Analysis of Cultivating a Software Engineering Community at DLR
(and some insights we gained)

Carina Haupt, Tobias Schlauch

Software Engineering Group
German Aerospace Center (DLR)
Software Development at DLR

Numbers
• More than 8500 employees
• ~20% of DLR employees involved in software development

Some Characteristics
• Variety of
  • Fields
  • Maturity
  • Software technologies
• Team sizes
• Backgrounds

We started the Software Engineering Initiative for DLR because we believe that our research results profit from better software!
Software Engineering Initiative for DLR

Policy Development
Guidelines and Tools
Trainings and Consulting
Knowledge and Experience Exchange
Knowledge Exchange Workshops

Regular knowledge exchange workshops are held to actively involve DLR scientists and to foster exchange.

Concept

• Intensive 1.5 day workshop
• Maximum of 60 participants
• Active involvement of the participants
• Knowledge, experience exchange and networking opportunities
• Results are shared via the SoftwareEngineering.Wiki
Research Questions

1. How stable is the represented part of the DLR software engineering community?

2. What is the influence of the workshop main topic on the workshop attendance?

3. Where are the workshop participants coming from?
**Analysis Approach**

**Idea:**
- Analysis of the participation data of the 6 WAWs on Software Engineering

**But it is just looking at attendance lists, right?**
- But: Unavailable or incomplete attendance lists, multiple registrations, name changes, location normalization, etc.

**Presented results are:**
- Based on (as good as possible) cleaned up and anonymized participation and location data sets
- Created with the help of great tools such as Python, Jupyter Notebook, Pandas, Matplotlib, Seaborn 😊

Data and analysis available at: https://zenodo.org/record/3403991
## Overview

<table>
<thead>
<tr>
<th>#</th>
<th>Topic</th>
<th># Participants</th>
<th>Date</th>
<th>Location</th>
<th># Employees at Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Kick-Off</td>
<td>57</td>
<td>November 2014</td>
<td>Braunschweig</td>
<td>1140</td>
</tr>
<tr>
<td>2</td>
<td>Tools and Processes</td>
<td>56</td>
<td>April 2015</td>
<td>Cologne</td>
<td>1581</td>
</tr>
<tr>
<td>3</td>
<td>Open and Inner Source</td>
<td>53</td>
<td>April 2016</td>
<td>Oberpfaffenhofen</td>
<td>1692</td>
</tr>
<tr>
<td>4</td>
<td>Software Architecture</td>
<td>52</td>
<td>April 2017</td>
<td>Berlin</td>
<td>649</td>
</tr>
<tr>
<td>5</td>
<td>Embedded Systems</td>
<td>47</td>
<td>May 2018</td>
<td>Bremen</td>
<td>183</td>
</tr>
<tr>
<td>6</td>
<td>Software Engineering for Data Science</td>
<td>55</td>
<td>May 2019</td>
<td>Jena</td>
<td>about 50</td>
</tr>
</tbody>
</table>

**Overall:**
- 320 participants
- 223 unique participants
- 53 participants per workshop in average
Typically, 16 (+/−3) participants attend the direct follow-up workshop.
How stable is the represented part of the DLR software engineering community?

Core group:
- 15% of unique participants

One-time participants:
- 78% of unique participants
- Including 34 persons of the last workshop that can become core group members.

Infrequent participants:
- 7% of unique participants

* Data of the last two workshops needs to stabilize!
How stable is the represented part of the DLR software engineering community?

<table>
<thead>
<tr>
<th>After Workshop</th>
<th># Core Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>27</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>33</td>
</tr>
</tbody>
</table>
How stable is the represented part of the DLR software engineering community?

<table>
<thead>
<tr>
<th></th>
<th># Core Group</th>
<th># New Core Group Cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td>After Workshop 2</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>After Workshop 3</td>
<td>25</td>
<td>3</td>
</tr>
<tr>
<td>After Workshop 4</td>
<td>27</td>
<td>6</td>
</tr>
<tr>
<td>After Workshop 5</td>
<td>29</td>
<td>2</td>
</tr>
<tr>
<td>After Workshop 6</td>
<td>33</td>
<td>7</td>
</tr>
</tbody>
</table>
How stable is the represented part of the DLR software engineering community?

<table>
<thead>
<tr>
<th></th>
<th># Core Group</th>
<th># New Core Group Cohort</th>
<th># All Workshops visited</th>
</tr>
</thead>
<tbody>
<tr>
<td>After Workshop 2</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>After Workshop 3</td>
<td>25</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>After Workshop 4</td>
<td>27</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>After Workshop 5</td>
<td>29</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>After Workshop 6</td>
<td>33</td>
<td>7</td>
<td>5</td>
</tr>
</tbody>
</table>
How stable is the represented part of the DLR software engineering community?

<table>
<thead>
<tr>
<th></th>
<th># Core Group</th>
<th># New Core Group Cohort</th>
<th># All Workshops visited</th>
</tr>
</thead>
<tbody>
<tr>
<td>After Workshop 2</td>
<td>20</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>After Workshop 3</td>
<td>25</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>After Workshop 4</td>
<td>27</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>After Workshop 5</td>
<td>29</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>After Workshop 6</td>
<td>33</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>After Workshop 7</td>
<td>?</td>
<td>0 .. 34 ?</td>
<td>?</td>
</tr>
</tbody>
</table>
What is the influence of the workshop main topic on the workshop attendance?

Workshop Topics:
1) Kick-Off
2) Tools and Processes
3) Open and Inner Source
4) Software Architecture
5) Embedded Systems
6) Software Engineering for Data Science

<table>
<thead>
<tr>
<th>Workshop</th>
<th>Total number of participants</th>
<th>Number of participants attending the next workshop</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>60</td>
<td>21 (35%)</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>16 (26%)</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>15 (30%)</td>
</tr>
<tr>
<td>4</td>
<td>50</td>
<td>13 (26%)</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>14 (35%)</td>
</tr>
</tbody>
</table>

Kick-Off & Follow-Up
Working on identified Topics
New and specialized Topics
Where are the workshop participants coming from?

- Attended from nearly every location (16 of 20)
- Big location, mainly administration
- Participation peak is at the “home” location
Detailed Participation Numbers of Workshop Locations

Workshop Topics:
1) Kick-Off
2) Tools and Processes
3) Open and Inner Source
4) Software Architecture
5) Embedded Systems
6) Software Engineering for Data Science
Research Questions

Some initial Answers

1. **How stable is the represented part of the DLR software engineering community?**
   
   In average, 36% community core members, 10% non-regular visitors, and 54% one-time visitors attend a workshop. Overall, the number of core group members is slowly increasing.

2. **What is the influence of the workshop main topic on the workshop attendance?**
   
   The main topic seems to be an important aspect for participation. We need to be aware with the main topic selection as it seems to influence workshop return rates and core group “onboarding”.

3. **Where are the workshop participants coming from?**
   
   Participants originate from nearly all DLR locations. Participants primarily attend workshops at their work location. A “smaller” workshop location does not necessarily reduce the overall participation.
Outlook, Ideas & Questions

Outlook & Ideas

• Focus on core SE topics again and RSE
• More involvement of former participants when selecting the main workshop topic
• Perform follow-up surveys and/or interviews

Some open questions

• Do we need a simpler definition of the core group?
• What additional information can we use to better interpret observed effects?
Thank You! - Questions?

Carina Haupt
Head of Software Engineering Group
Deputy Head of Department
DLR Intelligent and Distributed Systems
http://www.DLR.de/sc/ivs

Carina.Haupt@dlr.de
@caha42