

SMART-GS

歴史的文献研究のためのソフトウェアツール

久木田水生, 林晋, 大浦真
京都大学文学研究科

白眉センター&応用哲学・倫理学教育研究
センター共催セミナー

「デジタル・ヒューマニティーズの現在」

2012年12月10日

於 京都大学

- このスライドは2012年12月10日に行われた、京都大学白眉センター＆応用哲学・倫理学教育研究センター共催セミナー「デジタル・ヒューマニティーズの現在」で発表されたものに修正を加えたものです。



Images Notes

- root
- LectureNotesS
- 0: S09_001
- 1: S09_002
- 2: S09_003 [**]
- 3: S09_004 [**]
- 4: S09_005 [**]

SMART-GS

電子化されていない手書きの文献を
研究する歴史家のために
開発されているソフトウェアツール

Browse Edit Source Paste Mark up Text Search Highlight All Cancel

京都大学文学研究科の
林晋教授を中心に開発

\$5\$ Alles ト partielle Id. ヲモツ.

\$6\$ Scheler ノ欠点ハ Geist ヲ Leben ノ Hemmen u.

Compact Editor

Editor

SMART-GSはSourceforge.jpで開発、公開されている オンラインマニュアルもここで公開されている

SMART-GS (スマートジ - x

sourceforge.jp/projects/smart-gs/

Gmail: Google メ... SourceForge.JP: ... 京都現代哲学コロ... 久木田水生のページ Facebook 読書メーター - あ... Twitter

SOURCEFORGE.JP ダウンロード Magazine 開発 アカウント ヘルプ

ファイル改ざんや実行時アラートに対応する「コードサイニング証明書」を理解する
コードサイニング証明書（コード署名証明書）は10年以上前から利用されているが、マルウェア/アドウェアを仕込まれたコピー配布サイトの増加やWindowsにおける検証利用範囲の拡大などにより、近年では必要とされる場面が広がってきている。改めてコードサイニング証明書の現状を整理し、正確に把握していこう。
SOURCEFORGE.JP Magazine

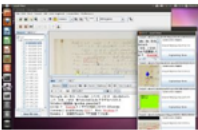
SourceForge.JP > ソフトウェアを探す > SMART-GS > 概要

SMART-GS

概要 ▼ ダウンロード ソースコード ▼ チケット ▼ 文書 ▼ コミュニケーション ▼ ニュース

MSDNを40%OFFで購入する方法 - 今ならもれなくAmazonギフト3000円！

プロジェクトの説明 Webページ 開発情報

 SMART-GS は歴史学、文献学などの人文学におけるテキスト研究用のツールです。SMART-GS は手書きを含むテキストの画像イメージを検索、マークアップ、リンクすることができます。SMART-GSは、形の類似性により検索を行うので、OCR とは異なり、手書き文書、古代語などの未知の言語のテキストでも検索を行うことができます。そのため歴史学、言語学、文献学などにおける文献研究に使用することができます。

お知らせ(2012年3月6日)

バージョン0.8.0に添付されているpreference-default.xmlには間違いがあります。このファイルの中には <entry key="dscFolderPath">./smart_gs/dsc/</entry> という行がありますが、これは新しいSMART-GSでは不要です。これはdscファイル(画像検索に利用するファイル)を置くフォルダを指定するものですが、現在のバージョンではdscファイルは対応する画像と同じフォルダに入れるようになっています。0.8.1をダウンロードしていただくか、生成されたpreference.xmlファイルからこの行を削除してください。詳しくはオンラインマニュアル「[行・行情報ファイル・DSCファイル](#)」の項目をご覧ください。0.8.0から0.8.1への変更点に関しては0.8.1のリリースノート、またWikiのページをご覧ください。

more

www.facebook.com を待機しています...

カテゴリ: ソフトウェア

Smart-GS

ツイート 30

「SPREADの再創造」
日本の業務アプリ開発における最強のデータグリッドを求めて SPREAD for WPF

期限	ロットNo.	在庫数	単位	単価	区分
27	2011092001	200	本	90	仕入
27	2011092003	400	本	90	仕入
27	00000001	490	本	90	製造
商品計		1090			
27	20111119001	10	本	120	仕入
27	123	4	本	100	製造
商品計		14			
27	20111119010	9	本		

ニュース

- 資料集を追加 (2012-12-11 01:27)
- 京都学派アーカイブ公式公開 (2012-12-11 01:12)

<http://sourceforge.jp/projects/smart-gs/>

プロの歴史家の研究を
補助するためのツール

専門家が現場で必要なことを
反映した設計

日常的なヘビーユースを想定

誰でも手軽に使える

<

使いこなすことで
大きな威力を発揮する

ただし可能な限りシンプルで
直観的なUIを目指す

開発の経緯

- SMART-SE の SML というマークアップ言語と処理系(河村、小林和晶作成)の一部を、SMART-GSを継承。
- 現在 Reasoning Web と呼ばれているものは、SMART-SEの要求仕様管理ツール(潘沂冰、薛世宗作成)を一部継承。
- SMART-GS の設計は、2006年ころ林を中心に京大情報・史料学専修により行われ、実装は小林和晶の2006年度修士論文研究の一環としておこなわれる。
- 北大(当時)の寺沢、猪村のチームによりSMART-GS 用のJava画像検索機能用モジュールが構築され組み込まれた。
- 小林和博を中心とする情報・史料学の学生のチーム(清水、田村、橋本)により、多くの改善・改良が行われた。

