

## A draft experience report on the use of StreamData for

# Property Testing

## A draft experience report on the use of StreamData for

# Property Jesting

in Elixir

## A draft experience report on the use of StreamData for

# Property lesting

## in Elixir at my last job



## Rob Howard @damncabbage http://robhoward.id.au





# What is **Property Testing?** (and StreamData)







## test "word-count" do text = "This is a test. And this too!"

assert WordCount.count(text) == 7
end

# est. And this too!"

test "word-count" do words = ["This", "is", "a", "test.", "And", "this", "too!", text = Enum.join(" ", words) assert WordCount.count(text) == 7 end

def word do
 Enum.random([
 "This", "is", "a",
 ])
end

test "word-count" do word\_count = Enum.random(1..10) text = (1..word\_count) |> Enum.map(fn \_ -> word() end) |> Enum.join(" ") assert WordCount.count(text) == word\_count end

def word do
 Enum.random([
 "This", "is", "a",
 ])
end

test "word-count" do
 word\_count = Enum.random(1..10)
 text = (1..word\_count)
 |> Enum.map(fn \_ -> word() end)
 |> Enum.join(" ")
 assert WordCount.count(text) == word\_count
end

def word do
 Enum.random([
 "This", "is", "a",
 ])
end

test "word-count" do word\_count = Enum.random(1..10) text = (1..word\_count) |> Enum.map(fn \_ -> word() end) |> Enum.join(" ") assert WordCount.count(text) == word\_count end

property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre>

do

text = words |> Enum.join(" ") assert WordCount.count(text) == count

end end



property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre>

do

text = words |> Enum.join(" ") assert WordCount.count(text) == count end end



property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre>

do

text = words |> Enum.join(" ") assert WordCount.count(text) == count end end



property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre>

do

text = words |> Enum.join(" ") assert WordCount.count(text) == count end

end



property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre>

do

text = words |> Enum.join(" ") assert WordCount.count(text) == count end end



property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre>

do

text = words |> Enum.join(" ") assert WordCount.count(text) == count

end end





# Stream Data





based testing is to find the properties we want our code to hold. Once a property is found, we can use those properties to complement our example-based tests.

At ElixirConf US 2017, we have announced that a property testing library will be part of Elixir v1.6. Our goal with this post is not to answer the technical questions behind StreamData but rather explain why it is being added to the language. For more information on property testing per se, the first three chapters of Fred's book 106 is a great starting point. To learn more about StreamData itself, see its announcement 76.



will have something out in October, having to wait until the next April to be able to talk about it in public, it is definitely too long.

For example, I announced StreamData for Elixir before it was part of master, and it turns out that it won't be part of Elixir core anyway. But the discussions that happened in the months after the announcement were very productive.

So I don't think doing the announcement before having something out is bad. We already have another thread about LiveView with @tmbb @qu/ch and Ldiscussing possible implementations. Lwould just



based testing is to find the properties we want our code to hold. Once a property is found, we can use those properties to complement our example-based tests.

At ElixirConf US 2017, we have announced that a property testing library will be part of Elixir v1.6. Our goal with this post is not to answer the technical questions behind StreamData but rather explain why it is being added to the language. For more information on property testing per se, the first three chapters of Fred's book (106) is a great starting point. To learn more about StreamData itself, see its announcement 76.

will nave sometning out in October, naving to wait until the next April to be able to talk about it in public, it is definitely too long.

For example, I announced StreamData for Elixir before it was part of master, and it turns out that it won't be part of Elixir core anyway. But the discussions that happened in the months after the announcement were very productive.

So I don't think doing the announcement before having something out is bad. We already have another thread about LiveView with @tmbh @anch and I discussing possible implementations. I would just





based testing is to find the properties we want our code to hold. Once a property is found, we can use those properties to complement our example-based tests.

At ElixirConf US 2017, we have announced that a property testing library will be part of Elixir v1.6. Our goal with this post is not to answer the technical questions behind StreamData but rather explain why it is being added to the language. For more information on property testing per se, the first three chapters of Fred's book (106) is a great starting point. To learn more about StreamData itself, see its announcement 76.

will nave sometning out in October, naving to wait until the next April to be able to talk about it in public, it is definitely too long.

