Business Process Mining and Conformance

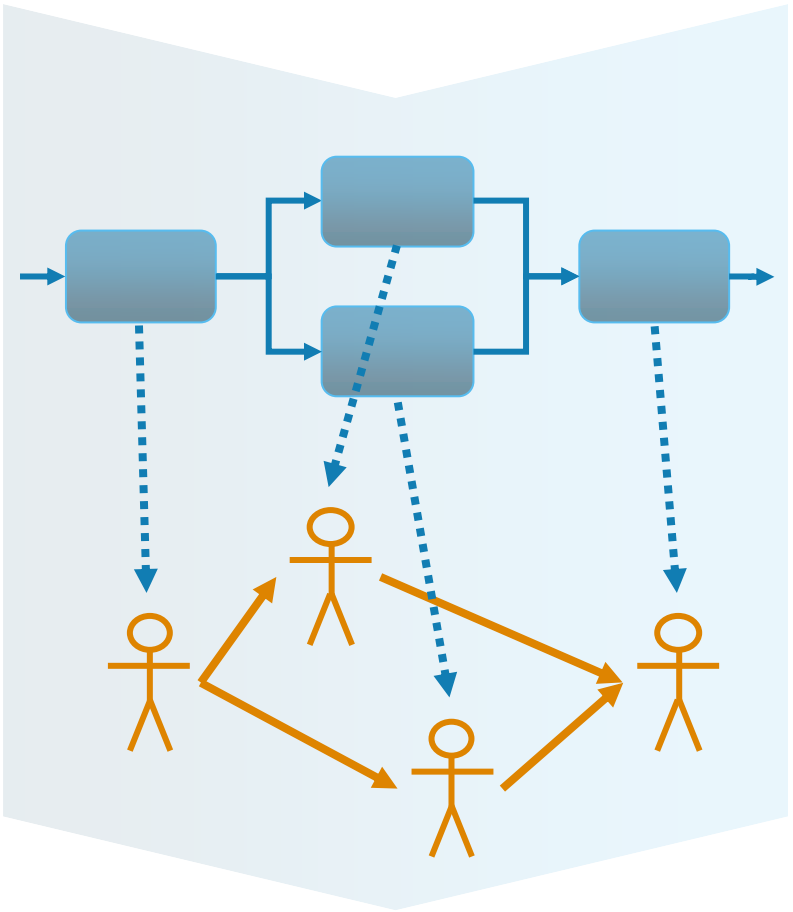
Diogo R. Ferreira

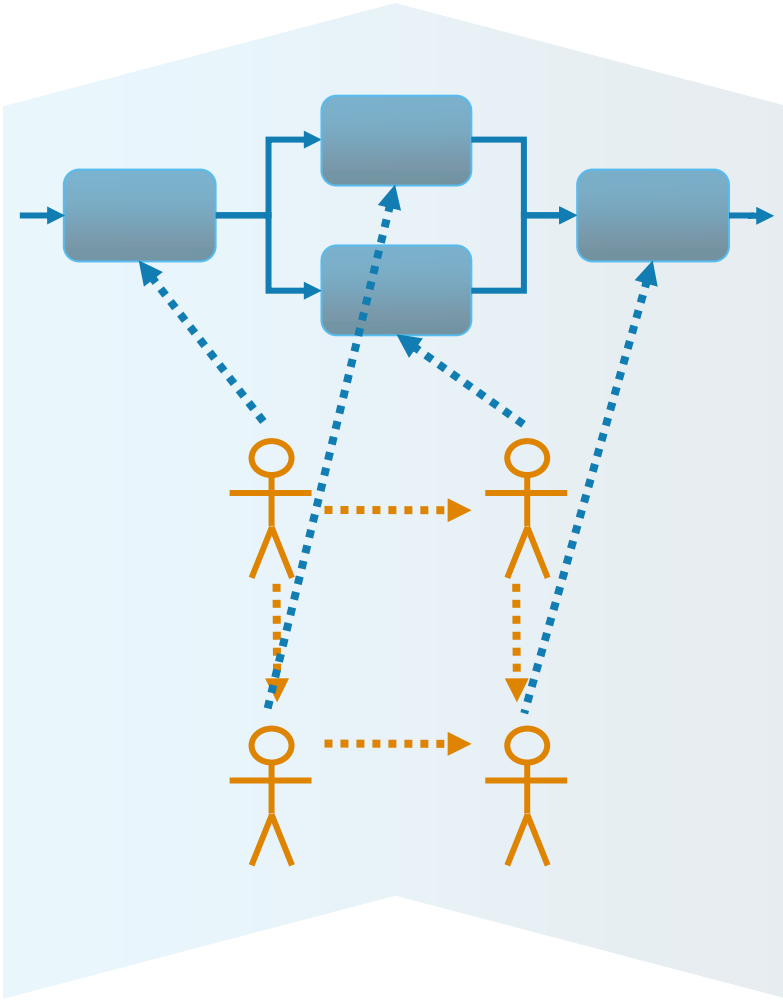
IST – Technical University of Lisbon

diogo.ferreira@tagus.ist.utl.pt

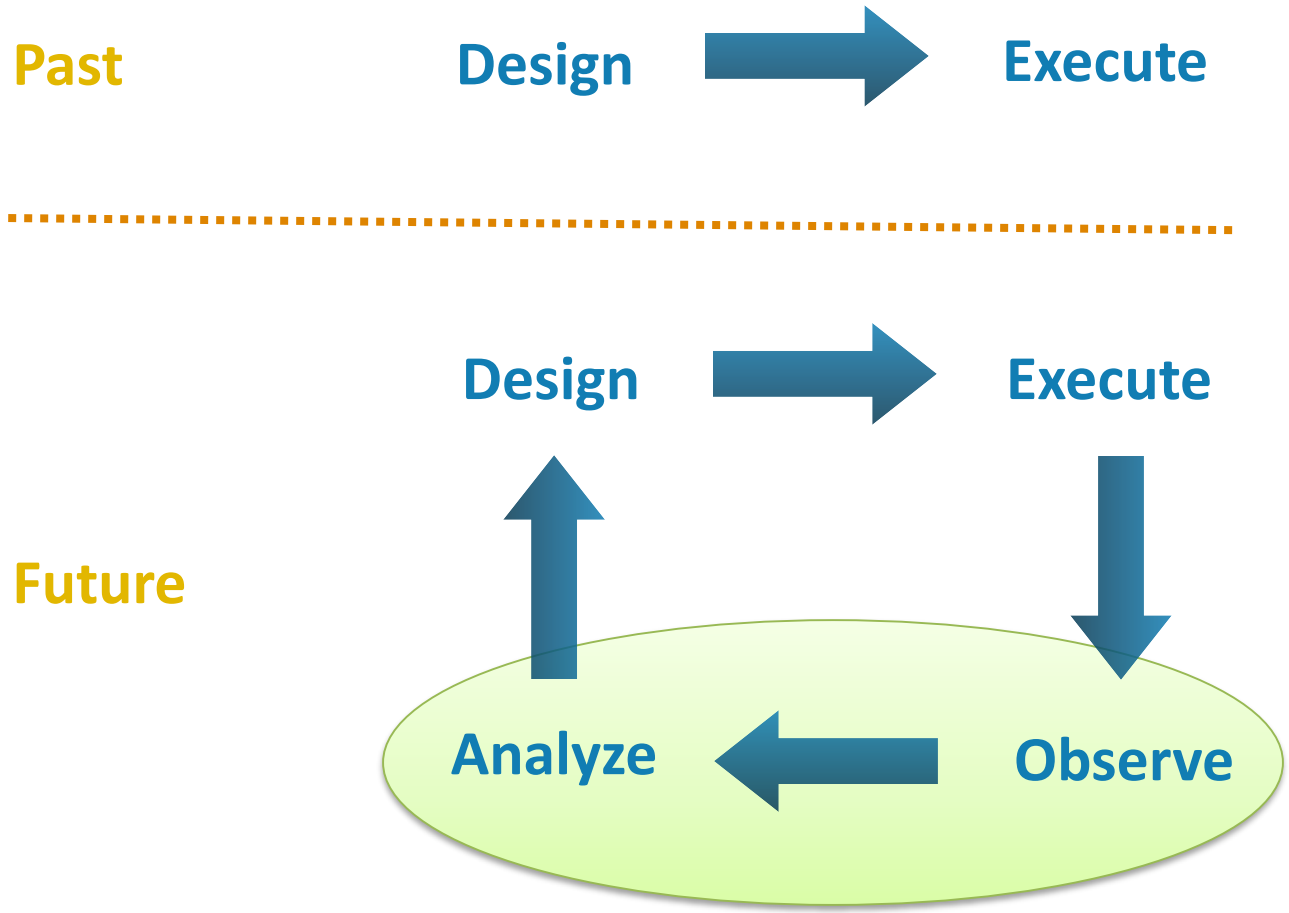
INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