開発の経緯

- 2010年度より相原健郎(NII)が参加して、SMART-GSをネット上協働ツールに進化させる人文学研究用の**情報基盤HCP**の開発と、そのためのSMART-GSの大幅な改造が科学研究費の補助による研究として開始された。
- 2010年夏より大浦真と久木田水生が、非常勤特別研究員として、このプロジェクトに参加。林自身がプログラマとしての活動を開始。
- 従来のSMART-GSに大幅な改善を加えて2010年、SMART-GS0.8、2011年、SMART-GS0.9を発表。
- この SMART-GS0.9を、相原が開発したHCP基盤のクライアント化することにより、**ネット上協働人文学研究用ツールを構築**する計画が進行中。

Digital HumanitiesとしてのSMART-GS

分析

デジタル化

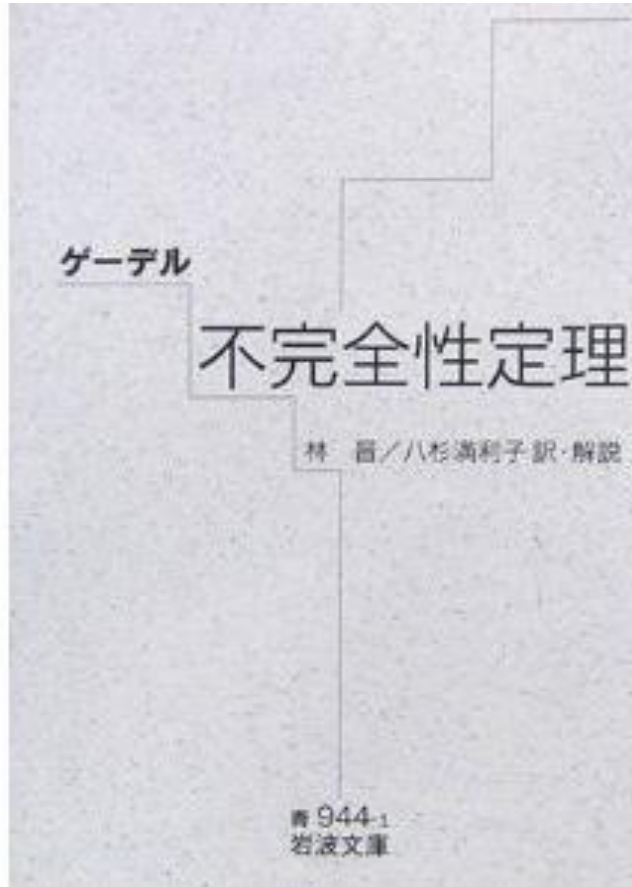
ツール開発

システム開発

教育

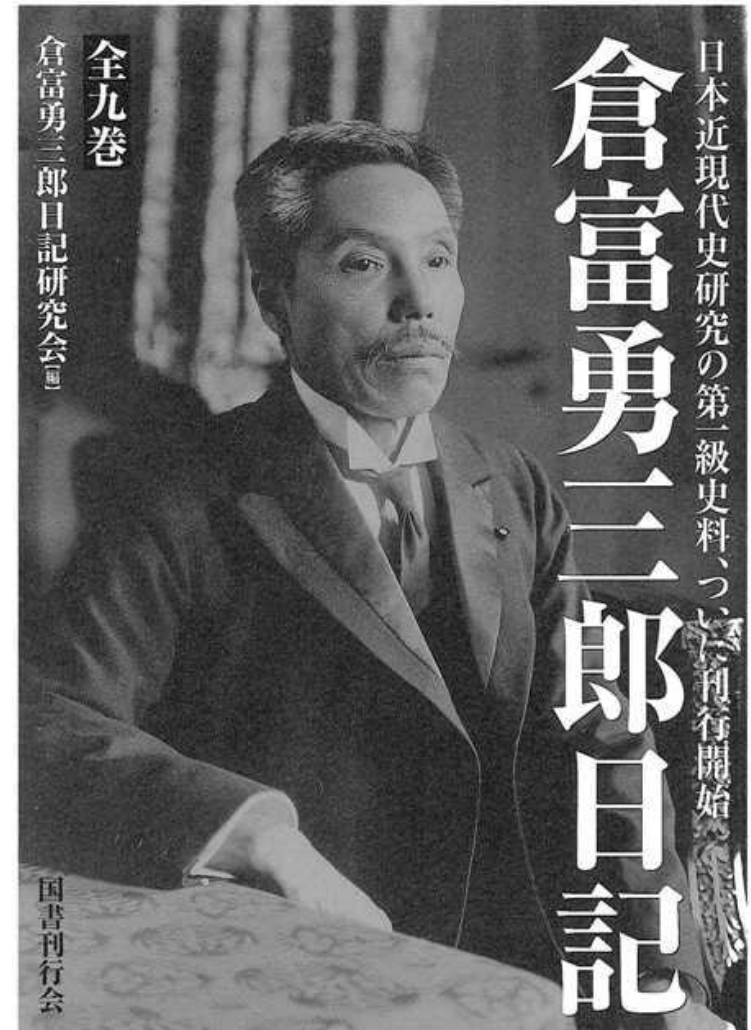
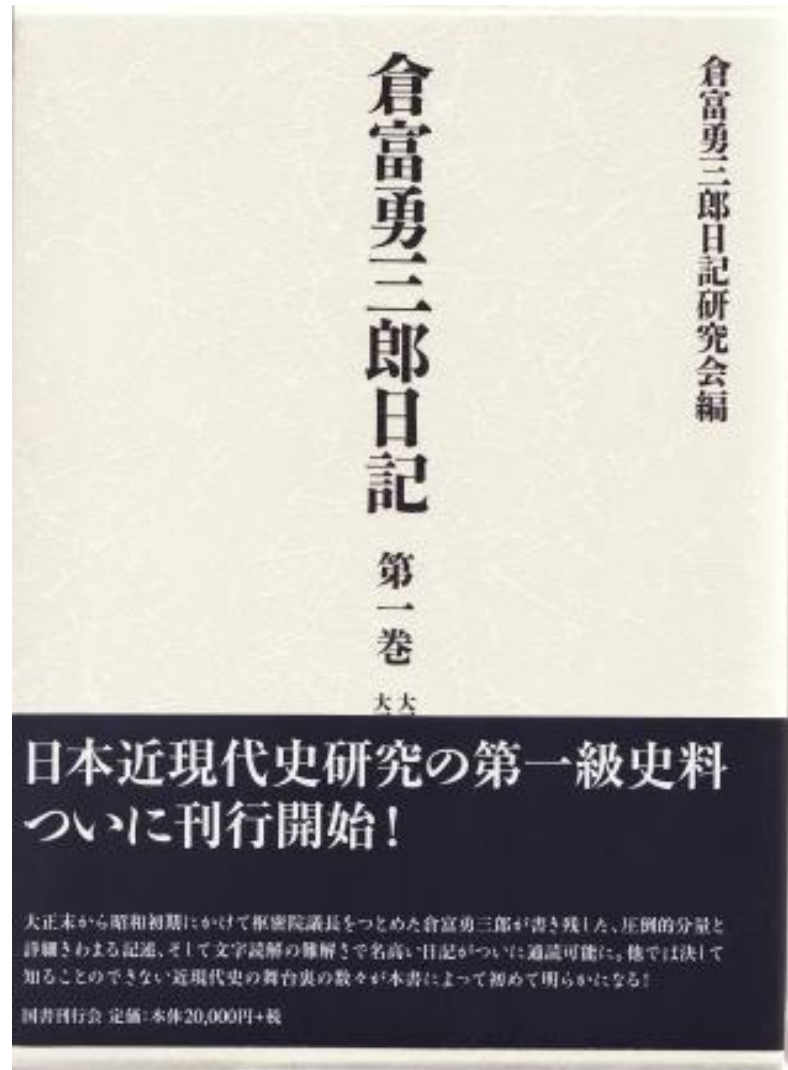
Cf. 永崎研宣、「人文学はデジタル技術に何を期待しうるか」

使用の実例：林教授によるDavid Hilbertの手稿研究



ヒルベルトの画像は<http://gallery.socionix.com/Gamma/Te-INTp/hilbert/>より
