Using biometric gadgets for express-tests in the UX/UI research

Dmitriy Kostiuk,
Anastasia Markina,
Alexandr Dubitsky

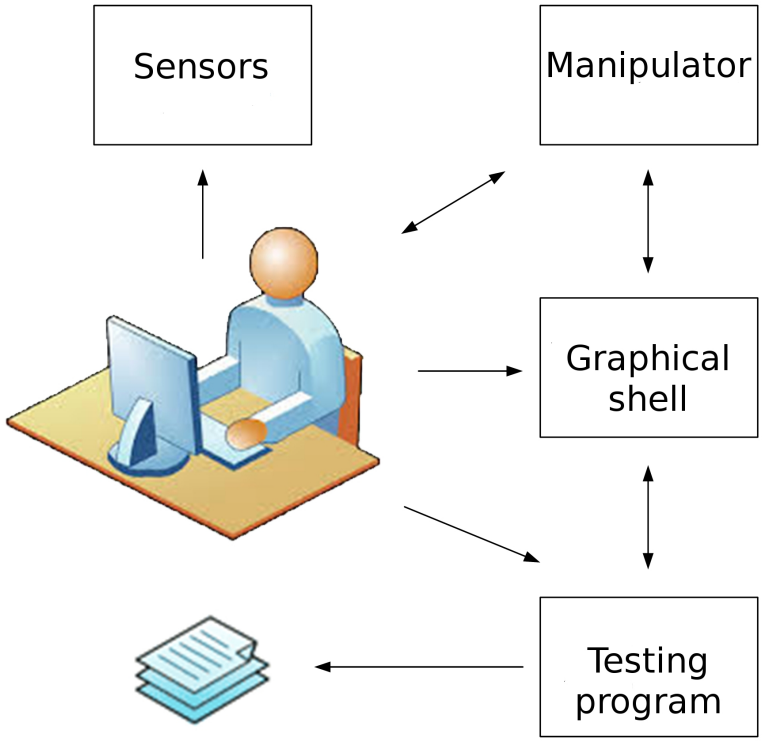
PERCONA

Biometrics in usability

- Biometric measuring tools have recently undergone a new wave of attention in the usability researches
- New powerful user-grade measuring gadgets are the reason
  - Mass-market production for the entertainment and fitness applications made them much cheaper
  - The precision they provide for contemporary games can be a valuable addition for the UX research

Scheme of Testing

- A User is typing the text, moving mouse, etc.
- Time and errors are taken into account
- Heart rate is measured by a fitness tracker
- Attention level and gaze direction, etc. are estimated by a consumer-grade devices: EEG, eye tracker, etc.



Available biometric parameters to measure

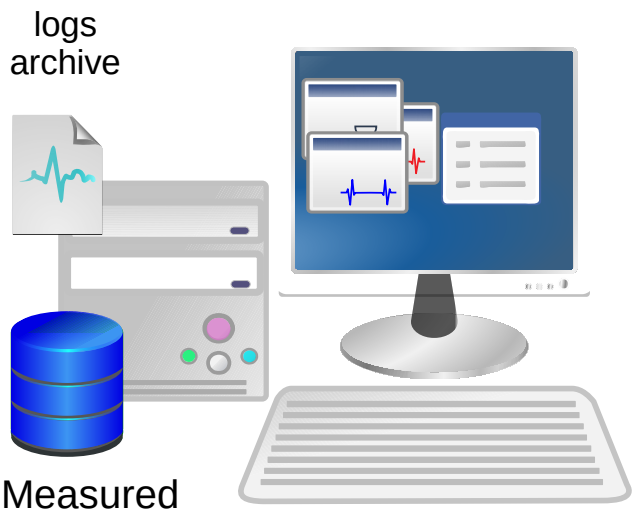
- Galvanic skin response
- Heart rate
- Blood pressure
- Electroencephalography waves (EEG)
- Kinematic activity
- Gaze direction

What we have used?

