



Appsilon
DATA SCIENCE

How to Set Up a Distributed Data Science Team

Olga Mierzwa-Sulima

olga@appsilon.com

28 July 2020 | RStudio Webinar

Olga Mierzwa-Sulima

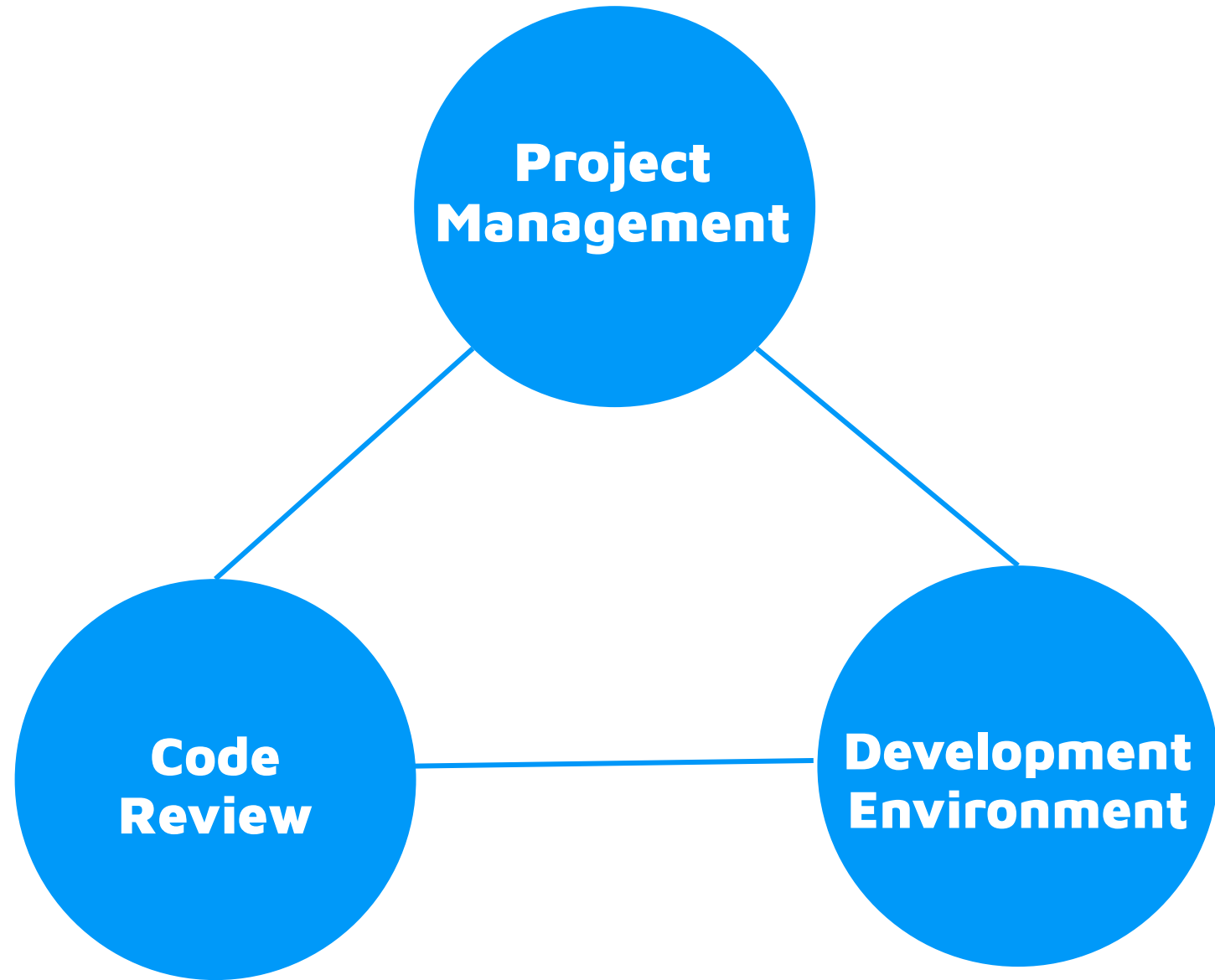
olga@appsilon.com

Senior Data Scientist and
Project Leader @ Appsilon



- Background in Econometrics
- My main tech stack is R
- Building R community in Warsaw