本の画像はAmazon.co.jpのページより

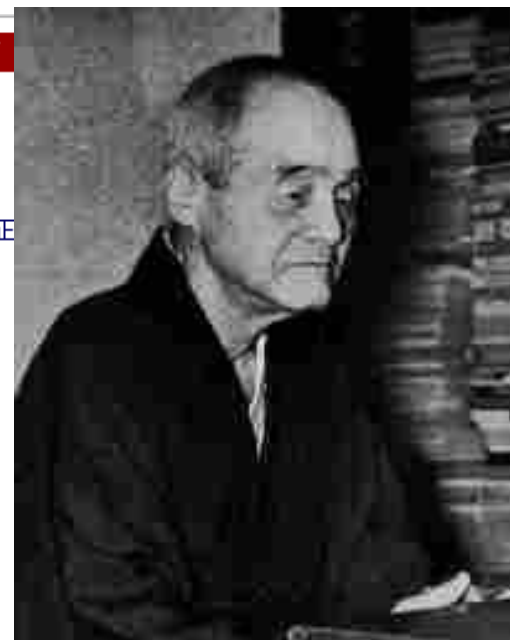
使用の実例：永井和教授らの研究グループによる 『倉富勇三郎日記』の翻刻



<http://nagaikazu.la.coocan.jp/kuratomi/kuratomi.html>

使用の実例：林教授による田辺元の手稿研究

思想 多様な知の交流の場として 2012年第1号 No.1053 目次 ■ バックナンバー ■ 購読のご案内 ■ 思想トップページ	
田辺元の思想——没後50年を迎えて	
思想の言葉	
— I —	
《対談》田辺元の思想	
死者と象徴 ——晩年の田辺哲学から——	
沈黙と偶然 ——田辺元『マラルメ覚書』をめぐって——	
近迫と渦流 ——田辺元・ハイデッガー対決が今私たちに突きつけているもの——	
田辺元のパロック哲学 ——絶対媒介の力学性と象徴性——	
「転換」の論理 ——田辺的思考の生成と〈倫理としての論理〉——	
田辺元の思索 ——「絶対無」の概念を中心に——	
田辺元のコミュニズム	
田辺元の「数理哲学」	林 晋 (197)
《資料》田辺元のカール・ヤスパース宛電報	
— II —	
ロゴスとアンチ・ロゴス	
「種の論理」における「種」とは何か	
「種の論理」の生成と構造 ——媒介としての生——	
われ、ヒトにあらず ——田辺元における個と人類——	
《資料紹介》情報の宝庫、二つの田辺元文庫	



