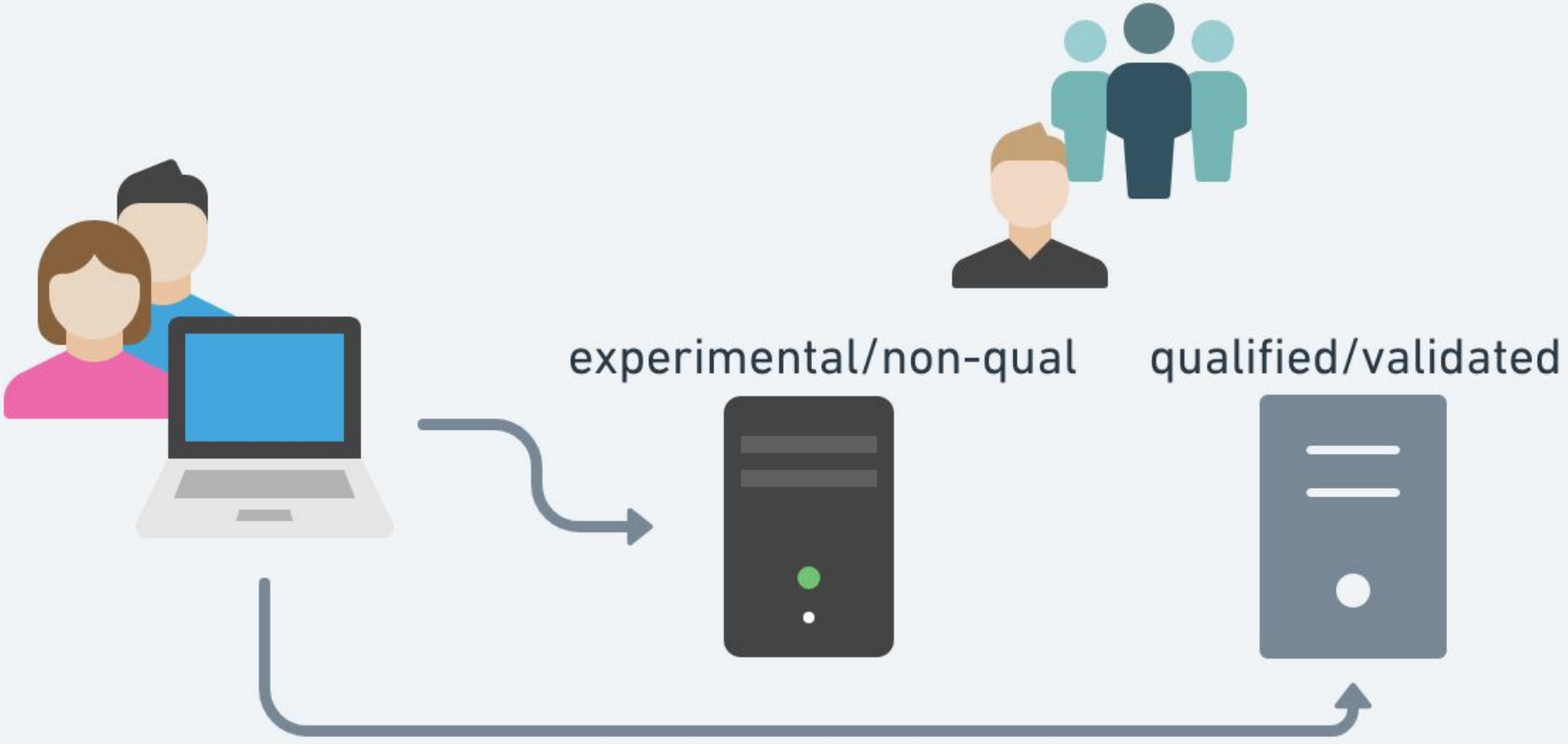
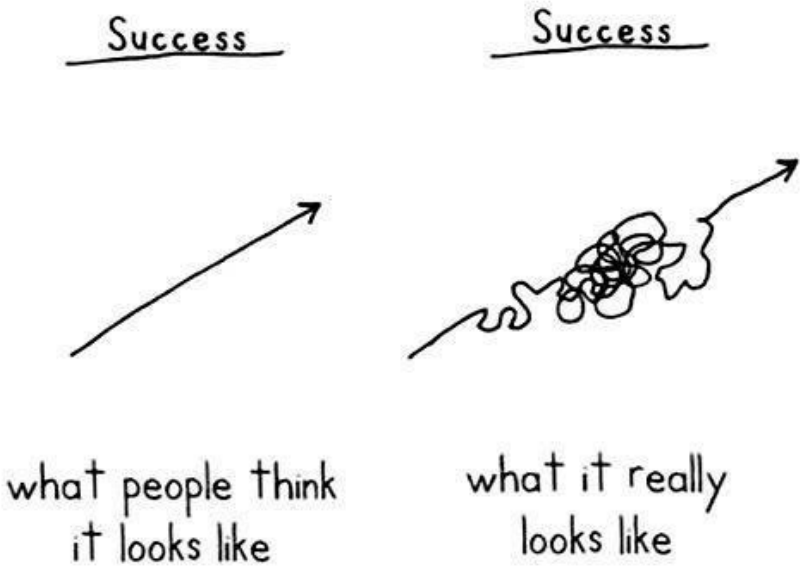
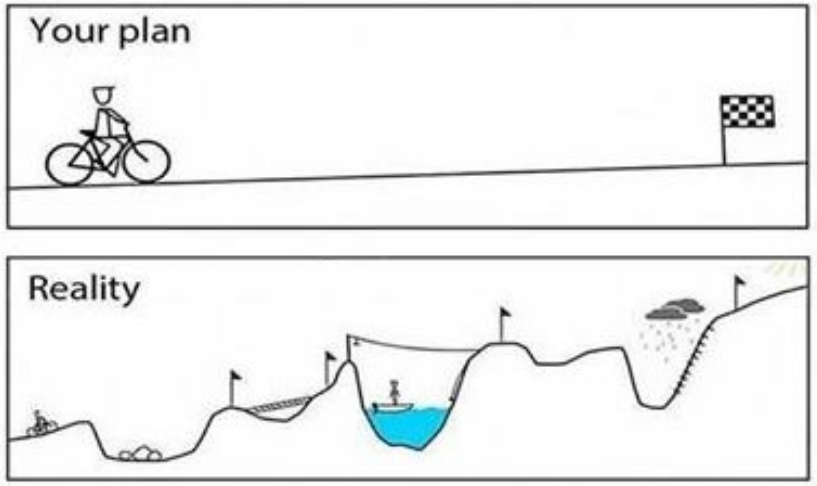
Moving Fast without Breaking Things