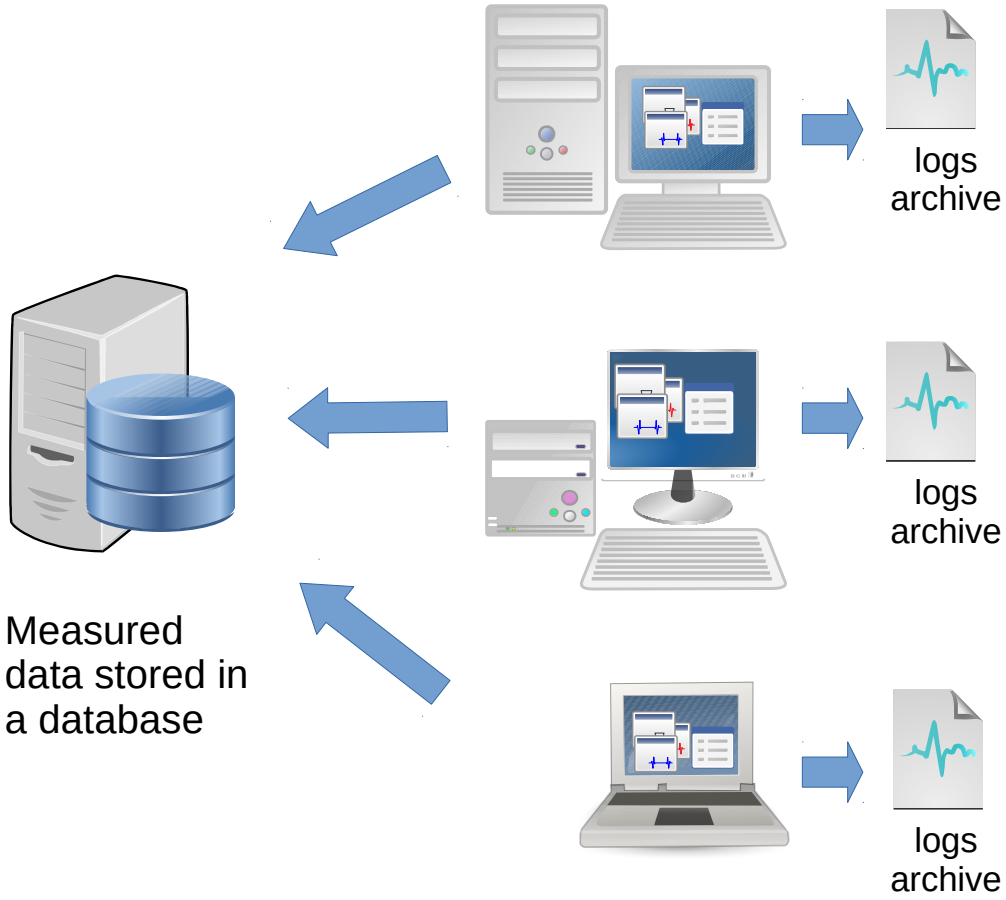
- EEG:
  - NeuroSky MindWave
  - NeuroSky MindSet
  - Emotiv EPOC/EPOC+
  - Emotiv Insight
- Fitness-trackers:
  - Fibit Charge HR
  - Xiaomi Mi Band 2/3
  - Amazfit Bip
- Tobii eye-trackers