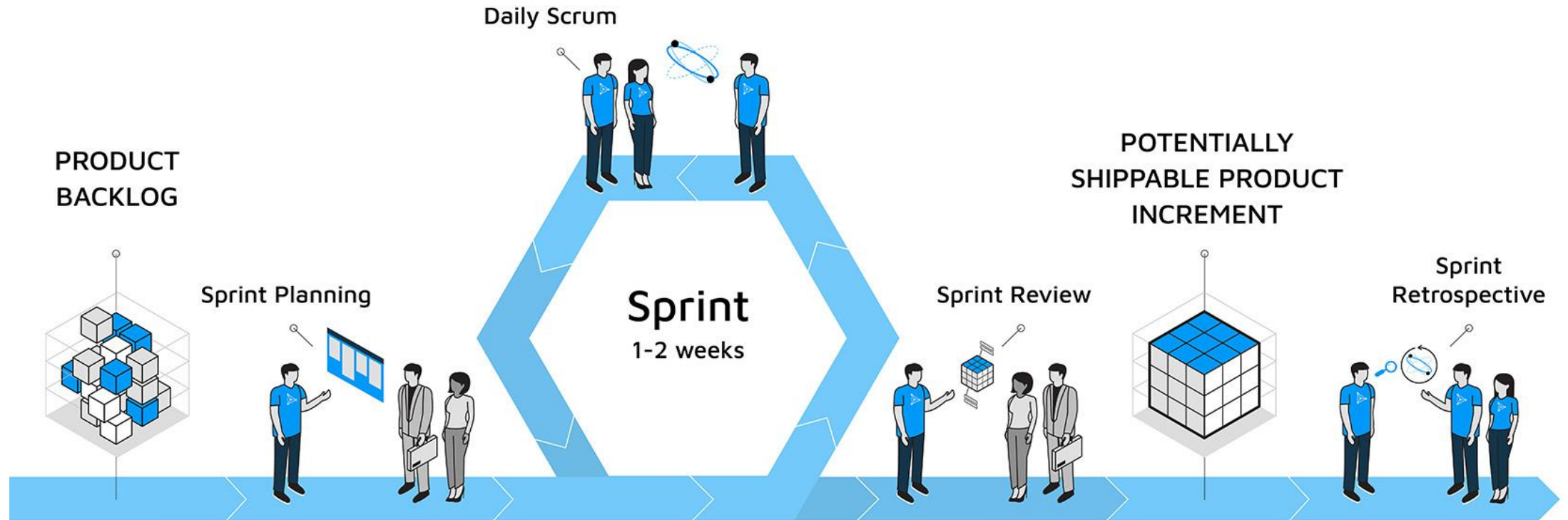




Project Management

Process

Scrum



- Dedicated proj-channel
- Status updates using online communicator
- 1-on-1 to discuss specific problems

Project board

Example

 shiny-app-showcase

Updated 9 minutes ago


 Filter cards

+ Add cards


 Fullscreen

 Menu


8 Backlog + ...

 Add plotly charts to pricing engine ...


Added by olgamie

 Design client pricing view ...


Added by olgamie

 Add autosave ...


Added by olgamie

 Add internationalization ...


Added by olgamie

 Add frontend tests ...

Added by olgamie


 Add login ...


Added by olgamie


 Deploy to RStudio Connect ...

Added by olgamie

2 ToDo + ...


 Create transactions chart ...

scRipts#1 opened by olgamie 

 Create a mockup ...


scRipts#2 opened by olgamie

1 Implementation Plan + ...

 Create pricing engine ...


scRipts#3 opened by olgamie


1 In Progress + ...


 Set up project dev env ...

scRipts#4 opened by olgamie


3 In Review + ...

 Add lintr ...

scRipts#5 opened by olgamie 

 Set up repository ...

scRipts#7 opened by olgamie

 Add CI ...

scRipts#6 opened by olgamie

Follow the Rules

→ WELL-ORGANIZED PROJECT BOARD

Overview of what is happening.

→ TASK FOR EVERYTHING

All work needs to be **clearly documented**. PR = task.

