

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Learning from errors: my academic career

Jesus M. Gonzalez-Barahona

Universidad Rey Juan Carlos
@jgbarah <http://github.com/jgbarah/presentations>

ICSME Doctoral Symposium
Madrid (Spain), September 25th 2018

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

*“Once upon a time
there was a wannabe researcher
in the Deep South of Madrid...”*

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Starting conditions



Learning from
errors

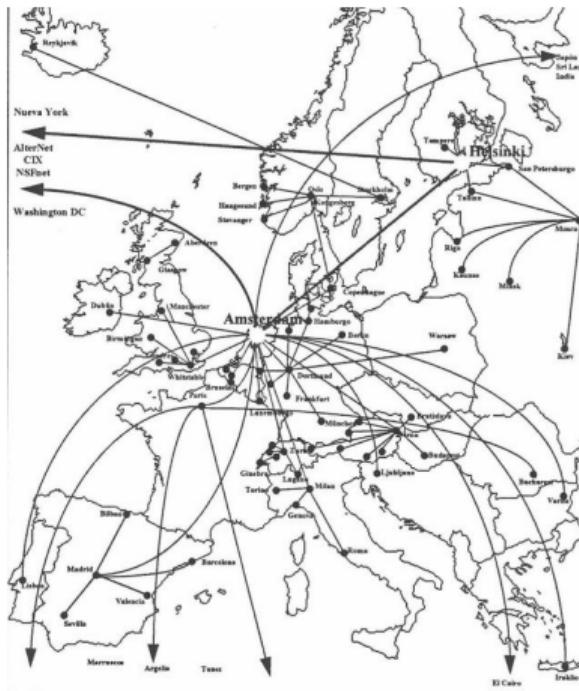
Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Being in the right place



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Being in the right place (lesson)

Look for exciting stuff
stuff that will change the world
even it doesn't seem that academic
and understand it

Look at the future, we're doing it

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Being at the right moment

```
Booting from Floppy...
Seek error 128, req = 0, at = -1
unit 0, type 0, sectrac 18, blknum 17
386BSD Release 0.1 by William and Lynne Jolitz. [0.1.24 07/14/92 19:07]
Copyright (c) 1989,1990,1991,1992 William F. Jolitz. All rights reserved.
Based in part on work by the 386BSD User Community and the
BSD Networking Software, Release 2 by UCB EECS Department.
pc0<color> at 0x60 irq 1 on isa
com1 fifo at 0x3f8 irq 4 on isa
com2 fifo at 0x2f8 irq 3 on isa
wd0 <ST1102AT> at 0x1f0 irq 14 on isa
fd0 drives 0: 1.2M at 0x3f0 irq 6 drq 2 on isa
ne0 ethernet address fe:fd:de:be:ef at 0x300 irq 9 on isa
npx0 at 0xf0 irq 13 on isa
changing root device to fd0a

warning: no swap space present (yet)
386BSD Distribution Installation Floppy (Tiny 386BSD) Release 0.1

Please read the installation notes (type 'zmore INSTALL.NOTES')
and registration information (type 'more REGISTRATION') before use.
To install on hard disk drive, type 'install'.

erase ^?, werase ^H, kill ^U, intr ^C
#
```

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Being at the right moment (lesson)

Play with cool stuff

Experiment

Build on the work of others

Have fun

Learn by doing

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Having the right mentors



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Having the right mentors (lesson)

Network with interesting people

You understand when you can explain

Details are (very) important

Yes, I can
(English skills are fundamental)

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Having the right colleagues



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Having the right colleagues (lesson)

Having great people around is very important
Playing in a team is much more fun
Together you can reach higher

