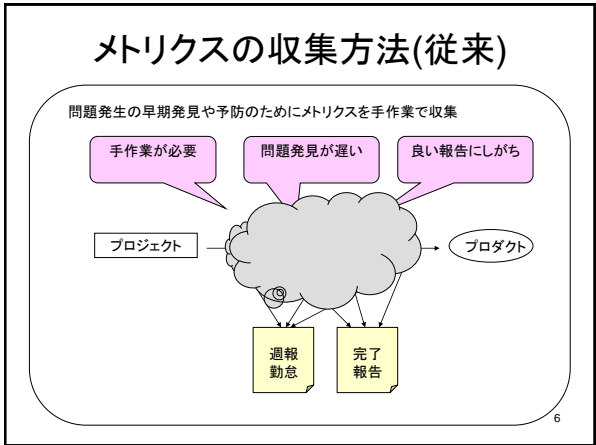
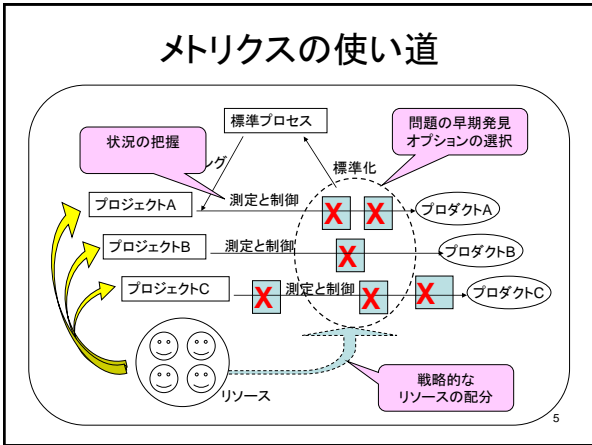
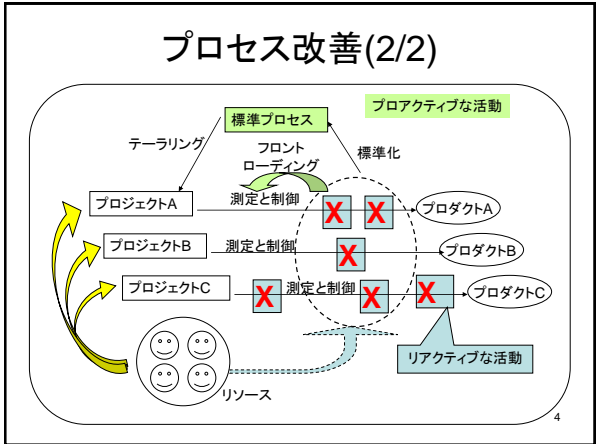
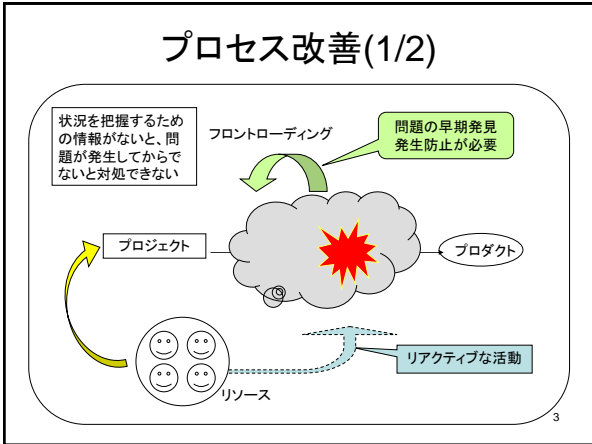


Empirical Project Monitor の紹介

SRA先端技術研究所
(EPM開発コンソーシアム)
阪井 誠

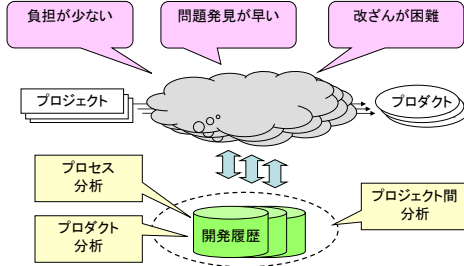
トヨタ流改善「見える化」

- 見える化は一言で言えば、問題点が常に「見える」ようにしておく工夫のことである。正常と異常の違いがすぐに分かる仕事場とか、仕事するうえであれこれ迷わずに済む現場のことを指すと言ってもいいかもしれない。
- IT業界もトヨタ流企業改革から始めてみては？
 - <http://itpro.nikkeibp.co.jp/free/ITPro/OPINION/20040123/138653/>
[TOYOTA](#)



