Microsoft

Microsoft Research Faculty Summit 2012

ADVANCING THE STATE OF THE ART

Software Engineering Productivity Tools (SWEPT)

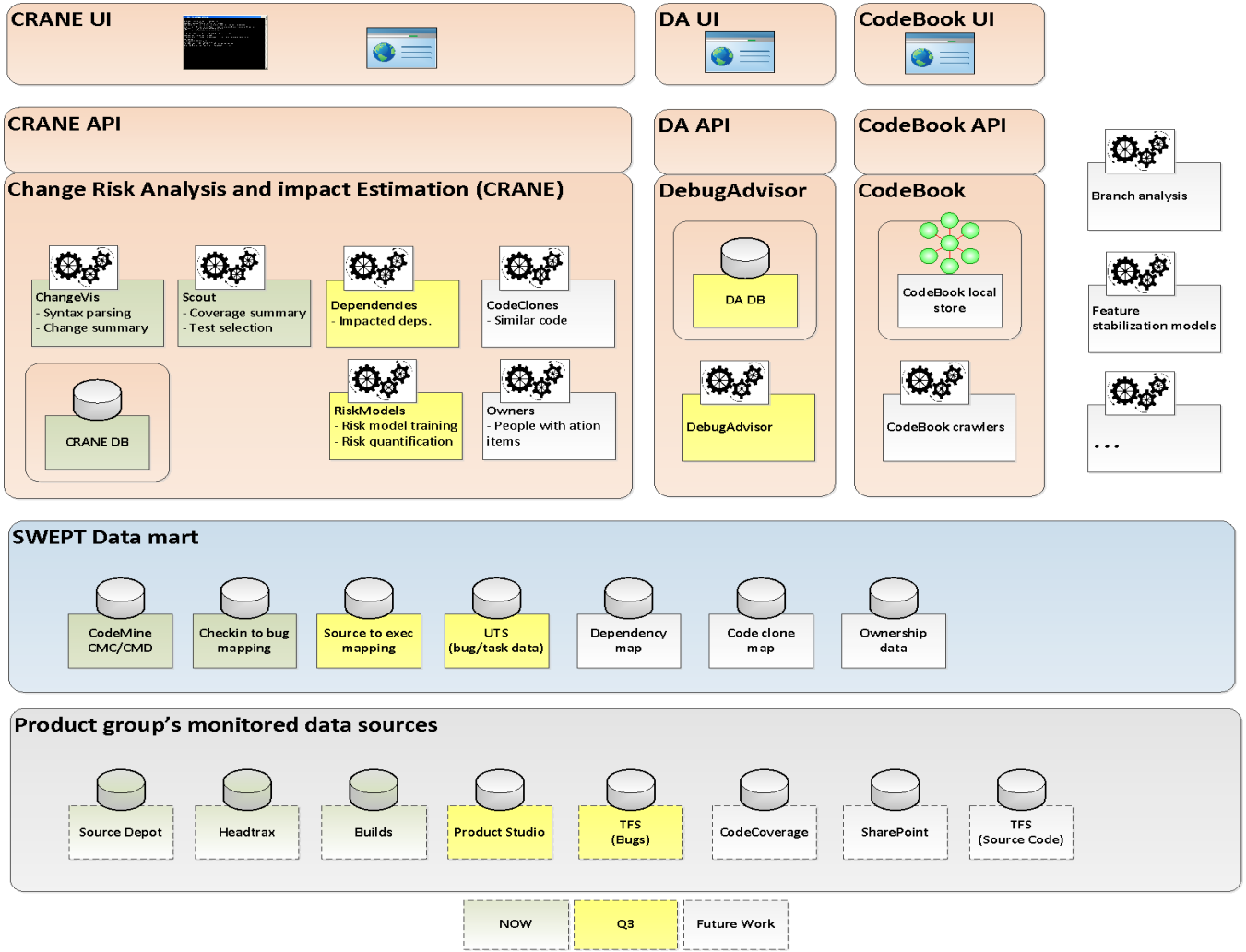
Robin Moeur
Principal Program Manager
Microsoft Research

SWEPT Data mart

- Set of data sources pertaining to product, engineering process and engineering organizations
- Provides consistency of data discovery and access across product groups
- Provides a standard platform for creating analytics
- Informs data driven decision making

Current Contents

- Collecting metadata on:
  - Source code
  - Change lists
  - People
  - Daily builds
  - Tasks/Bugs
- and relationships among them


CRANE UI
DA UI
CodeBook UI
CRANE API
DA API
CodeBook API
Change Risk Analysis and impact Estimation (CRANE)
DebugAdvisor
CodeBook
Branch analysis
ChangeVis
- Syntax parsing
- Change summary
Scout
- Coverage summary
- Test selection
Dependencies
- Impacted deps.
CodeClones
- Similar code
DA DB
CodeBook local store
Feature stabilization models
CRANE DB
RiskModels
- Risk model training
- Risk quantification
Owners
- People with ation items
DebugAdvisor
CodeBook crawlers
...
SWEPT Data mart
CodeMine CMC/CMD
Checkin to bug mapping
Source to exec mapping
UTS (bug/task data)
Dependency map
Code clone map
Ownership data
Product group's monitored data sources
Source Depot
Headtrax
Builds
Product Studio
TFS (Bugs)
CodeCoverage
SharePoint
TFS (Source Code)
NOW
Q3
Future Work