Top-down vs. bottom-up

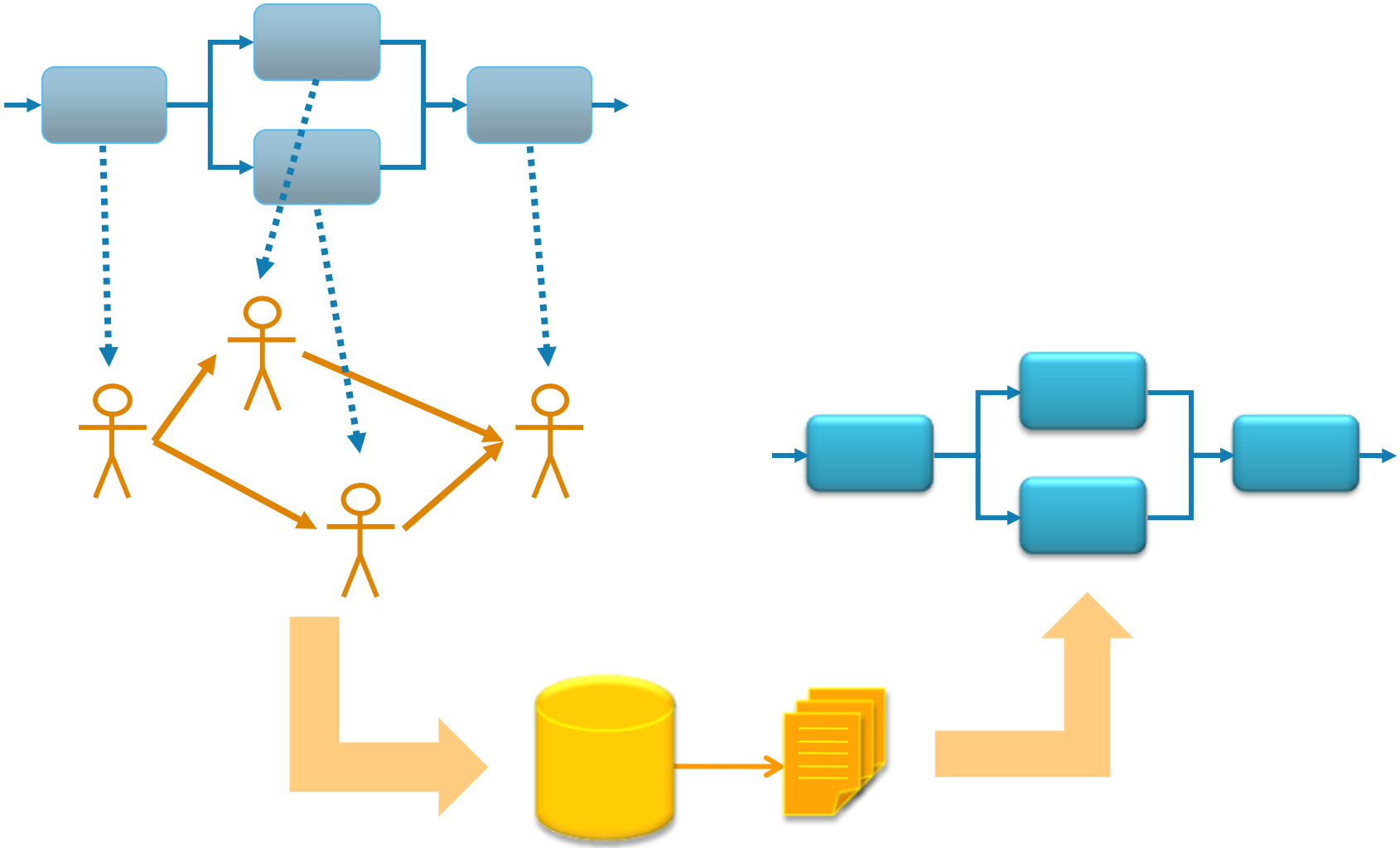




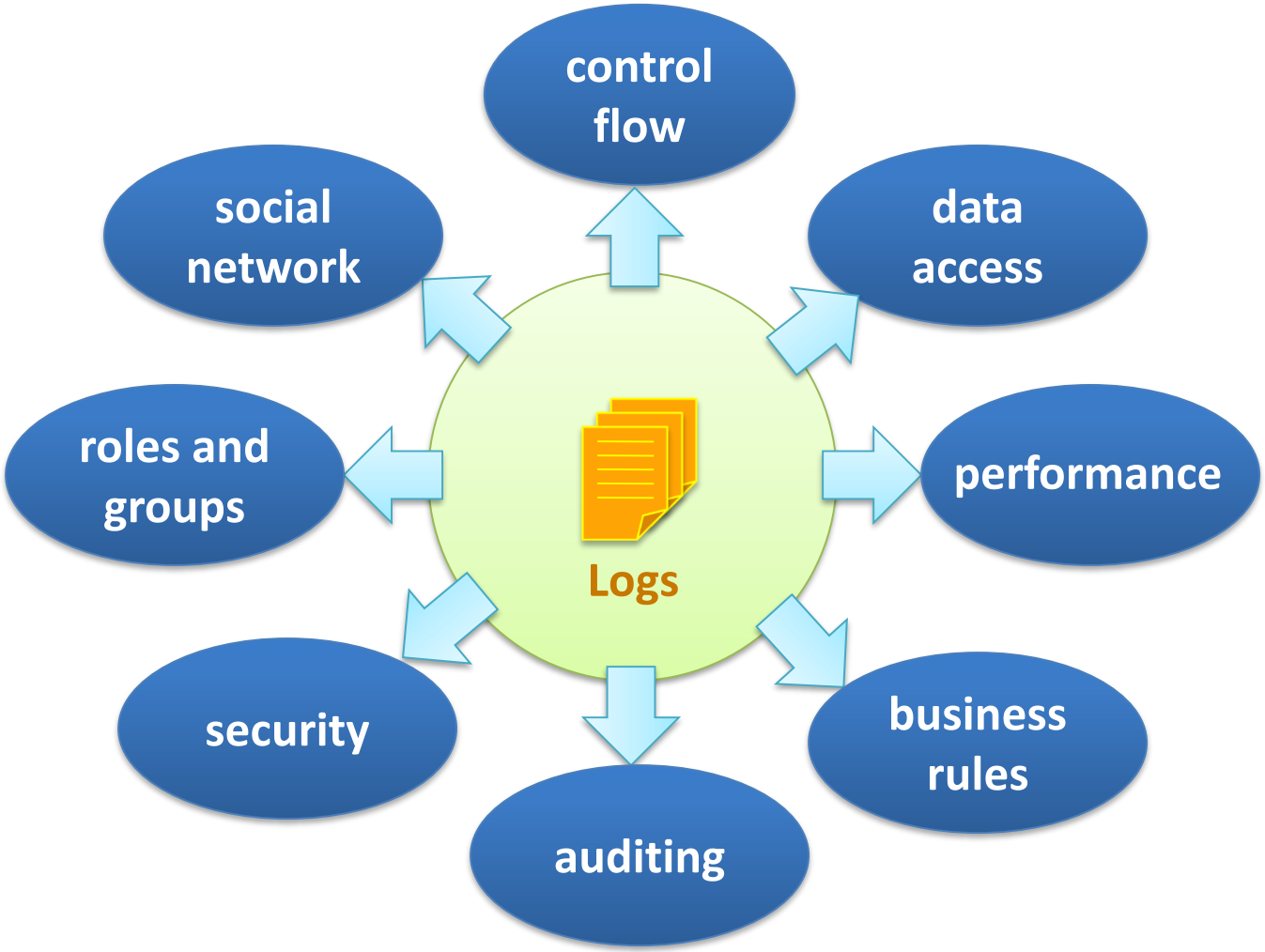
Life cycle


Past
Design
Execute
Future
Design
Execute
Analyze
Observe

Process mining



Mining potential



Need for conformance

- internal control (e.g. Sarbanes-Oxley)
- risk management (e.g. Basel II)
- best practices (e.g. ITIL)
- etc.

- *are processes being performed according to plan?*
- *do processes adhere to common best practices?*
- *…*

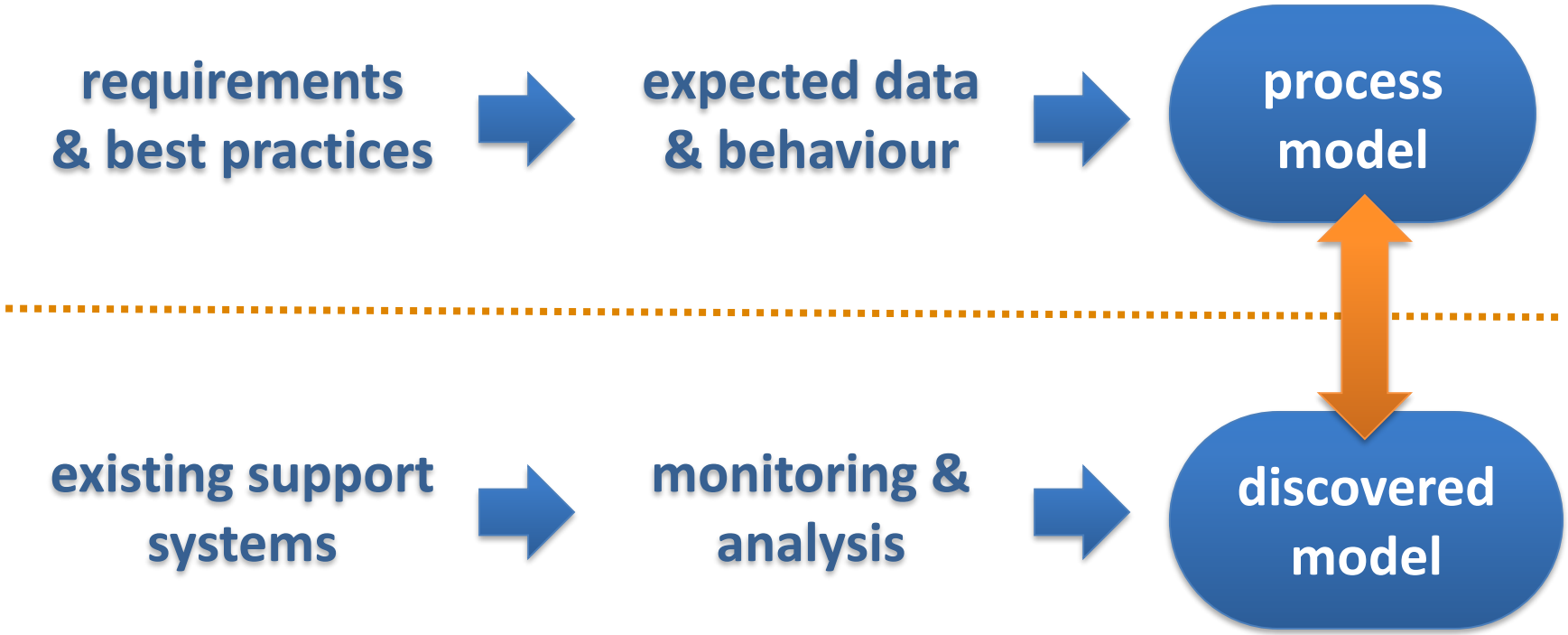
ITIL

- ITIL = Information Technology Infrastructure Library
  - comprehensive framework of best practices for ITSM
    - service level management
    - capacity management
    - incident management
    - configuration management
    - release management
    - change management
    - etc.
- customer-focus, quality, availability, reliability, cost,…