Testing schemes

Individual testing mode

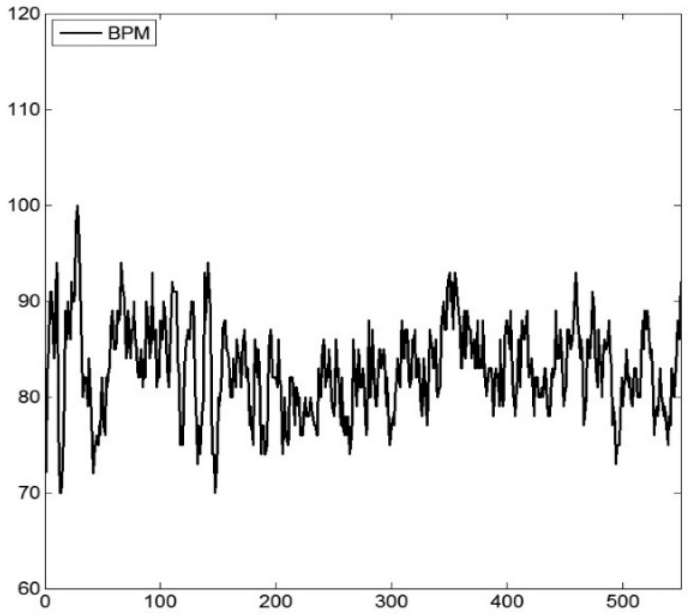


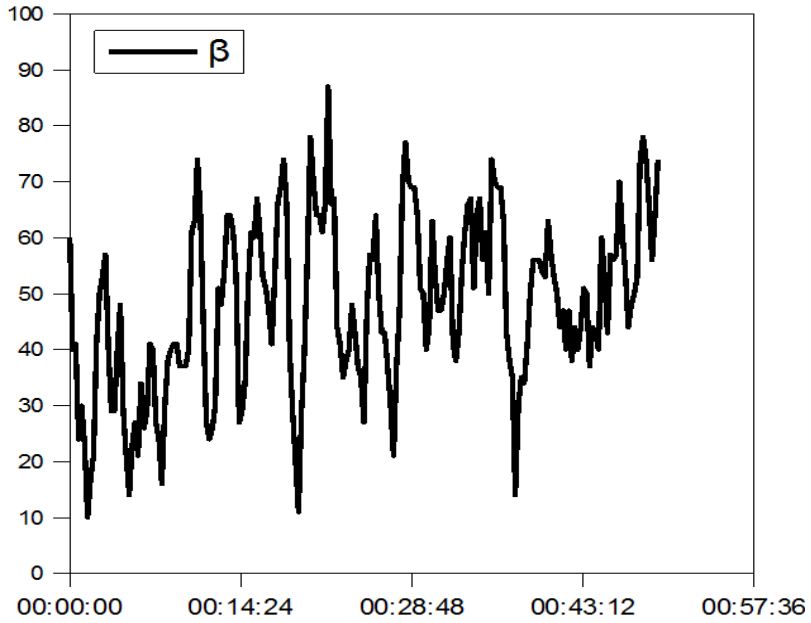
Parallel testing mode



Our own UXDump project bitbucket.org/AsyaAliset/uxdump
or some shell script to run the data acquisition software in parallel

Heart rate, EEG





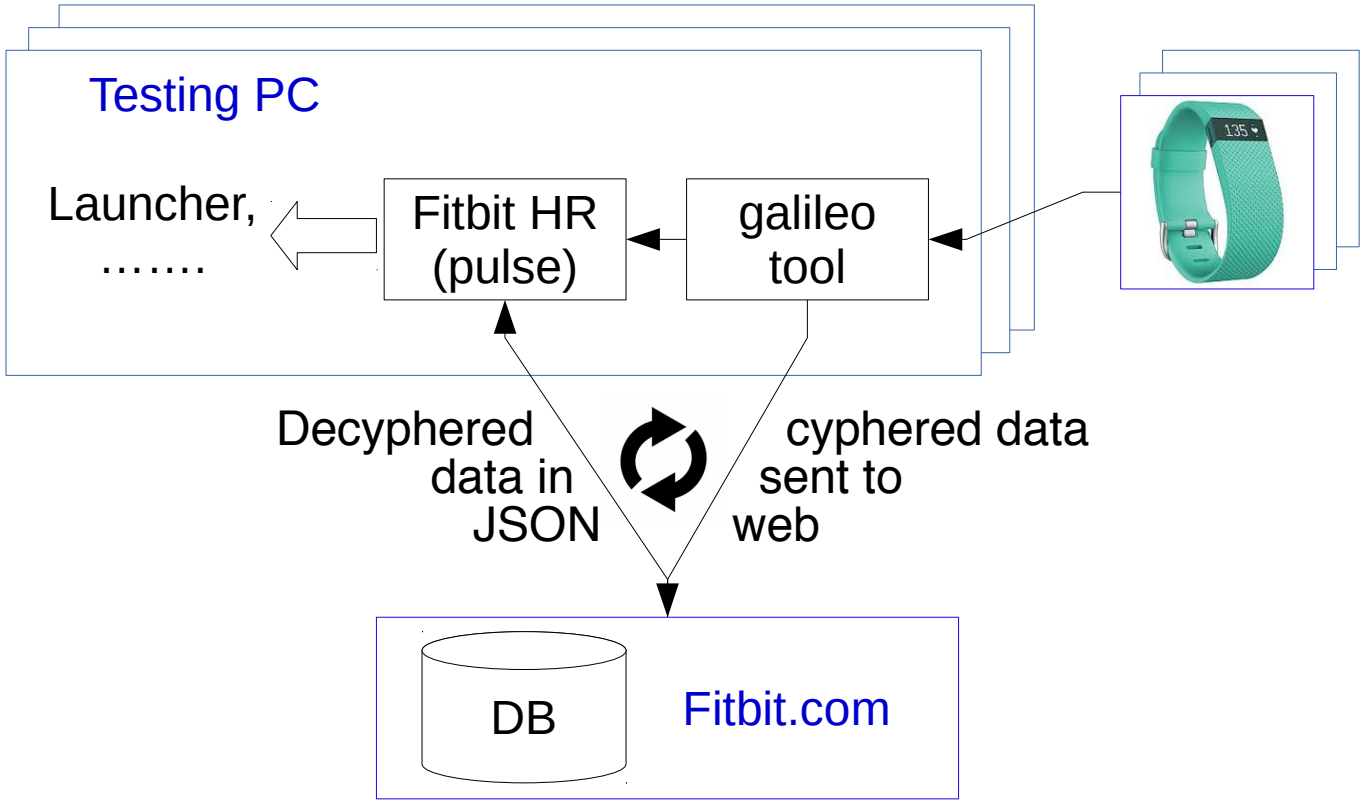
How to get data from device

- Using universal API
  - Bluez/QtBluetooth, etc.
- API from the device vendor
  - Fitbit web API
  - Tobii SDK
  - Special licenses to process personal data are possible
- Special open source tools per device
  - Mindwave (EEG) – PuzzleBox Synapse project
  - … a lot more
- File abstraction
  - Garmin fitness trackers pretend to be just a flash drive with logs

If the gadget does not provide biometrics data (especially in realtime) by itself, the open source may come into the deal

Getting heart rate with the remote access API ...