For example, I announced StreamData for Elixir before it was part of master, and it turns out that it won't be part of Elixir core anyway. But the discussions that happened in the months after the announcement were very productive.

So I don't think doing the announcement before having something out is bad. We already have another thread about LiveView with @tmbh @aruch and Ldiscussing possible implementations. I would just







based testing is to find the properties we want our code to hold. Once a property is found, we can use those properties to complement our example-based tests.

At ElixirConf US 2017, we have announced that a property testing library will be part of Elixir v1.6. Our goal with this post is not to answer the technical questions behind StreamData but rather explain why it is being added to the language. For more information on property testing per se, the first three chapters of Fred's book (106) is a great starting point. To learn more about StreamData itself, see its announcement 76.

will nave sometning out in October, naving to wait until the next April to be able to talk about it in public, it is definitely too long.

For example, I announced StreamData for Elixir before it was part of master, and it turns out that it won't be part of Elixir core anyway. But the discussions that happened in the months after the announcement were very productive.

So I don't think doing the announcement before having something out is bad. We already have another thread about LiveView with @tmbh @anch and I discussing possible implementations. I would just





# Generators

example = constant("Hello, World")

example = constant("Hello, World")

example = constant("Hello, World")
# example :: StreamData.t(String.t)

example = constant("Hello, World")
# example :: StreamData.t(String.t)

example = constant("Hello, World")

example = constant("Hello, World")

example = constant("Hello, World")
word = member\_of([ "Hello", "World", "Tokyo", "Perth", check all a\_word <- word do

assert String.length(a\_word) == 5 end

word = member\_of([ "Hello", "World", "Tokyo", "Perth", ])

check all a\_word <- word do assert String.length(a\_word) == 5 end

word = member\_of([...])
words = list\_of(word)

# check all original <- words do

word = member\_of([...])
words = list\_of(word)

# check all original <- words do

word = member\_of([...]) words = list\_of(word) sentence = list\_of(words) |> ...

# check all original <- words do

word = member\_of([...]) words = list\_of(word) sentence = list\_of(words) |> ...

# check all original <- words do



def word, do: member\_of([...]) def words, do: list\_of(word()) def sentence do

















def word, do: member\_of([...]) def words, do: list\_of(word()) def sentence do gen all words <- words()</pre> ending <- member\_of(~W[. ! ? ?])</pre>

do

words

|> Enum.join(" ") >> String.capitalize() > Kernel.<>(ending) end end

def word, do: member\_of([...]) def words, do: list\_of(word()) def sentence do gen all words <- words()</pre> ending <- member\_of(~W[. ! ? ?])</pre>

do

## words

|> Enum.join(" ") >> String.capitalize() > Kernel.<>(ending) end end

```
def naive_datetime() do
  gen all date <- date(),</pre>
           time <- time()</pre>
  do
    {:ok, ndt} = NaiveDateTime.new(date, time) do
  end
end
def date do
  gen all year <- integer(2010..2040),
           \{:ok, jan1\} = Date.new(year, 1, 1),
           days <- integer(0..(if Date.leap_year?(jan1), do: 365, else: 364))</pre>
  do
    Date.add(jan1, days)
  end
end
def time do
  gen all hour <- integer(0..23),</pre>
           min <- integer(0..59),</pre>
           sec <- integer(0..59),</pre>
           usec <- integer(0...999_999)</pre>
  do
    {:ok, time} = Time.new(hour, min, sec, usec)
    time
  end
end
```

## **StreamData** v0.4.2

Q search

PAGES MODULES **EXCEPTIONS** 

ExUnitProperties

Тор Summary + Functions

StreamData

Тор Summary

+ Types + Functions Interest della deconante este grient predicate fantettent

fixed\_list(datas)

 $\equiv$ 

fixed\_map(data)

integer()

integer(range)

iodata()

iolist() Generates iolists

list\_of(data)

map(data, fun)

Generates a list of fixed length where each element is generated from the corresponding generator in data