ITIL Assessment

- compare:
  - the way the ITSM is currently performed
  - the ITIL guidelines
- possible gaps
  - process, people, technology, services supplied by third parties, etc.
    - a gap exists if not all incidents are recorded
    - a gap exists if incidents are not classified
    - etc.
- identify requirements for ITIL implementation

Process conformance



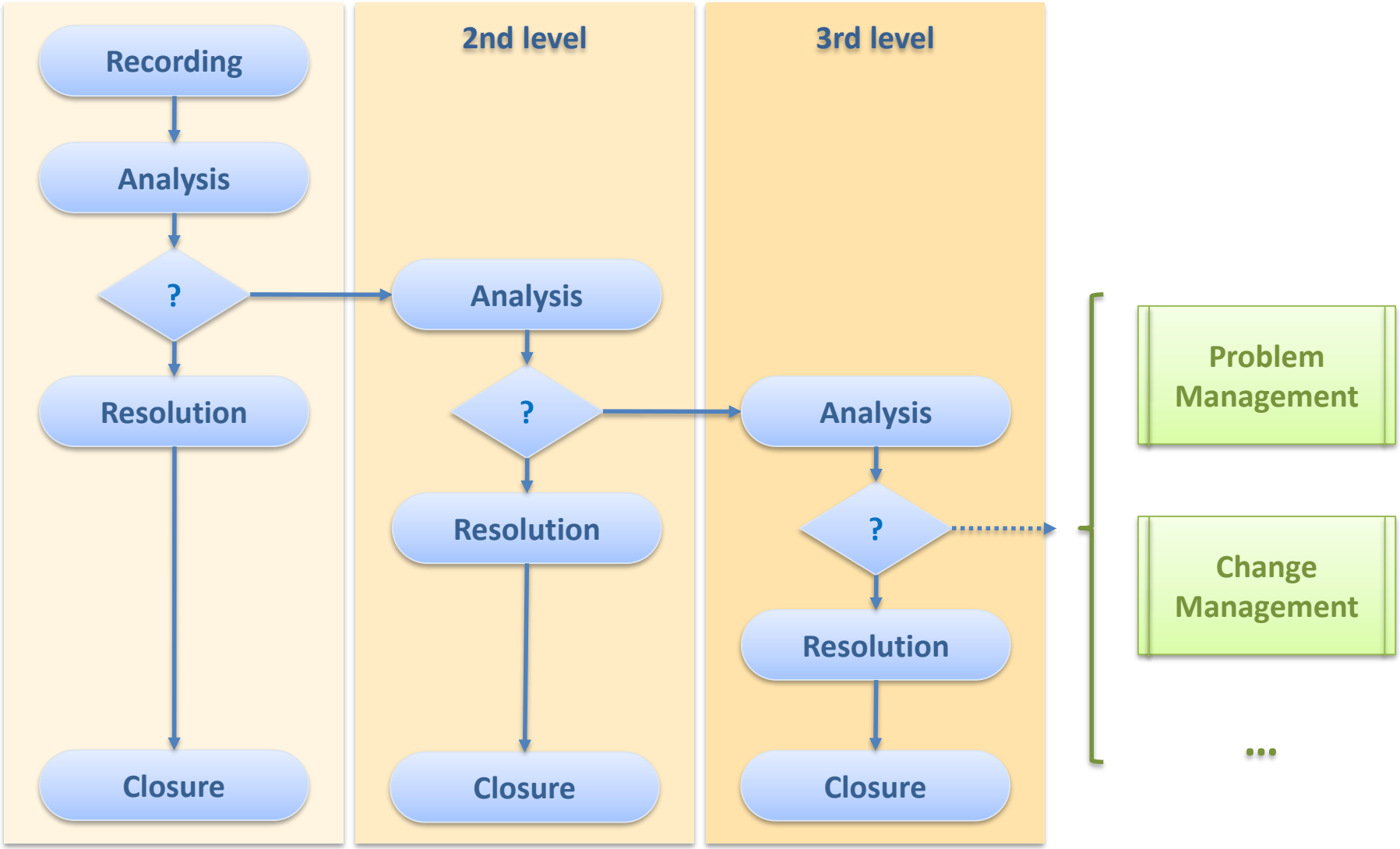
Case Study

Incident Management

ITIL Incident Management

- definition of incident
  - any event causing interruption or reduction in QoS
- purpose of incident management
  - restore normal service level as soon as possible
- comprises
  1. Recording
  2. Classification
  3. Matching
  4. Diagnosis
  5. Resolution
  6. Closure

