

u^b

The Shared Language of Crowds: A Crowd-Sourced Approach to Mapping RSE and SER Terminology

Timo Kehrer

University of Bern

Robert Haines

University of Manchester

Guido Juckeland

Helmholtz-Zentrum Dresden

Shurui Zhou

University of Toronto

David E. Bernholdt

Oak Ridge National Laboratory



u^b

The Origin: Dagstuhl Seminar 24161



Research Software Engineering: Bridging Knowledge Gaps

(Apr 14 - Apr 19, 2024)

Organizers:

- **Stephan Druskat** (German Aerospace Center (DLR), Berlin)
- **Lars Grunske** (HU Berlin)
- **Caroline Jay** (University of Manchester)
- **Daniel S. Katz** (University of Illinois Urbana-Champaign)

u^b

The Origin: Dagstuhl Seminar 24161



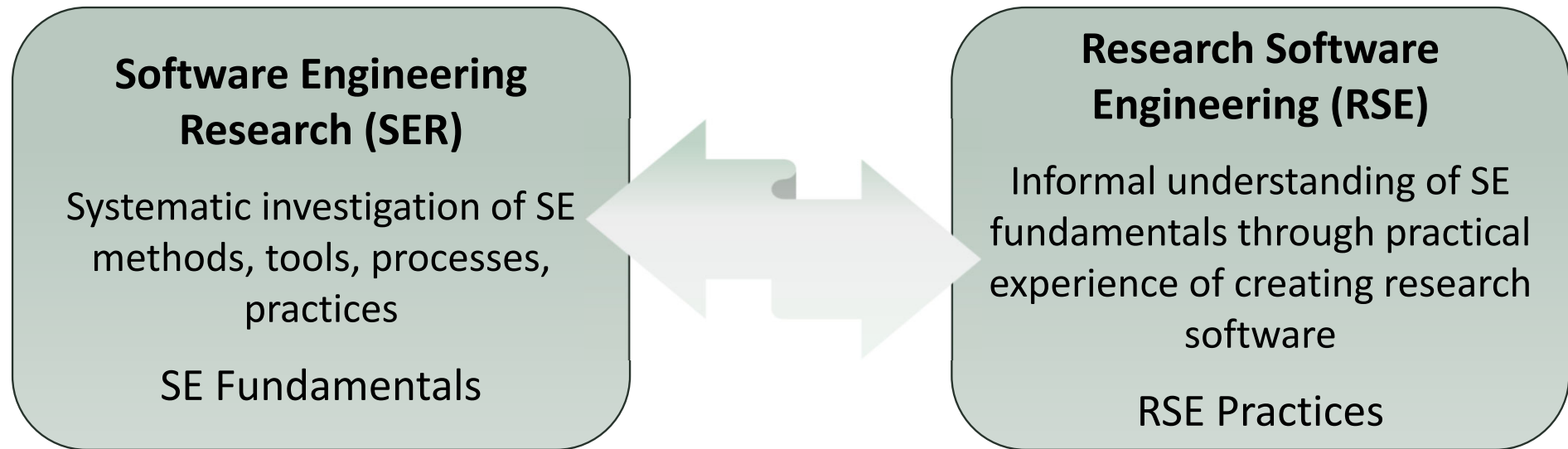
Working Group: Developing a Common Language

Members:

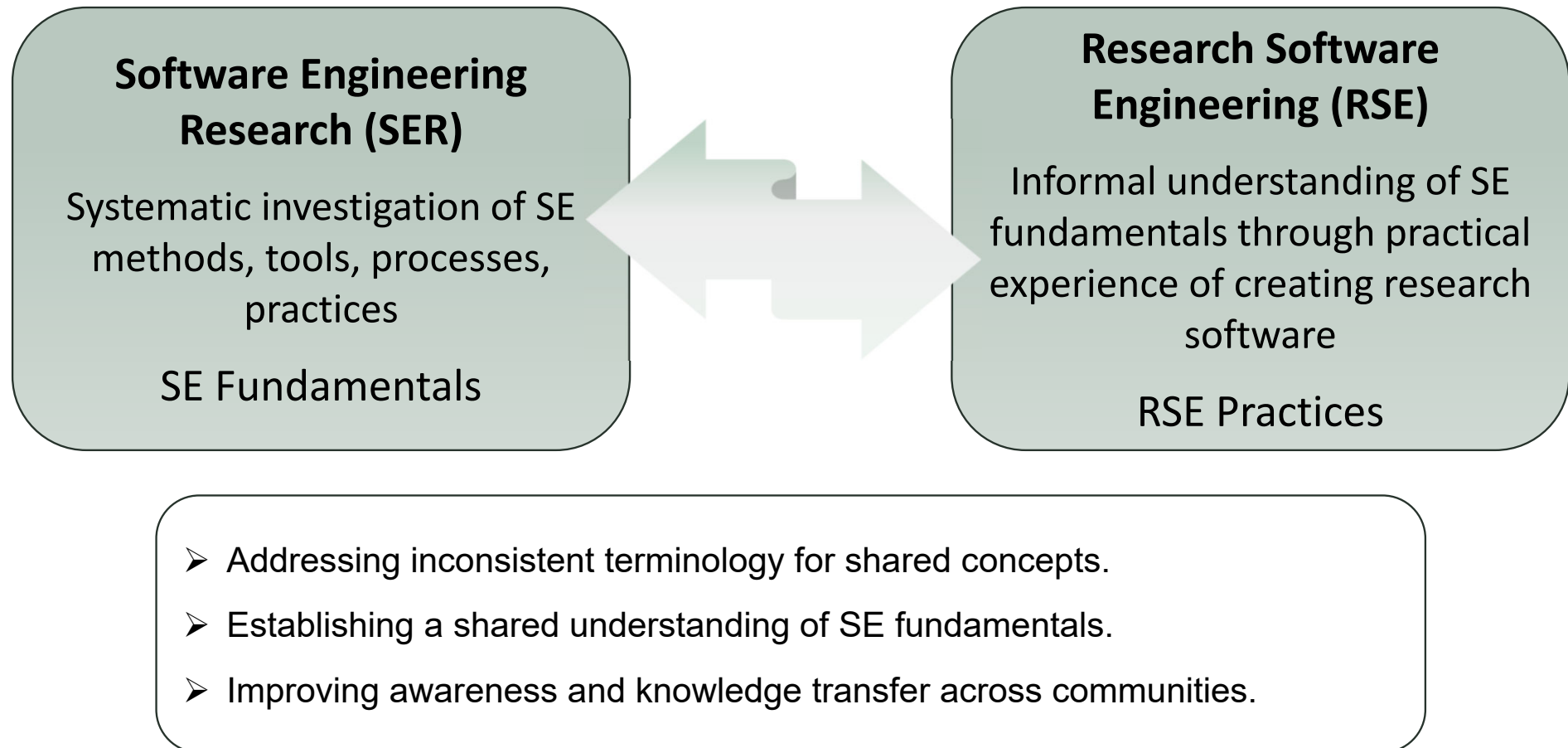
- **David E. Bernholdt** (Oak Ridge National Laboratory)
- **Robert Haines** (University of Manchester),
- **Guido Juckeland** (Helmholtz-Zentrum Dresden-Rossendorf),
- **Timo Kehrer** (Universität Bern)
- **Shurui Zhou** (University of Toronto)



u^b Motivation and Goals



u^b Motivation and Goals



u^b

Emerging Questions from Initial Discussions



u^b

Emerging Questions from Initial Discussions

Scope & Boundaries

- What is the scope, where to draw boundaries?



u^b

Emerging Questions from Initial Discussions

Scope & Boundaries

- What is the scope, where to draw boundaries?

Fundamentals & Core Concepts

- What counts as fundamental?



u^b

Emerging Questions from Initial Discussions

Scope & Boundaries

- What is the scope, where to draw boundaries?

Fundamentals & Core Concepts

- What counts as fundamental?

Abstraction & Granularity

- What level of granularity is appropriate?



u^b

Emerging Questions from Initial Discussions

Scope & Boundaries

- What is the scope, where to draw boundaries?

Fundamentals & Core Concepts

- What counts as fundamental?

Abstraction & Granularity

- What level of granularity is appropriate?

Consensus & Alignment

- How to reach consensus?



u^b

Approach

- Fundamentals & Core Concepts?
- Scope & Boundaries?
- Abstraction & Granularity?
- Consensus & Alignment?