Devin Pastoor
Senior Director Engineering Services

METRUM
RESEARCH GROUP


experimental/non-qual
qualified/validated

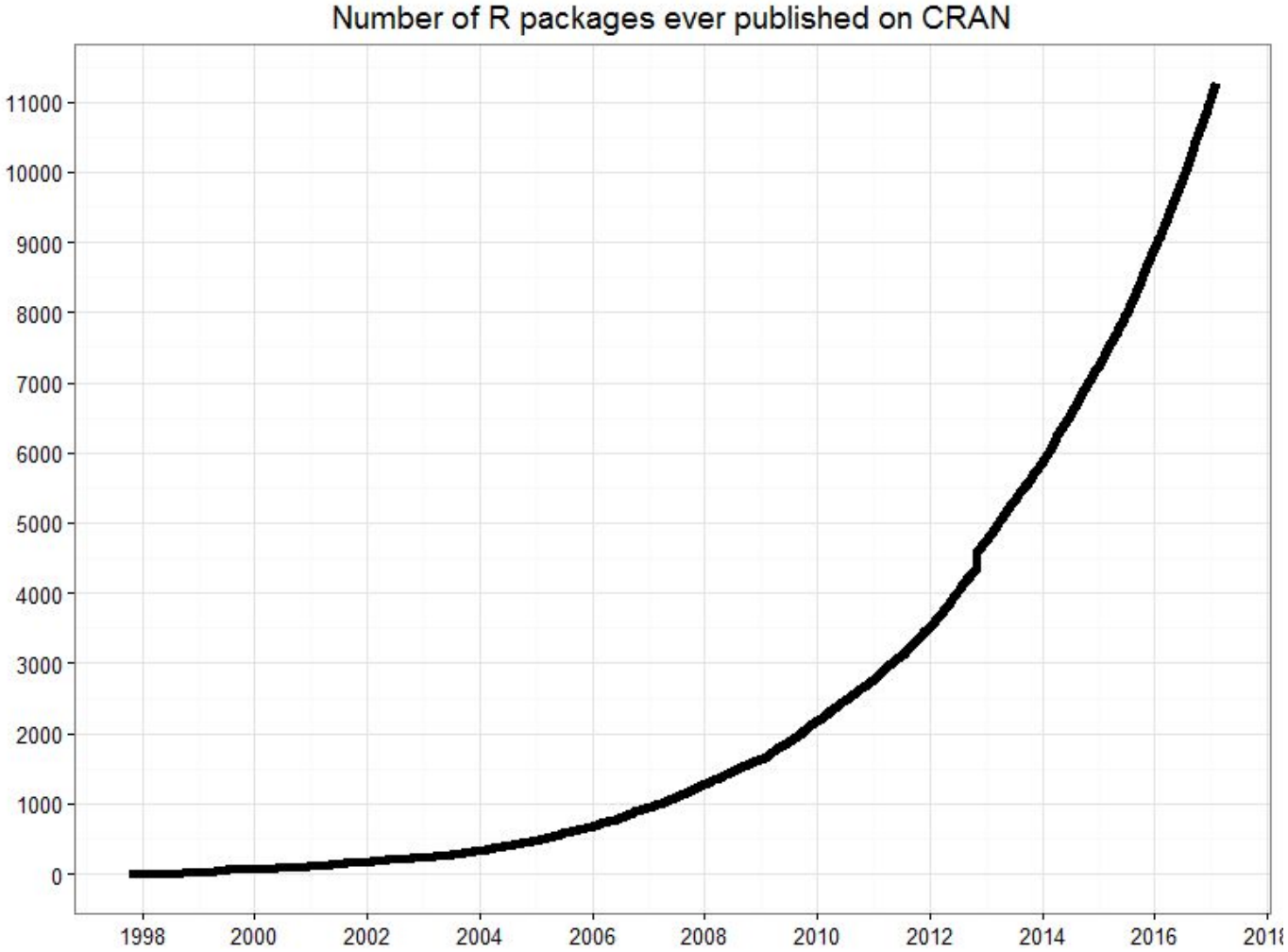




Many “validated” environments are inflexible, monolithic, and fall out of date rapidly

Moving ~~Fast~~ without Breaking Things

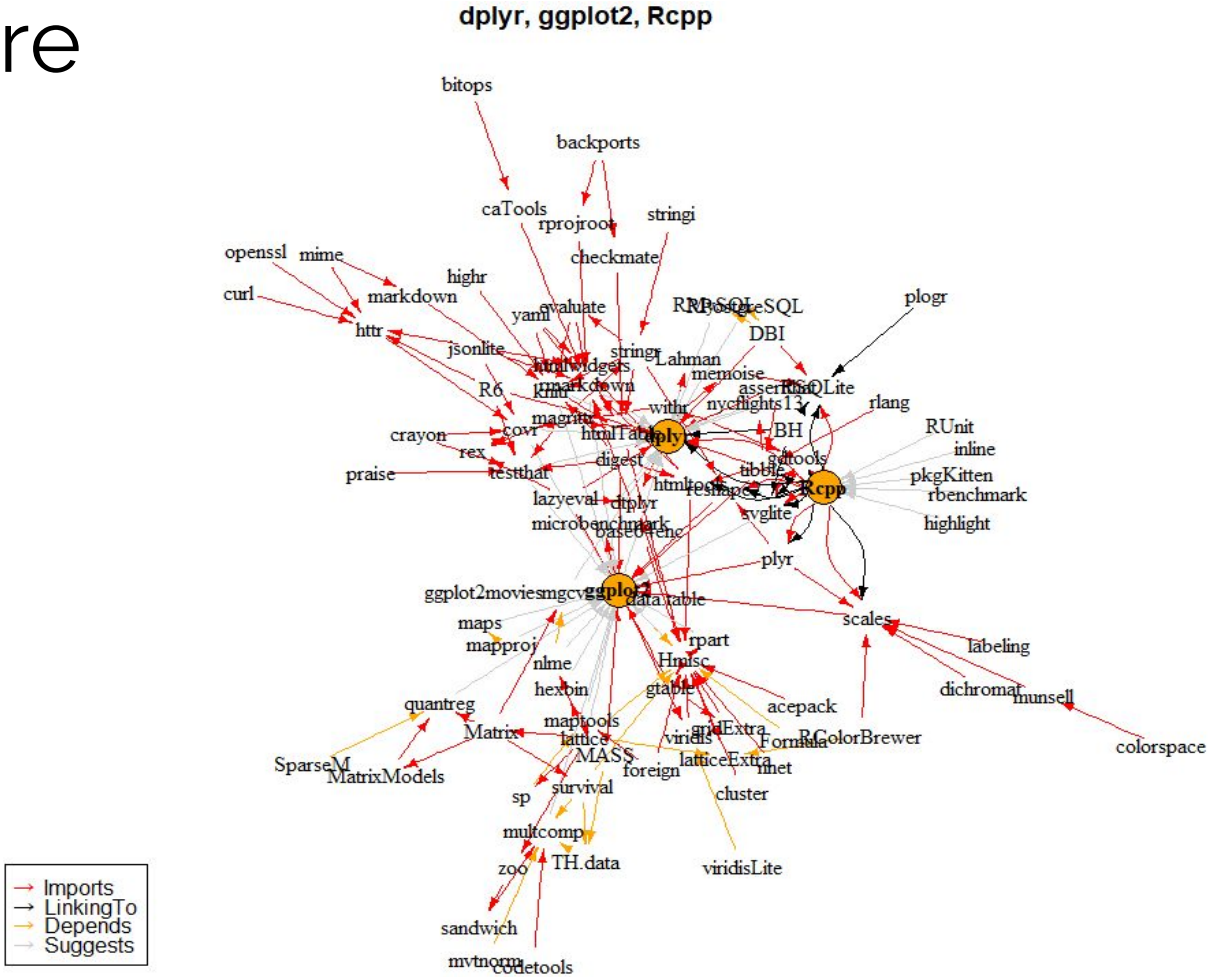
How is the R ecosystem designed?