Let's try it

Learning from
errors

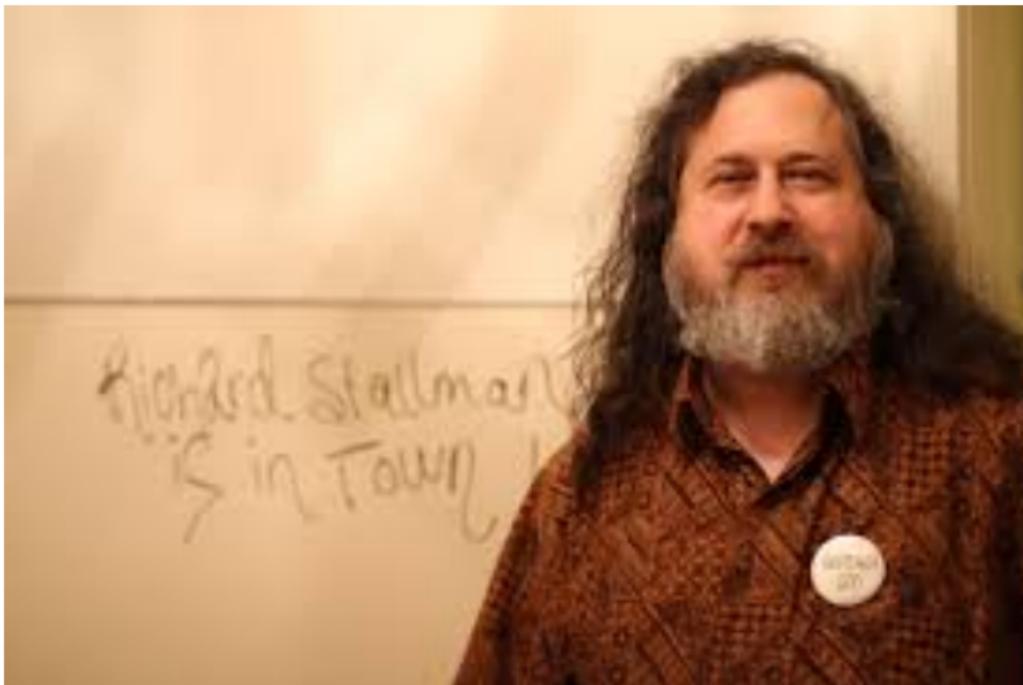
Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Having the path paved



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Having the path paved (lesson)

The importance of sharing

The importance of ethics

The importance of tools

Have principles, but have also a plan

Learning from
errors

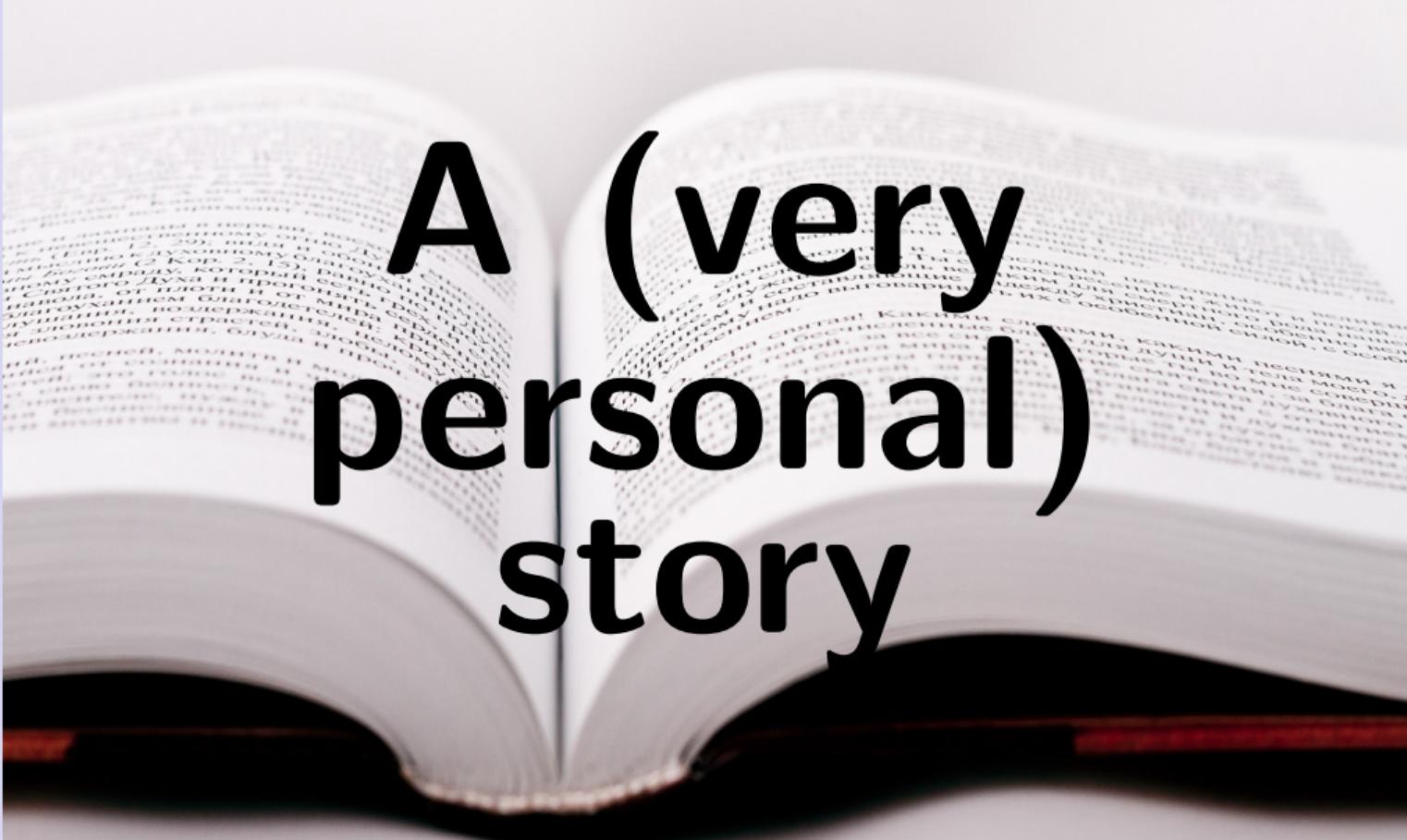
Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