メトリクスの収集方法(EASE)

開発履歴からメトリクスを収集



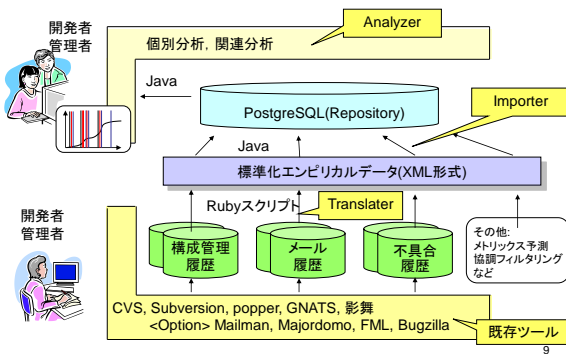
7

EPMの分析機能

- 履歴分析
 - 履歴の可視化(行数、更新/参照数、メール数、障害数/滞留時間)
 - SRGM(信頼度成長モデル)
 - ロジカルカップリング(同時更新ファイルの検出)
 - プロジェクト遅延リスク検出モデル出力(不安定要求/不完全設計/劣悪コード)
- 統計分析
 - バレート図
 - クロス分析
 - 散布図(はずれ値の分析)

8

Empirical Project Monitorのシステム構成



9

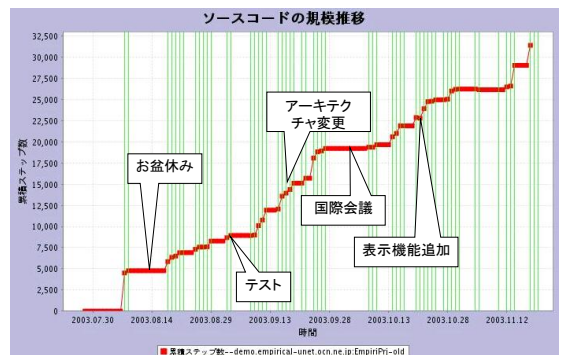
Empirical Project Monitorの特長

- (プロダクト生成以前に)リアルタイムに開発者の振る舞いを知ることができる
- オープンソース開発にも使われているツールを利用
 - データ収集の負荷がない(透過的)
- 人為的操作の少ないデータを分析

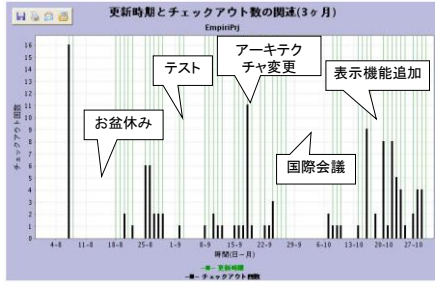
10

機能紹介

ソースコードの規模推移



更新時期とチェックアウト数の関連

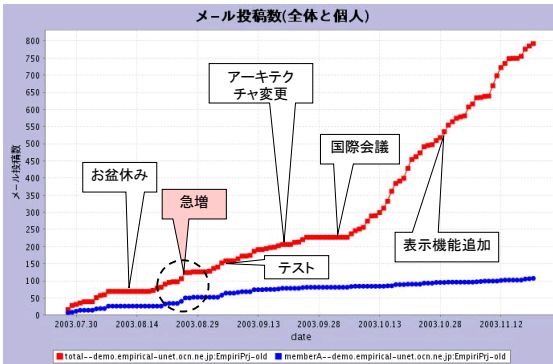


ソースコードの規模

- 開発の進捗に最も関わるメトリクス
- 一般に右上がりになり、生産性が高いほど傾きが急になる
- コーディングの工程で規模が増えることは、開発が順調であると考えられる
- テスト工程で規模が増えているなら、仕様や構造の問題が発生しているかもしれない
- テスト終盤で、ソースコードが更新されているなら、詳細情報を確認したほうが良い

