



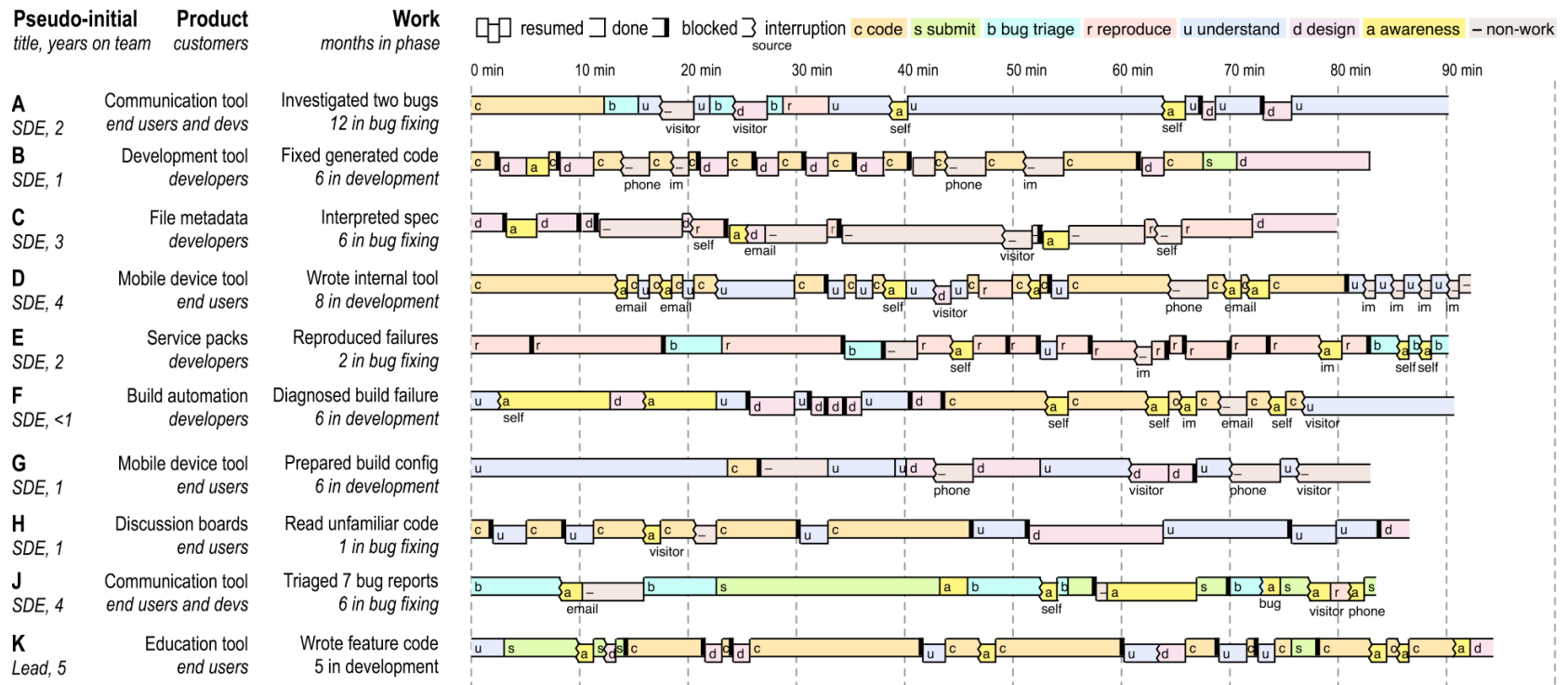
# Informationsbedürfnisse in Softwareentwickler-Teams

1. Vergleich zweier Studien zu Informationsbedürfnissen und -beschaffung von Softwareentwicklern
2. Codebook – Ein Lösungsansatz

# 1. Informationsbedürfnisse und -beschaffung: Studie Eins

Studie von Andrew Ko (Carnegie Mellon University), Robert DeLine (Microsoft Research) und Gina Venolia (Microsoft Research):

## Beobachten und Auswerten der Arbeitsweise von 17 Kollegen:



# 1. Informationsbedürfnisse und -beschaffung: Studie Eins

**Auswertung:**

- c: Code schreiben
- s: Änderungen vornehmen und submitten
- b: Fehler sichten
- r: Fehler reproduzieren
- u: fremden Code verstehen
- d: Design
- a: Informieren über die Arbeit von Kollegen (Awareness)