Number of R packages ever published on CRAN
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016

Misconception #1

Packages are isolated entities

Packages are interwoven



```
Description: A fast, consistent tool for working with data frame like objects,
    both in memory and out of memory.
License: MIT + file LICENSE
URL: http://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr
BugReports: https://github.com/tidyverse/dplyr/issues
Depends:
    R (>= 3.1.2)
Imports:
    assertthat (>= 0.2.0),
    bindrcpp (>= 0.2.0.9000),
    glue (>= 1.1.1),
    magrittr (>= 1.5),
    methods,
    pkgconfig (>= 2.0.1),
    R6 (>= 2.2.2),
    Rcpp (>= 0.12.18),
    rlang (>= 0.2.1),
    tibble (>= 1.3.1),
    tidyselect (>= 0.2.3),
    utils
Suggests:
    bit64 (>= 0.9.7),
    callr,
    covr (>= 3.0.1),
    DBI (>= 0.7.14),
```

DESCRIPTION

- Depends – **required** + loaded
- Imports - **required**
- Suggests – optional for some funcs
- LinkingTo – **required** (usually c++ related)
- SystemRequirements - **required**

Packages can have system dependencies

GitHub, Inc. [US] | https://github.com/r-spatial/sf/blob/master/DESCRIPTION#L61

```
58          rmarkdown
59  LinkingTo: Rcpp
60  VignetteBuilder: knitr
61  SystemRequirements: C++11, GDAL (>= 2.0.1), GEOS (>= 3.4.0), PROJ (>= 4.8.0)
62  License: GPL-2 | MIT + file LICENSE
63  URL: https://github.com/r-spatial/sf/
```

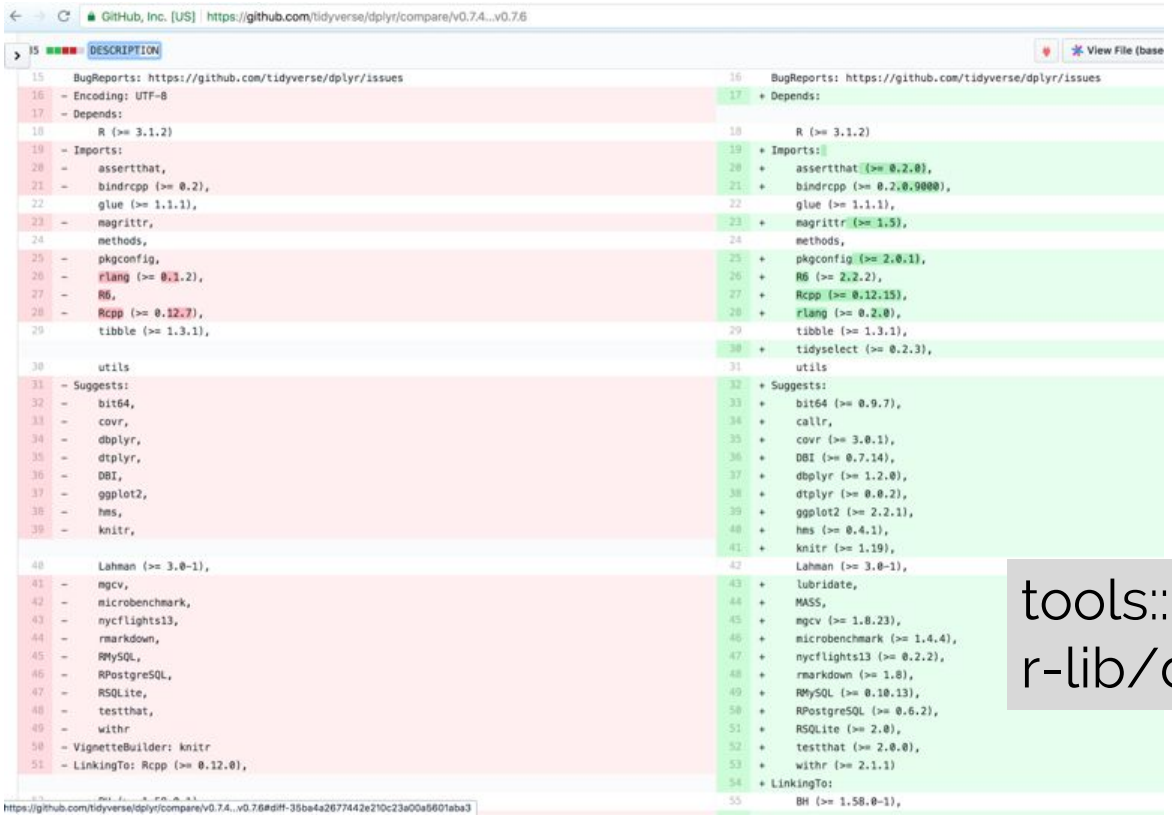
how to install them is NOT communicated in the DESCRIPTION file

Takeaway #1

Think of the surface area change when adding/changing package(s)

github.com/<org>/<repo>/**compare/<start>…<end>**

github.com/tidyverse/dplyr/compare/v0.7.4…v0.7.6



tools::packageDependencies()
r-lib/desc

Misconception #2

Packages on CRAN all work together

```
✔  checking for left-over files
✔  checking index information
✔  checking package subdirectories ...
✔  checking R files for non-ASCII characters ...
✔  checking R files for syntax errors ...
✔  checking whether the package can be loaded ...
✔  checking whether the package can be loaded with stated dependencies ...
✔  checking whether the package can be unloaded cleanly ...
✔  checking whether the namespace can be loaded with stated dependencies ...
✔  checking whether the namespace can be unloaded cleanly ...
✔  checking loading without being on the library search path ...
✔  checking dependencies in R code ...
✔  checking S3 generic/method consistency (623ms)
✔  checking replacement functions ...
✔  checking foreign function calls ...
✔  checking R code for possible problems (1.8s)
✔  checking Rd files ...
✔  checking Rd metadata ...
✔  checking Rd cross-references ...
✔  checking for missing documentation entries ...
✔  checking for code/documentation mismatches (459ms)
✔  checking Rd \usage sections (792ms)
✔  checking Rd contents ...
✔  checking for unstated dependencies in examples ...
✔  checking line endings in C/C++/Fortran sources/headers
✔  checking line endings in Makefiles
✔  checking compilation flags in Makevars ...
✔  checking for GNU extensions in Makefiles
✔  checking for portable use of $(BLAS_LIBS) and $(LAPACK_LIBS)
✔  checking compiled code ...
✔  checking examples (490ms)
✔  checking for unstated dependencies in 'tests' ...
─  checking tests ...
✔  Running 'testthat.R'
─  checking PDF version of manual ...
```

CRAN Package Check Results for Package dplyr

Last updated on 2018-08-15 15:46:57 CEST.

Flavor	Version	$T_{install}$	T_{check}	T_{total}	Status	Flags
r-devel-linux-x86_64-debian-clang	0.7.6	316.61	124.08	440.69	OK	
r-devel-linux-x86_64-debian-gcc	0.7.6	307.18	96.59	403.77	OK	
r-devel-linux-x86_64-fedora-clang	0.7.6			640.96	NOTE	
r-devel-linux-x86_64-fedora-gcc	0.7.6			637.13	NOTE	
r-devel-windows-ix86+x86_64	0.7.6	588.00	217.00	805.00	NOTE	
r-patched-linux-x86_64	0.7.6	358.81	114.24	473.05	OK	
r-patched-solaris-x86	0.7.6			546.80	NOTE	
r-release-linux-x86_64	0.7.6	363.78	113.02	476.80	OK	
r-release-windows-ix86+x86_64	0.7.6	619.00	242.00	861.00	NOTE	
r-release-osx-x86_64	0.7.6				NOTE	
r-oldrel-windows-ix86+x86_64	0.7.6	819.00	286.00	1105.00	NOTE	
r-oldrel-osx-x86_64	0.7.6				NOTE	

Reverse depends: braQCA, censusr, chunked, crplyr, dggridR, dplyr.teradata, efreadr, egor, EpiCurve, EpiStats, etl, eyetrackingR, flatr, GADMTools, geotoolsR, ggmcmc, ggraptR, implyr, kerasformula, manifestoR, merTools, mnreadR, mosaic, mosaicModel, msigdbr, nitrcbot, nlshelper, norris, PogromcyDanych, poplite, prodest, PWFSLSmoke, quickpsy, radiant.data, recipes, reinsureR, rmdHelpers, sdStaf, SEERaBomb, sergeant, sfc, sourceR, spdplyr, staRdom, stranger, summariser, surveybootstrap, tcR, TeachBayes, toolmaRk, treeplyr, turfR, VWPre