14

メール投稿数と更新時期の関連



メール詳細情報一覧

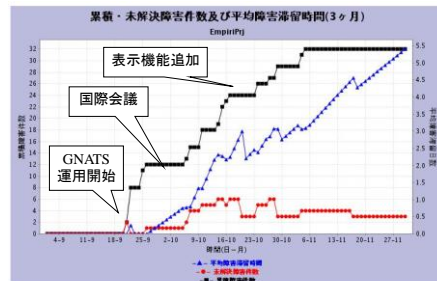
a List of Top Mail Topics
Project ID : demo.empirical-unet.ocn.ne.jp/EmpiriPrj-oid
All the affair number of cases: 262

Number	begin_time	end_time	topic	count
1	2003-07-28 13:38:49	2003-11-20 13:45:09		132
2	2003-10-14 12:10:49	2003-11-01 17:02:24	EASE2003 登録人数	36
3	2003-10-27 17:53:58	2003-11-06 16:51:42	VIP List	20
4	2003-08-21 23:27:00	2003-09-26 14:59:08	Japanese Characters of MailMan	18
5	2003-09-12 11:44:37	2003-09-15 23:33:36	Mail Archives is not in the server	15
6	2003-11-10 11:34:00	2003-11-10 22:11:20	お礼と今後	15
7	2003-11-11 21:27:21	2003-11-12 14:12:43	Pictures	12
8	2003-11-11 15:14:37	2003-11-12 18:56:41	Questionnaire	11
9	2003-10-11 21:09:27	2003-10-15 14:49:33	参加者の所属	11
10	2003-10-16 11:28:09	2003-10-20 17:56:23	定員に関するコメント	10
11	2003-11-04 13:21:30	2003-11-04 21:08:15	参加者確定版	10
12	2003-10-10 10:17:24	2003-10-11 11:30:19	シンポジウムの件	10
13	2003-10-21 11:54:15	2003-10-23 11:18:08	8日の件	9
14	2003-10-15 23:13:24	2003-10-16 12:17:46	Webページ進捗	9
15	2003-07-29 18:18:25	2003-09-01 00:41:16	xml	8

メール情報

- プログラムを更新時、障害発生時には、情報交換のメールの投稿数が増える(ことが多い)
- メールデータはソース更新中や障害対応検討中でもリアルタイムに変化する
- ソースや障害に直接関係のない問題であっても補えられる可能性がある
- サブジェクトや関係者を特定できる

更新・未解決障害件数及び平均障害滞留時間



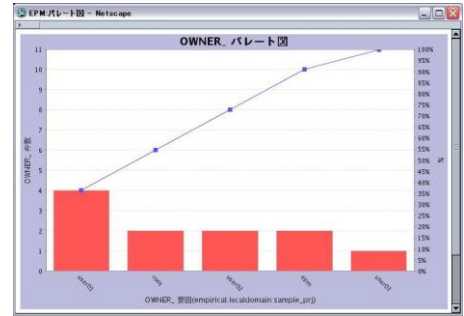
17

障害情報

- 累積障害件数
 - 傾きが急な場合は短期間に多くの障害が発見されている
 - テスト工程の終盤で傾きが水平にならない場合は、信頼性に問題があるかもしれない
- 未解決障害件数
 - 解決できていない不具合を示す
 - 未解決の障害が多数残っているなら、試験項目消化後も多くの作業が必要になるかもしれない
- 平均障害滞留時間
 - 障害を解決するために要した一件あたりの時間を示す
 - 解決できない問題があると大きくなる

19

パレート図



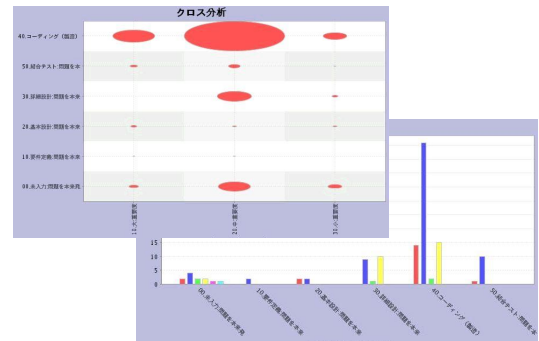
20

パレート図

- QC7つ道具のひとつ。
 - パレートの法則(一般に不良全体の80%は、20%の原因に由来する)に基づく
- 左の目盛りと棒グラフは件数、右の目盛りと折れ線グラフは累積の比率を示す
 - 件数の多いものから並んでいる
- 重要な要因が明確になり、対策をとると効果的な要因がわかる
- CVSやメールから、作業が誰に集中しているかがわかる
- 障害データからは、障害の傾向がわかる