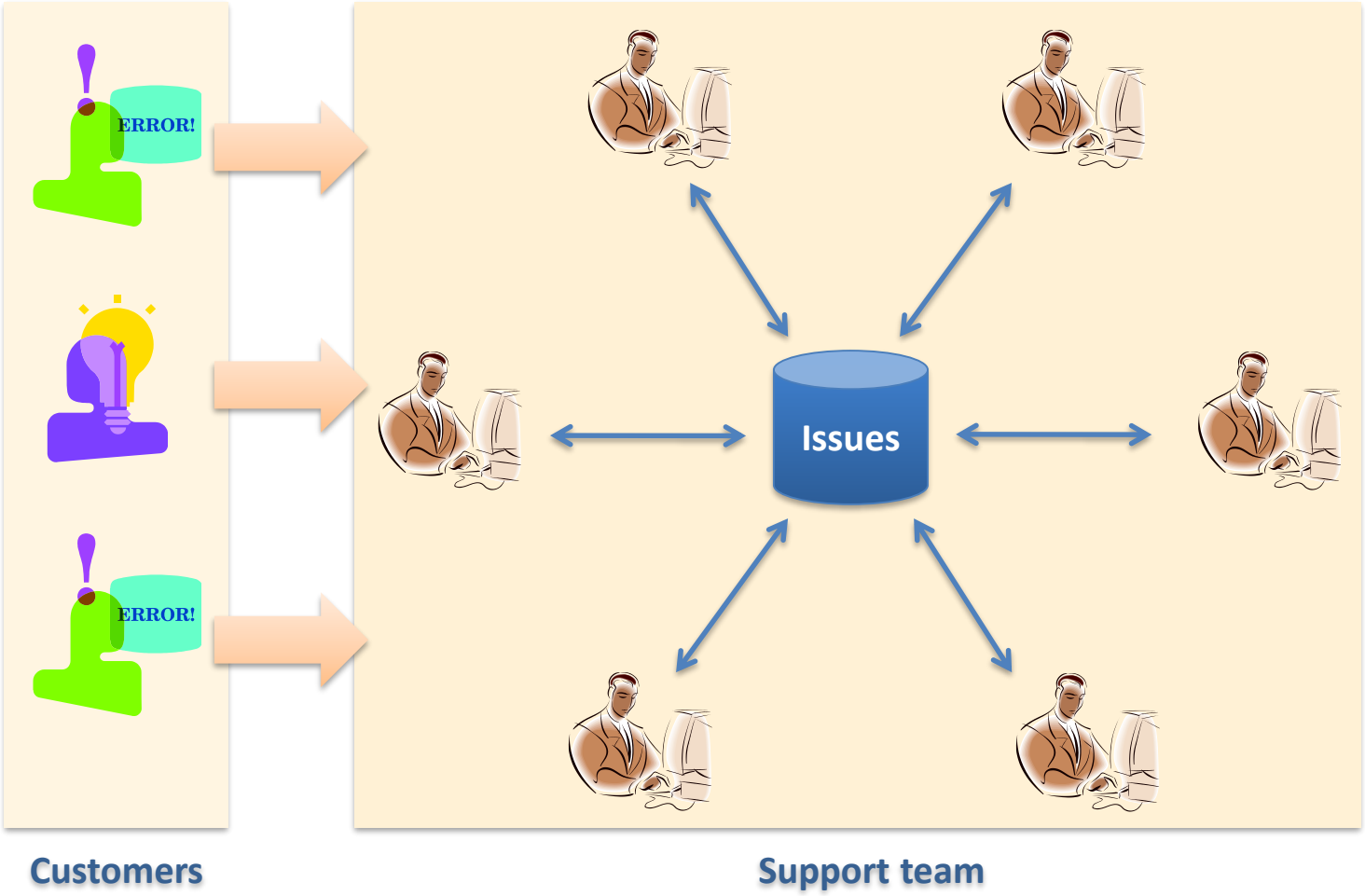
ITIL Incident Management – *Escalation*



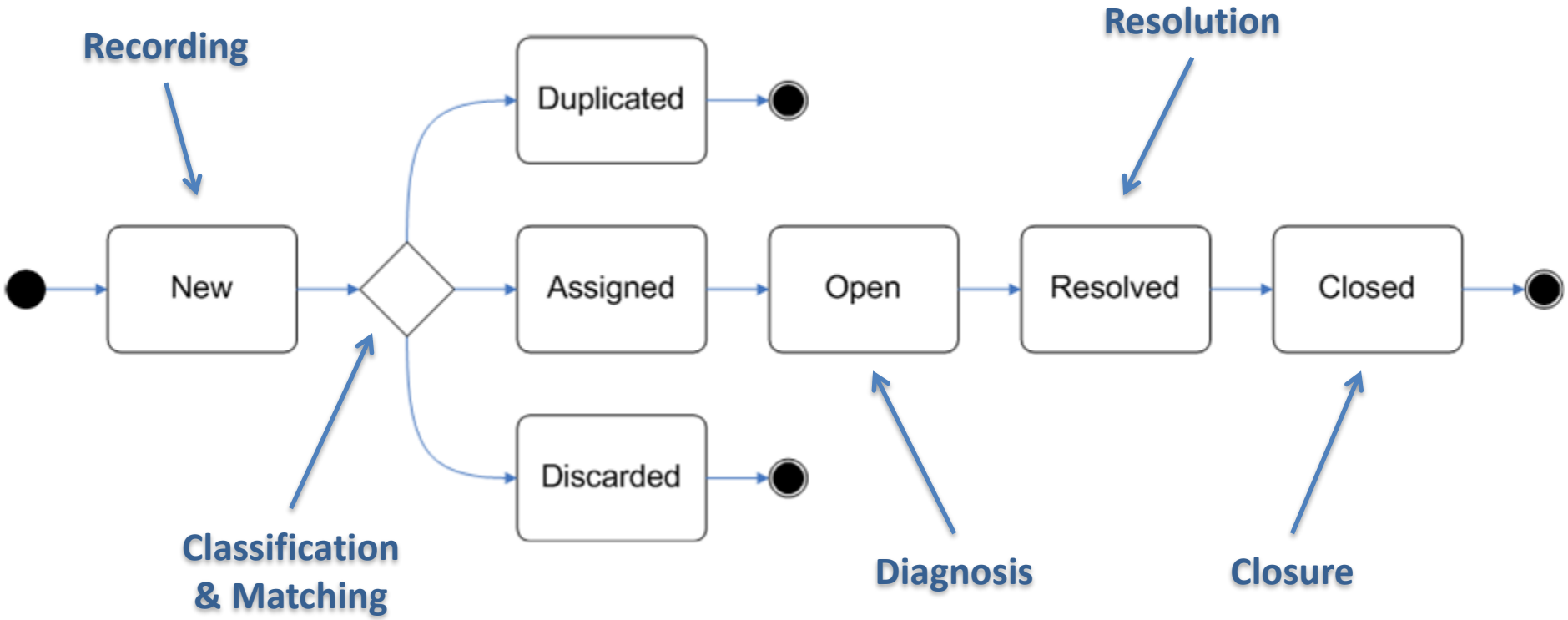
Case study scenario

- IT company
  - complex software platform for custom business solutions
  - continuously improved by successive release versions
  - extensive in-house testing
    - both manual and automated
  - users play active role
    - desired improvements
    - problems to be solved
  - issue handling process & supporting system