A (very personal) story



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Labs based on free software (1993-2005)



Libre Software for Computer Science Classes

In 1993, the authors began using 386BSD to teach computer science classes at Madrid's Carlos III University, in Spain. Seven years later, NetBSD and GNU/Linux are the operating systems of choice for

Jesus M. Gonzalez-Barahona, Pedro de-las-Heras-Quiros, José Centeno-González, Vicente Matellán-Olivera, and Francisco J. Ballesteros, King Juan Carlos University, Spain

During 1991 and 1992, the landscape of libre software, and of software development in general, was ready for a change. In two different communities, two very exciting events were taking place: 386BSD, a libre derivative of the BSD code, was born, and Linus Torvalds distributed the first versions of the Linux kernel. Soon after, the software community had two libre operating systems at its disposal (for some basic references to online information on libre products and distribution, see the "Online Resources" and "Terminology" sidebars).

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Labs based on free software (lesson)

Cool stuff may need a lot of time
a lot
a whole lot

But it may pay the effort

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

PhD thesis & Lower_Layer (1995-2002)

Teaching Network Programming
with Ada and *Lower_Layer**

Jesús M. González-Barahona, José Centeno-González, Pedro de las Heras-Quirós
Francisco J. Ballesteros-Cámara, Luis López-Fernández

Grupo de Sistemas y Comunicaciones (GSyC)

Universidad Carlos III de Madrid

Phone no: +(34-91) 624-9497. Fax no: +(34-91) 624-9430

Butarque 15, E-28911 Leganés (Madrid). Spain.

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

PhD thesis & Lower_Layer (lesson)

Details are important,
but getting things done too
Doing is important
but focus too

It's difficult to identify the main result

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Free software in Europe (1997-2005)



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Free software in Europe (lesson)

You can do cool stuff
at many levels
with cool people

It may count as zero for your academic life

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Counting potatoes (2000-2001)

Counting Potatoes: the Size of Debian 2.2

Jesús M. González-Barahona, Miguel A. Ortuño Pérez, Pedro de las Heras Quirós, José Centeno González and Vicente Matellán Olivera

Debian is the largest Free Software distribution, with well over 2,800 source packages in the latest stable release (Debian 2.2) and more than 4,000 source packages in the release currently in preparation. But, how large is “the largest”? In this paper, we use David Wheeler’s sloccount system to determine the number of physical source lines of code (SLOC) of Debian 2.2 (aka Potato). We show that Debian 2.2 includes over 56,000,000 physical SLOC (almost twice than Red Hat 7.1, released about 8 months later), showing that the Debian development model (based on the work of a large group of voluntary developers spread around the world) is at least as capable as other development methods (like the more centralized one, based on the work of employees, used by Red Hat or Microsoft) to manage distributions of this size.