Reverse imports: abjutils, ACDm, ActisoftR, actogrammr, adaptalint, adegenet, adjustedcranlogs, admixturegraph, ADMMsigma, ADPclust, aemo, afmToolkit, AGread, AIG, aire.zmvm, alakazam, alfred, alphabetr, alphavantager, AmesHousing, AMR, amt, anchoredDistr, AnglerCreelSurveySimulation, anomalize, anomaly, anomalyDetection, anyflights, apa, apaTables, apaText, ARTool, assertr, atlantistools, auctestr, augmentedRCBD, auk, automagic, AutoModel, aws.alexa, BALCONY, ballr, BANEScarparkinglite, banR, banter, baseballDBR, basictabler, BatchGetSymbols, BAwiR, bayesammi, bayesplot, baystability, beadplexr, benthos, BETS, bib2df, binneR, binomen, Bioi, bioinactivation, biomartr, bioOED, bioset, BIS, bjscrapeR, bkmr, blastula, blkbox, blorr, blscrapeR, BMSC, bomrang, bootnet, boxr, BradleyTerryScalable, BrailleR, brazilmaps, breathtestcore, breathteststan, broom, bsam, bulletr, bupaR, c3, caffsim, cancensus, CARBayesST, carpenter, CaseBasedReasoning, catcont, catenary, catlearn, cattonum, CausalImpact, cbar, cdcfluview, CDECRetrieve, cepR, CGPfunctions, childesr, childhoodmortality, childsds, choroplethr, chorrrds, chromer, circumplex, ciTools, CityWaterBalance, civis, cleanNLP, clustRcompaR, clustrd, clustree, coalitions, cocktailApp, codebook, coefplot, cofeatureR, CollapsABEL, CollapseLevels, colorednoise, cometExactTest, compareDF, comperank, comperes, concaveman, condformat, congressbr, Conigrave, coopProductGame, coreSim, corrr, Countr, countyfloods, countytimezones, countyweather, covTestR, coxed, cpr, cr17, CRANsearcher, crawl, cRegulome, cricketr, cromwellDashboard, CrossClustering, crosswalkr, crossword.r, CRPClustering, crsra, crypto, cutpointr, CVglasso, cytominer, d3r, d3Tree, dat, DataCombine, dataCompareR, datadogr, datadr, dataMeta, dataRetrieval, datastepr, datasus, dbfaker, dbplot, dbplyr, ddpcr, DecisionAnalysis, DeLorean, dendroTools, denovolyzeR, DescriptiveStats.OBeu, descriptr, desctable, detrendr, dexter, dextergui, dexterMST, DiagrammeR, dials, diceR, didrooRFM, diffrprojects, diffrprojectswidget, discord, DisimForMixed, DisImpact, dissever, disto, distrr, DiversityOccupancy, dkanr, DLMtool, dlookr, dMod, dmutate, DMwR2, doBy, docxtools, docxtractr, dotwhisker, drake, driftR, DSAIDE, DSAIRM, dsr, dtplyr, dtwclust, dynfrail, dynutils, eAnalytics, easyformatr, echor, ecoengine, econet, eda4treeR, edeaR, eechidna, eemR, eesim, EFDR, electionsBR, electoral, emil, emuR, engsoccerdata, epicontacts, epidata, episheet, epubr, ERSA, esc, estatapi, ESTER, ether, europepmc, eurostat, evaluator, EventStudy, exifr, ExPanDaR, explor, extdplyr, exuber, ezec, ezsummary, facerec, factorMerger, fastLink, fastqcr, fastR2, fbar, febr, FedData, fedregs, feedeR, finalfit, fingertipscharts, fingertipsR, finreportr, fitteR, flora, ForecastFramework, forestControl, forestmodel, forestr, frequencies, freqweights, FRK, FSA, ftDK, funModeling, funrar, futureheatwaves, fuzzyjoin, gaiah, gastempt, gbfs, gdns, genBaRcode, gender, genderBR, genogeographer, GenomicMating, geomnet, geoparser, geoSpectral, GerminaR, getCRUCLdata, GetDFPData, GetHFData, GetITRData, GetLattesData, getTBinR, GFE, gganimate, ggalt, ggdag, ggdistribute, ggedit, ggeffects, ggfan, ggfocus, ggforce, ggfortify, ggguitar, gghighlight, ggiraphExtra, gglogo, ggmosaic, ggmuller, ggpage, ggplotAssist, ggpmisc, ggpol, ggpubr, ggQC, ggQQunif, ggquickeda, ggquiver, ggraph, ggsom, ggspectra, ggstatsplot, ggvis, gibble, giphyr, gistr, gitgadget, gitlabr, glmmfields, glycanr, goldi, googleAnalyticsR, googlenlp, googlesheets, GRANBase, graphicalVAR, graphTweets, groupdata2, groupedstats, GSODR, gunsales, gutenbergr, gvcR, gwdegree, hansard, happybiRthday, harrietr, healthcareai, heatwaveR, heemod, highcharter, HMMoce, homologene, hpiR, HTSSIP, HURDAT, hurricaneexposure, HydeNet, hydrolinks, hypothesisr, iadf, IAT, IATscores, idbr, IDE, idealstan, idefix, imaginator, imdbapi, imfr, IMP, imputeTestbench, incadata, incgraph, incR, inctools, IncucyteDRC, INDperform, infer, inferr, influxdbr, InformativeCensoring, InjurySeverityScore, inlmisc, internetarchive, interplot, iotables, ipeaData, ipft, ipumsr, iRF, IRISMustangMetrics, ITNr, jaccard, janitor, jpndistrict, jsonstat, jstor, kendallRandomWalks, keyholder, kntnr, kokudosuuchi, konfound, KraljicMatrix, labelled, LAGOSNE, Lahman, laketemps, landscapemetrics, landscapetools, lans2r, LBSPR, LendingClub, lifelogr, lilikoi, linkspotter, live, livechatR, lmeresampler, LocFDRPois, loggit, longurl, lookupTable, loopr, lpirfs, lplyr, lvnet, M2SMF, macleish, MAGNAMWAR, makeFlow, malariaAtlas, mapedit, mapfuser, mase, mason, MazamaSpatialUtils, MazamaWebUtils, mbgraphic, mdsr, meetupapi, meltt, mem, memapp, messaging, metacoder, MetaComp, MetamapsDB, metaplot, metaviz, metricTester, mgcViz, mglR, MIAmaxent, mice, mixOmics, MIBayesOpt, mlbgameday, mlVAR, MLZ, mnis, Modelcharts, modeldb, modelgrid, modelr, moderndive, momentuHMM, Momocs, monkeylearn, morse, mosaicCore, mousetrap, mplot, mRchmadness, mrgsolve, mscstexta4r, mtconnectR, mudata2, muHVT, muir, multistateutils, mvMonitoring, MXM, myTAI, nandb, naniar, nasadata, ncdump, ncmeta, ndjson, nesRdata, NestedCategBayesImpute, net.security, NetworkExtinction, networkreporting, NetworkRiskMeasures, neuropsychology, neutralitytestr, NFP, NLMR, nls.multstart, nlstimedist, nmfem, Nmisc, noaastormevents, NOAAWeather, nomisr, nonmemica, nos, notifyme, NPC, nscprepr, nullabor, observer, oec, olsrr, omu, openadds, openair, opencage, opendotaR, opensensmapr, OptimClassifier, OutliersO3, PAC, packagetrackr, padr, PakPC2017, parlitools, parsemsf, parSim, particles, patrick, patternplot, pccc, pcr, peptider, perccalc, performanceEstimation, PetfinderR, petrinetR, PHEindicatormethods, phenomap, PhenotypeSimulator, philentropy, photobiology, photobiologyInOut, phylopath, pivot, pivottabler, pixiedust, pkggraph, PKNCA, PKPDmisc, Plasmidprofiler, plater, platetools, pleiades, plotly, plotrr, pmatch, pmc, PNADcIBGE, pointblank, poio, pollen, pool, pooling, PopED, popprr, postal, ppcSpatial, PPforest, prcr, prepdat, pRF, prisonbrief, processanimateR, processmapR, processmonitR, prophet, proPubBills, proustr, prozor, PSLM2015, psychmeta, psycho, psychReport, ptstem, purrrlyr, qdap, qicharts2, qqplotr, qsub, quadmesh, quantable, queuecomputer, quickReg, quokar, qwraps2, r2glmm, r4lineups, r511, radiant.basics, radiant.design, radiant.model, radiant.multivariate, randomForestExplainer, rattle, RBesT, rbgm, rbison, rccmisc, rchess, rcicr, RClickhouse, rclimateca, RCMIP5, RColetum, rcongresso, rcrossref, RCrypto, rcv, rdefra, rdfp, Rdrools, rdrop2, readabs, rebird, reclin, REDCapR, RedditExtractoR, REddyProc, refuge, refund.shiny, replyr, reproducible, rerddap, reReg, rERR, resumer, revengc, rfacebookstat, rfishbase, rfm, rfPermute, rgho, rhierbaps, rhmmer, Rilostat, rLandsat, rmapzen, RmarineHeatWaves, rmcfs, RMCriteria, rmweather, RNeXML, RNHANES, Rnightlights, rnoaa, roadoi, roahd, robis, rollmatch, rolypoly, ropenaq, ropercenter, rpcdsearch, rplos, rPraat, rPref, rprev, Rraven, rrr, rsample, rscopus, rscorecard, RSDA, rsinaica, rsoi, Rspotify, rSQM, RSSL, RSwissMaps, rtimes, rtimicropem, rtrek, rtrends, rubias, ruler, rusk, rvertnet, rwalkr, rwunderground, sabre, saeSim, salesforcer, sambia, sampler, SanFranBeachWater, SanzCircos, sars, scanstatistics, scholar, SCORPIUS, SCPME, sealr, segclust2d, sejmRP, seplyr, Seurat, sfdct, shazam, shinyHeatmaply, ShinyTester, SIBER, sicegar, sidrar, simglm, simmer.plot, simPH, SimplifyStats, SimRVPedigree, simTool, SingleCaseES, SIRItoGTFS, sitar, sjlabelled, sjmisc, sjPlot, sjstats, skimr, skynet, slackr, solrium, solvebio, sophisthse, sorvi, SpaCCr, SpaDES.core, sparkavro, sparklyr, SpatialBall, SpatialEpiApp, spbabel, spellcheckr, splithalf, spotifyr, srvyr, ss3sim, Sstack, stability, stacomiR, statar, states, statesRcontiguous, statsDK, statsr, stminsights, stormwindmodel, stplanr, StratifiedRF, STRMPS, SubgrPlots, sugrrants, sunburstR, superheat, surveydata, survminer, survsup, survtmle, survutils, SVMMaj, sweep, swmnr, SWMPr, SWMPrExtension, syllabifyr, syuzhet, taber, tabr, tabularaster, taxa, taxizedb, tbl2xts, teachingApps, telefit, temperatureresponse, tetraclasse, textfeatures, textmining, textreuse, textstem, theseus, thinkr, tibbletime, tidybayes, tidyboot, tidycensus, tidygenomics, tidygraph, tidyhydat, tidyimpute, tidyLPA, tidymodels, tidyposterior, tidypredict, tidyquant, tidyr, tidyRSS, tidystats, tidystringdist, tidytext, tidytransit, tidytree, tidyverse, tigger, tigris, timelineR, timelineS, timetk, Tmisc, togglr, toxplot, TrafficBDE, tree.bins, trelloR, trialr, tribe, tricolore, trread, tsibble, tuber, TwoRegression, uaparserjs, ubci, ubeR, ukbtools, uncmbb, understandBPMN, unjoin, unpivotr, useful, utilsIPEA, valr, vcfR, vdmR, vip, visdat, voteogram, vpc, vqtl, waccR, wakefield, walkalytics, walker, wallace, walmartAPI, wand, weathercan, webTRISr, wec, weibulltools, wfindr, whereport, WHO, widyr, WikidataQueryServiceR, windfarmGA, woeR, wordbankr, WordR, worldmet, wrangle, WRTDStidal, WufooR, xesreadR, XGR, XKCDdata, xmrr, xpose, xpose4, xray, xtractomatic, yardstick, yorkr, Zelig, ZeligChoice, ZeligEI, zeligverse, zFactor, ZipRadius, ztype