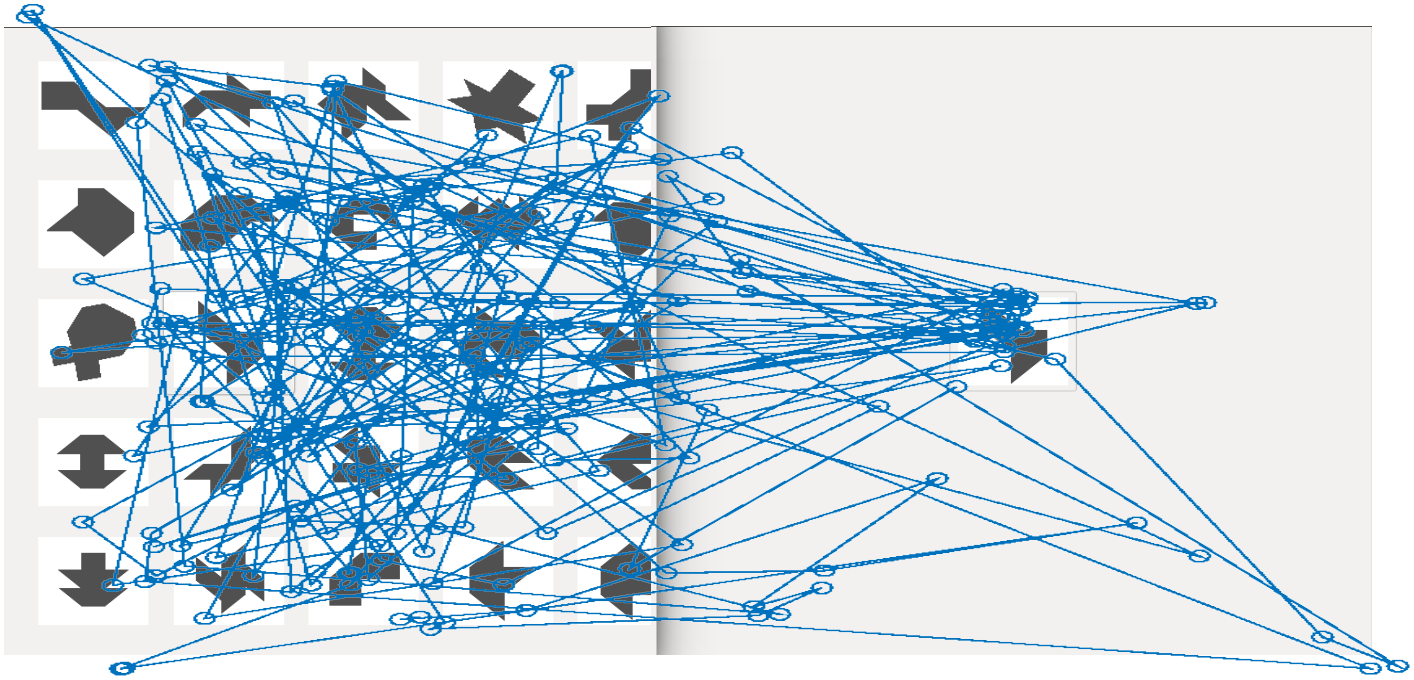
- … may be ridiculously complicated:



Evaluation criteria

- duration of the given actions
- number of errors
- heart rate
- attention level
  - calculated based on beta to alpha waves ratio
  - pre-calculated metrics (like «Attention» by NeuroSky) are good enough for practical needs
- emotions?