- Systematic mapping based on SWEBOK as reference
- Crowdsourced collection across RSE and SER
- Web-based platform for community contributions
- Interactive discussion forum for qualitative insights

u^b

Starting Point: SWEBOK 4.0

Section	Heading	Page	Mapping
01	Software Requirements	1-1	
01.01	Software Requirements Fundamentals	1-2	
01.01.01	Definition of a Software Requirement	1-2	
01.01.02	Categories of Software Requirements	1-3	



Section	Heading	Page	Mapping
05.02	Test Levels	5-6	
05.02.01	The Target of the Test	5-6	
05.02.01.01	Unit Testing	5-6	mapping
05.02.01.02	Integration Testing	5-7	mapping
05.02.01.03	System Testing	5-7	mapping
05.02.01.04	Acceptance Testing	5-7	mapping
05.02.02	Objectives of Testing	5-7	
05.02.02.01	Conformance Testing	5-7	
05.02.02.02	Compliance Testing	5-8	
05.02.02.03	Installation Testing	5-8	mapping
05.02.02.04	Alpha and Beta Testing	5-8	mapping
05.02.02.05	Regression Testing	5-8	mapping

u^b Term Mapping Schema

```

1 # array of synonyms or adjacent concepts
2 se_fundamental:
3
4 # Short (1 sentence) description of the SE fundamental
5 fundamental_description:
6
7 # array of section identifiers
8 swebok_section:
9
10 # array of synonyms or adjacent concepts
11 rse_concept:
12
13 # Text, a brief description of the typical realizations of the fundamental, in RSE practice
14 rse_practice:
15
16 # General level of awareness of the fundamental in the research software community
17 # integers 0-3, 0=effectively no awareness, 3=widespread awareness
18 rse_awareness:
19 rse_awareness_source:
20
21 # General level of usage of the fundamental in the research software community
22 # integers 0-3, 0=effectively no usage, 3=widespread use
23 rse_usage:
24 rse_usage_source:
25
26 # Potential for SE research to improve use in research software
27 # integers 0-3, 0=effectively no opportunity, 3=significant SE research beneficial
28 ser_potential:
29 ser_potential_source:
30
31 # Reasons/opportunities for the SE research
32 ser_opportunities:
33
34 # References (external links, papers, etc., that may provide useful connections)
35 references:
36
37 # Date of last review by the editorial board (YYYY-MM-DD)
38 last_reviewed:

```

Term: The oracle problem

Source: [SWEBOK section 05.01.02.07](#)

SE Fundamental	RSE Equivalent
The oracle problem	- Checking with the science lead - Checking prior papers - Gold result
Fundamental description	RSE practice
An oracle can be any human or mechanical agent that decides whether the system under test behaved correctly. If the oracle cannot decide, the test output is classified as inconclusive.	RSEs are not generally aware of the the SER terminology, but do practice the oracle concept. Often through an appeal to expert knowledge or prior results.
SER potential (source) and opportunities	RSE awareness (source) and usage (sources)
3 (expert judgement)	1 (expert judgement)
RSEs *should* know the term 'the oracle problem' and should be thinking about it more specifically as they work. The oracle problem will be increasingly important/challenging in AI-related settings, for example. Strategies for dealing with inconclusive outcomes are also of increasing importance.	3 (expert judgement)

u^b Term Mapping Schema

```

1 # array of synonyms or adjacent concepts
2 se_fundamental:
3
4 # Short (1 sentence) description of the SE fundamental
5 fundam
6
7 # arra
8 swebo
9
10 # arra
11 rse_cl
12
13 # Tex
14 rse_pr
15
16 # General level of awareness of the fundamental in the research software community
17 # integers 0-3, 0=effectively no awareness, 3=widespread awareness
18 rse_awareness:
19 rse_awareness_source:
20
21 # General level of usage of the fundamental in the research software community
22 # integers 0-3, 0=effectively no usage, 3=widespread use
23 rse_usage:
24 rse_usage_source:
25
26 # Potential for SE research to improve use in research software
27 # integers 0-3, 0=effectively no opportunity, 3=significant SE research beneficial
28 ser_potential:
29 ser_potential_source:
30
31 # Reasons/opportunities for the SE research
32 ser_opportunities:
33
34 # References (external links, papers, etc., that may provide useful connections)
35 references:
36
37 # Date of last review by the editorial board (YYYY-MM-DD)
38 last_reviewed:

```

Source (e.g., SWEBOOK Section)

