

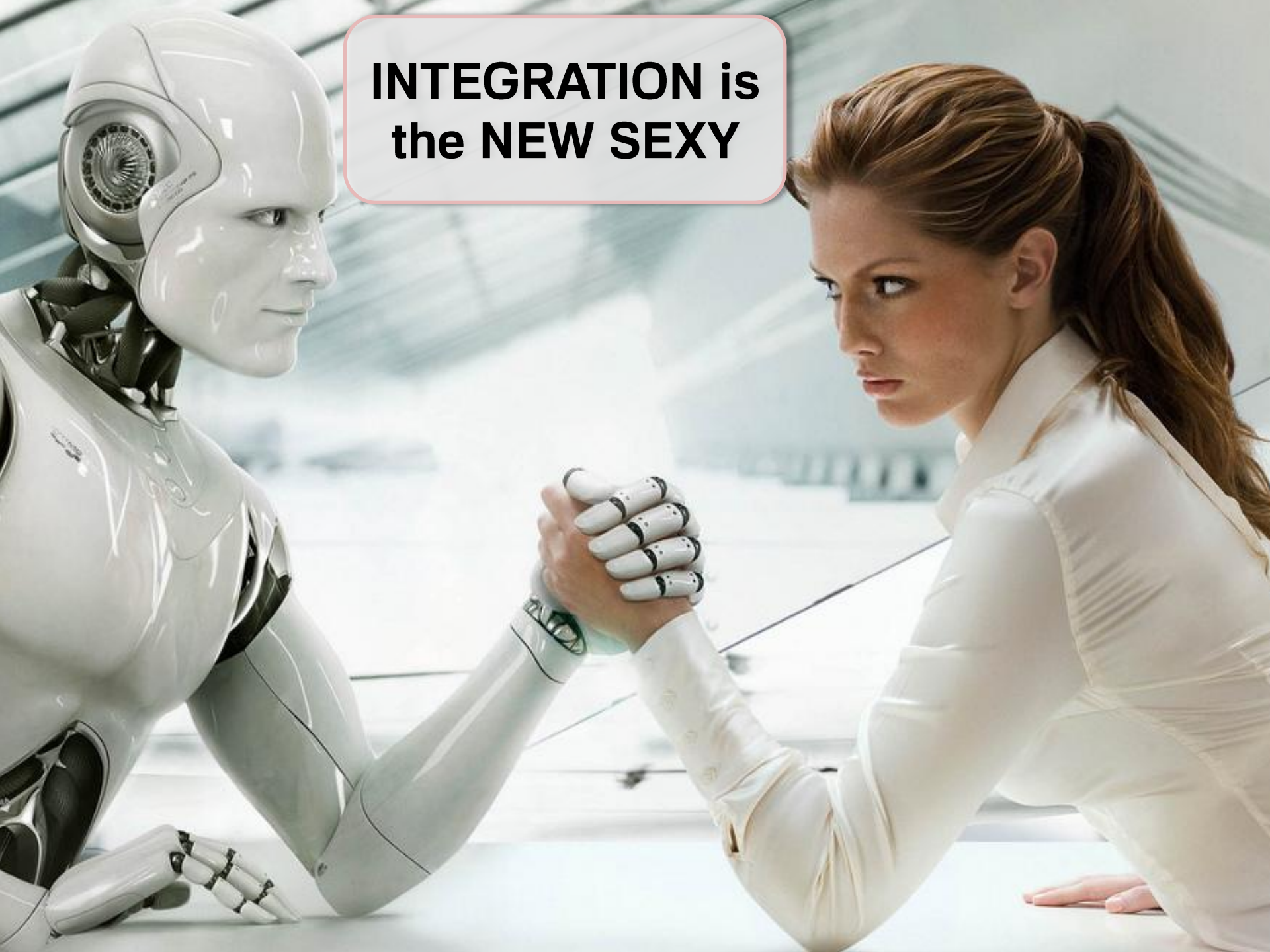
UGENE Workflow Designer.