Gaze detection

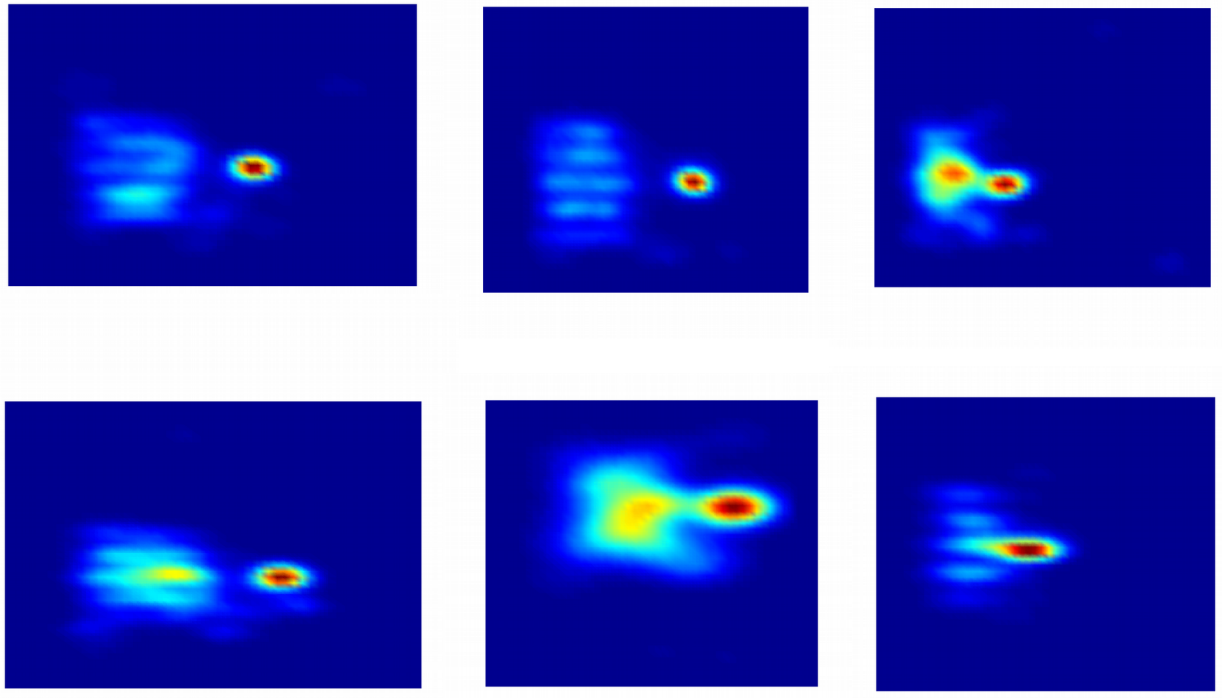


GNU octave turned out to be the easiest tool to plot gaze movements and gaze heatmaps

```
m=csvread('/tmp/sdf.txt');
x=m(:,1); y=m(:,2); z=m(:,3);
plot(x,y,"o-");
```

This allows to calculate the ratio between gaze fixations on the object of interest (e.g. tools panel) and the irrelevant areas

Gaze tracking heatmaps: Same task executed by six users



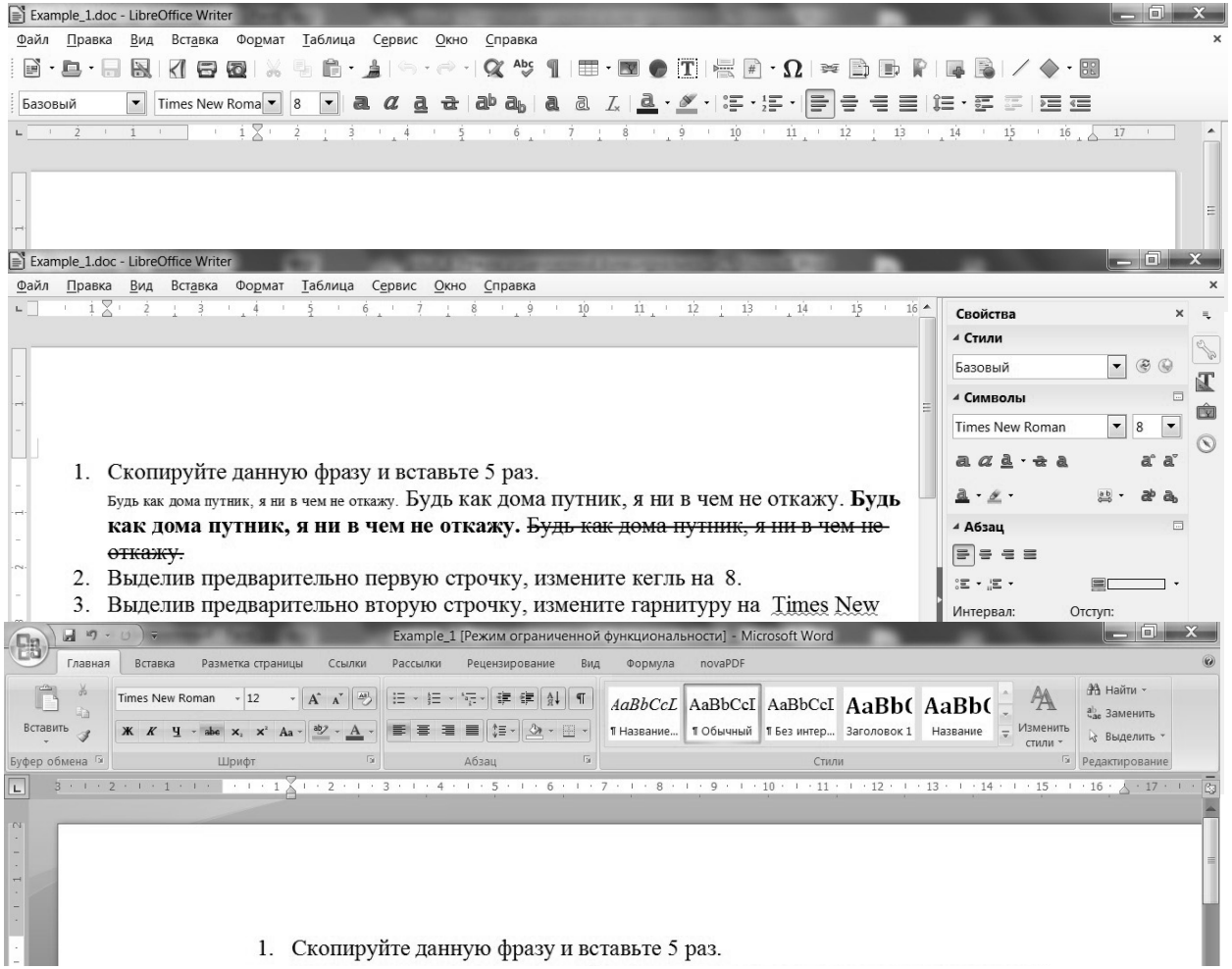
Heatmaps in GNU octave:

```
m=csvread('sdf.csv');
x=m(:,1); y=m(:,2);
p= gkde2([x'; y']);
figure(1);
surf(p.x,p.y,p.pdf);
colormap(jet);
shading interp;
grid off; axis off;
view(2);
saveas (1,'sdf.png');
```