Reverse suggests: adaptMT, afex, agridat, airGR, ALA4R, alluvial, ameco, archivist, assertive.types, bannerCommenter, BAS, bayesGDS, beezdemand, bench, BiDimRegression, bigrquery, billboard, binford, bisect, bmlm, bodenmiller, bossMaps, BSDA, CAISEr, canvasXpress, caret, causaldrf, cbsodataR, checkr, climbeR, clustermq, cna, codemetar, codingMatrices, codyn, comtradr, cowplot, ctmle, cvequality, cytofan, DALEX, dalmatian, datamaps, DataVisualizations, DeclareDesign, decoder, deeplr, DepthProc, describer, diffdf, DirectEffects, directlabels, dirichletprocess, EbayesThresh, ecb, ecotox, edgarWebR, epiphy, epitab, europop, eyelinker, feather, fiftystater, filesstrings, fishdata, foghorn, forwards, fourierin, fractional, frailtyEM, fredr, FSelectorRcpp, fueleconomy, furniture, furrr, gapminder, geex, geoknife, gestalt, ggconf, ggenealogy, ggformula, ggtinnards, ggiraph, ggmap, ggplot2, ggRandomForests, ggridges, ggspatial, ggswissmaps, ggthemes, ggvoronoi, glue, googledrive, graphframes, grattan, Greg, growthcurver, gscounts, hdme, hei, highlightHTML, hipread, hrbrthemes, htmlTable, hutils, huxtable, hydrostats, imager, importar, inlabru,

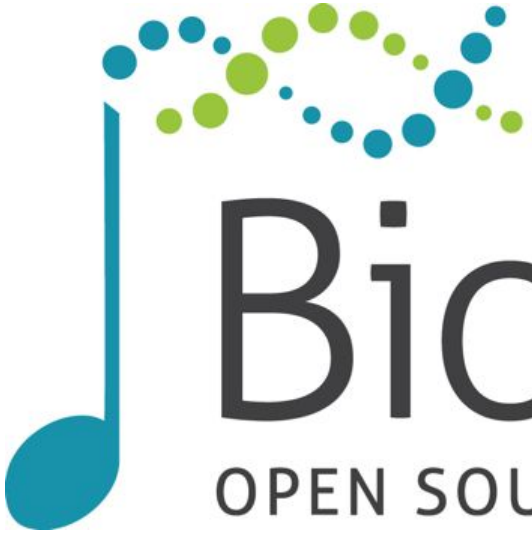
Correction:

The most recent version of all packages on CRAN (should) work together

Takeaway(s) #2

Think in terms of package cohorts.