21

クロス分析



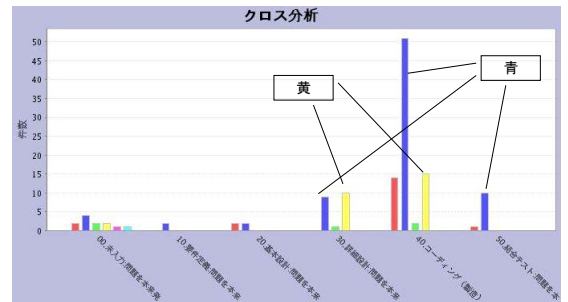
22

クロス分析

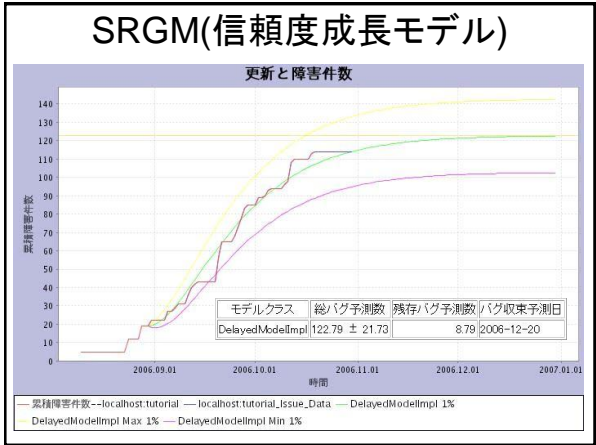
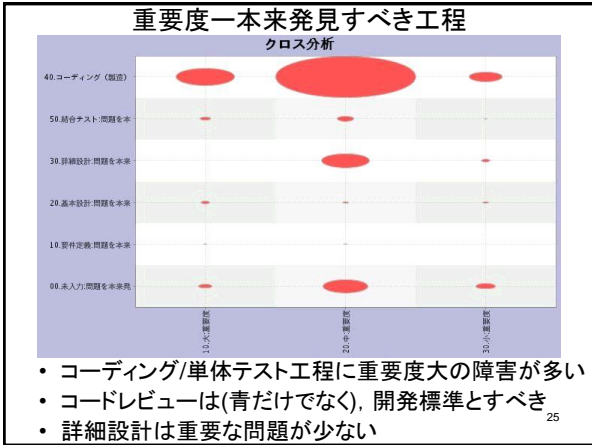
- 2次元の要因ごとに、棒グラフ、折れ線グラフ、バブルグラフで示す
- 要因間の関係がわかるので、対策の検討が容易になる

23

サブシステムー本来発見すべき工程



- 青はコーディング工程で発見すべき障害が多い
- 黄はコーディングと詳細設計で発見すべき障害が多い
- コードレビューまたは詳細設計でレビュー強化すべき



SRGM(信頼度成長モデル)

- “Software Reliability Growth Model”の略
- ソフトウェア信頼性を評価するためのモデルの一つ
- 総バグ予測数, 残存バグ予測数, バグ収束予測日を算出することができる
- 試験項目に対して, 障害(バグ)がほぼ均質に分布することが前提になっている

ロジカルカップリング検索

パターン検出条件

同時更新間隔時間 15 分

同時更新回数 2 回以上

検索ファイル条件 SVN_manager.rb

検索母集団条件 指定なしで全件

※条件は、Java用正規表現で入力

ロジカルカップリング検索

同時更新回数	支持度	ファイルリスト
3	0.023809523809523808	EmpiriPri/src/ruby/translator/Product_SVN_Line.rb
2	0.015873015873015872	EmpiriPri/src/ruby/translator/SVN_export.rb
2	0.015873015873015872	EmpiriPri/src/ruby/translator/SVN_export.rb EmpiriPri/src/ruby/translator/Product_SVN_Line.rb

ロジカルカップリング検索

- 同時刻にソースの更新を行った関連が深いと思われるファイルを探す
- 改造するファイルを「検索ファイル条件」にすれば, 予想外の問題が少なくなる