Flexible control and extension of pipelines with scripts

Yuriy Vaskin

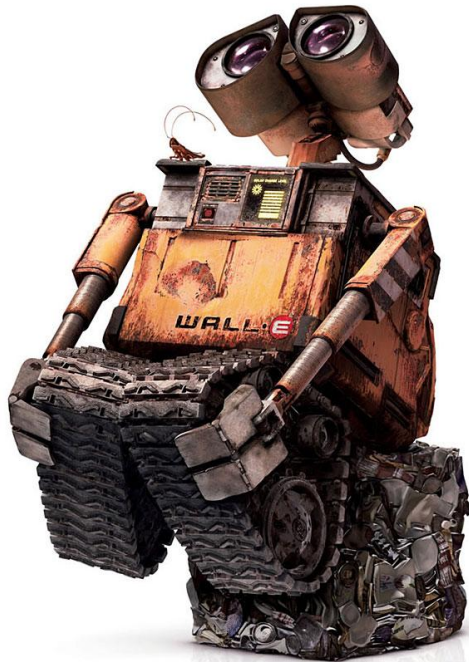


**INTEGRATION is
the NEW SEXY**



Software development in Bioinformatics

Yeah! I've come up with a new tool



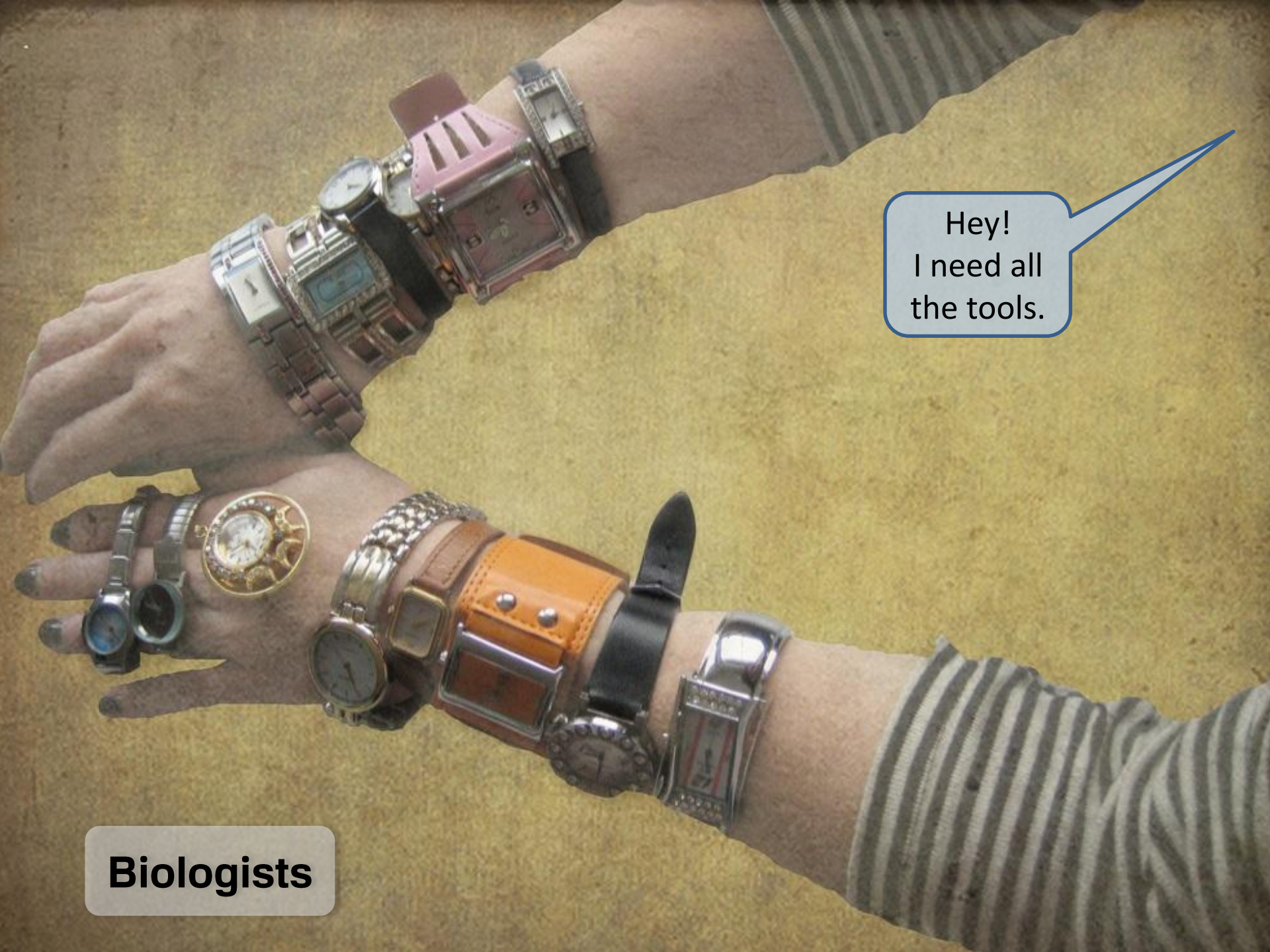
Web Images Maps More ▾

About 68,700 results (0.27 seconds)



Ooh... It's already there





Hey!
I need all
the tools.

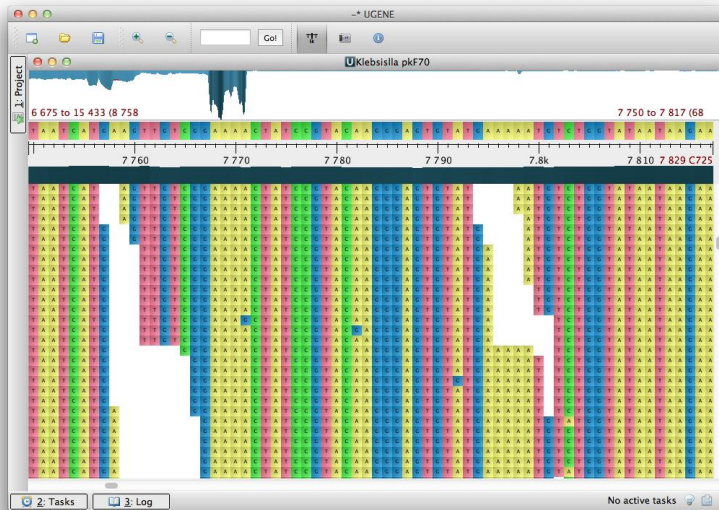