information type	search times			% agreed info is...			frequency and outcome of searches				frequency of sources
	min	mid	max	import.	unavail.	inacc.	acquired	deferred	gave up	beyond obs.	br = bug report, dbug = debugger
s1 Did I make any mistakes in my new code?	0	1	6	59	7	12	.....				dbug 10 compile 26 intuition 6 unit test 4
a2 What have my coworkers been doing?	0	1	11	17	10	10	.....				<b>coworker</b> 20 email 13 tool 4 bug alert 4 im 2
u3 What code caused this program state?	0	2	21	90	49	32	.....				dbug 16 <b>br</b> 3 intuition 3 log 3 tools 3 code 2 <b>coworker</b> 1
r2 In what situations does this failure occur?	0	2	49	80	32	20	.....				<b>br</b> 8 <b>coworker</b> 8 inference 5 tools 3 dbug 2 comment 1
d2 What is the program supposed to do?	0	1	21	93	29	29	.....				spec 13 <b>coworker</b> 9 docs 5 email 1
a1 How have resources I depend on changed?	0	1	9	41	15	15	.....				tools 12 <b>coworker</b> 6 email 4 <b>br</b> 2 code 1
u1 What code could have caused this behavior?	0	2	17	73	20	22	.....				<b>coworker</b> 5 intuition 4 log 4 <b>br</b> 4 dbug 2 im 1 code 1 spec 1
c2 How do I use this data structure or function?	0	1	14	71	20	29	.....				docs 11 code 5 <b>coworker</b> 4 spec 1
d3 Why was this code implemented this way?	0	2	21	61	37	39	.....				code 4 intuition 4 history 3 <b>coworker</b> 2 dbug 2 tools 2 comment 1 <b>br</b> 1
b3 Is this problem worth fixing?	0	2	6	44	10	20	.....				<b>coworker</b> 12 email 2 <b>br</b> 1 intuition 1
d4 What are the implications of this change?	0	2	9	85	44	49	.....				<b>coworker</b> 13 log 1
d1 What is the purpose of this code?	1	1	5	56	24	29	.....				intuition 5 code 2 dbug 2 tools 2 spec 1 docs 1
u2 What's statically related to this code?	0	1	7	66	27	27	.....				tools 8 intuition 2 email 1
b1 Is this a legitimate problem?	0	1	2	49	17	34	.....				<b>br</b> 5 <b>coworker</b> 1 log 1
s2 Did I follow my team's conventions?	0	7	25	41	10	15	.....				docs 2 tools 2 memory 1
r1 What does the failure look like?	0	0	2	88	24	23	.....				<b>br</b> 3 screenshot 2
s3 Which changes are part of this submission?	0	2	3	61	7	5	.....				tools 2 memory 2
c3 How I can coordinate this with this other code?	1	1	4	75	28	30	.....				docs 2 code 1 <b>coworker</b> 1
b2 How difficult will this problem be to fix?	2	2	4	41	15	32	.....				code 1 <b>coworker</b> 1 screenshot 1
c1 What can be used to implement this behavior?	2	2	2	61	27	22	.....				memory 1 docs 1
a3 What information was relevant to my task?	1	1	1	59	15	13	.....				memory 2

**Figure 3.** Types of information developers sought, with search times, perceptions importance availability, and accuracy, frequencies and outcomes of searches, and sources. The most common sources are in bold.

## Zusammenfassung:

- **Meistgenutzte Quelle: Arbeitskollegen** (und Bug Reports)
- **Meistgesuchte Informationen haben mit Fehlern im Code und mit der Arbeit der Kollegen zu tun**
- **Kein Zusammenhang zwischen Suchhäufigkeit und der Wichtigkeitsbewertung!**

Studie von Andrew Begel (Microsoft Research), Khoo Yit Phang (University of Maryland) und Thomas Zimmermann (Microsoft Research):

Umfrage bei Microsoft:

- 1. Verantwortlichen zu gegebenem Feature, API, Produkt oder Service finden (83%)**
- 2. Experten zu gegebenem Feature, API, Produkt oder Service finden (67%)**
3. Informationen zu einem Feature, API, Produkt oder Service aus einem anderen Team finden (64%)
4. Herausfinden, warum Code geändert wurde (62%)
5. Darüber informiert werden, dass Code geändert wurde (60%)

6. Herausfinden, wer von meinen Änderungen betroffen wird (57%)
- 7. Herausfinden, von wem ein bestimmter Code stammt oder wer damit gearbeitet hat (56%)**
- 8. Herausfinden, von wem eine bestimmte Spezifikation stammt oder wer am meisten darüber weiß (56%)**
- 9. Herausfinden, von wem ein benötigtes Feature, Produkt oder Service stammt (53%)**
10. Herausfinden, wer außerhalb meines Teams mein Feature, API, Produkt oder Service benötigt (50%)