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Counting potatoes (lesson)

Any detail may change a life

In some moments you're ready for a change
and you feel you have been preparing for it

Switching fields may be rebooting from scratch

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Using data from CVS (lesson)

Studying the evolution of libre software projects using publicly available data

Gregorio Robles-Martínez, Jesús M. González-Barahona,
José Centeno-González, Vicente Matellán-Olivera, and Luis Rodero-Merino
GSyC, Universidad Rey Juan Carlos
{grex,jgb,jcenteno,vmo,lrodero}@gsyc.escet.urjc.es

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Using data from CVS (2001-2003)

The are niches of opportunity
that you can exploit better than others
but quickly you need to learn new stuff

Starting in a new research field
is not always easy

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

CVSAnalY (2002-2015)

Remote analysis and measurement of libre software systems by means of the CVSAnalY tool

Gregorio Robles
Universidad Rey Juan Carlos
grex@gsyc.escet.urjc.es

Stefan Koch
Wirtschaftsuniversität Wien
stefan.koch@wu-wien.ac.at

Jesús M. González-Barahona
Universidad Rey Juan Carlos
jgb@gsyc.escet.urjc.es

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

CVSAnalY (lesson)

Tools are important and cool

Tools help you learn the details

Tools can be done in collaboration

Good tools don't guarantee publication

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Macro-evolution (2005-2008)

Empir Software Eng (2009) 14:262–285
DOI 10.1007/s10664-008-9100-x

**Macro-level software evolution: a case study
of a large software compilation**

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Macro-evolution (lesson)

A solid work may take a little while...
...with a little help from friends
and your past experience

But you need to find the right approach

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Teaching about free software (2002-2008)

Introducción al software libre

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Teaching about free software (lesson)

Learning by teaching

Gaining perspective

Observing the landscape

Having a looooot of fun

Teaching takes a looooot of time

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

FLOSSMetrics (2006-2009)

Project funded by the European Comission under contract number FP6-033982

FM is part of the FLOSSQuality initiative



Free/Libre and Open Source Software Metrics

Definition

What's It?
Details

Project News

First version of the data web interface
Deliverable D3.2 "Database" released

The FLOSSMetrics project

FLOSSMetrics stands for Free/Libre Open Source Software Metrics.

The main objective of FLOSSMETRICS is to construct, publish and analyse a large scale database with information and metrics about libre software development coming from several thousands of software projects, using existing methodologies, and tools already developed. The project will also provide a public platform for validation and industrial exploitation of results.

FLOSSMetrics targets

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

FLOSSMetrics (lesson)

Funded projects bring resources
They allow for implementing ambitious ideas
They put together interesting teams

Funded projects may be (very) distracting

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Studying Wikipedia (2007-2009)

Proceedings of the 41st Hawaii International Conference on System Sciences - 2008

On The Inequality of Contributions to Wikipedia

Felipe Ortega, Jesus M. Gonzalez-Barahona and Gregorio Robles

Libresoft Group, Universidad Rey Juan Carlos

Tulipan, s/n 28933. Mostoles.

Madrid. SPAIN.

Email: {jfeliпе,jgb,grex}@gsyc.es

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Studying Wikipedia (lessons)

You can explore new fields
with some advantage
because you have the right background
and the right people around

Did I say that new fields may be like
rebooting from scratch?

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Reproducibility (2008-2012)

Empir Software Eng (2012) 17:75–89
DOI 10.1007/s10664-011-9181-9

**On the reproducibility of empirical software
engineering studies based on data retrieved
from developer repositories**

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Reproducibility (lesson)

You can step back for a moment
look at things from a new perspective
and find a new use for what you know

Literature surveys may be fun, but hard too

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Evolution based in data (2008-2013)

Studying the evolution of libre software projects using publicly available data

Gregorio Robles-Martínez, Jesús M. González-Barahona,
José Centeno-González, Vicente Matellán-Olivera, and Luis Rodero-Merino
GSyC, Universidad Rey Juan Carlos
{grex,jgb,jcenteno,vmo,lrodero}@gsyc.escet.urjc.es

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Evolution based in data (lesson)