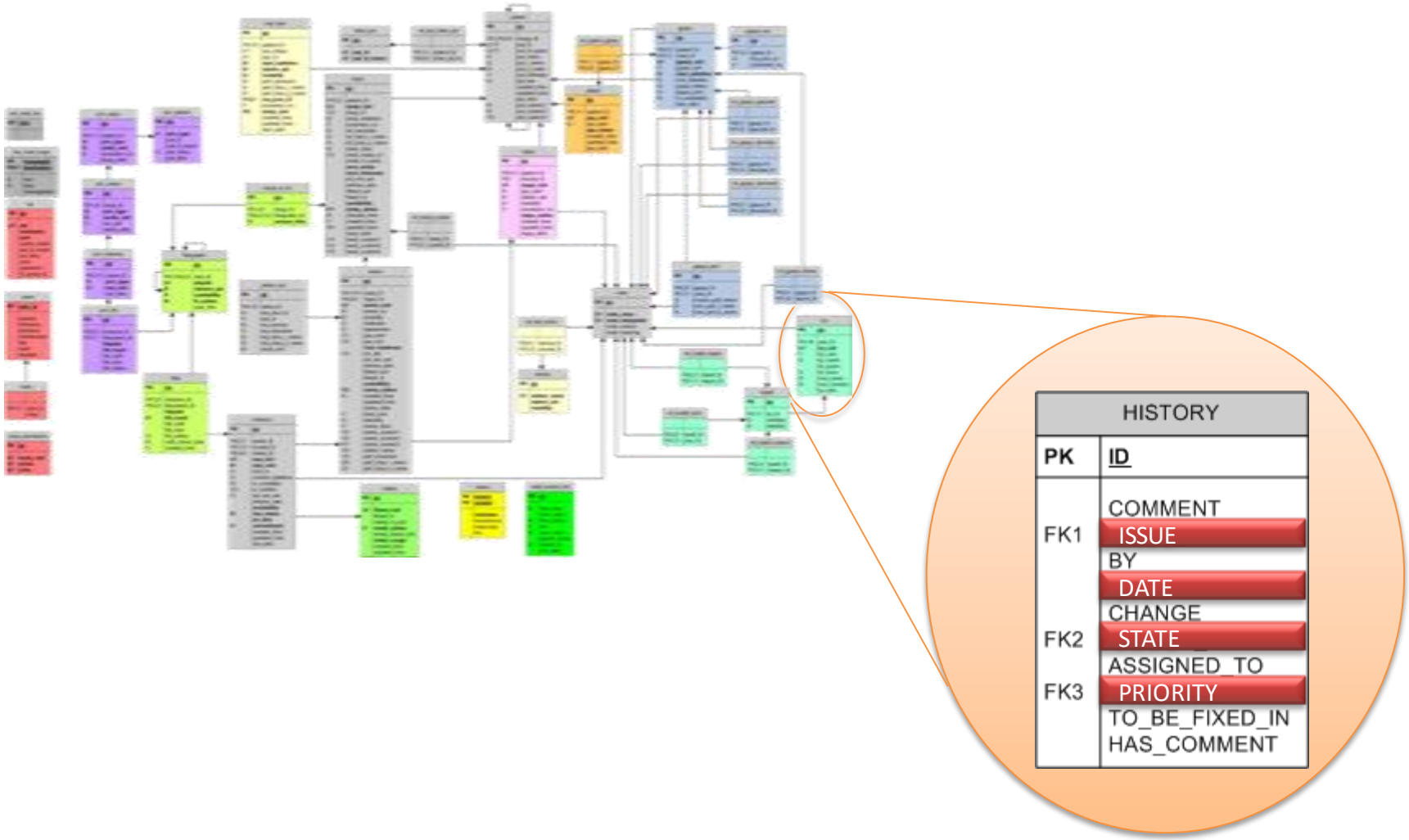
Case study scenario



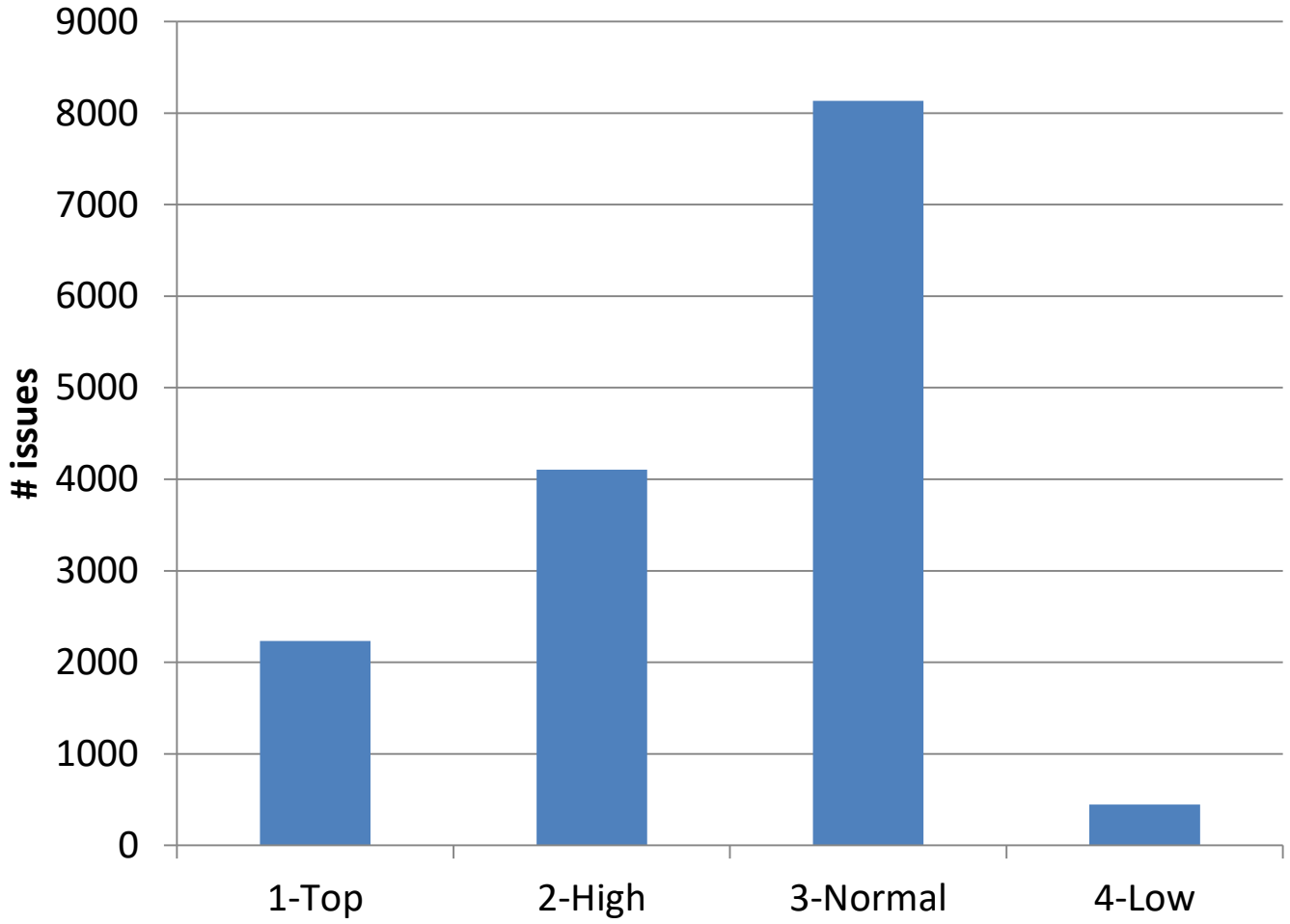
Case study scenario



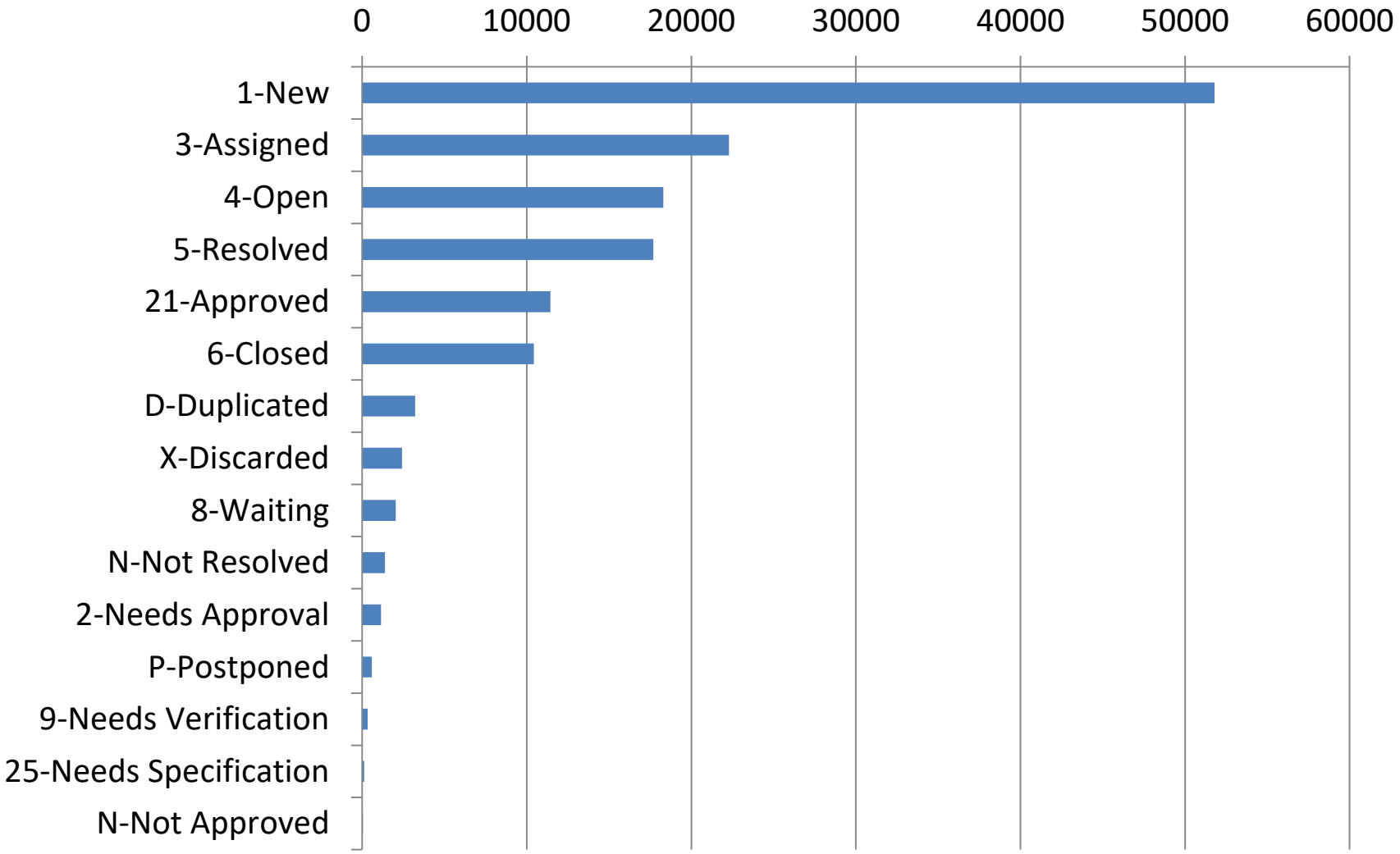
System database



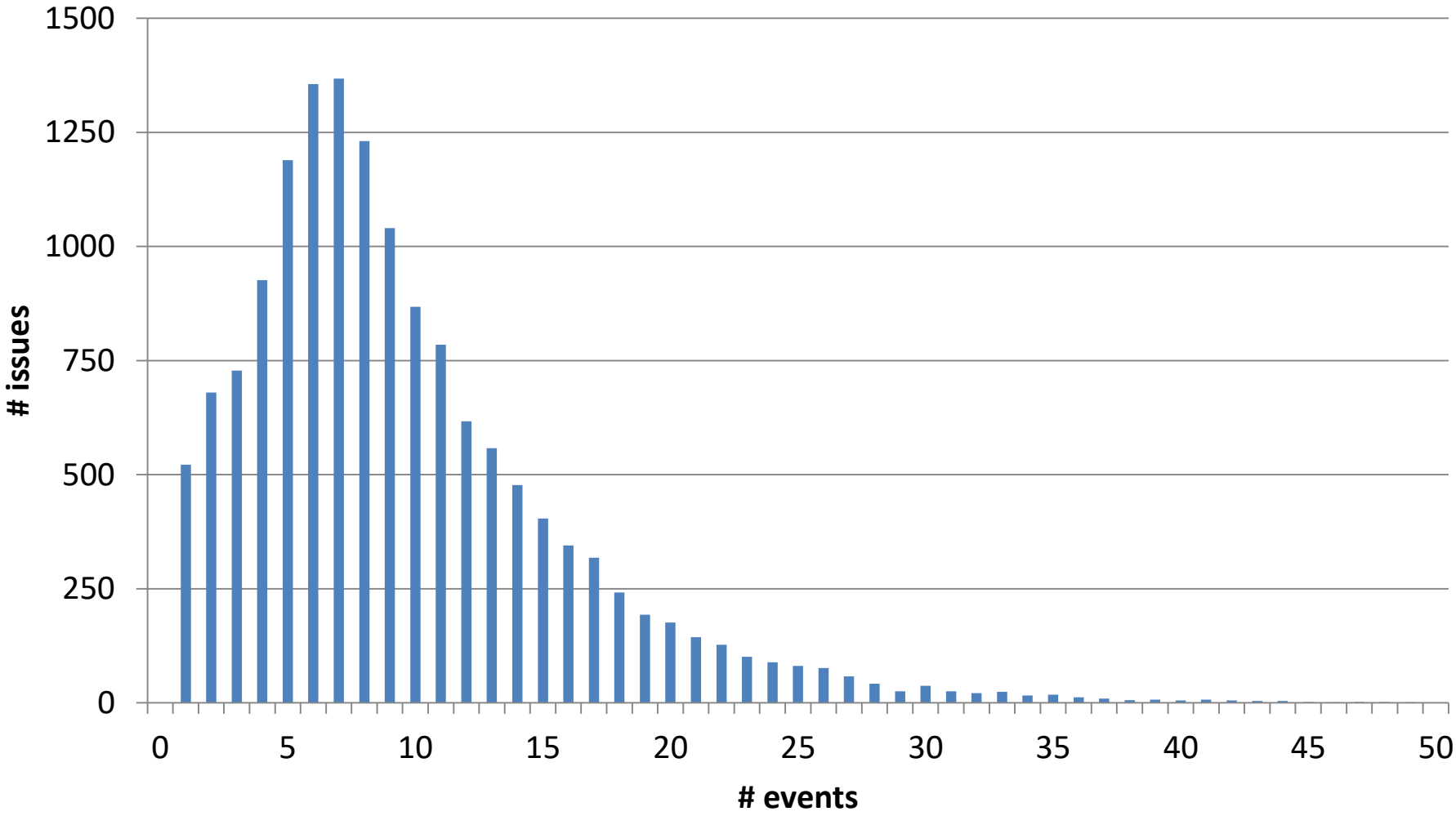
Priority


9000
8000
7000
6000
5000
4000
3000
2000
1000
0
issues
1-Top
2-High
3-Normal
4-Low

State


0
10000
20000
30000
40000
50000
60000
1-New
3-Assigned
4-Open
5-Resolved
21-Approved
6-Closed
D-Duplicated
X-Discarded
8-Waiting
N-Not Resolved
2-Needs Approval
P-Postponed
9-Needs Verification
25-Needs Specification
N-Not Approved

Sequence length



Sequence length

- Shortest sequence:

  New **(1 event)**

- Longest sequence:

  New => New => New => New => Assigned => Assigned => Assigned => Assigned => Approved => Approved => Approved => Approved => Approved => Approved => Waiting => Assigned => Assigned => Assigned => Open => Waiting => NotResolved => NotResolved => New => New => New => New => Assigned => Open => Open => Waiting => Open => Open => Waiting => Waiting => Resolved => Resolved => Resolved => Resolved => Resolved => Resolved => Closed **(75 events)**

- Average length: 143 220 / 14 982 ≈ 10 events

Input data

- 14 982 sequences in total
- 7164 different sequences!
  - on average, each sequence repeats only once (2x)
- 5909 unique sequences!
  - 40% of all behavior never repeats itself!

Strong component of ad-hoc behavior!