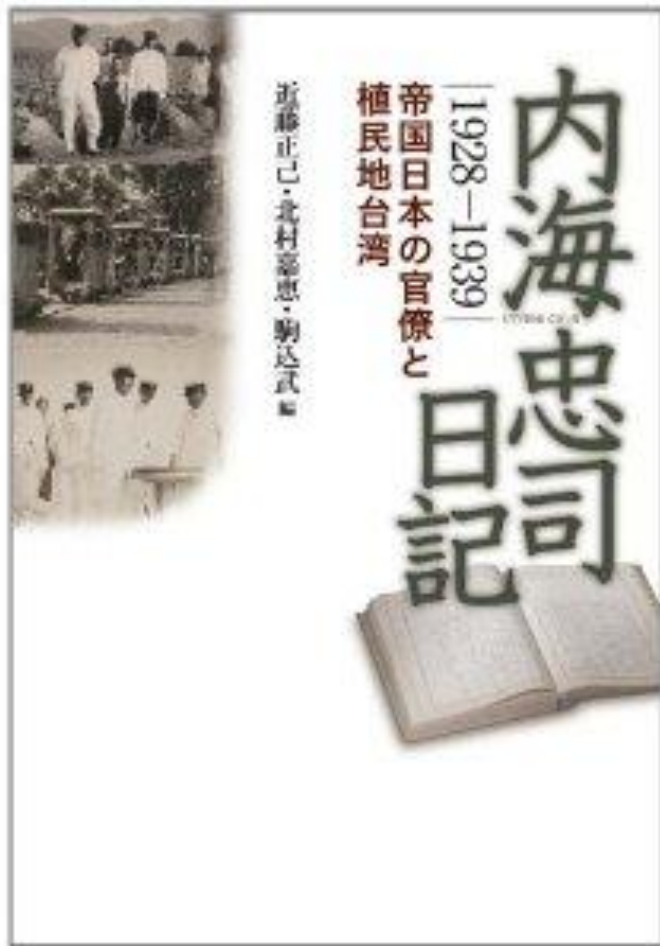
合田正

<http://www.iwanami.co.jp/shiso/>

田辺の画像はhttp://www.bun.kyoto-u.ac.jp/japanese_philosophy/jp-tanabe_guidance/より転載

使用の実例：

近藤正己教授による「内海忠司日記研究 SMART-GS ツールを
応用したテキストデータ化と分析」(基盤研究C、課題番号 24520784)



参考画像 (Amazon.jpのページより)