Reproducing past research is fun
and interesting
and helps to advance the state of the art
It may lead to new lines

Reproducing is not always easy

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Companies and communities (2010-2013)

Understanding How Companies Interact with Free Software Communities

progress of new analytics techniques to study them. In turn, these techniques have started to produce useful results for both practitioners and other stakeholders.²

In particular, studies of company participation in large projects have started to raise industrial interest: FLOSS foundations want to learn about company participation in their projects, and the companies want to know more about corporate activity in the projects on which they rely. Here, we present two studies in this area—by LibreSoft, a research group special-

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Companies and communities (lesson)

The real world is there, and it's nice
If you can explain a part of it
that will be appreciated

The real world is dirty, complex,
difficult to understand

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

MetricsGrimoire

Tools for software & development
community analytics



Update: Check the new GrimoireLab
platform

MetricsGrimoire (2002-2015)

MetricsGrimoire

[Update]: Currently, main development related with *MetricsGrimoire* has stopped. Main contributors are working on a new platform: **GrimoireLab**

MetricsGrimoire (pronounced /'metrɪksgrɪm'wər/) is a toolset to obtain data from repositories related to software development: source code management (aka *version control*), issue tracking (aka *bug reporting*) systems, mailing lists, etc. Data and metadata about the software development processes is retrieved from those repositories (information about commits, ticket management, communication in mailing lists, etc.), and then organized and stored into SQL databases that can later be mined for specific patterns or summaries of activity.

MetricsGrimoire tools support many kinds of repositories, including those provided by GitHub (git

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

MetricsGrimoire (lesson)

Tools keep you anchored in reality

Tools give you a competitive advantage

Tools let you learn the details

Tools may be distracting,
difficult to “convert” in papers

Learning from
errors

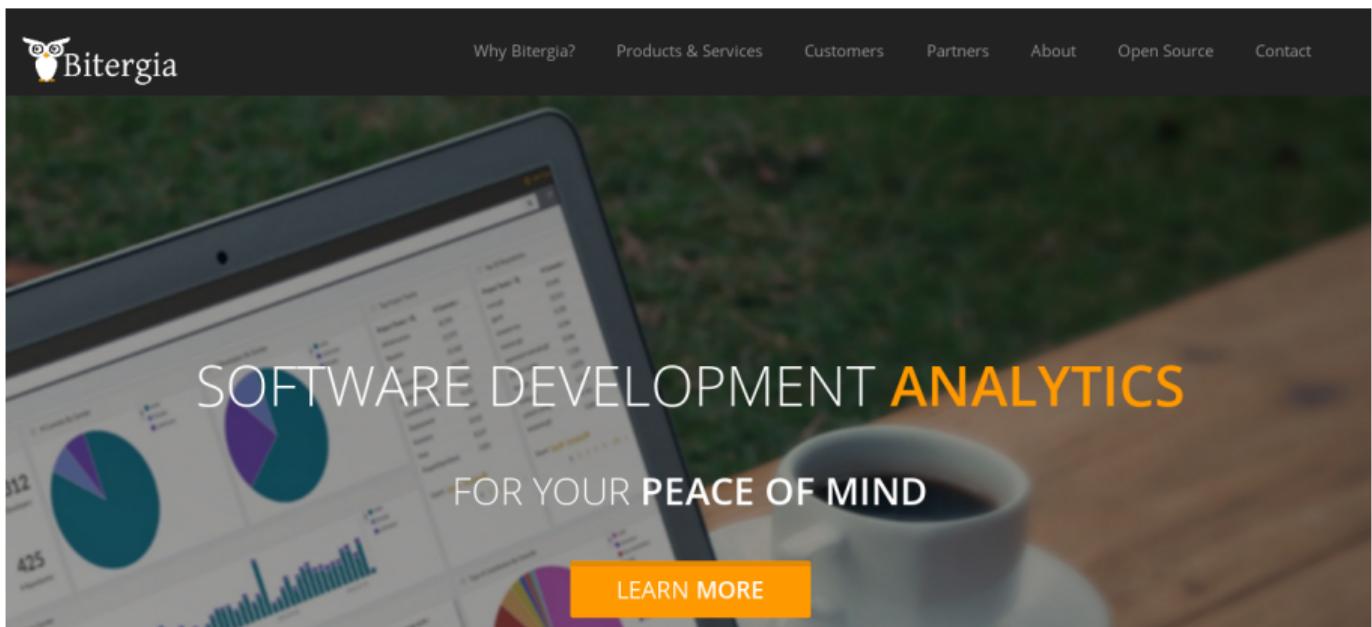
Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Bitergia (2012-)