1. Drop consecutive repeated events

New => Assigned => Assigned => Open => Resolved => Resolved => Closed

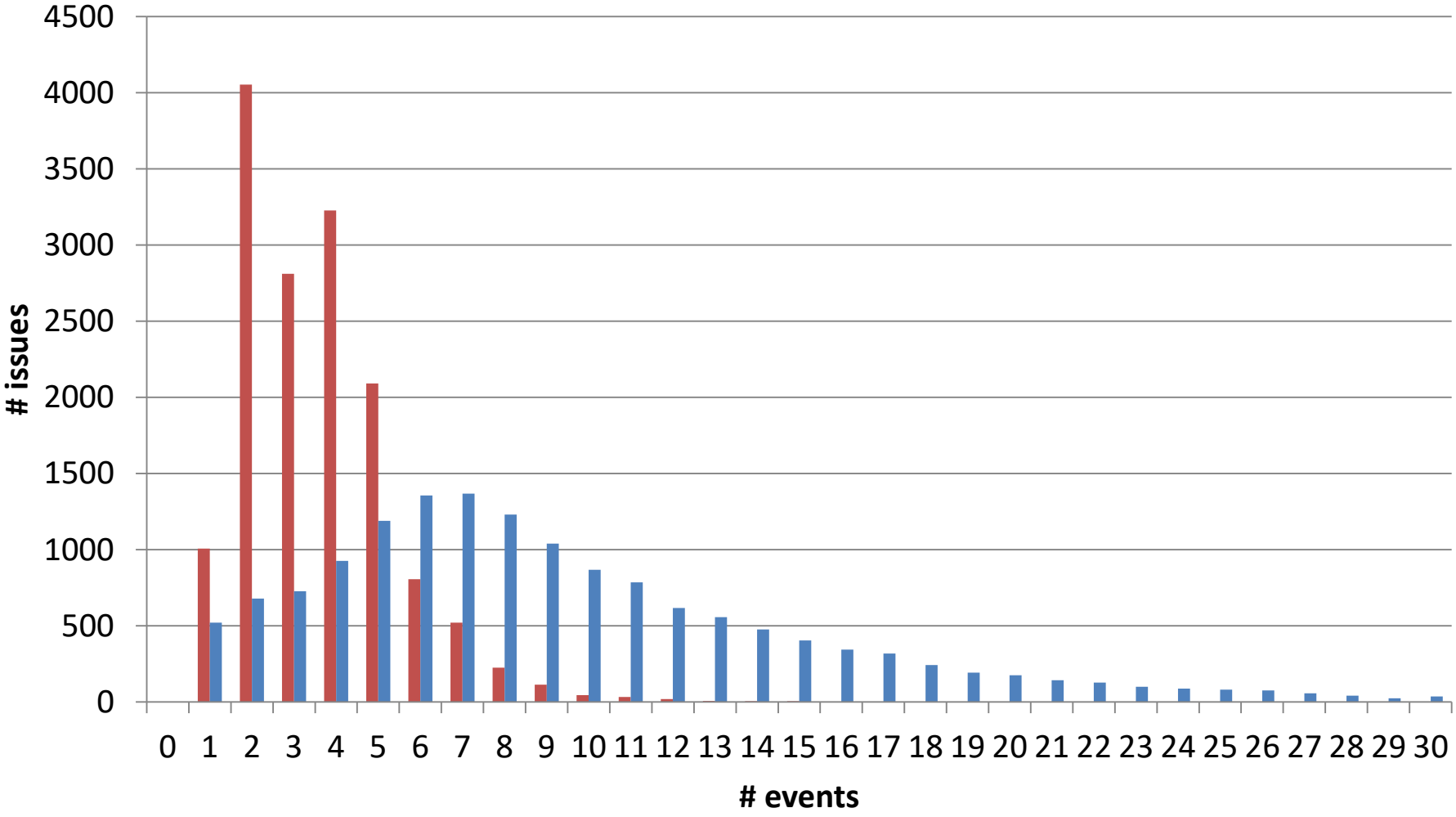
New => Assigned => Open => Resolved => Closed

New => New => New => New => Duplicated

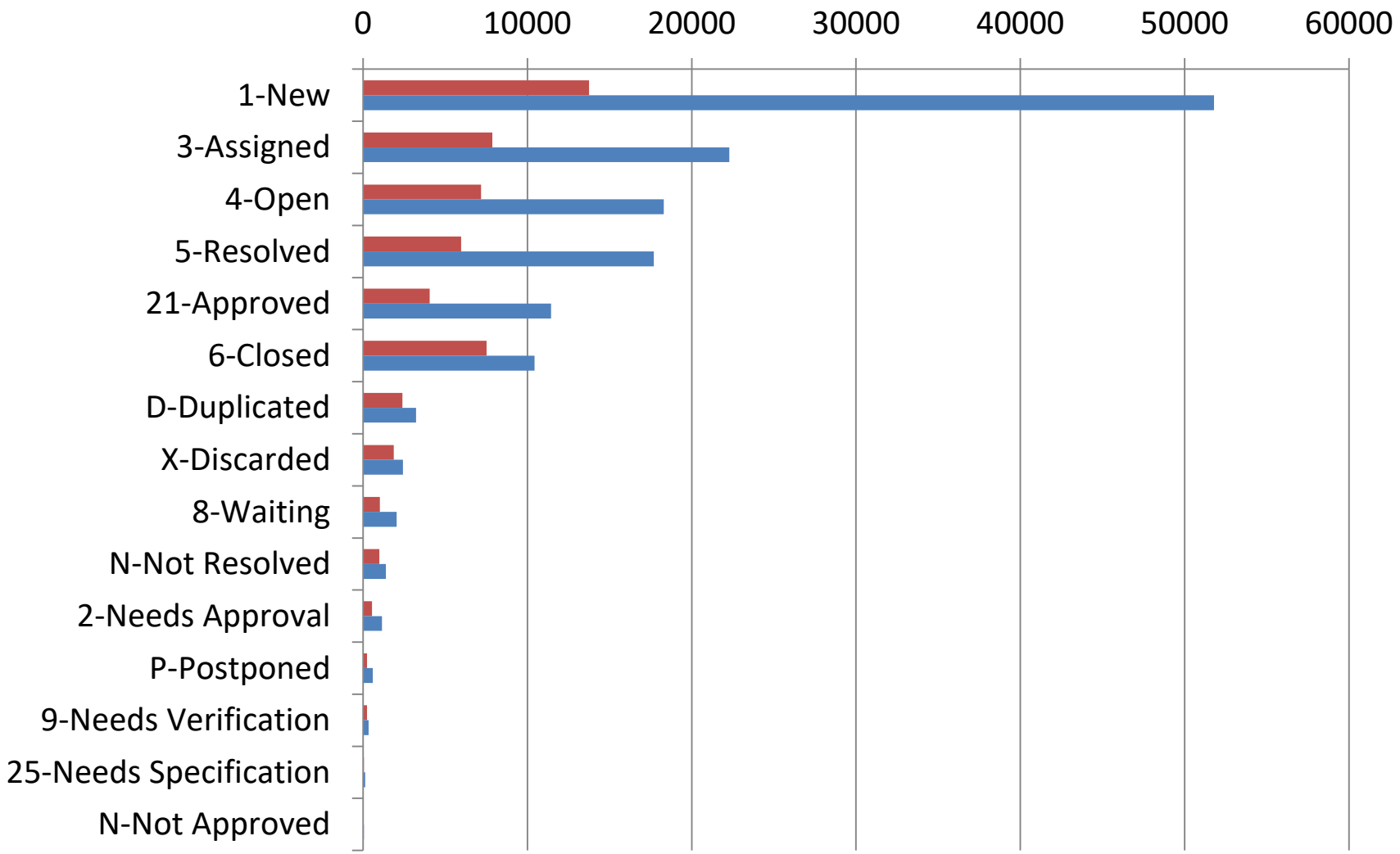
New => Duplicated

Eliminated 89 577 of 143 220 events (62.5%)

Sequence length



State


0
10000
20000
30000
40000
50000
60000
1-New
3-Assigned
4-Open
5-Resolved
21-Approved
6-Closed
D-Duplicated
X-Discarded
8-Waiting
N-Not Resolved
2-Needs Approval
P-Postponed
9-Needs Verification
25-Needs Specification
N-Not Approved

2. Drop single-event sequences

- sequences with a single step

New	(419)
Closed	(358)
Discarded	(128)
Duplicated	(83)
Approved	(9)
Assigned	(4)
NotResolved	(4)
Resolved	(2)

Eliminated 1007 of 14 982 sequences (6.7%)