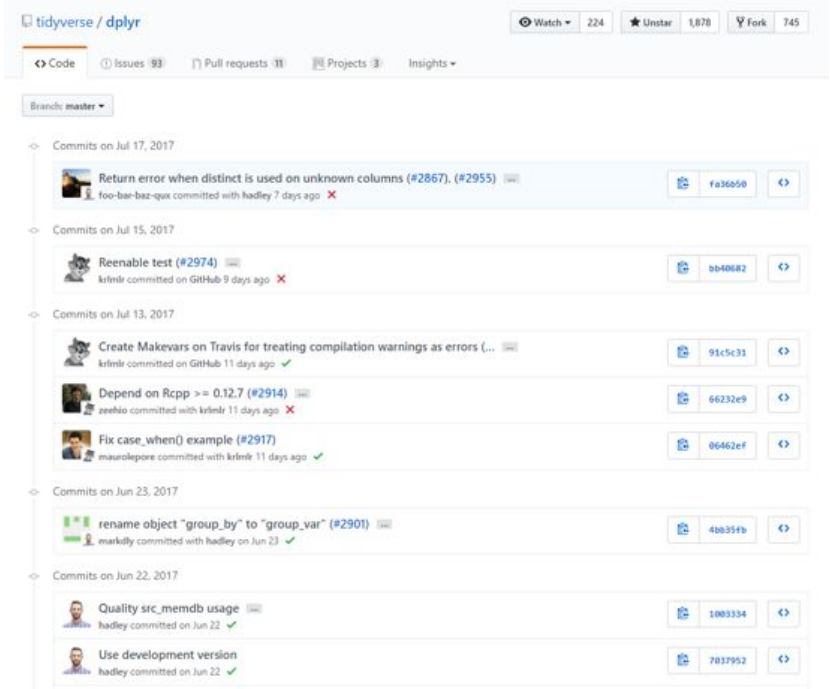
Avoid mismatch in package sets over time

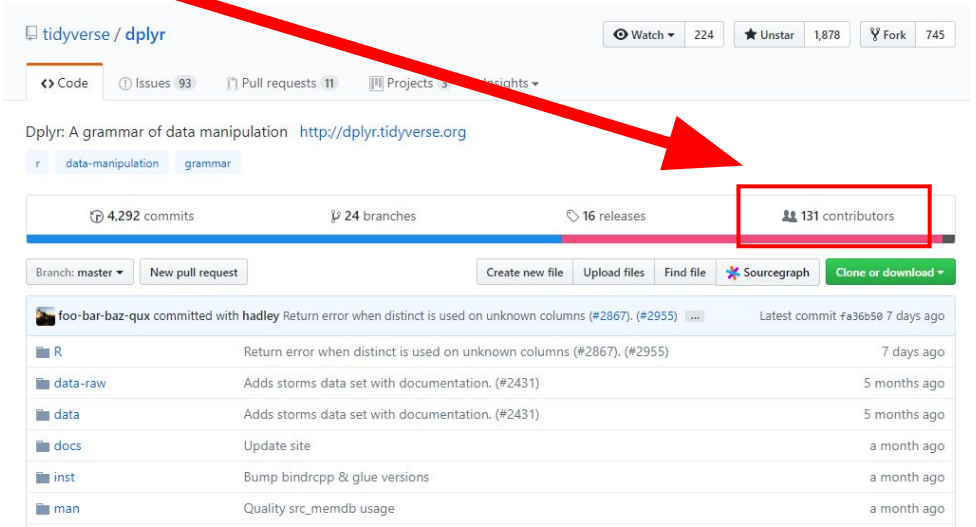


Bioconductor

OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Github – a collaborative, social space to share code





github
GAME OF THRONES
ASCENT

GitHub, Inc. [US] | https://github.com/tidyverse/ggplot2/blob/master/DESCRIPTION

```
9         rmarkdown,
0         sf (>= 0.3-4),
1         svglite (>= 1.2.0.9001)
2     Remotes: hadley/scales,
3         hadley/svglite,
4         jimhester/withr,
5         r-lib/rlang
6     Enhances: sp
```

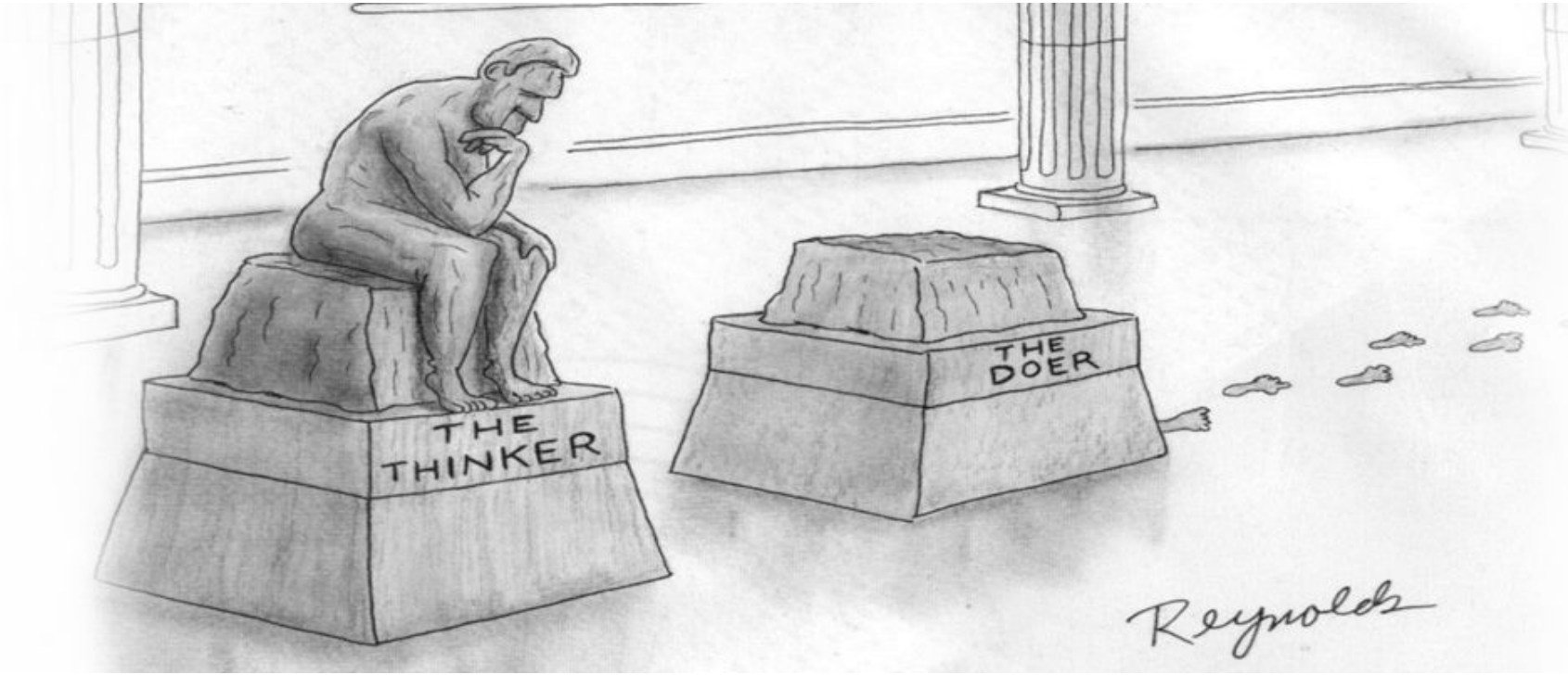
Upgrading packages in production

Takeaway(s) #3

When pulling in a ”single” github package, identify what else will come in from outside CRAN

Takeaway #1

Think of the surface area change when adding/changing package(s)


THE THINKER
THE DOER
Reynolds

https://github.com/metrumresearchgroup/pkgcheck

```
rstudio@8c913778e534:~/tarballs$ pkc check . --config=../crancheck.yml
INFO[0000] running check for package: tidyr
INFO[0000] running check for package: PKPDmisc
INFO[0000] running check for package: dplyr
INFO[0033] tidyr v0.8.1 CHECK RESULTS: 0 ERRORS, 0 WARNINGS, 1 NOTES
INFO[0033] tidyr v0.8.1 TEST RESULTS: 323 OK, 0 Skipped, 0 Failed
INFO[0048] PKPDmisc v2.1.1 CHECK RESULTS: 0 ERRORS, 0 WARNINGS, 1 NOTES
INFO[0048] PKPDmisc v2.1.1 TEST RESULTS: 48 OK, 0 Skipped, 0 Failed
INFO[0210] dplyr v0.7.5 CHECK RESULTS: 0 ERRORS, 0 WARNINGS, 2 NOTES
INFO[0210] dplyr v0.7.5 TEST RESULTS: 2733 OK, 6 Skipped, 0 Failed
INFO[0210] running on 8 threads took 3m30.048510975s
```