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Bitergia (2012-)

Startups are a lot of fun
They link you to reality,
and see a new side of reality

Startups are a huge sink for your time

Learning from
errors

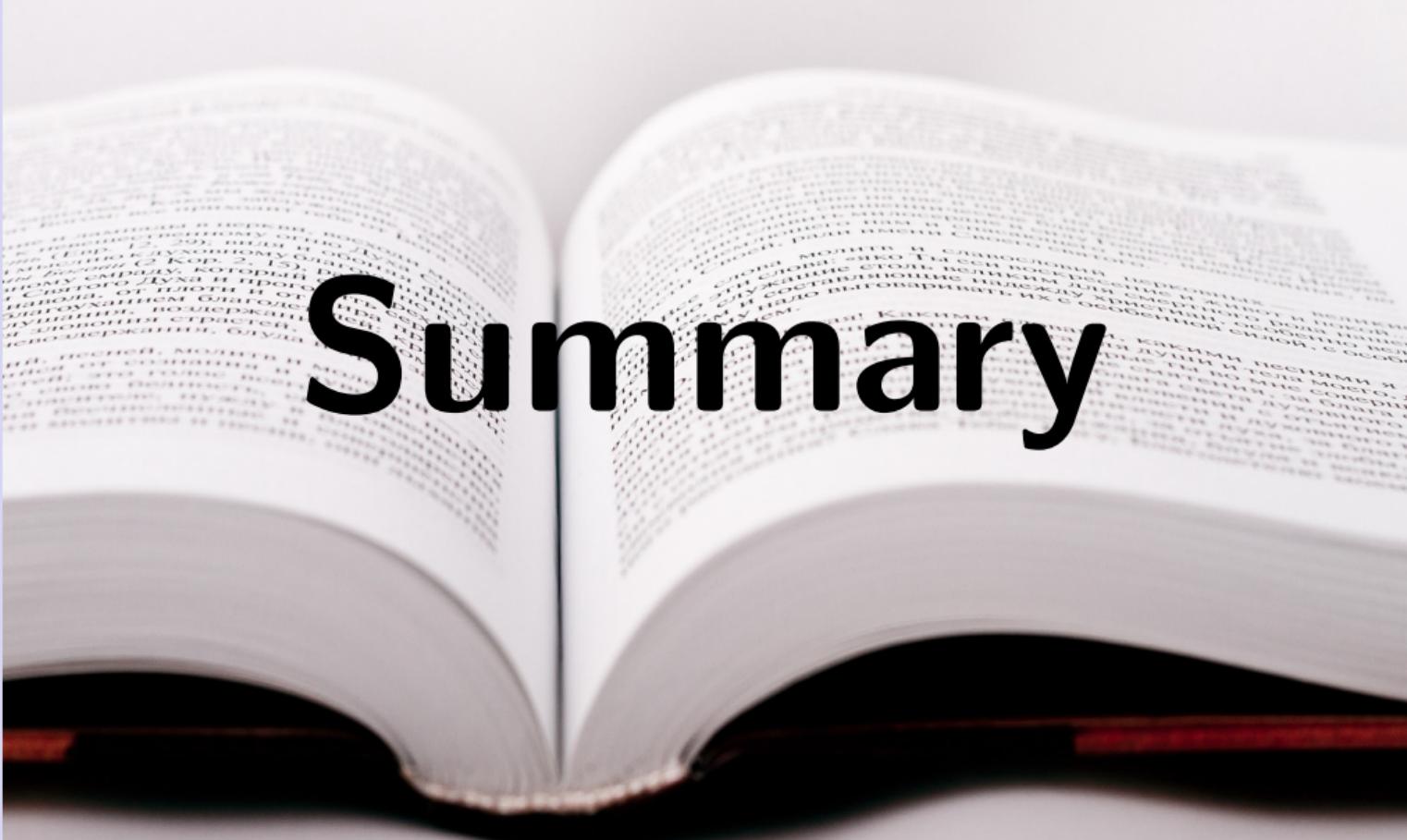
Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