3. Drop sequences with repeated events

New => Assigned => Open => Resolved => Assigned => Resolved => Closed

New => Assigned => Approved => Open => Approved => NotResolved

New => Duplicated => New => Duplicated => New => Duplicated

Closed => Assigned => NeedsApproval => Approved => Resolved => Closed

New => Assigned => Open => Discarded => New => Approved => Discarded

New => Open => NotResolved => Open => NotResolved => Resolved => Closed

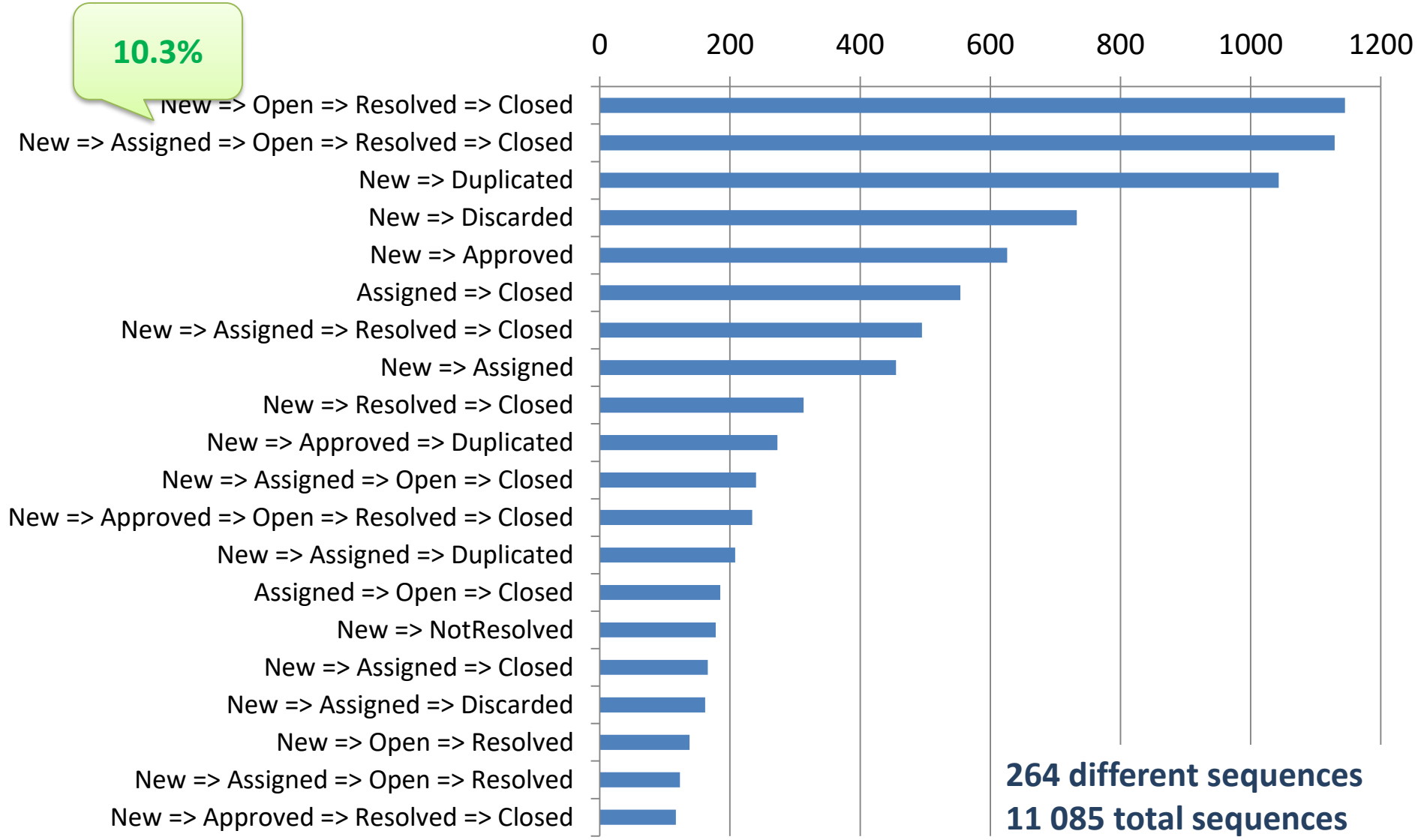
Eliminated 2584 of 13 975 sequences (18.5%)

4. Drop unique sequences

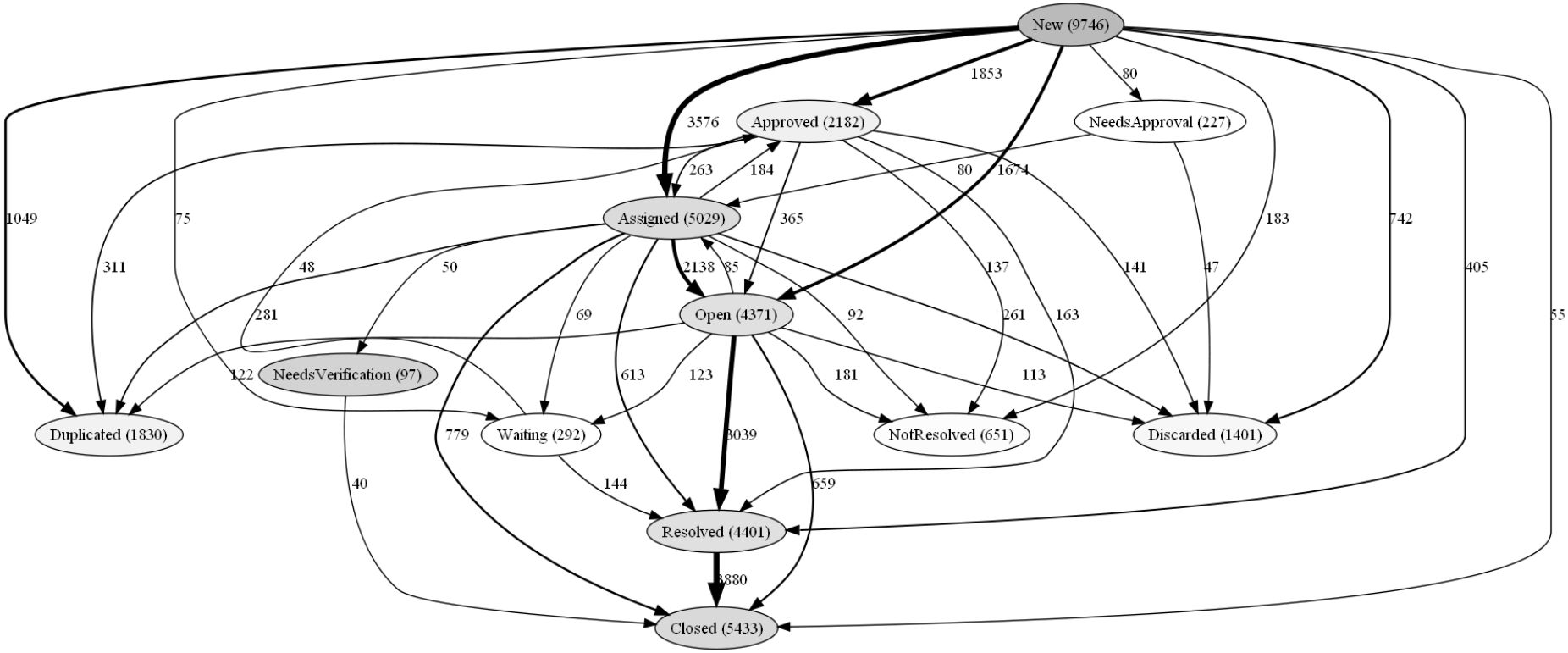
- from very short ones:
  - Closed => Discarded
  - Waiting => Approved
  - Postponed => New
- to longer ones:
  - New => NeedsApproval => NeedsSpecification => Approved => Assigned => Resolved => Closed
  - New => Discarded => Approved => Waiting => Assigned => Resolved => Closed => Duplicated

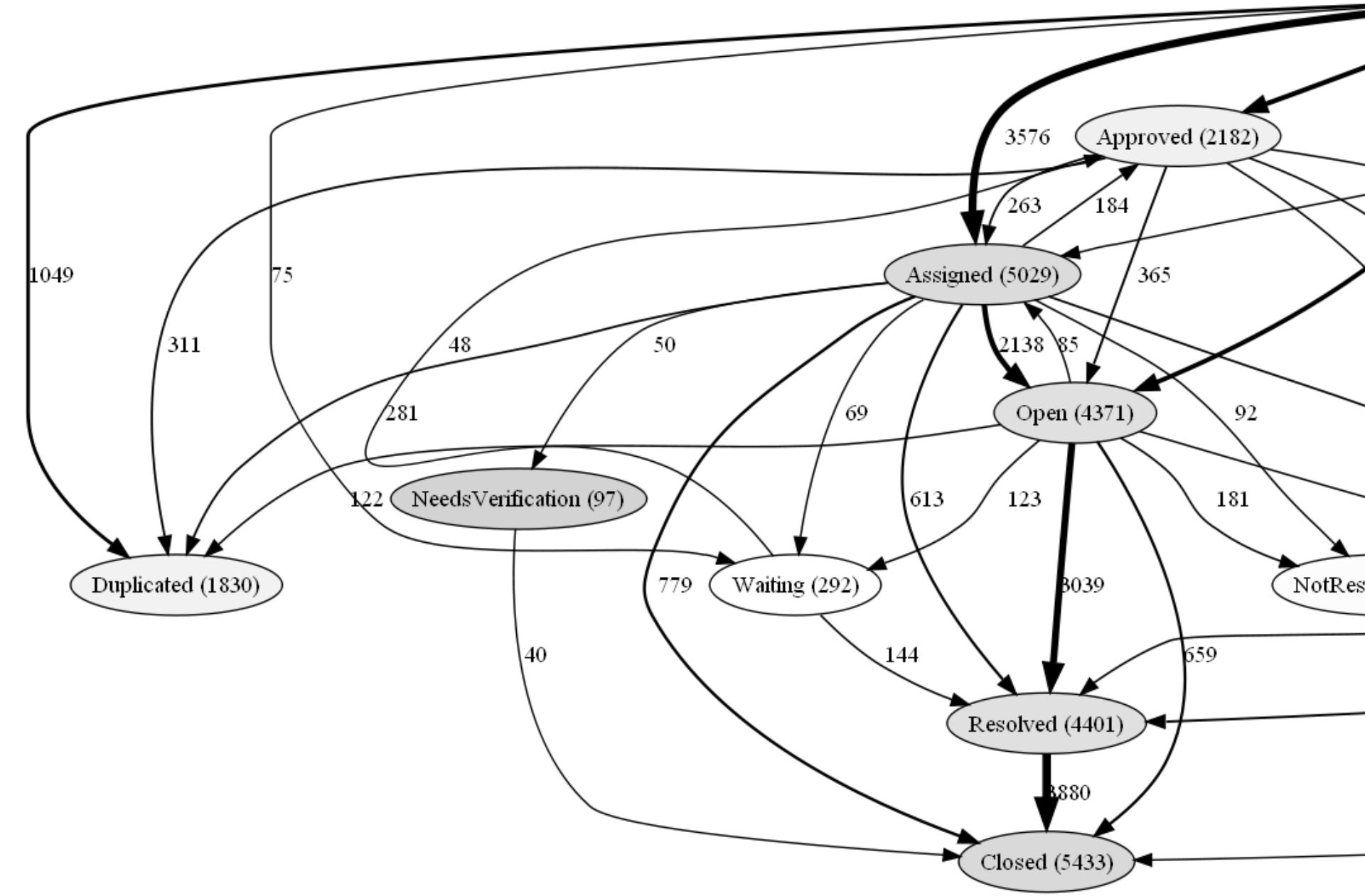
Eliminated 306 of 11 391 sequences (2.7%)