→ WRITE IMPLEMENTATION PLAN

Make sure that any proposed change **makes sense**.

→ SATISFY ALL TASK REQUIREMENTS

Reviewers make sure the **solution corresponds to task requirements**.

Code Review

Checklist

Version Control

1 ● ● ● ● ● ●

All code **peer-reviewed**.

2 ● ● ● ● ● ●

Continuous integration checks
(linter, unit tests, integration tests) configured and passed.

3 ● ● ● ● ● ●

Any added or modified code **follows style-guide**.

4 ● ● ● ● ● ●

Change has been **tested**
(manually or with automated tests).

5 ● ● ● ● ● ●

No new errors or warnings.

6 ● ● ● ● ● ●

README, documentation,
and code comments updated.

Continuous Integration

Github Actions

→ TRIGGER ON EVENT

Automatically run e.g. unit tests after push and check the build on different operating systems

→ READY TO USE TOOLS

Find what you want in the marketplace or build your own

Marketplace / Search results

Types

Apps

Actions

Categories

API management

Chat

Code quality

Code review

Continuous integration

Dependency management

Deployment

IDEs

Learning

Localization

Mobile

Monitoring









Project management

Search for apps and actions

Actions

An entirely new way to automate your development workflow.

3934 results filtered by Actions

- **ansible-lint**
By ansible ✓
Run Ansible Lint
62 stars
- **Trigger Buildkite Pipeline**
By buildkite ✓
A GitHub Action for triggering a build on a Buildkite pipeline
14 stars
- **GitHub Action for Cloudflare Workers**
By cloudflare ✓
Deploy a Cloudflare Worker with the Serverless Framework
59 stars
- **Velocity deploy action**
By codeclimate ✓
A GitHub Action for sending deployment information to Velocity
16 stars
- **Codecov**
By codecov ✓
GitHub Action that uploads coverage reports for your repository to codecov.io
325 stars
- **Coveralls GitHub Action**
By coverallsapp ✓
Send test coverage data to Coveralls.io for analysis, change tracking, and notifications
121 stars
- **Glo Parse Card Links**
By AxoSoft ✓
GitHub Action to parse links to Glo Boards cards
- **Glo Add Label To Cards**
By AxoSoft ✓
GitHub action to add a label to Glo Boards cards



Github Templates

ProTip

collabrica / shakecast

Watch ▾

88

★ Star

247

🍴 Fork

14

<> Code

🔔 Issues 1

🔗 Pull requests 0

▶ Actions

📁 Projects 0

📖 Wiki

🛡 Security

📊 Insights

⚙ Settings

Options

[Collaborators](#)

[Branches](#)

[Webhooks](#)

[Notifications](#)

[Integrations & services](#)

[Deploy keys](#)

[Secrets](#)

[Moderation](#)

[Interaction limits](#)

Settings

Repository name

shakecast

Rename

☒ Template repository

Template repositories let users generate new repositories with the same directory structure and files. Indicate if collabrica/shakecast can be used as a template for creating other repositories.

Social preview

Upload an image to customize your repository's social media preview.

Images should be at least 640×320px (1280×640px for best display).

[Download template](#)



Development Environment

Lack of Development Environment

antipattern



CODE ERRORS

The same code can give different results on different machines.



DIFFICULT DEPLOYMENTS

Establishing infrastructure is challenging if not tracked.