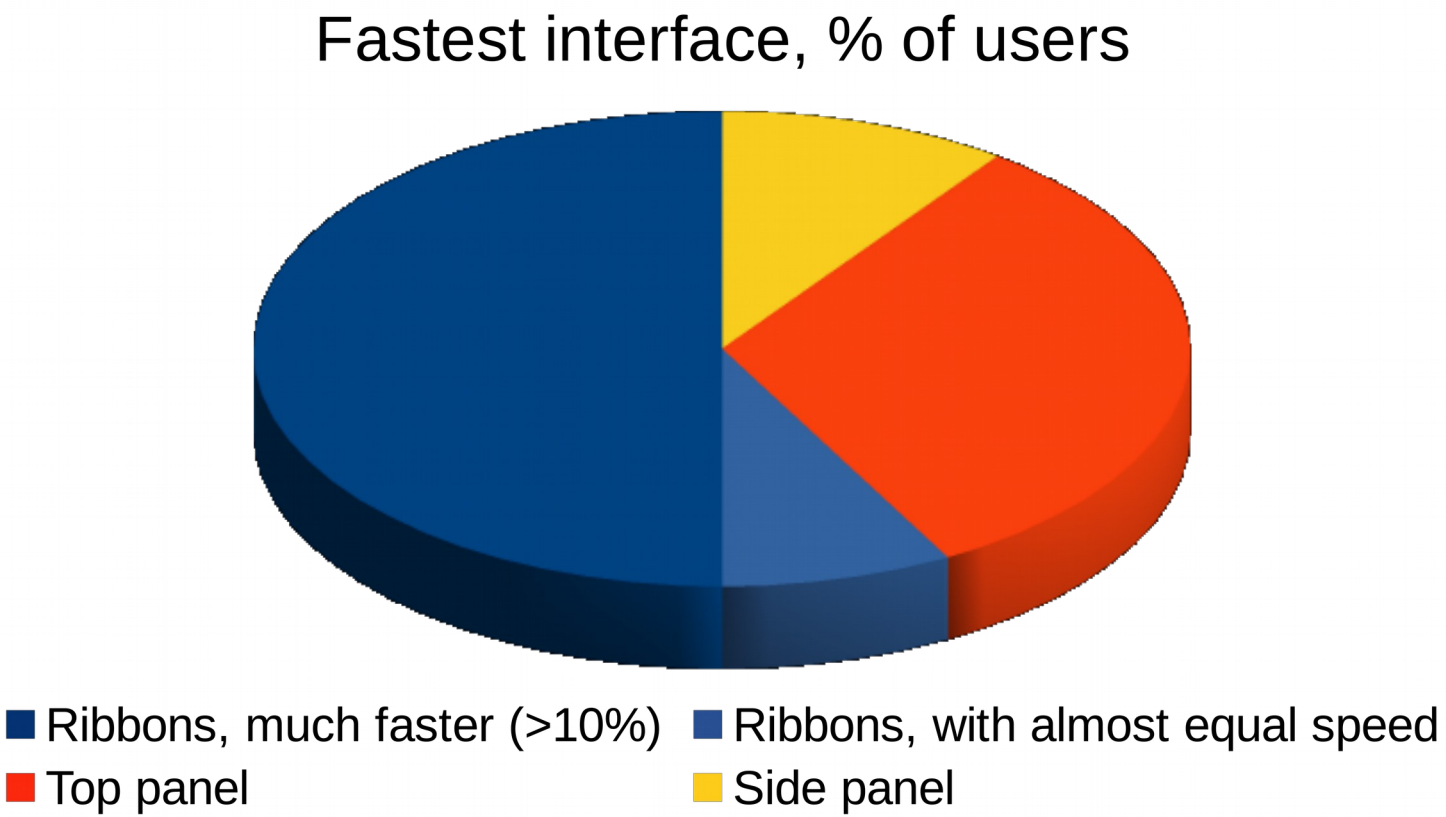
Two types of test tasks

- Series of different type operations in one program
  - user is supplied with a set of tasks under one general thematic direction
    - e.g. the working with a word processor
  - assess how the overall layout and dynamics of the application interface affect the user
- Long sequence of routine operations
  - large amount of the same type tasks
  - each task involves several applications, or parts