SMART-GSの基本的な機能

画像とテキストを統合

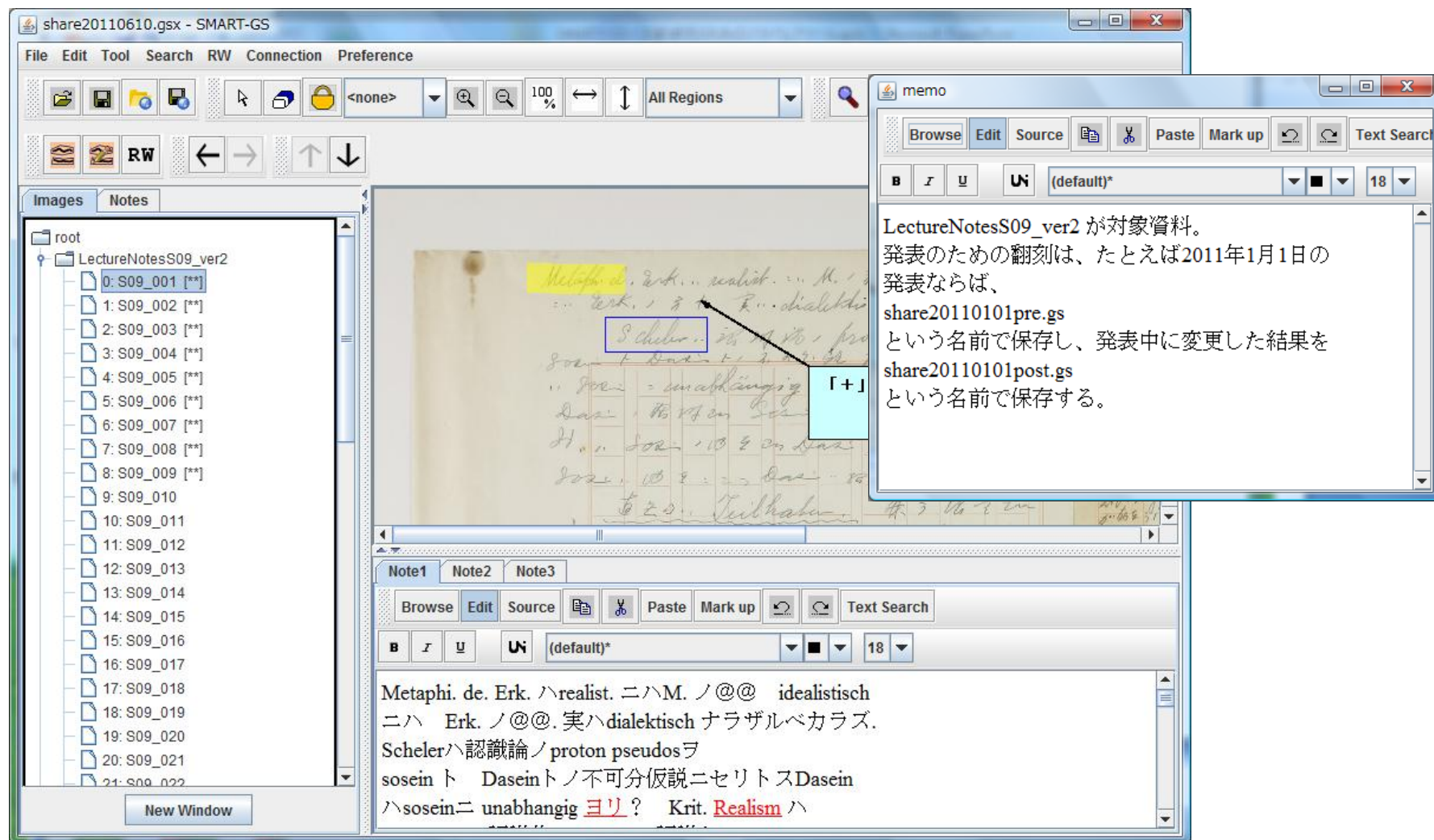
画像上のマークアップ、
アノテーション

マークアップ同士のリンク

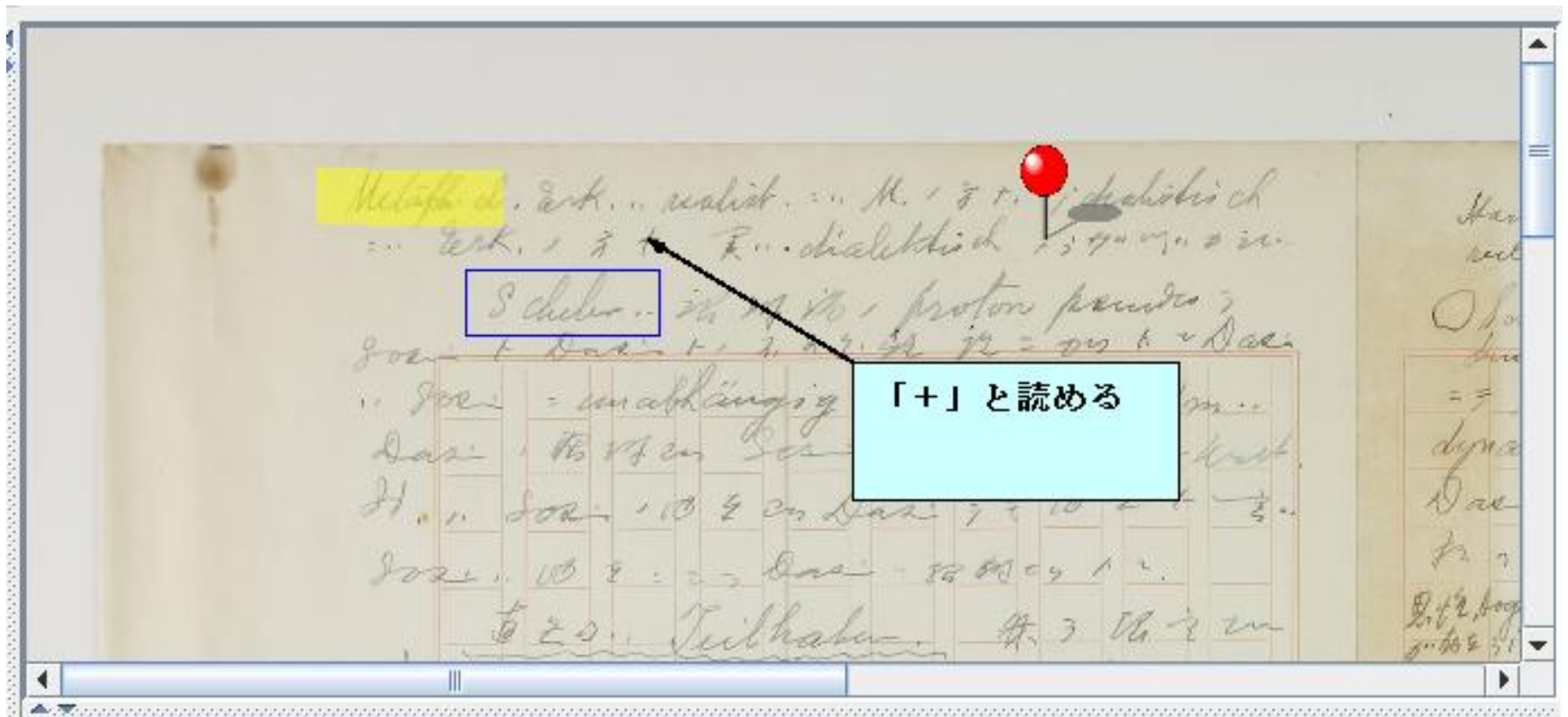
手書き文字の検索

詳しくはオンラインマニュアルを参照。

画像とテキストを統合



画像上のマークアップ



マークアップ同士のリンク

Sg Editor: S09_001

Image Note1 Note2 Note3

部分の読みの「候補」冒頭の「方ナ。」に見える「ところ」は、それ

「+」と読む？

Local View

Source: smart://segment/shayashi/20111007001611/^usersNote

Connection Name
shayashi@20111007001646

Explanatory Note

Target: smart://segment/mkobayashi/20111027103616/^second

ページ冒頭（0行目と1行目）
なかった部分

Connection Name
mkobayashi@20111027103642

Explanatory Note

Target: smart://rect/aaa/20110708025925/^spread/S09_005/aaa

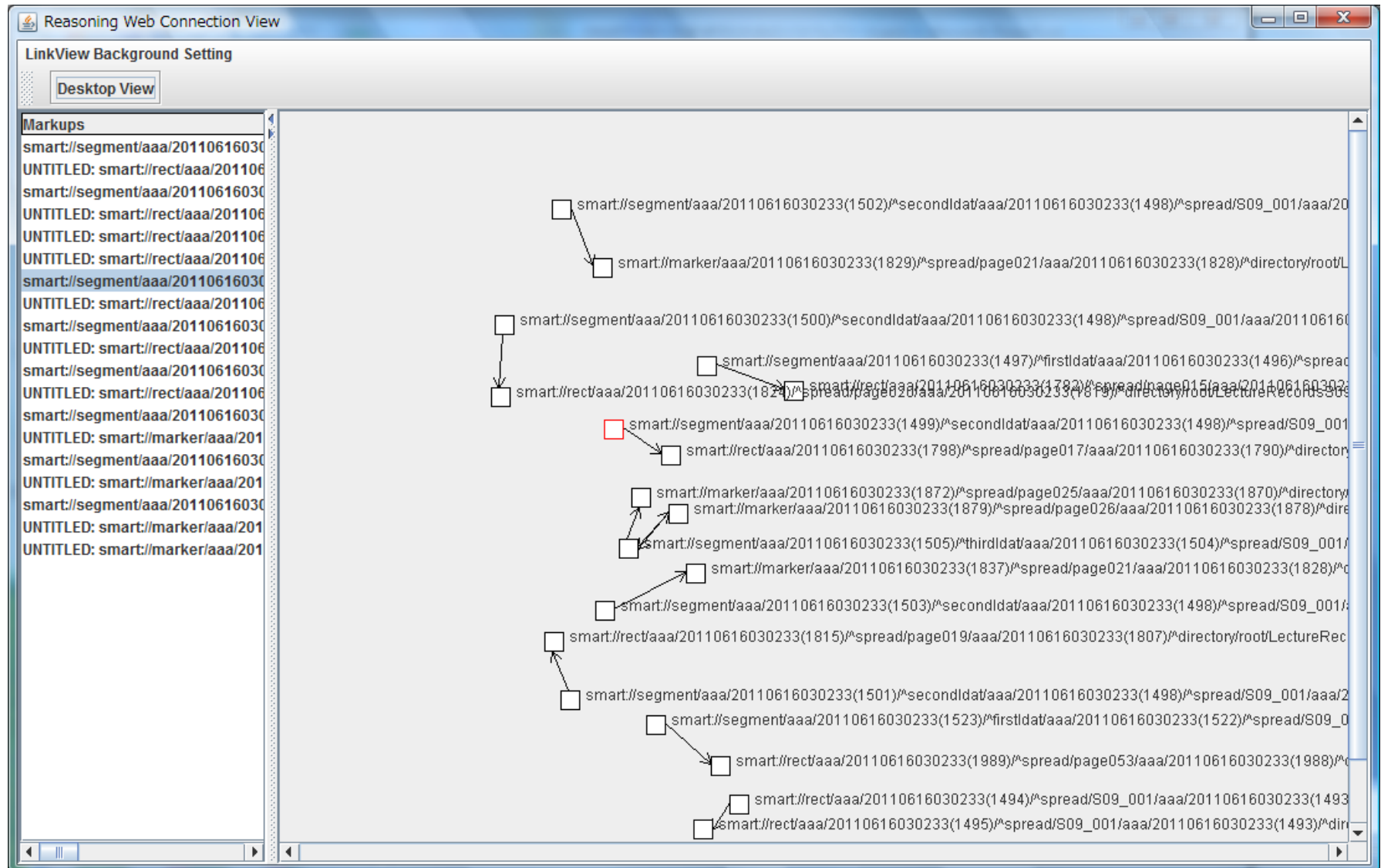
Kind: Rectangle
User: aaa
Time: 20110708115925 (Asia/Tokyo)
Location Kind: Image
Spread Name: S09_005
Path: /root/LectureNotes/S09_ver2/S09

Explanatory Note

OK Cancel

The screenshot displays the Sg Editor interface with a handwritten document. Annotations include a blue circle, a blue rectangle, a green rectangle, and a yellow rectangle. A Local View dialog box is open, showing the source and target of the annotations. The source is a smart://segment/shayashi/20111007001611/^usersNote and the target is a smart://segment/mkobayashi/20111027103616/^second. The dialog box also shows the connection name, explanatory note, and target details.