Takeaway #3

check compatibility and track full dependency trees

Cohort management

```
packrat
…
Package: hms
Source: CRAN
Version: 0.4.2
Hash: b4096a4f6a6736138e9a825c2baaacf0
Requires: pkgconfig, rlang

Package: htmltools
Source: CRAN
Version: 0.3.6
Hash: b24df7ea0856eab6618f6a56016d940d
Requires: Rcpp, digest

Package: htmlwidgets
Source: CRAN
Version: 1.2
Hash: de18b75f31630089b22e30d4b188cfbe
Requires: htmltools, jsonlite, yaml
…
```

- switchr
- jetpack
- packrat
- containerit

Tooling evolution - Cohort generation builds on **r-lib/pkgdepends** (very experimental)

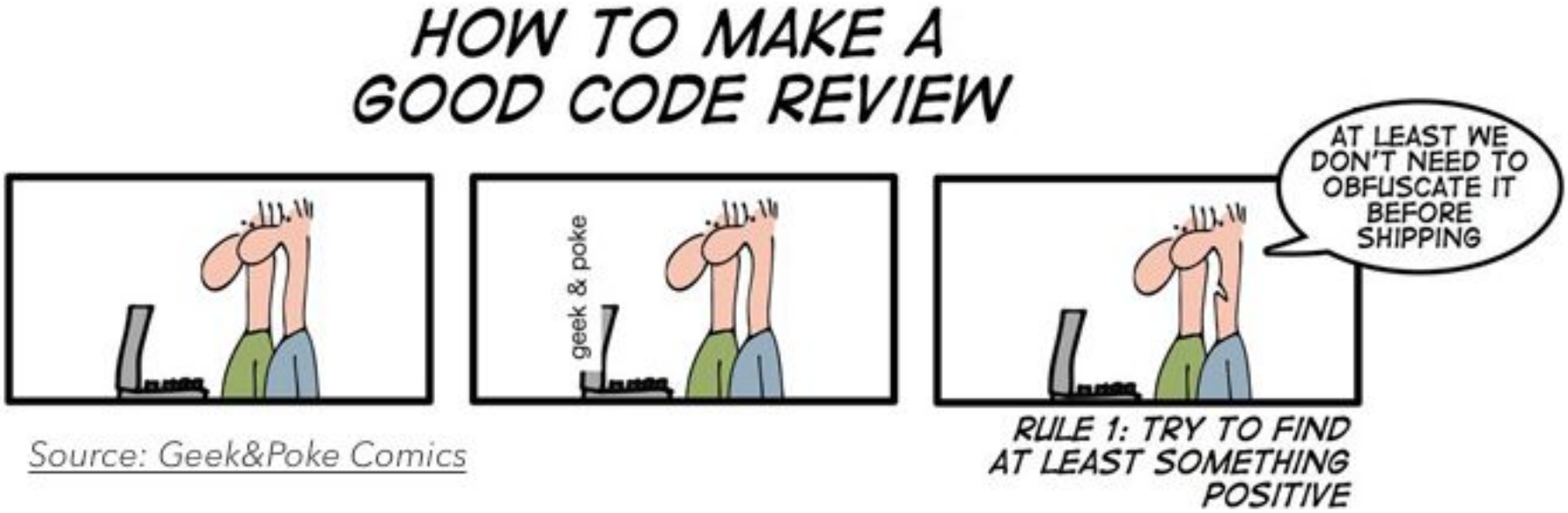
```
gen_runtime_description(
  c("dplyr",
    "ggplot2",
    "flexdashboard",
    "purrr",
    "sparklyr (>= 0.7.0-9020)",
    "pool (>= 0.1.4)"),
  github = c("rstudio/pool", "rstudio/sparklyr")
)
```

Takeaway #4

consider building cohorts immutably, check compat, then use lockfiles to create “production” environments

How to Move Fast(er)?

Invest in training and process improvement



Source: Geek&Poke Comics

Take Advantage of Tools and Add-ons

- usethis
- datapasta
- remedy
- sinew
- covr
- custom scaffolding functions

```
use_package("MASS", "Suggests")
#> ✔ Adding 'MASS' to Suggests field in DESCRIPTION
#> ● Use `requireNamespace("MASS", quietly = TRUE)` to test if package is installed
#> ● Then use `MASS::fun()` to refer to functions.

use_dev_package("callr")
 #> ✔ Adding 'callr' to Imports field in DESCRIPTION
 #> ✔ Adding 'r-lib/callr' to Remotes field in DESCRIPTION

use_roxygen_md()
#> ✔ Setting Roxygen field in DESCRIPTION to 'list(markdown = TRUE)'
#> ✔ Setting RoxygenNote field in DESCRIPTION to '6.1.0'
#> ● Run `devtools::document()`

use_rcpp()
#> ✔ Adding 'Rcpp' to LinkingTo field in DESCRIPTION
#> ✔ Adding 'Rcpp' to Imports field in DESCRIPTION
#> ✔ Creating 'src/'
#> ✔ Adding '*.o', '*.so', '*.dll' to 'src/.gitignore'
#> ● Include the following roxygen tags somewhere in your package
#> #' @useDynLib mypkg, .registration = TRUE
#> #' @importFrom Rcpp sourceCpp
#> ● Run `devtools::document()`
```

```r
library(ggplot2)
library(rlang)
my_awesome_plot <- function(df, x, y) {
  xvar <- enquo(x)
  yvar <- enquo(y)
  ggplot(df, aes(!!xvar, !!yvar)) +
    geom_point()
}
```