AFFECTING OTHER PROJECTS

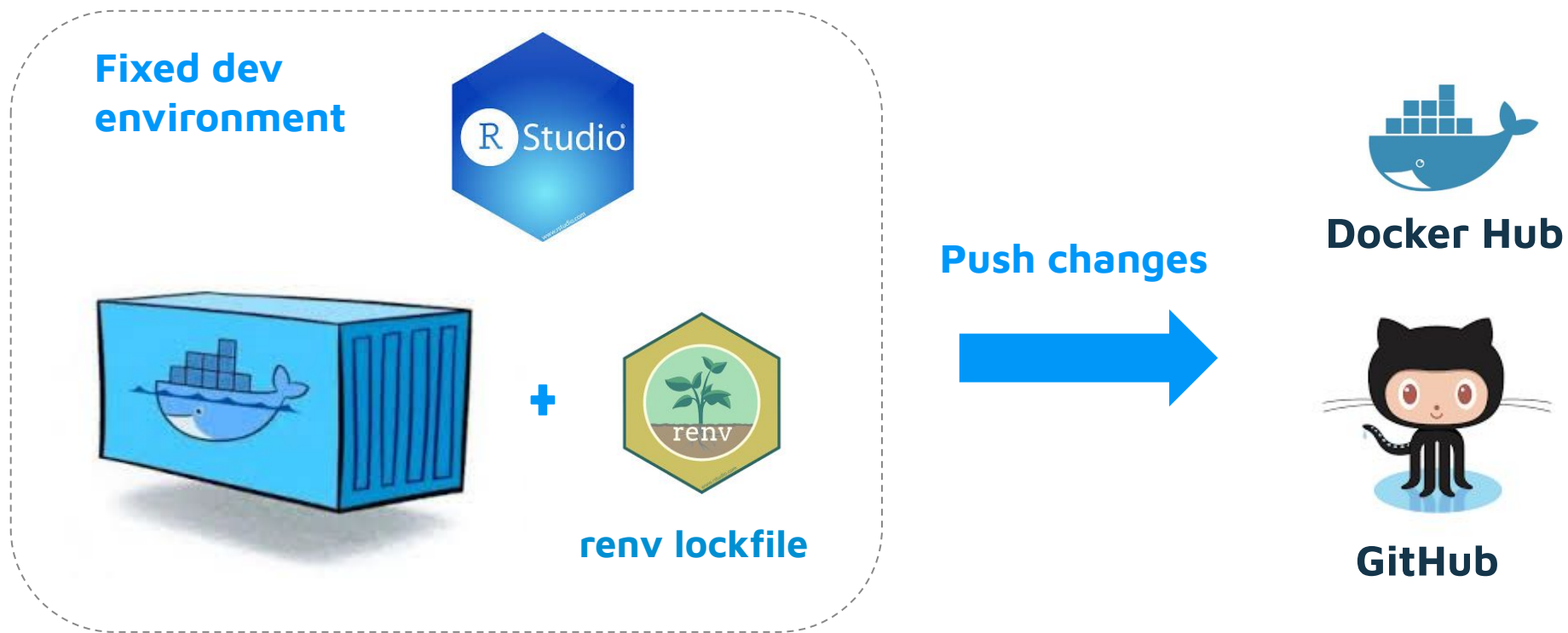
In a global environment, shared packages can change and crash unrelated projects.



PAINFUL COLLABORATION

Team will waste time setting up a new environment rather than jumping straight to work.

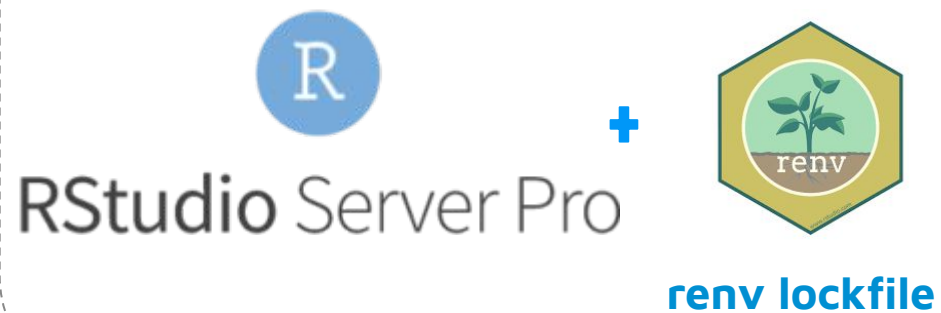
Docker + renv to the rescue!



Dedicated to a project

Team at scale

**Secure, flexible
and collaborative**



Push changes



GitHub

Takeaways

- USE SCRUM & WELL-ORGANIZED PROJECT BOARD
- MAKE PROPER CODE REVIEW PART OF YOUR WORKFLOW
- WORK IN ISOLATED & SHAREABLE DEVELOPMENT ENVIRONMENT

Thank you!



@olgamie



olga@appsilon.com

appsilon.com