  - evaluate the contribution of the GUI auxiliary elements

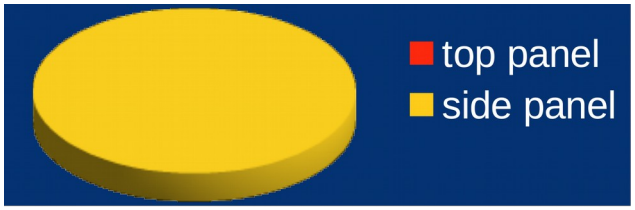
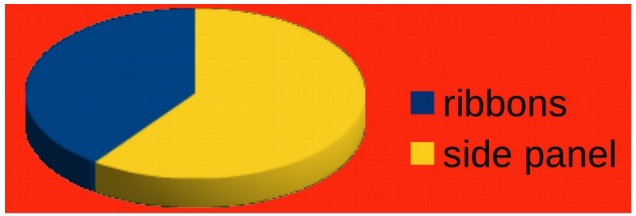
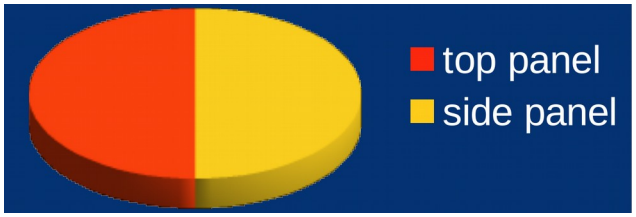
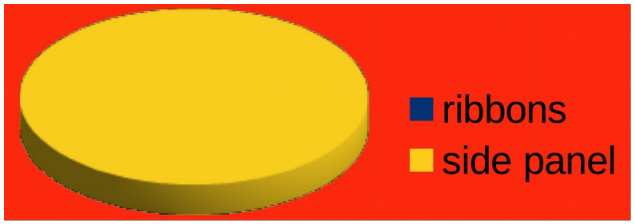
Tests example (office interfaces)



Office interfaces / Results



Office interfaces / Results

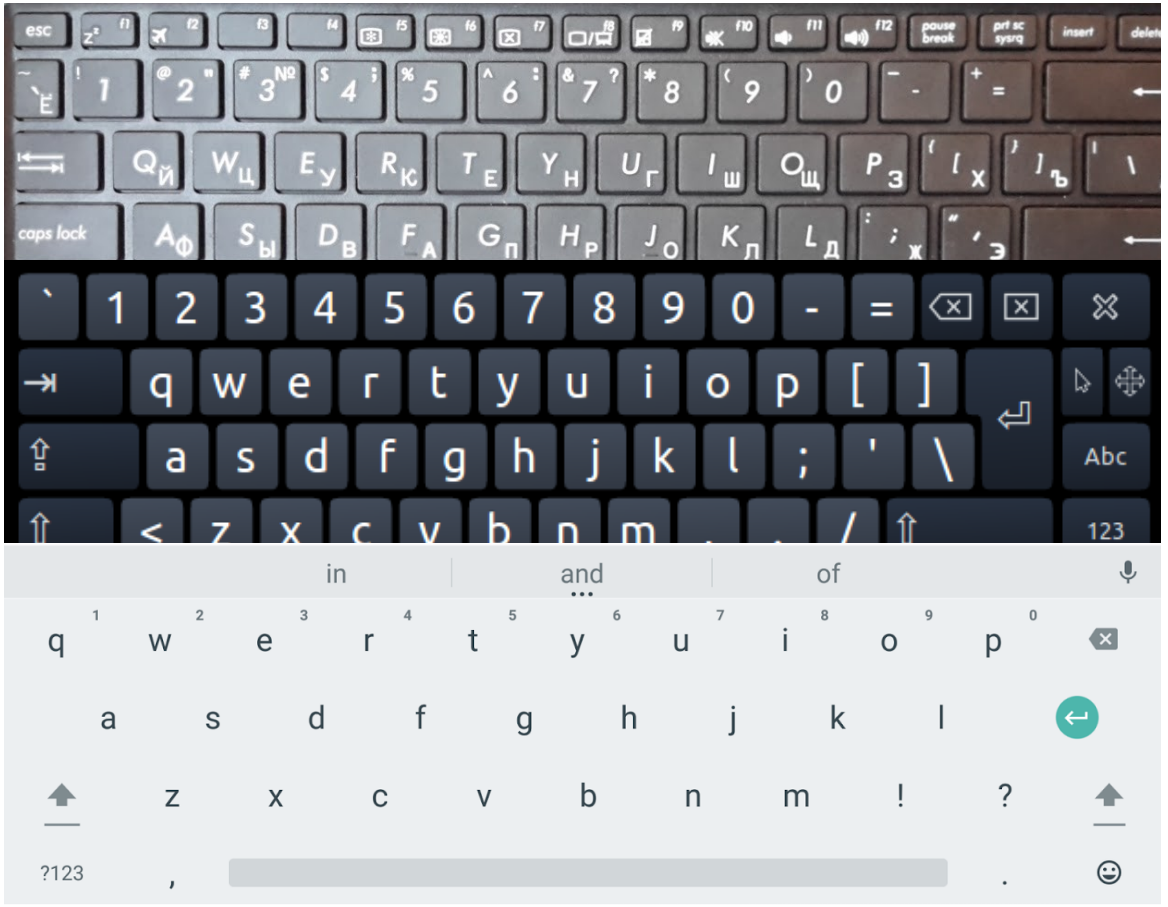
	Users who work best with ribbons	Users who work best with top panel
max. GSR,%		
max. HR,%		