Term: The oracle problem

Source: SWEBOOK section 05.01.02.07

SE Fundamental	RSE Equivalent
The oracle problem	- Checking with the science lead - Checking prior papers - Gold result
Fundamental description	RSE practice
An oracle can be any human or mechanical agent that decides whether the system under test behaved correctly. If the oracle cannot decide, the test output is classified as inconclusive.	RSEs are not generally aware of the the SER terminology, but do practice the oracle concept. Often through an appeal to expert knowledge or prior results.
SER potential (source) and opportunities	RSE awareness (source) and usage (sources)
3 (expert judgement)	1 (expert judgement)
RSEs *should* know the term 'the oracle problem' and should be thinking about it more specifically as they work. The oracle problem will be increasingly important/challenging in AI-related settings, for example. Strategies for dealing with inconclusive outcomes are also of increasing importance.	3 (expert judgement)

u^b Term Mapping Schema

```
1 # array of synonyms or adjacent concepts
2 se_fundamental:
3
4 # Short (1 sentence) description of the SE fundamental
5 fundamental_description:
6
7 # array
8 swebo
9
10 # arra
11 rse_c
12
13 # Tex
14 rse_p
15
16 # Ger
17 # inte
18 rse_awa
19 rse_awareness_source:
20
21 # General level of usage of the fundamental in the research software community
22 # integers 0-3, 0=effectively no usage, 3=widespread use
23 rse_usage:
24 rse_usage_source:
25
26 # Potential for SE research to improve use in research software
27 # integers 0-3, 0=effectively no opportunity, 3=significant SE research beneficial
28 ser_potential:
29 ser_potential_source:
30
31 # Reasons/opportunities for the SE research
32 ser_opportunities:
33
34 # References (external links, papers, etc., that may provide useful connections)
35 references:
36
37 # Date of last review by the editorial board (YYYY-MM-DD)
38 last_reviewed:
```

Collection of synonyms or adjacent concepts

Term: The oracle problem

Source: [SWEBOK section 05.01.02.07](#)

SE Fundamental	RSE Equivalent
The oracle problem	- Checking with the science lead - Checking prior papers - Gold result
Fundamental description	RSE practice
An oracle can be any human or mechanical agent that decides whether the system under test behaved correctly. If the oracle cannot decide, the test output is classified as inconclusive.	RSEs are not generally aware of the the SER terminology, but do practice the oracle concept. Often through an appeal to expert knowledge or prior results.
SER potential (source) and opportunities	RSE awareness (source) and usage (sources)
3 (expert judgement)	1 (expert judgement)
RSEs *should* know the term 'the oracle problem' and should be thinking about it more specifically as they work. The oracle problem will be increasingly important/challenging in AI-related settings, for example. Strategies for dealing with inconclusive outcomes are also of increasing importance.	3 (expert judgement)

u^b Term Mapping Schema

```
1 # array of synonyms or adjacent concepts
2 se_fundamental:
3
4 # Short (1 sentence) description of the SE fundamental
5 fundamental_description:
6
7 # array of section identifiers
8 swebok_section:
9
10 # array of synonyms or adjacent concepts
11 rse_concept:
12
13 # Text
14 rse_p
15
16 # Ger
17 # inte
18 rse_av
19 rse_av
20
21 # Ger
22 # inte
23 rse_us
24 rse_usage_source:
25
26 # Potential for SE research to improve use in research software
27 # integers 0-3, 0=effectively no opportunity, 3=significant SE research beneficial
28 ser_potential:
29 ser_potential_source:
30
31 # Reasons/opportunities for the SE research
32 ser_opportunities:
33
34 # References (external links, papers, etc., that may provide useful connections)
35 references:
36
37 # Date of last review by the editorial board (YYYY-MM-DD)
38 last_reviewed:
```

Definition vs. typical realization

Term: The oracle problem

Source: [SWEBOK section 05.01.02.07](#)