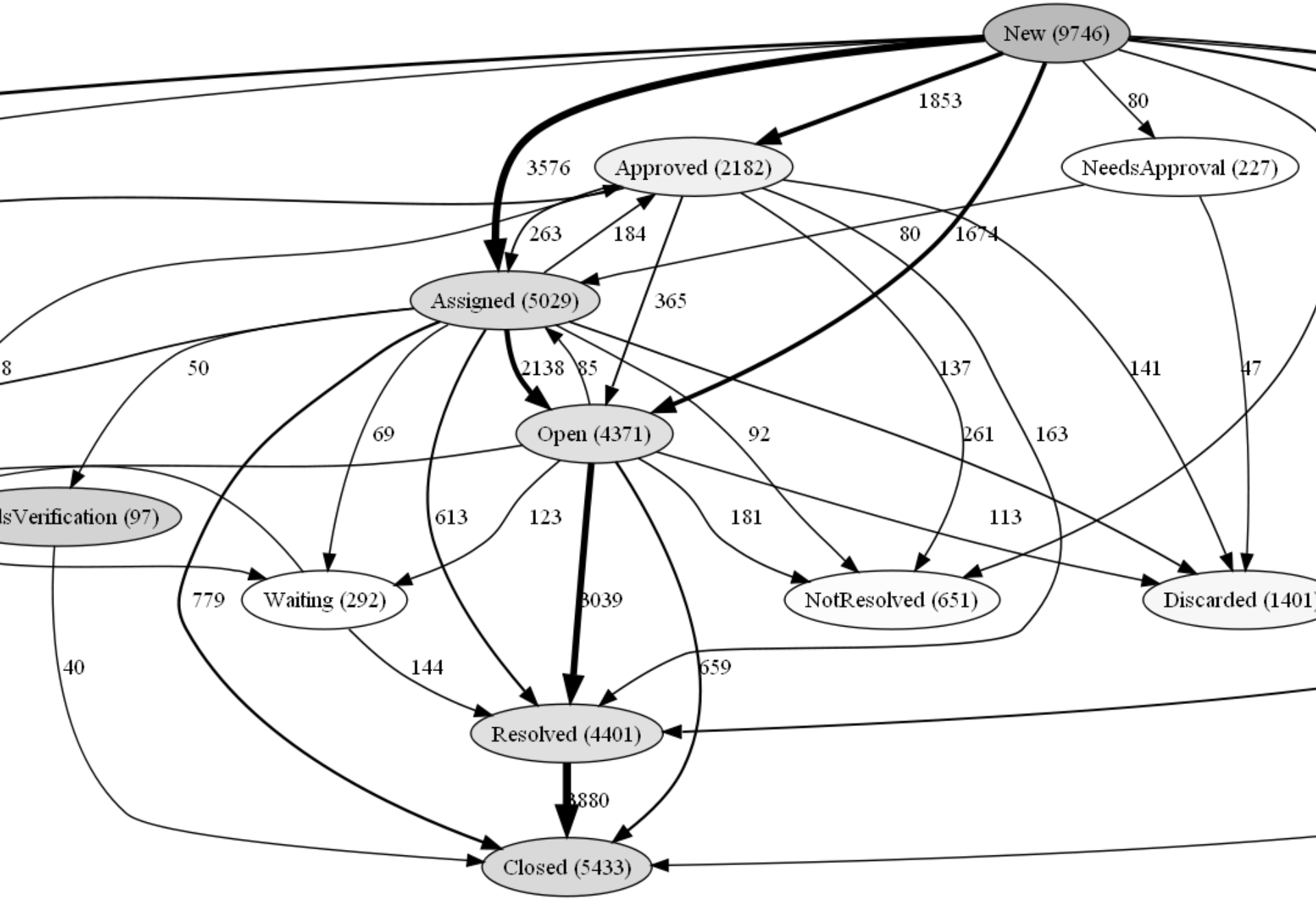
Most frequent sequences (top 20)

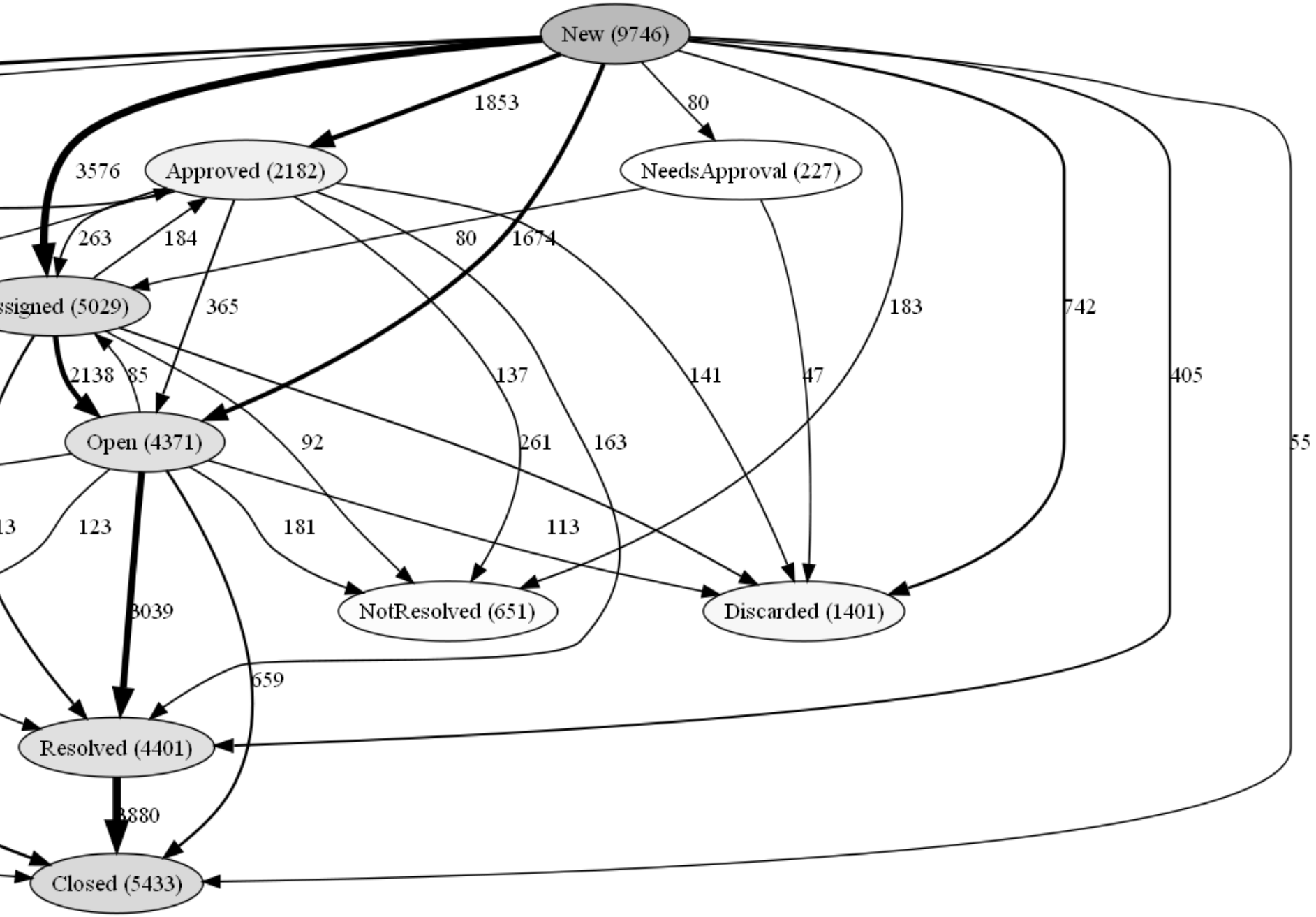


Global behavioral model




Approved (2182)
Assigned (5029)
Open (4371)
NeedsVerification (97)
Duplicated (1830)
Waiting (292)
Resolved (4401)
Closed (5433)
NotRes
3576
263
184
1049
75
365
311
48
50
2138
85
281
69
92
122
613
123
181
779
3039
40
144
659
880


New (9746)
1853
80
3576
Approved (2182)
NeedsApproval (227)
263
184
80
1674
Assigned (5029)
365
8
50
2138
85
137
141
47
69
Open (4371)
92
261
163
sVerification (97)
613
123
181
113
779
Waiting (292)
3039
NotResolved (651)
Discarded (1401)
40
144
659
Resolved (4401)
880
Closed (5433)


New (9746)
1853
80
3576
Approved (2182)
NeedsApproval (227)
263
184
80
1674
ssigned (5029)
365
183
742
2138
85
137
141
47
405
Open (4371)
92
261
163
55
13
123
181
113
3039
NotResolved (651)
Discarded (1401)
659
Resolved (4401)
3880
Closed (5433)

Explaining the results

- New => Approved (?)
  - issue approval is no longer being used
  - approval is implicit in New => Assigned
- New => Open (?)
  - support team skips steps when solution obvious
- Open => Assigned (?)
  - arbitrary loops / wrong assignment / escalation
- Assigned / Open => Duplicated / Discarded (?)
  - decision comes later when more data is available

Conclusion

- case study
  - overall behaviour is close to Incident Management
  - unusual behavior can now be further investigated

- in general
  - process mining facilitates extraction and analysis
  - valuable tool for process conformance
  - relies on available systems and data
  - focus on behavioral perspective

Additional info & resources

- Process Mining Group: *http://www.processmining.org/*
- ProM Framework: *http://prom.sourceforge.net/*
- Process Mining TV: *http://ga1717.tm.tue.nl/user/christian/pmtv/*
- W.M.P. van der Aalst and A.J.M.M. Weijters, *Process Mining: A Research Agenda*, Computers in Industry, vol.53, no.3, pp.231-244, 2004
- A. Rozinat, W. van der Aalst, *Conformance checking of processes based on monitoring real behavior,* Information Systems, vol.33, no.1, 2008
- D. Ferreira, M. Zacarias, M. Malheiros, P. Ferreira, *Approaching Process Mining with Sequence Clustering: Experiments and Findings,* 5th International Conference on Business Process Management (BPM 2007), LNCS 4714, Springer, 2007
- ITIL: *http://www.itil-officialsite.com/*