Summary



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary



Summary

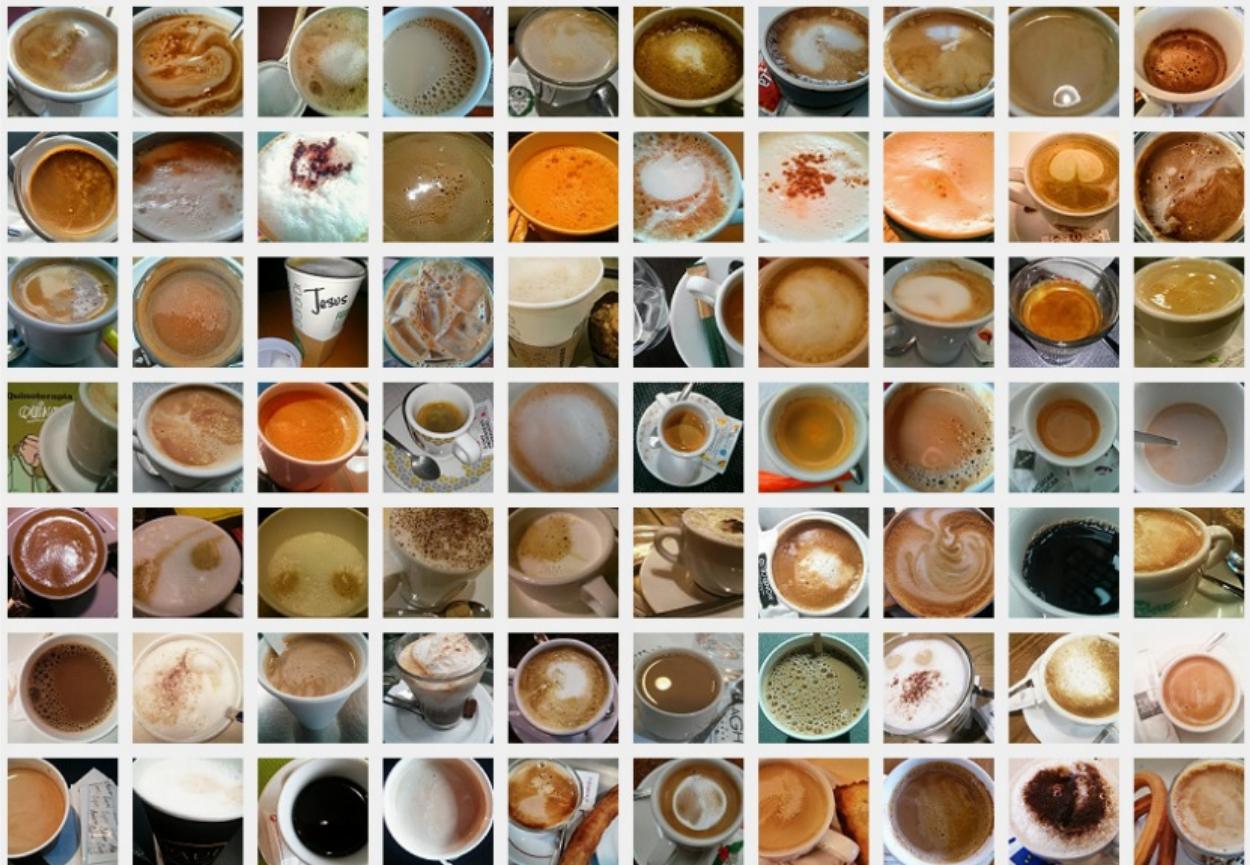
Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary



Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary

*Finding a balance
between having fun
and having a career*

Learning from
errors

Jesus M.
Gonzalez-Barahona

Starting conditions

A (very personal)
story

Summary



©2018 Jesus M. Gonzalez-Barahona.

Some rights reserved. This document is distributed under the terms of the Creative Commons License "Attribution-ShareAlike 4.0", available in

<http://creativecommons.org/licenses/by-sa/4.0/>

This document (including source) is available from
<https://github.com/jgbarah/presentaciones>