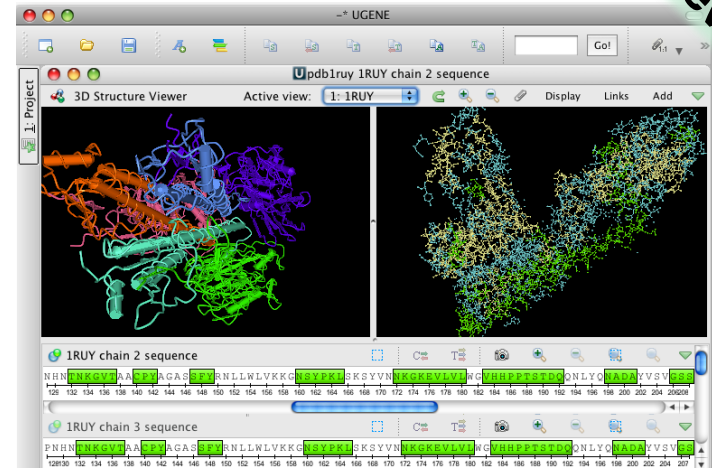
Biologists



UGENE

>5 years of development
Open-source

- Multiple Alignment: MUSCLE, Kalign, Clustal, Mafft, T-Coffee
- Ultra fast repeat finder



- HMMER, ORF Finder, Restriction sites
- Remote databases queries (BLAST, CDD, DAS, etc.)
- Phylogenetic trees
- TFBSs
- > 20 data formats

UGENE Workflow Designer

New Project UGENE - [Workflow Designer - Marking by length]