SE Fundamental	RSE Equivalent
The oracle problem	- Checking with the science lead - Checking prior papers - Gold result
Fundamental description	RSE practice
can be any human or mechanical agent that decides whether the system under test behaved correctly. If the oracle cannot decide, the test output is classified as inconclusive.	RSEs are not generally aware of the the SER terminology, but do practice the oracle concept. Often through an appeal to expert knowledge or prior results.
SER potential (source) and opportunities	RSE awareness (source) and usage (sources)
3 (expert judgement)	1 (expert judgement)
RSEs *should* know the term 'the oracle problem' and should be thinking about it more specifically as they work. The oracle problem will be increasingly important/challenging in AI-related settings, for example. Strategies for dealing with inconclusive outcomes are also of increasing importance.	
	3 (expert judgement)

u^b Term Mapping Schema

```
1 # array of synonyms or adjacent concepts
2 se_fundamental:
3
4 # Short (1 sentence) description of the SE fundamental
5 fundamental_description:
6
7 # array of section identifiers
8 swebok_section:
9
10 # array of synonyms or adjacent concepts
11 rse_concept:
12
13 # Text, a brief description of the typical realizations of the fundamental, in RSE practice
14 rse_practice:
15
16 # General level of awareness of the fundamental in the research software community
17 # integers 0-3, 0=effectively no awareness, 3=widespread awareness
18 rse_awareness:
19 rse_awareness_source:
20
21 # General level of usage of the fundamental in the research software community
22 # integers 0-3, 0=effectively no usage, 3=widespread use
23 rse_usage:
24 rse_usage_source:
25
26 # Potential for the fundamental to be used in the research software community
27 # integers 0-3, 0=effectively no potential, 3=widespread potential
28 ser_potential:
29 ser_potential_source:
30
31 # Realization of the fundamental in the research software community
32 ser_realization:
33
34 # Reference to the fundamental in the research software community
35 ser_reference:
36
37 # Date of last review by the editorial board (YYYY-MM-DD)
38 last_reviewed:
```

General level of awareness

- 0=effectively no awareness;
- 3=widespread awareness

Term: The oracle problem

Source: [SWEBOK section 05.01.02.07](#)

SE Fundamental	RSE Equivalent
• The oracle problem	• Checking with the science lead • Checking prior papers • Gold result
Fundamental description	RSE practice
An oracle can be any human or mechanical agent that decides whether the system under test behaved correctly. If the oracle cannot decide, the test output is classified as inconclusive.	RSEs are not generally aware of the the SER terminology, but do practice the oracle concept. Often through an appeal to expert knowledge or prior results.
SER potential (source) and opportunities	RSE awareness (source) and usage (sources)
	1 (expert judgement)
RSEs *should* know the term 'the oracle problem' and should be thinking about it more specifically as they work. The oracle problem will be increasingly important/challenging in AI-related settings, for example. Strategies for dealing with inconclusive outcomes are also of increasing importance.	3 (expert judgement)

u^b Term Mapping Schema

```
1 # array of synonyms or adjacent concepts
2 se_fundamental:
3
4 # Short (1 sentence) description of the SE fundamental
5 fundamental_description:
6
7 # array of section identifiers
8 swebok_section:
9
10 # array of synonyms or adjacent concepts
11 rse_concept:
12
13 # Text, a brief description of the typical realizations of the fundamental, in RSE practice
14 rse_practice:
15
16 # General level of awareness of the fundamental in the research software community
17 # integers 0-3, 0=effectively no awareness, 3=widespread awareness
18 rse_awareness:
19 rse_awareness_source:
20
21 # General level of usage of the fundamental in the research software community
22 # integers 0-3, 0=effectively no usage, 3=widespread use
23 rse_usage:
24 rse_usage_source:
25
26 # Potential for the fundamental to be used in the research software community
27 # integers 0-3, 0=effectively no potential, 3=widespread potential
28 ser_potential:
29 ser_potential_source:
30
31 # Realization of the fundamental in the research software community
32 ser_realization:
33
34 # Reference to the fundamental in the research software community
35 references:
36
37 # Date of last review by the editorial board (YYYY-MM-DD)
38 last_reviewed:
```

General level of usage

- 0=effectively no usage;
- 3=widespread usage

Term: The oracle problem

Source: [SWEBOK section 05.01.02.07](#)

