MANGO SOLUTIONS

An Operating Model for R

Mark Sellors - Mango Solutions
Email: msellors@mango-solutions.com
Twitter: @sellorm

About me

- Head of Data Engineering
- Founded Data Engineering London
- Automation
- Productionisation
- API's
- Pipelines
- Industrialisation
- Mango Labs

Warning!

This presentation contains some sweeping generalisations!


I build bridges

Bridges are about:

- The unknown
- Seeing what's on the other side (exploration)
- Getting from here to there
- Efficiency - moving people and goods **over** rather than **around**
- Joining things together (communities)

I build bridges in two ways

I sit between Data Science and IT. I speak both languages well enough to understand the needs of both and to effectively communicate between the two.

I build bridges in two ways

Getting any piece of data science to run in a production setting is an act of bridge building. Any solution must get from A (where we are now) to B (where we want to be) as **efficiently**, **repeatably** and **supportably** as possible

Real Bridges come with constraints

- The need to carry a certain traffic type
- The need to carry a certain volume of traffic
- The need to span a particular distance
- The need to allow certain things to pass underneath (eg tall ships)

What is an Operating Model?

- Business strategy
- Operational design that implements that strategy

Why is it useful?

- Helps to define the implementation strategy within the business
- Provides a clean map for how things should work within the business

MANGO SOLUTIONS

Tips to implement an Operating Model for R


Start small


Think about what you need

Grow your operating model


Focus on
repeatability


Strong support


Plan for change


Share

So what are we doing?

- Driving business adoption
- Making R supportable in a business
- Providing ‘standards’
- Ensuring everyone understands the target environment
- Keeping the ops teams happy
  - They need to know how to support this stuff
  - Not everyone has a dedicated DE team


What are we
not doing?

Shiny Adoption at Mango


Become a bridge builder

The three P's of a successful Operating Model

- Policy
  - What to do
- Procedure
  - How to do it
- People
  - Skills, training and Community

Other important things

- Must be a place for experimentation - either locally or on a server - full on bleeding edge stuff
  - That's how you advance the state of the art
  - That's the 'science' bit
- Have an upgrade plan - doesn't matter what it is
  - When will you upgrade to newer versions of R or packages?
  - What does the roll-out plan look like?
  - Beware of trying to support lots of different versions of things
  - Deprecation strategy

Other important things

- Must also consider how to advance legacy systems
  - It's almost always the case that not enough consideration is given to maintaining legacy systems
  - As soon as a data product is implemented it's a legacy tool and requires an upgrade plan/path
- Be prepared for change
  - Adaptability is the key to survival

Thanks!