Generates maps with fixed keys and generated values

## float(options \\ [])

Generates floats according to the given options

## frequency(frequencies)

Generates values from different generators with specified probability

## Generates integers bound by the generation size

Generates an integer in the given range

## Generates iodata

## keyword\_of(value\_data)

Generates keyword lists where values are generated by value\_data

Generates lists where each values is generated by the given data

## list\_of(data, options)

Generates lists where each values is generated by the given data

# Positives, Pains, and Sundry General Experiences

# **Generator Pitfalls**



# 1) Size of Generated Data



word = member\_of([...]) words = list\_of(word) sentence = list\_of(words)

# check all data <- sentence, text = List.flatten(data) > Enum.join(" ")

# sentence = list\_of( list\_of(member\_of([...]))

# check all data <- sentence,

end

text = List.flatten(data) > Enum.join(" ")

sentence = list\_of( list\_of(member\_of([...]))

# check all data <- sentence,

end

text = List.flatten(data) > Enum.join(" ")















words sentence = smallish(list\_of(words))



# = smallish(list\_of(word()))



## = smallish(list\_of(word())) words sentence = smallish(list\_of(words))





## = smallish(list\_of(word())) words sentence = smallish(list\_of(words))

# def biggish(generator) do generator >> SD.scale(fn size -> size \* 99 end) end



# 2) Reconstructing information after the fact is tough. So don't try to do that.



def word do Enum.random(["Hello", "World"]) end def sentence do (1..Enum.random(1..10)) |> Enum.map(fn \_ -> word() end) > Enum.join(" ") |> (&(&1 <> ".")).() # help end

test "word-count" do sentence = sentence() # random sentence! assert WC.count(sentence) == ... # now what? end

def word do Enum.random(["Hello", "World"]) end def sentence do (1..Enum.random(1..10)) |> Enum.map(fn \_ -> word() end) > Enum.join(" ") |> (&(&1 <> ".")).() # help end

test "word-count" do end

# sentence = sentence() # random sentence! assert WC.count(sentence) == ... # now what?

def word do member\_of(["This", "is", "a", ...]) end

property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre> do text = words |> Enum.join(" ") assert WordCount.count(text) == count end end



def word do member\_of(["This", "is", "a", ...]) end

property "word-count" do check all count <- positive\_integer()</pre> words <- list\_of(word, length: count)</pre> do text = words |> Enum.join(" ") assert WordCount.count(text) == count end end



# 3) Timex and/or Timezones



```
def naive_datetime() do
  gen all date <- date(),</pre>
           time <- time()</pre>
  do
    {:ok, ndt} = NaiveDateTime.new(date, time) do
  end
end
def date do
  gen all year <- integer(2010..2040),
           \{:ok, jan1\} = Date.new(year, 1, 1),
           days <- integer(0..(if Date.leap_year?(jan1), do: 365, else: 364))</pre>
  do
    Date.add(jan1, days)
  end
end
def time do
  gen all hour <- integer(0..23),</pre>
           min <- integer(0..59),</pre>
           sec <- integer(0..59),</pre>
           usec <- integer(0...999_999)</pre>
  do
    {:ok, time} = Time.new(hour, min, sec, usec)
    time
  end
end
```

def naive\_datetime() do gen all date <- date(),</pre> time <- time()</pre> do end end

def potentially\_ambiguous\_datetime() do gen all naive <- naive\_datetime()</pre> do Timex.to\_datetime(naive, "America/Los\_Angeles") end end

# {:ok, ndt} = NaiveDateTime.new(date, time) do
Failed with generated values (after 100 runs): • • •

left: %{timestamp: #DateTime<</pre> 2024-11-03 01:20:57-07:00 PDT America/Los\_Angeles >}

right: %{timestamp: #<Ambiguous(</pre> #DateTime<</pre> 2024-11-03 01:20:57-07:00 PDT America/Los\_Angeles > ~ #DateTime<</pre> 2024-11-03 01:20:57-08:00 PST America/Los\_Angeles >)>





alias Example.Health.Json

describe "JSON round-trips" do property "without IDs" do do

end end

property "with IDs" do report\_gen = Gen.Health.health\_report(nil) |> Gen.Id.with\_id() check all reports <- list\_of(report\_gen) do assert Json.parse(Json.generate(reports)) == {:ok, reports} end end

end

### check all reports <- list\_of(Gen.Health.health\_report(nil))</pre>

assert Json.parse(Json.generate(reports)) == {:ok, reports}



### def parse(list\_of\_maps) do

# {:ok, [%HealthReport{
 # or
 # {:error, ...}
end

def generate(reports) do

# [%{"id" => ..., "attributes" => ...}, ...]
end

### # {:ok, [%HealthReport{}, %HealthReport{}, ...]}

# **Idempotency**



- 1. Generate data to send. 2. Make controller request. 3. Make controller request. 4. Retrieve models.
- 5. ... Clean up! ! (Remove DB rows) 6. Then assert your expected result.