File Actions Settings Tools Window Help

100% Element style Run mode Scripting mode

Elements Samples

- Local BLAST search
- Local BLAST+ search
- Merge annotations
- ORF Marker
- Remote BLAST
- Smith-Waterman Search
- Data Converters
- HMMER2 tools
- Multiple Sequence Alignment
- Transcription Factor
- Utils
 - DNA statistics
 - Generate DNA
- Custom Elements with Script
- Custom Element Command Tools
- Data Flow
 - Filter
 - Sequence Marker

1: Project

```
graph TD; A[Read Sequences] -- Sequence --> B[Sequence Marker]; B -- Marked sequence --> C[Filter 'Short' Sequences]; B -- Marked sequence --> D[Filter 'Long' Sequences']; C -- filtered-data --> E[Write 'Short' Sequences']; D -- filtered-data --> F[Write 'Long' Sequences'];
```

Read Sequences
Reads sequence(s) from the list of files.

Sequence

Sequence Marker
Adds one or several marks to the input sequence.

Marked sequence

Marked sequence

Filter 'Short' Sequences
Filters input data from Sequence Marker by value(s) short.

filtered-data

Write 'Short' Sequences
Save all sequences from Read Sequences to short.fa.

filtered-data

Filter 'Long' Sequences
Filters input data from Sequence Marker by value(s) long.

Write 'Long' Sequences
Save all sequences from Read Sequences to long.fa.

Property Editor
Element name: ong' Sequences

Write sequence : Writes all supplied sequences to file(s) in selected format.

To configure the parameters of the element go to "Parameters" area below.

Iterations

Parameters

Name	Value
Accumulation objects	True
Document format	fasta
Output file	long.fa
Existing file	Rename

Input data

Set of annotations	<Listes>
Sequence	Seq...es)
Location	<e...y>

2: Tasks

3: Log

No active tasks



UGENE Workflow Designer Features

On **local** machine

Share files with schemes

SQLite 


Imports data to internal DB

Or passes file URLs

Extension with
external instruments

Wizards to setup the parameters

U Call Variants Wizard



SAMtools *mpileup* parameters

Parameters

Count anomalous read pairs	False
Disable BAQ computation	False
Mapping quality downgrading coefficient	0
Max number of reads per input BAM	250
Extended BAQ computation	False
BED or position list file	...
Pileup region	
Minimum mapping quality	0
Minimum base quality	13

Additional
Show additional parameters +

U GENE

< Back Next > Cancel

Dashboard for results and statistics

Output Files

File	Producer
res_seq.gb	Write Sequence
Open containing directory	
Open by operating system	