SE Fundamental	RSE Equivalent
• The oracle problem	• Checking with the science lead • Checking prior papers • Gold result
Fundamental description	RSE practice
An oracle can be any human or mechanical agent that decides whether the system under test behaved correctly. If the oracle cannot decide, the test output is classified as inconclusive.	RSEs are not generally aware of the the SER terminology, but do practice the oracle concept. Often through an appeal to expert knowledge or prior results.
SER potential (source) and opportunities	RSE awareness (source) and usage (sources)
3 (expert judgement)	1 (expert judgement)
RSEs *should* know the term 'the oracle problem' and should be thinking about it as they work. The oracle problem will be increasingly important/challenging in AI-related settings, for example. Strategies for dealing with inconclusive outcomes are also of increasing importance.	3 (expert judgement)

u^b Term Mapping Schema

```
1 # array of synonyms or adjacent concepts
2 se_fundamental:
3
4 # Short (1 sentence) description of the SE fundamental
5 fundamental_description:
6
7 # array of section identifiers
8 swebok_section:
9
10 # array of synonyms or adjacent concepts
11 rse_concept:
12
13 # Text, a brief description of the typical realizations of the fundamental, in RSE practice
14 rse_practice:
15
16 # General level of awareness of the fundamental in the research software community
17 # integers 0-3, 0=effectively no awareness, 3=widespread awareness
18 rse_awareness:
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37 #
38 last_reviewed:
```

Potential for SER to improve usage in RSE

- 0=effectively no opportunity;
- 3=significant SE research beneficial

Term: The oracle problem

Source: [SWEBOK section 05.01.02.07](#)

SE Fundamental	RSE Equivalent
• The oracle problem	• Checking with the science lead • Checking prior papers • Gold result
Fundamental description	RSE practice
An oracle can be any human or mechanical agent that decides whether the system under test behaved correctly. If the oracle cannot decide, the test output is classified as inconclusive.	RSEs are not generally aware of the the SER terminology, but do practice the oracle concept. Often through an appeal to expert knowledge or prior results.
SER potential (source) and opportunities	RSE awareness (source) and usage (sources)
3 (expert judgement)	1 (expert judgement)
RSEs *should* know the term 'the oracle problem' and should be thinking about it more specifically as they work. The oracle problem will be increasingly important/challenging in AI-related settings, for example. Strategies for dealing with inconclusive outcomes are also of increasing importance.	
	3 (expert judgement)

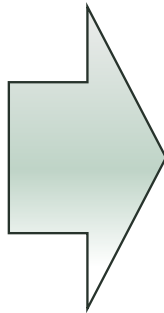
u^b

Initial Experience

- Pilot crowdsourcing sessions with mixed RSEs and SERs
- Participant-driven term selection and open discussions
- Insights on terminology granularity and SWEBOK structure
- Identification of shared, differing, and overlapping concepts

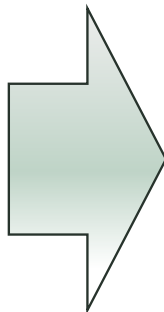
u^b Envisioned Outcome

SE fundamentals
unfamiliar to the
RSE community



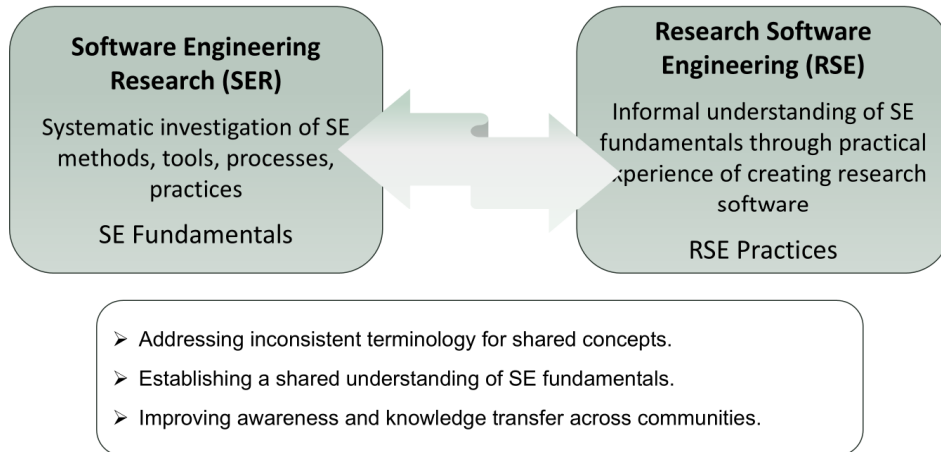
opportunities for education,
training, and outreach to
promote adoption by the RSE
community;