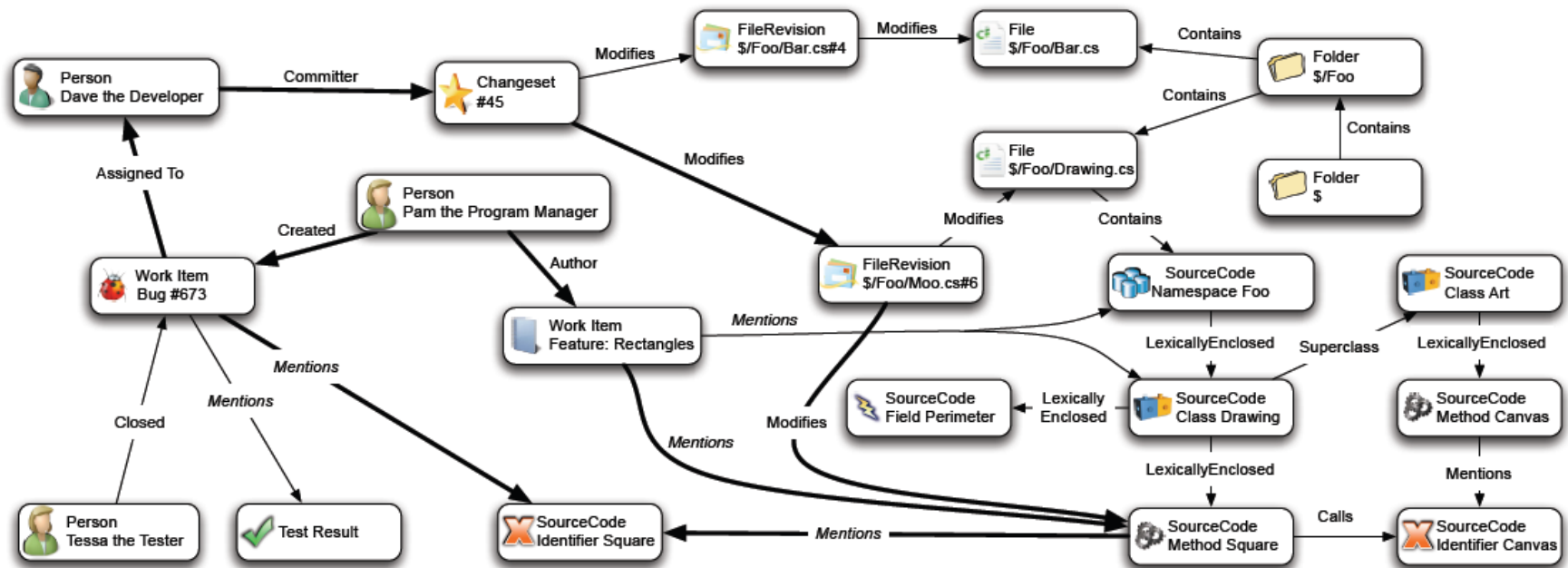
# 1. Informationsbedürfnisse und -beschaffung: Vergleich

	Studie Eins (Andrew Ko)	Studie Zwei (Andrew Bagel)
Umfeld	Microsoft	Microsoft
Befragte	17 bzw. 42	<b>1000</b>
Schwerpunkt	<b>rein wissenschaftlich</b>	<b>praxisbezogen</b>
Art des Vorgehens	<b>Beobachten</b> und Befragen	Befragen
Umfang	umfangreich und komplex	wenige Details
Informationen, die als „wichtig“ eingestuft wurden	Fehlerbehandlung, <b>Konsequenzen von Änderungen</b> , Programmverhalten	Personensuche, über Änderungen und deren Grund informiert werden, <b>Konsequenzen von Änderungen</b>
Gemeinsamkeiten der Ergebnisse	meistgenutzte Informationsquelle: <b>Kollegen</b>	wichtigste Informationsbedürfnisse haben mit <b>Personensuche zu tun</b>

