

Why Getting Good Test Data is Hard

(and what to do about it)

Tina Fletcher

@fletchertinam

Tina Fletcher

@fletchertinam

@D2L - Senior Test Strategist

@KWSQA - President



Test Data Challenges

Not
Glamorous
Avoided

Sounds
Easy
Underestimated

Not
“Testing”
Forgotten

Test Data Consequences

Not
Glamorous
Avoided

Necessary
prerequisite

Sounds
Easy
Underestimated

Surprisingly
challenging

Not
“Testing”
Forgotten

Unanticipated
costs



MEDIA PLAYER

Record and play back
audio/video files



EMAIL CLIENT

Send and receive
emails



LEARNING PLATFORM

Offer and participate in
online courses

Introducing our assistants.

Applications I have worked on with interesting test data requirements

10 Useful Test Data Questions to Ask

**When Starting a New
Software Project**

1.

**What is test data, in
our context?**

1. What is test data, in our context?

Define test data
units



Media Files



Emails



Courses, Users

1. What is test data, in our context?

Examples can help to identify scope and variables



<1 sec 128kbps mp3 file



Email sent to 2 recipients, 3 image attachments, HTML table in body, AS 12.1 provider



Course with >1500 students, 6 quizzes and heavy group discussion usage



User registered as a student in 3 courses and as a teaching assistant in another

2.

**Why does test data
matter to us?**

2. Why does test data matter to us?

Reasons it matters in general...

- Need test data in order to execute tests
- Want to test under conditions that resemble production
- Many team members benefit (developers, testers, designers, doc writers, product owners, sales, etc)

2. Why does test data matter to us?

For our project, we want to be confident that...



...users can play all common file types without errors



...users can read and interact with any email message they receive



...the learning platform's performance is not impacted by large class sizes



...permissions in the learning platform are properly restricted according to roles

3.

**What should our test
data look like?**

3. What should our test data look like?

Random

VS

Controlled



**mailing list subscription
or
crafting specific emails**

3. What should our test data look like?

Realistic

VS

Simplified



test combinations of features
or
one thing at a time

3. What should our test data look like?

**Understand
functionality and
performance
expectations**



- **support claims for file types**
- **load time before playback starts**

3. What should our test data look like?

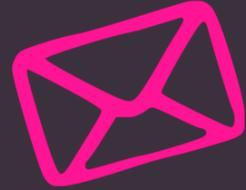
**Assess risk and
prioritize
accordingly**



**emails that have been forwarded
across multiple providers**

3. What should our test data look like?

**Decide what is in
and out of scope**



top 25 email providers

3. What should our test data look like?

Understand your customers



common configurations of permission settings

4.

**How can we get test
data?**

4. How can we get test data?

Use multiple methods to maximize coverage and variation

- **Test Data Management suites**
(generation, analysis, cleaning)
- **Copy from existing source**
 - test data repository
 - production



- **Generate with free tools**
 - generatedata.com
 - randomuser.me
 - mockaroo.com



4. How can we get test data?

Use multiple methods to maximize coverage and variation

- **Build a custom tool**
 - email spammer (volume)



- **Mock APIs**
 - getsandbox.com
 - mockable.io



- **Use your own product**
 - corporate email on phones
 - internal training courses



4. How can we get test data?

Use multiple methods to maximize coverage and variation

- **Generate via automated tests**
 - simulate student activity



- **Create manually**
 - compose emails
 - record videos
 - log in as student users



5.

**What are our test
data challenges?**

5. What are our test data challenges?

Sometimes, you just can't win!

- **Fake data is too fake!**
 - inconsistencies
 - old timestamps
 - harder to spot issues



- **Real data is too real!**
 - privacy
 - masking
 - copyright



5. What are our test data challenges?

Sometimes, things
cost money



copyrighted files



paid accounts



hosting costs

6.

**Where will we keep
test data?**

7.

**How will we move
and share test data?**

6. Where will we keep test data?

7. How will we move and share test data?

Copy from a central location



DRM restrictions



security/privacy concerns



created against older version

6. Where will we keep test data?

7. How will we move and share test data?

Shared accounts or test environments

- return to known state
- avoid irreversible change
- simultaneous usage conflicts



**6. Where will we
keep test data?**

**7. How will we move
and share test data?**

**Back up your test
data**



8.

**Is our test data
effective?**

8. Is our test data effective?

**More questions to
ask yourself**

Are you finding bugs?

Are they valid?

Are they important?

Are other people finding bugs?

Can you test everything you want to?

9.

**What test data do we
not have?**

9. What test data do we not have?

Understand and mitigate

Beta or crowd source testing

More unit tests

Review usage patterns

Communicate

New ideas from unexpected sources

Prepare stakeholders

10.

**What test data
legacy will we leave?**

10. What test data legacy will we leave?

Help your future self, and others

- What you have (& why)
- What you don't have (& why)
- Where to find it
- How to use it
- How to add to it
- How to change it

Conclusion

**Why is getting
good test data
hard?**

- Many variables, data characteristics
- Many sources, generation methods
- Balance realism, privacy, \$\$, time
- Tricky to move, store, share, tweak

Conclusion

**What can you do
about it?**

- Plan ahead
- Ask the 10 Questions
- Keep improving



Over to you!

What other questions would you add to this list?

What creative test data solutions are you proud of?

What challenges have you encountered related to test data?

What else did I miss?

THANKS

@fletchertinam

Credits

Shapes & Icons

Vectorial Shapes in this Template were created by [SlidesPPT](#)

<http://www.flaticon.com/packs/hipster-style-2>

Vector Design by Vecteezy

Images

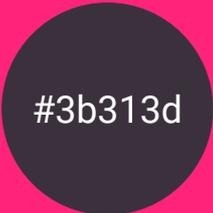
The photos in this template have been downloaded from www.gratisography.com

Fonts

The fonts used in this template are taken from Google fonts. (Quicksand, Anton)

Download the fonts from the following url:
<https://www.google.com/fonts/>

Colors



#3b313d



#514454



#ffffff



#ff2479