マークアップ同士のリンク

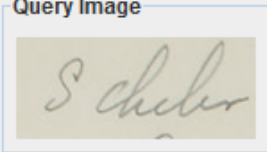


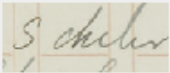
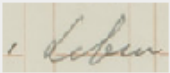
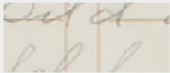
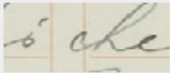
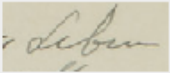
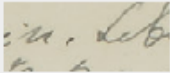
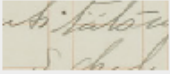
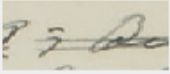
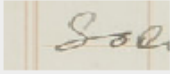
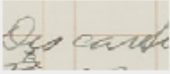
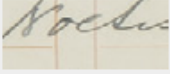
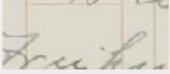
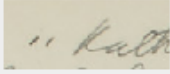
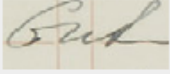
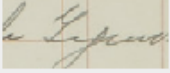
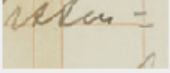
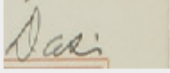
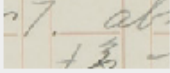
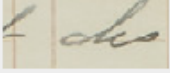
画像に基づく サーチ