- 3. Retrieve models.

# 4. ... Clean up! ! (Remove DB rows) 5. Then assert your expected result.

- 3. Retrieve models.

# 4. ... Clean up! ! (Remove DB rows) 5. Then assert your expected result.

- 3. Retrieve models.

# 4. ... Clean up! / (Remove DB rows) 5. Then assert your expected result.

# (Shrinking)

## Shrinking replays a test, backtracking on the input data, until the test starts passing again.

## But ExUnitProperties runs inside a test case. DB data is not reset between runs.



- 3. Retrieve models.

# 4. ... Clean up! / (Remove DB rows) 5. Then assert your expected result.





created\_at: nil, day\_starting\_at: #DateTime<2021-04-01 23:53:29.047627Z>, exercise\_seconds: nil, id: nil, runwalk\_meters: 45576, steps\_count: nil, updated\_at: nil, user: #Ecto.Association.NotLoaded<association :user is not loaded>, user\_id: nil, weights: [%Example.HealthReport.WeightMeasurement{\_\_meta\_\_: #Ecto.Schema.Metadata<:built, "health\_report\_weights">, created\_at: nil, health\_report: #Ecto.Association.NotLoaded<association :health\_report is not loaded>, health\_report\_id: nil, id: nil, timestamp: #DateTime<2023-10-01 02:50:52.461705+11:00 AEDT Australia/ Sydney>, updated\_at: nil, weight\_grams: 2992005}, %Example.HealthReport.WeightMeasurement{\_\_meta\_\_: #Ecto.Schema.Metadata<:built, "health\_report\_weights">, created\_at: nil, health\_report: #Ecto.Association.NotLoaded<association :health\_report is not loaded>, health\_report\_id: nil, id: nil, timestamp: #DateTime<2013-05-31 06:38:36.865466Z>, updated\_at: nil, weight\_grams: 2286711}]}, %Example.HealthReport{\_\_meta\_\_: #Ecto.Schema.Metadata<:built, "health\_reports">, burnt\_joules: 142, created\_at: nil, day\_starting\_at: #DateTime<2037-10-27 20:32:03.518323Z>, exercise\_seconds: nil, id: nil, runwalk\_meters: 24141, steps\_count: nil, updated\_at: nil, user: #Ecto.Association.NotLoaded<association :user is not loaded>, user\_id: nil, weights: []}, %Example.HealthReport{\_\_meta\_\_: #Ecto.Schema.Metadata<:built, "health\_reports">, burnt\_joules: 20942, created\_at: nil, day\_starting\_at: #DateTime<2034-08-20 10:59:11.639846Z>, exercise\_seconds: 11893, id: nil, runwalk\_meters: 17910, steps\_count: 24013, updated\_at: nil, user: #Ecto.Association.NotLoaded<association :user is not loaded>, user\_id: nil, weights: [%Example.HealthReport.WeightMeasurement{\_\_\_meta\_\_: #Ecto.Schema.Metadata<:built, "health\_report\_weights">, created\_at: nil, health\_report: #Ecto.Association.NotLoaded<association :health\_report is not loaded>, health\_report\_id: nil, id: nil, timestamp: #DateTime<2022-09-25 04:24:59.049975-07:00 PDT America/ Los\_Angeles>, updated\_at: nil, weight\_grams: 2891106}, %Example.HealthReport.WeightMeasurement{\_\_meta\_\_: #Ecto.Schema.Metadata<:built, "health\_report\_weights">, created\_at: nil, health\_report: #Ecto.Association.NotLoaded<association :health\_report is not loaded>, health\_report\_id: nil, id: nil, timestamp: #DateTime<2019-03-12 21:30:26.737857-07:00 PDT America/Los\_Angeles>, updated\_at: nil, weight\_grams: 1042956}, %Example.HealthReport.WeightMeasurement{\_\_meta\_\_: #Ecto.Schema.Metadata<:built,</pre> "health\_report\_weights">, created\_at: nil, health\_report: #Ecto.Association.NotLoaded<association :health\_report is not loaded>, health\_report\_id: nil, id: nil, timestamp: #DateTime<2029-02-22 20:36:40.061115Z>, updated\_at: nil, weight\_grams: 34965}, %Example.HealthReport.WeightMeasurement{\_\_meta\_\_: #Ecto.Schema.Metadata<:built, "health\_report\_weights">, created\_at: nil, health\_report: #Ecto.Association.NotLoaded<association :health\_report is not loaded>, health\_report\_id: nil, id: nil, timestamp: #DateTime<2019-04-14 03:56:08.321487-07:00 PDT America/ Los\_Angeles>, updated\_at: nil, weight\_grams: 458541}]}, %Example.HealthReport{\_\_meta\_\_: #Ecto.Schema.Metadata<:built. "health reports">. burnt joules: nil. created at: nil. dav starting at:



## Factories VS Generators VS both and a second secon







# Macros are **Composition-Resistant**



check all thing <- thingy(:thing),</pre> max\_runs: 10 do expensive!(thing) assert thing == thing end

check all thing <- thingy(:thing),</pre> max\_runs: 10 do expensive!(thing) assert thing == thing end

slowcheck thing <- thingy(:thing)</pre> do expensive!(thing) assert thing == thing end

## defmacro slowcheck(...), do: ... slowcheck thing <- thingy(:thing)</pre> do expensive!(thing) assert thing == thing end





# this one is a bit tough to give a useful answer for

# It's a design tool.

# 1) JSON round-tripping (Parsing bug, generator bug, parsing bug, ...)

## 2) Idempotent actions (Finding bugs with the conflict-resolution for inserting.)



# 3) Ruby example: zips

property\_of { char, integer }.check { lchar,sizel file = File.join(tmpdir, "testfile-#{size}.bin") zip = File.join(tmpdir, "testfile-#{size}.zip")

filename = char \* size

```
data_read = nil
Zip::File.open(zip) {|f|
 data_read = f.first.get_input_stream.read
```

expect(data\_write).to == data\_read

data\_write = char \* size # size-length string, all char.

File.open(file, 'wb') { If I f.write(data\_write) } Zip::File.open(zip, CREATE) {If | f.add(filename, file) }

Size: 65535 - Gen'd, Written, equals read data. Size: 65536 - Gen'd, Written, Zipped, /Users/rhoward/code/

invalid stored block lengths (Zlib::DataError)

- Zipped, Unzipped. Written data
- experiments/p7zip/rubyzip/lib/ zip/inflater.rb:44:in `inflate':

### \$ 7z x testfile-65536.zip 7-Zip [64] ...

Processing archive: testfile-65536.zip

Errors: Headers Error Errors: Unconfirmed start of archive archive

- Warnings: There are data after the end of

Extracting testfile-65536: Segmentation fault

\$ 7z x testfile-65536.zip 7-Zip [64] ...

Processing archive: testfile-65536.zip

Errors: Headers Error Errors: Unconfirmed start of archive Warnings: There are data after the end of archive

Extracting testfile-65536: Segmentation fault

# so what's the answer then

## Yes, if you have someone on your team who's already used property testing before.

<u>Tentative yes</u>, if you don't have that person. Keep it to isolated cases, to try it out, so you can rip it out later.

# No, if you don't want to go off the beaten path.

## Start with things that are easy to generate data for, or you want to test the crap out of.

### How to draw an owl



### 1. Draw some circles



2. Draw the rest of the fucking owl

at my last job

## Rob Howard **Odamncabbage** http://robhoward.id.au



A draft experience report on the use of StreamData for Property Testing in Elixir

 Proper Testing https://propertesting.com/

StreamData docs

# Things To Read



### https://hexdocs.pm/stream\_data/StreamData.html