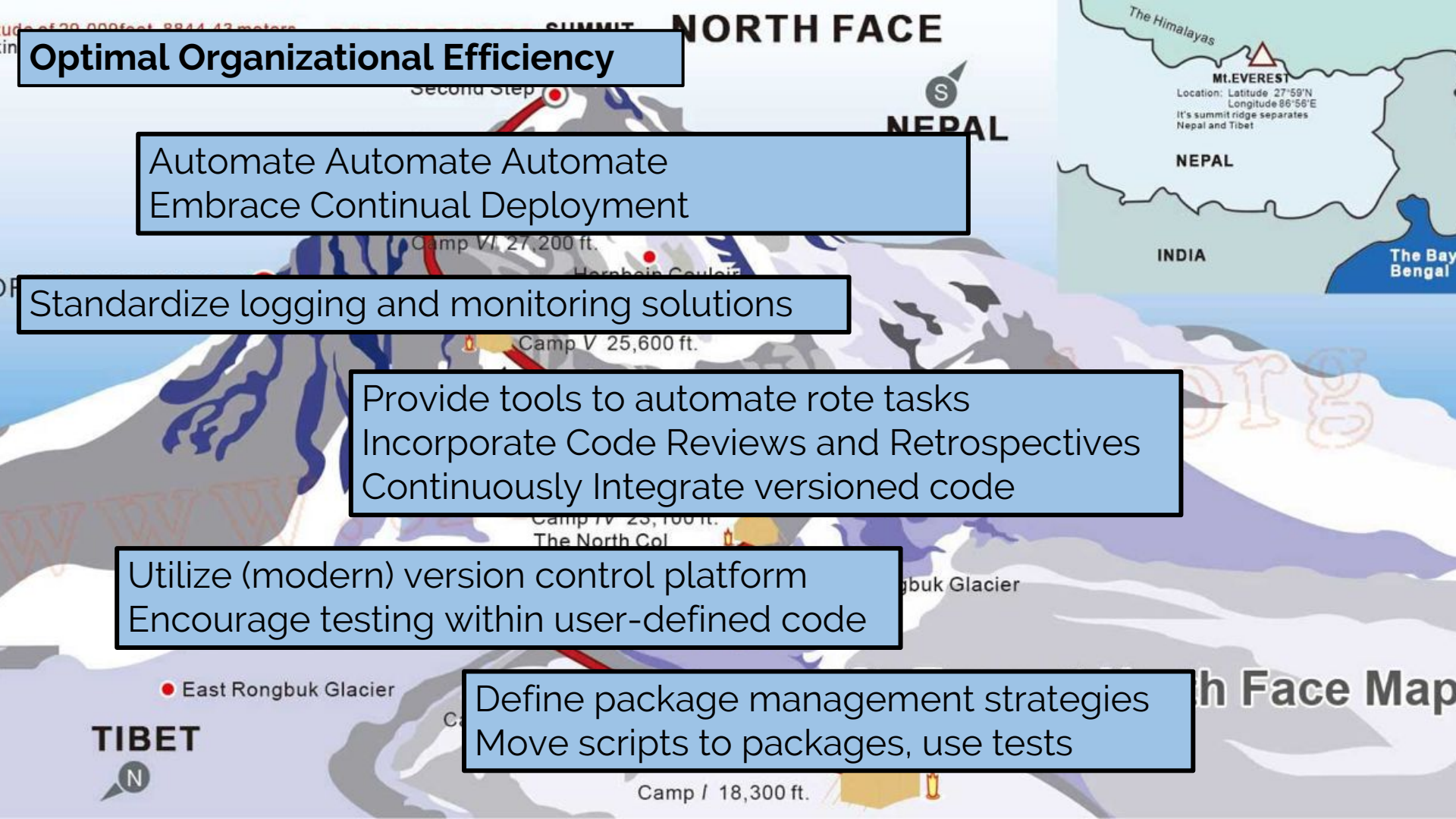
```
> sinew::pretty_namespace("example_plot.R")

functions changed in 'example_plot.R':
  base::library       (line: 1 cols: [01-07])
  base::library       (line: 2 cols: [01-07])
  dplyr::enquo        (line: 7 cols: [11-15])
  dplyr::enquo        (line: 8 cols: [11-15])
  ggplot2::ggplot     (line: 9 cols: [03-08])
  ggplot2::aes        (line: 9 cols: [14-16])
  ggplot2::geom_point (line: 9 cols: [37-46])

library(ggplot2)
library(rlang)



my_awesome_plot <- function(df, x, y) {
  xvar <- dplyr::enquo(x)
  yvar <- dplyr::enquo(y)
  ggplot2::ggplot(df, ggplot2::aes(!!xvar, !!yvar)) + ggplot2::geom_point()
}
```



Optimal Organizational Efficiency

Automate Automate Automate
Embrace Continual Deployment

Standardize logging and monitoring solutions

Provide tools to automate rote tasks
Incorporate Code Reviews and Retrospectives
Continuously Integrate versioned code

Utilize (modern) version control platform
Encourage testing within user-defined code

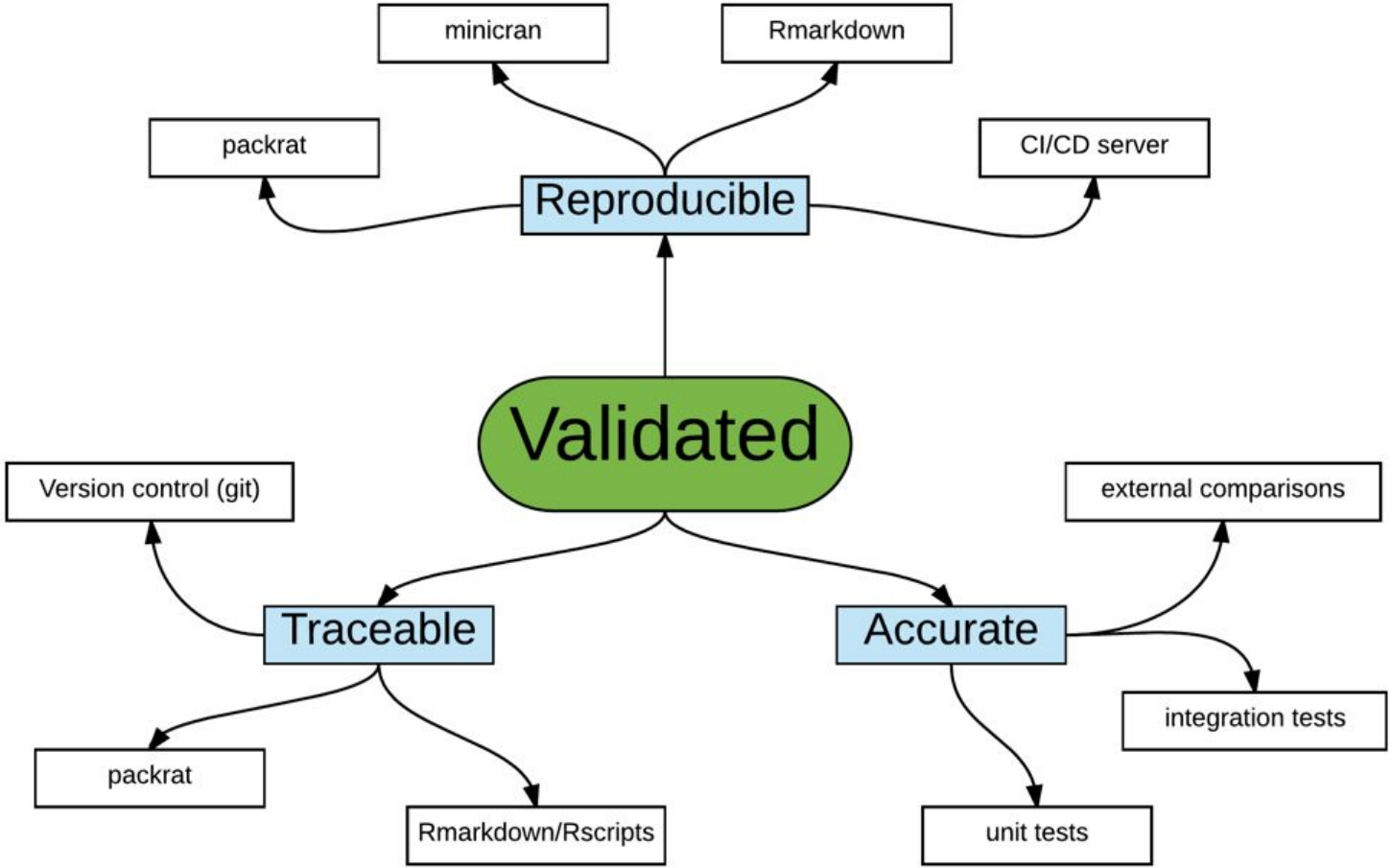
Define package management strategies
Move scripts to packages, use tests

devinp@metrumrg.com

@dpastoor on **github**

@devinpastoor on **twitter**

backup


minicran
Rmarkdown
packrat
CI/CD server
Reproducible
Validated
Version control (git)
external comparisons
Traceable
Accurate
integration tests
packrat
Rmarkdown/Rscripts
unit tests

Package (version) Assessment Checklist

- **What is the package needed for?** Upgrading an inflight project vs new project.
- **Where is the package coming from?**
  - If not from CRAN, how much new development has been done?
  - What is the additional packaging surface area? (new dependencies)
    - Are there any new and/or remote dependencies?
- **What does the package do?**
  - Methodology? Data manipulation?
  - How easy will it be to tell if the package is not accurate?
- **What is the package pedigree?**
  - Who created it?
  - How active/responsive are they?
  - Have they communicated future plans about features/maintenance?
- **What checks are in place** to prevent regressions and demonstrate the package will continue to do what it says it does?
  - Tests
  - Community Engagement