SE fundamentals
perceived as less
useful by the
RSE community

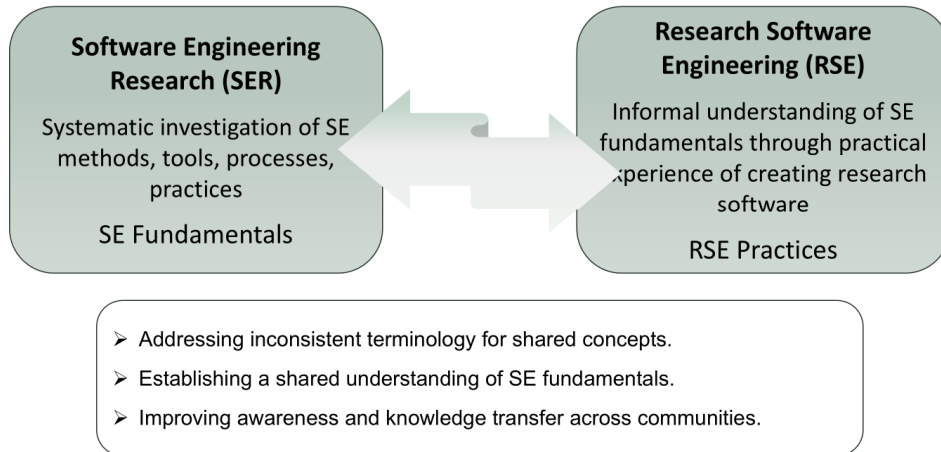


opportunities for adaptation
and refinement by the SER
community to make SE
fundamentals more suitable
for RS contexts.

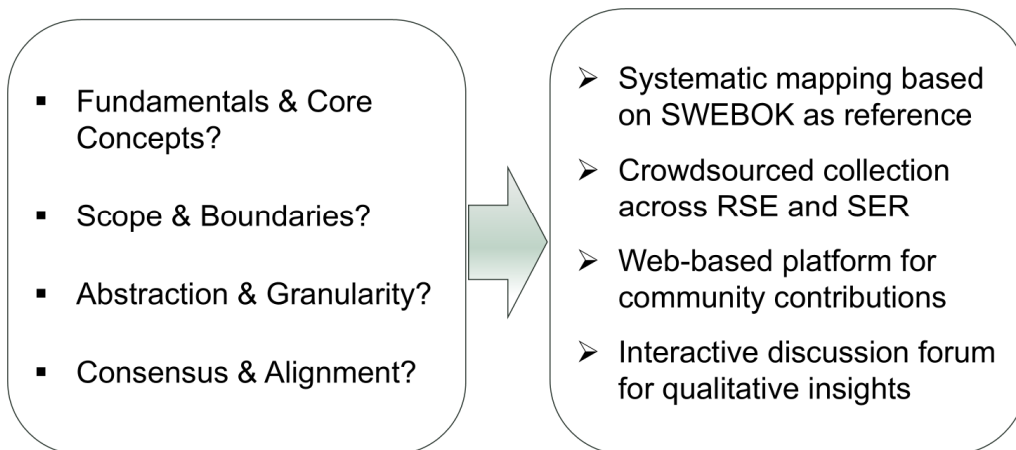
u^b Motivation and Goals



u^b Motivation and Goals



u^b Approach



u^b Emerging Questions from Initial Discussions

Scope & Boundaries

- What is the scope, where to draw boundaries?

Fundamentals & Core Concepts

- What counts as fundamental?

Abstraction & Granularity

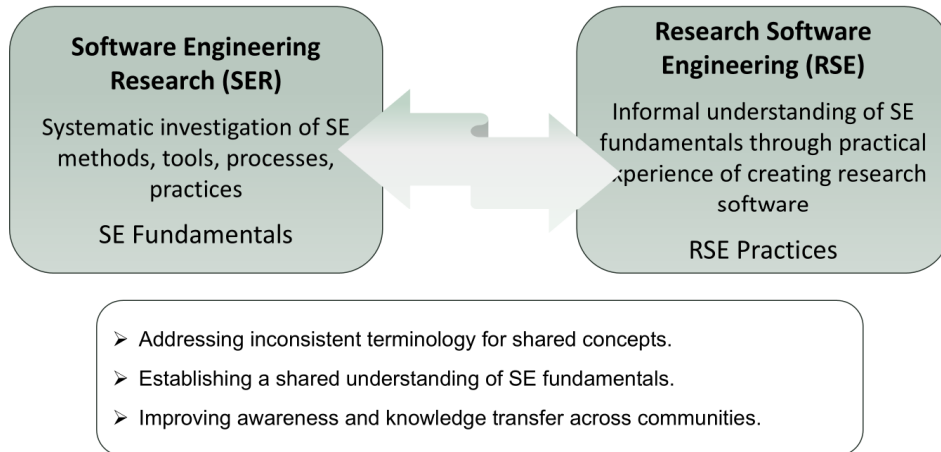
- What level of granularity is appropriate?