Tests / Keyboards

Hardware keyboard by Asus

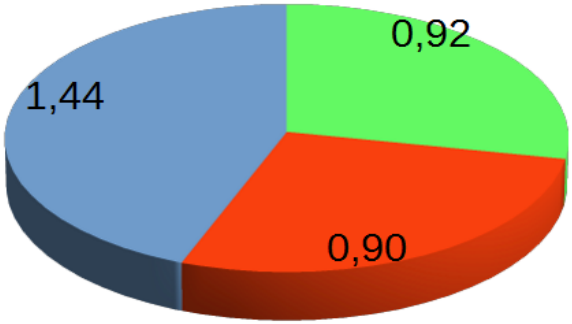
Onboard from Ubuntu 16.04

Default AOSP Keyboard from Android-x86

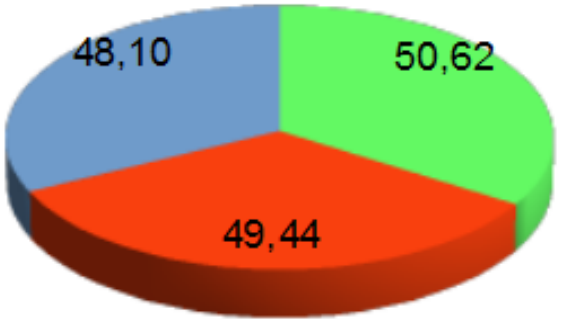


Keyboards / Results

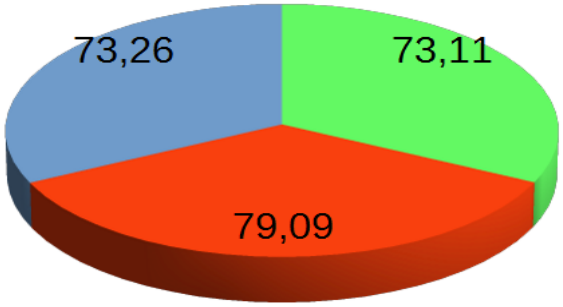
speed, chars/s



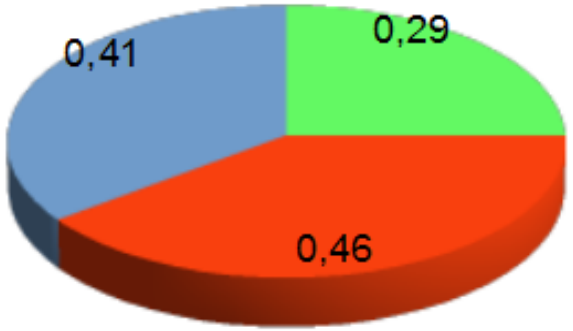
mind concentration, a.u.



pulse, beats/min



errors, %



Hardware keyboard

Onboard (Ubuntu)

AOSP Keyboard (Android)

Conclusions

- Consumer-grade biometrics turned out to be enough mature to be used in the UI/UX comparison
  - There are enough open source friendly devices on the mass market
- Biometric indicators are really useful in practical evaluation of the humans physical and mental load
  - Based on the test results, the average values for the listed parameters and the maximum deviation of the parameter from the average component can be easily calculated
- Each of the presented criteria allows to reduce the time series to a single value
  - This value reflects the nature of the work of a specific user in a particular test