## 2. Codebook – Ein Lösungsansatz

**Codebook:** Datenbanksystem das nach Informationen durchsucht werden kann

Struktur dahinter: Netzwerk von Objekten

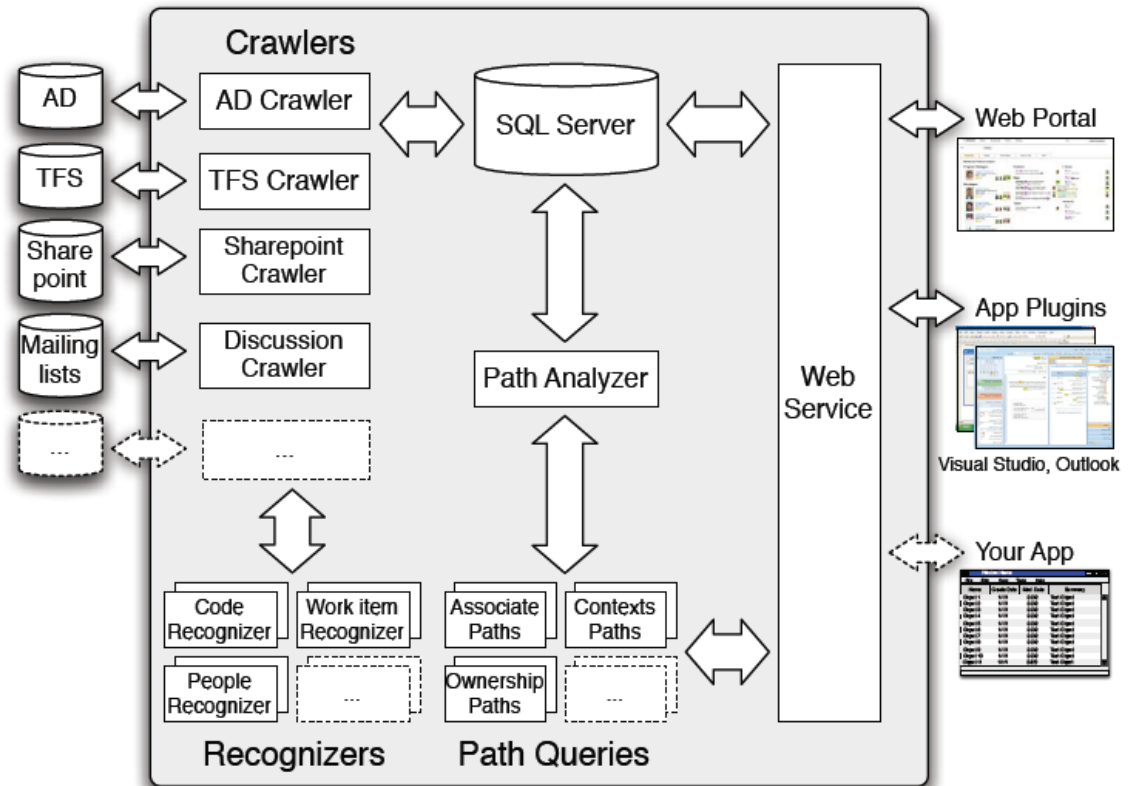


**Figure 2:** A canonical graph of possible relationships between objects in Codebook. “Mentions” edge labels are written in italics to indicate that they may derive from an incorrect inference. The bolded paths help explain a scenario written in the prose.



## 2. Codebook – Ein Lösungsansatz

Crawlers und Recognizers	Knoten mit Metadaten
SQL Server	Datenbank mit Knoten
Web Service	Schnittstelle
Applications	Zugriff auf Codebook - Struktur
Path Analyzer und Path Queries	Verbindungen (aus regulären Ausdrücken)



# 2. Codebook – Hoozizat

Codebook Home Bookmarks Notes Settings

Foo

Foo

Overview People Work Items Source Code Files

(Showing top 5 items per category)

## Program Managers

 **Gargiulo Abbruzzese**  
Program Manager, Department A  
Office-20167

Associates



## Developers

 **Rudolf Cuchares**  
Senior Developer, Department B  
Office-8081

Associates



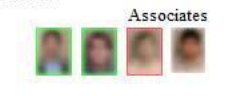
 **Bookard Ostrum**  
Developer, Department B  
Office-8079

Associates



 **Lindenberg Turnbo**  
Development Lead, Department B  
Office-8090

Associates



 **Leffel Westergard**  
Senior Developer, Department C

## Features

#42: **Foo** cloud storage feature  
Department B has requested that **Foo**...

## Bugs

#294 **Bug**: **Foo** service silently failed  
It happened a couple of times now...

#293 **Bug**: **Foo** gets 0 rows upon restart  
With no updates, **Foo** gets 0 rows after...

#302 **Bug**: **Foo** node failed to connect to index  
at startup  
At the startup, there is nothing in the DB. **Foo**...

## Tasks

#156 **Task**: Sample code for **Foo**  
No description

## Classes



**Foo**.Node  
class SaveTask

**Foo**.Core  
class FooReader

**Foo**.Node  
class Task

 **Neese Barkema**  
Tester, Department D  
Office-1337  
[fysh@microsoft.com](mailto:fysh@microsoft.com)  
[Send IM](#)

## Methods

**Foo**.Core  
Get()

**Foo**.Importers  
StartImporting()

**Foo**.FooService  
Import()

- 1. Verantwortlichen zu gegebenem Feature, API, Produkt oder Service finden (83%)**
- 7. Herausfinden, von wem ein bestimmter Code stammt oder wer damit gearbeitet hat (56%)**
- 8. Herausfinden, von wem eine bestimmte Spezifikation stammt oder wer am meisten darüber weiß (56%)**
- 9. Herausfinden, von wem ein benötigtes Feature, Produkt oder Service stammt (53%)**

1. Studie von Andrew Begel  
**Codebook: Discovering and Exploiting Relationships in Software Repositories**
2. Studie von Andrew Ko  
**Information Needs in Collocated Software Development Teams**