Consensus & Alignment

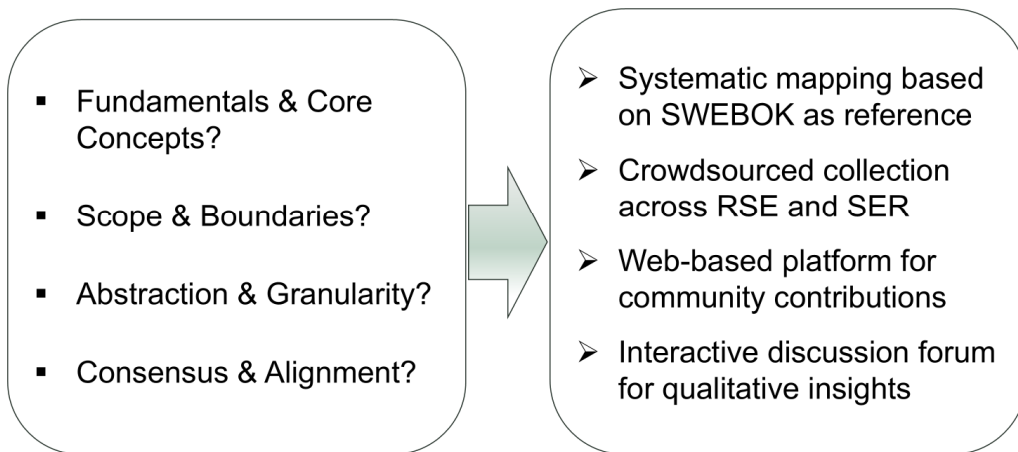
- How to reach consensus?



u^b Motivation and Goals



u^b Approach



u^b Emerging Questions from Initial Discussions

Scope & Boundaries

- What is the scope, where to draw boundaries?

Fundamentals & Core Concepts

- What counts as fundamental?

Abstraction & Granularity

- What level of granularity is appropriate?

Consensus & Alignment

- How to reach consensus?



u^b Envisioned Outcome

SE fundamentals unfamiliar to the RSE community



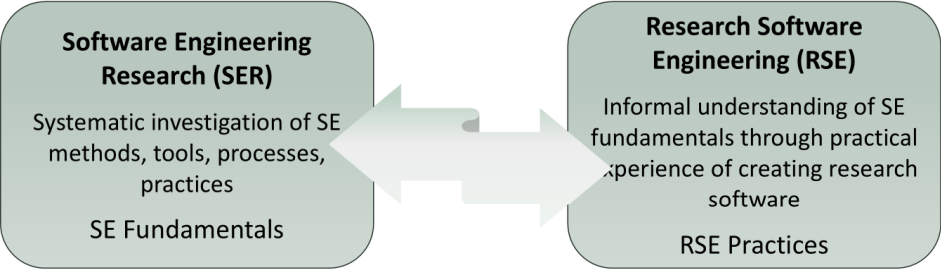
opportunities for education, training, and outreach to promote adoption by the RSE community;

SE fundamentals perceived as less useful by the RSE community



opportunities for adaptation and refinement by the SER community to make SE fundamentals more suitable for RS contexts.

u^b Motivation and Goals



u^b Emerging Questions from Initial Discussions

Scope & Boundaries

- What is the scope, where to draw boundaries?

Fundamentals & Core Concepts

- What counts as fundamental?

Abstraction & Granularity



the language?

ation,
ch to

Call for Contribution and Collaboration



- Fundamentals & Core Concepts?
- Scope & Boundaries?
- Abstraction & Granularity?
- Consensus & Alignment?

- on SWEBOK as reference
- Crowdsourced collection across RSE and SER
 - Web-based platform for community contributions
 - Interactive discussion forum for qualitative insights

unfamiliar to the RSE community

SE fundamentals perceived as less useful by the RSE community

promote adoption by the RSE community;

opportunities for adaptation and refinement by the SER community to make SE fundamentals more suitable for RS contexts.