Uses of data – typical questions

- Change data
  - Size, significance, frequency of changes
  - Changed functions, classes, namespaces
  - Executables affected by changes
  - Tests to run on a change; parts of a change without test coverage
- Churn data
- Branch Analysis
  - Code velocity
  - Branching structure
- Ownership
  - Organizational
  - Individual
    - code reviews, ship room,…
- Release Management
  - Dashboard to influence triage decisions
- Risk Analysis
  - Complexity, regression history, coverage %
- QA
  - Test prioritization, redundancy

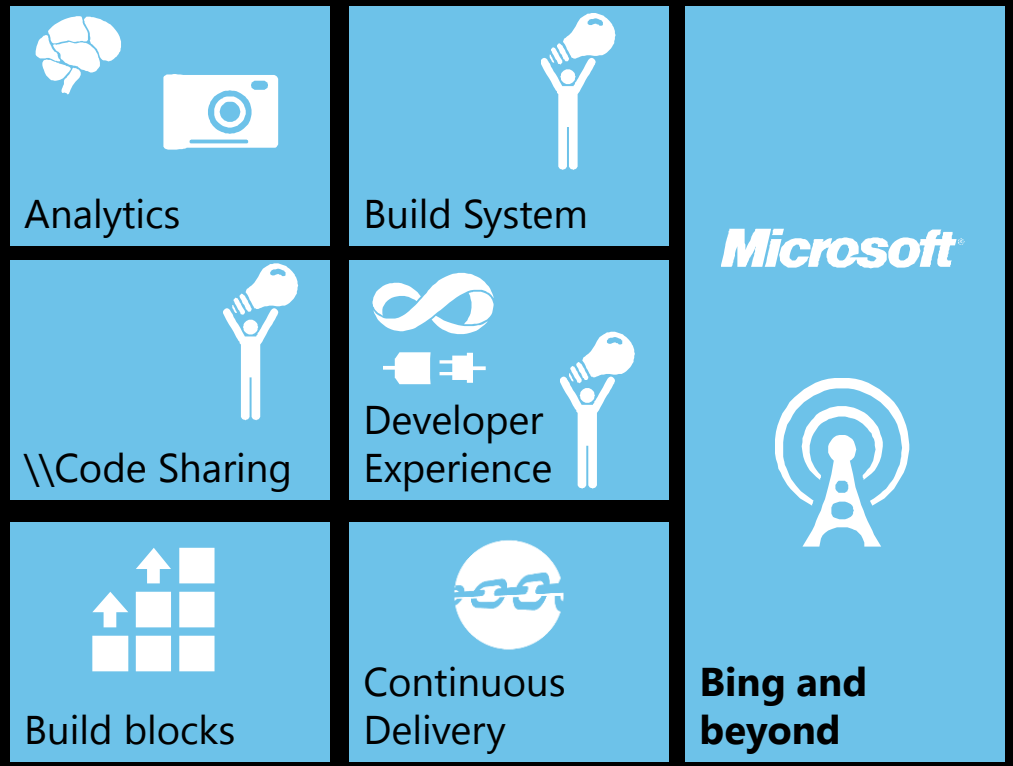
ENGINEERING ANALYTICS AT BING

Ranga Narasimhan

Principal SDE Lead

OSD Engineering Fundamentals (EFun)

EFun – Software Engineering @ scale



Engineering Analytics

- Primary focus on engineer behavior to encourage right behavior and drive better experiences in addition to computing KPIs using past data

  Why?
  - Teams in Bing operate in a decentralized manner and we want to make fundamental changes so that it can up level the whole system
- We treat data as a strategic asset to
  - Validate existing initiatives
  - Support new initiatives
  - Provide direction
- In the past year we have leveraged SWEPT to
  - Come up with new business metrics
  - Deep dive on engineer behavior
  - Come up with new Engineering KPIs

Business metrics

- Checkin Frequency (frequent and small)
  - The measurement of gap between the current process and continuous delivery
- Code Velocity (time from earliest checkin to going live)
  - Not necessary the quickest, the best. Ideally you checkin and it is shipped the next day

Business metrics

- New hire ramp-up time (time from joining Bing to engineering activities)
  - Time to first checkin
  - Time to reach the checkin maturity level of an experienced engineer
- Bing Engineering Culture study
  - Correlating to new hire ramp-up by geo location
    - Ratio of junior to senior engineers
    - % of org undertaking engineering activities

Role of SWEPT

- Enables us to compute KBIs (Key Behavior Indicators) and KPIs (Key Performance Indicators) in an agile manner
- Provides a common schema and this is allowing us to reuse the analytics work for other parts of the OSD Division (e.g. Ads) and the company

© 2012 Microsoft Corporation. All rights reserved. Microsoft, Windows, and other product names are or may be registered trademarks and/or trademarks in the U.S. and/or other countries.
The information herein is for informational purposes only and represents the current view of Microsoft Corporation as of the date of this presentation. Because Microsoft must respond to changing market conditions, it should not be interpreted to be a commitment on the part of Microsoft, and Microsoft cannot guarantee the accuracy of any information provided after the date of this presentation. MICROSOFT MAKES NO WARRANTIES, EXPRESS, IMPLIED OR STATUTORY, AS TO THE INFORMATION IN THIS PRESENTATION.