Workflow Task

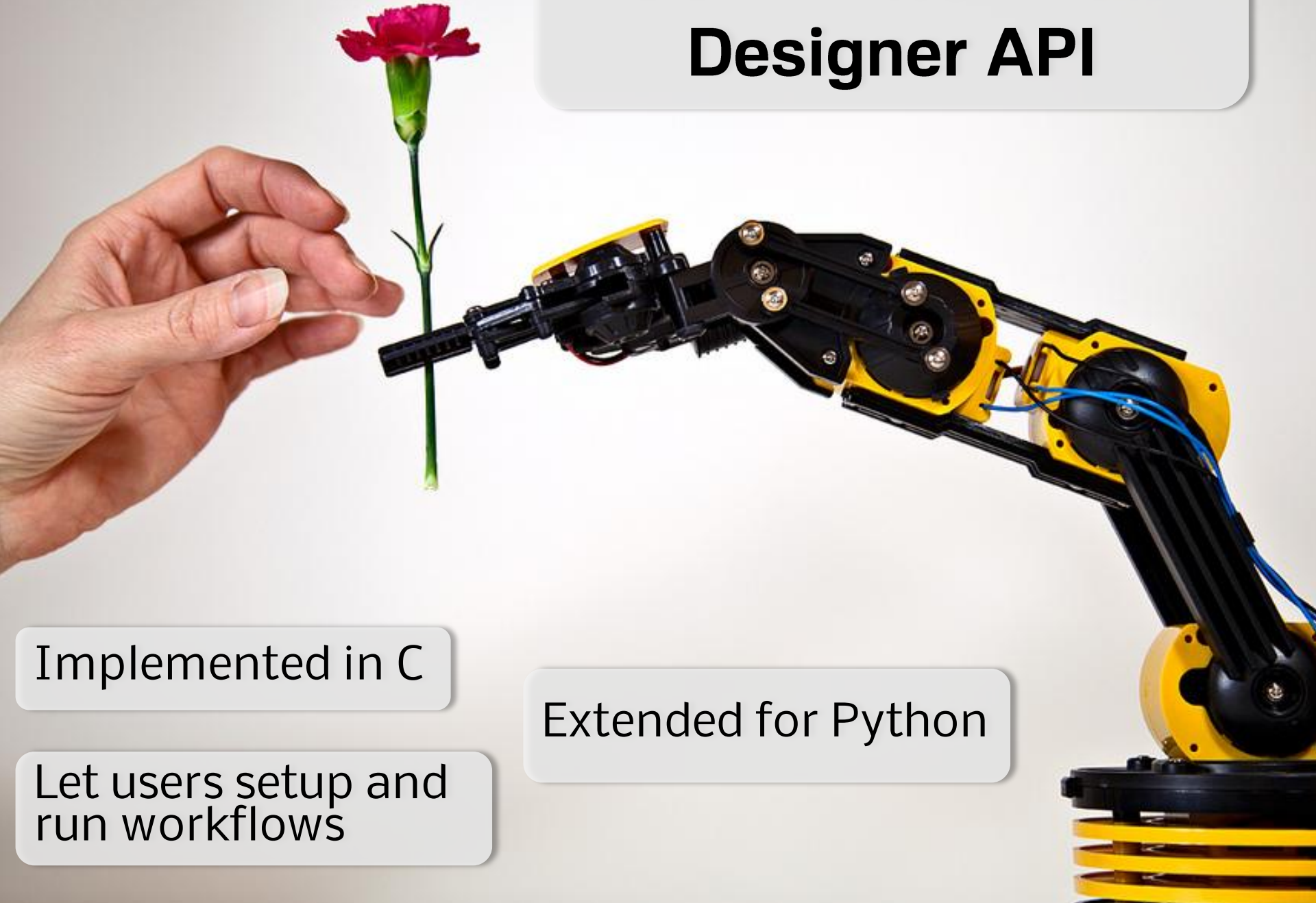
Time: 00:00:01

The workflow task has been finished successfully!

Common Statistics

Element	Elapsed time	Output messages
Filter	00:00:00	1
Sequence Marker	00:00:00	1
Read Sequence	00:00:00	1
Find Pattern	00:00:00	1
Write Sequence	00:00:00	0

UGENE Workflow Designer API



Implemented in C

Let users setup and
run workflows

Extended for Python

Running Single Algorithm Scheme

```
import u2py
u2py.initContext( './' )

inputFile = '../..//data/samples/FASTA/chrUn_gl000231.fa'
outputFile = '../..//data/samples/FASTA/ann.gb'

u2py.Scheme.launchSas( 'repeats-search', inputFile, outputFile )

u2py.releaseContext( )
```

BioPython Objects

Object inputs

```
from Bio.Align import MultipleSeqAlignment
from Bio import AlignIO
import u2py
u2py.initContext( './' )

#read an alignment
inputAlignments = []
inputFile = open( '../..data/samples/CLUSTALW/COI_copy1.sto', 'rU' )
inputAlignments.append( AlignIO.read( inputFile, 'stockholm' ) )
inputFile.close( )

#create a scheme
scheme = u2py.Scheme( 'muscle', inputAlignments )
scheme.setElementAttribute( 'Write Alignment', 'document-format', 'stockholm' )

#run the scheme
outputAlignments = scheme.launch( )

scheme.cleanUp( )
u2py.releaseContext( )
```

A cardboard robot, resembling a droid, stands on a rocky surface next to a dandelion. The robot is made of brown cardboard and has a simple face with two dots for eyes and a triangle for a mouth. The background is a blurred natural setting with green plants and grey rocks.

UGENE Integrations

Instruments

**GUI for
biologists**

**API for
technicians**

Thank you!

Contacts

UGENE

ugene@unipro.ru

ugene.unipro.ru

Yuriy Vaskin

vaskin@unipro.ru