Search Result

View Option

Query Image

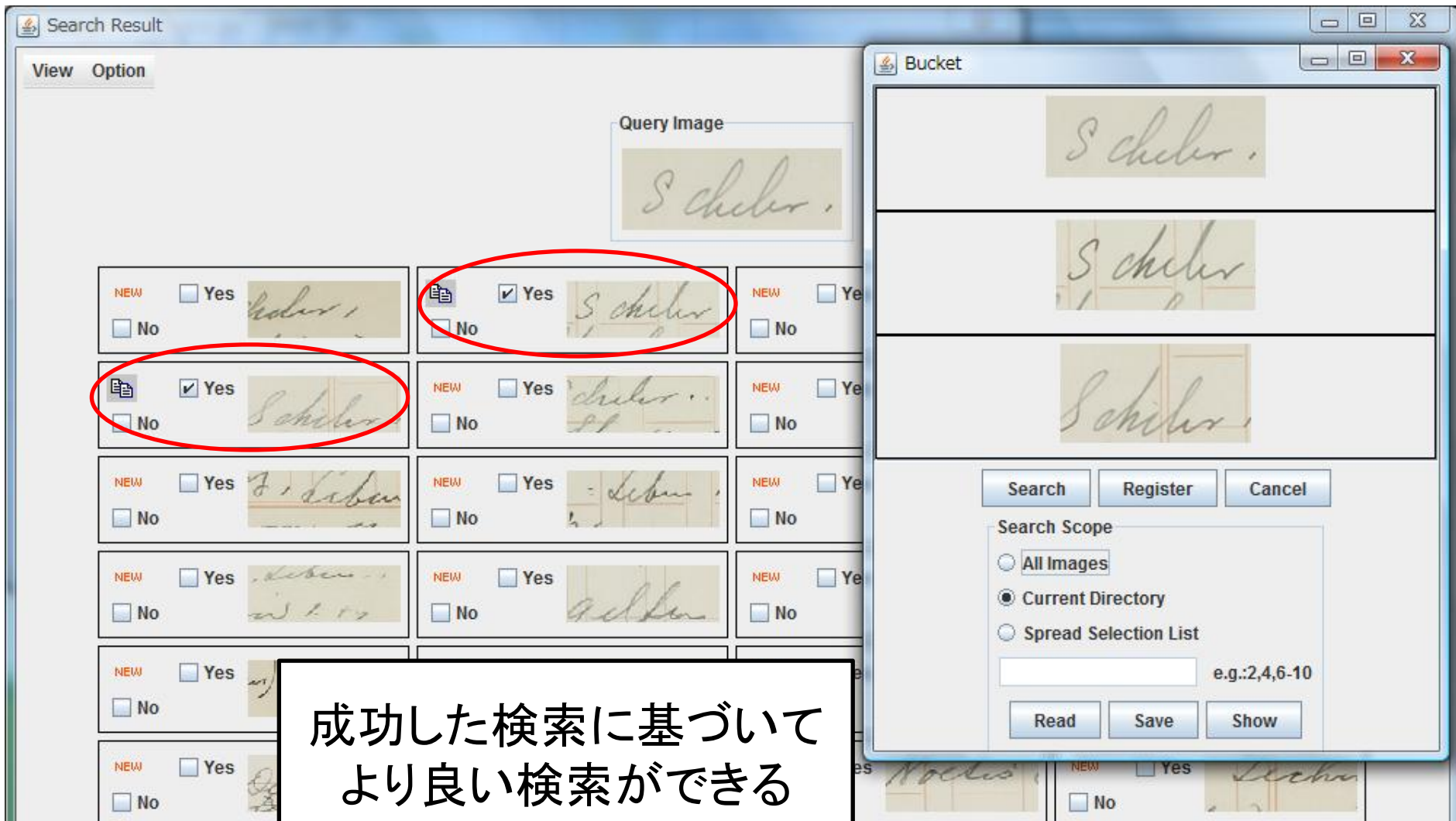


NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No	
NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No	
NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No	
NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No	
NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No	
NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No		NEW ☐ Yes ☐ No	

back next

※サーチができるためには画像に行の情報が付加されている必要がある。
詳しくはオンラインマニュアルを参照。

画像に基づく サーチ



SMART-GSの基本的な機能

画像とテキストを統合

画像上のマークアップ、
アノテーション

マークアップ同士のリンク

手書き文字の検索

これらは個人として
史料研究を行う際に
有用な機能

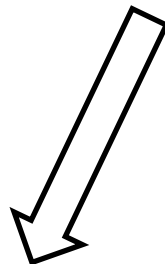
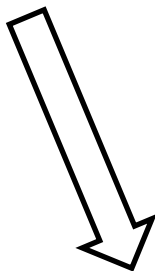
編集の記録が残される

オリジナルの画像は
変化させていない

OSに依存しない

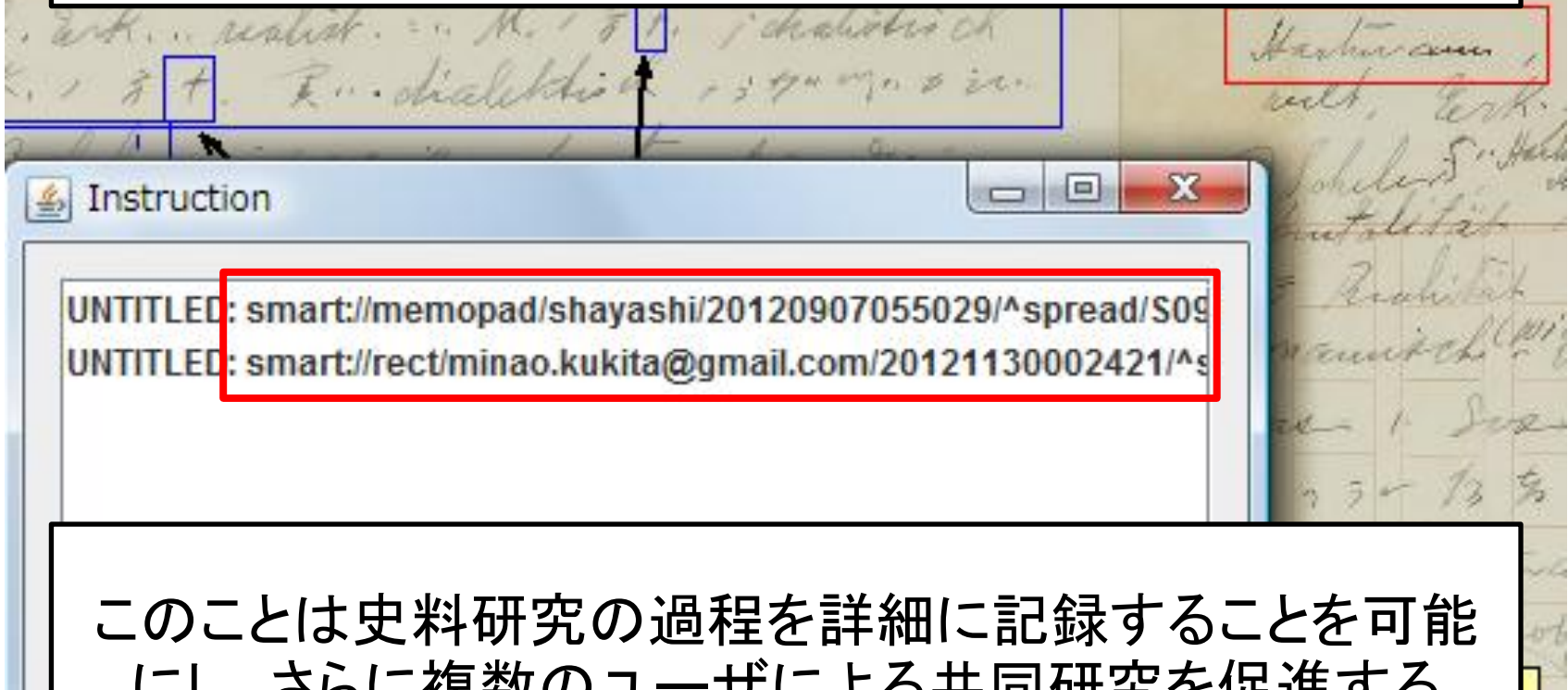
ファイルのやり取り
が容易

共同作業の促進



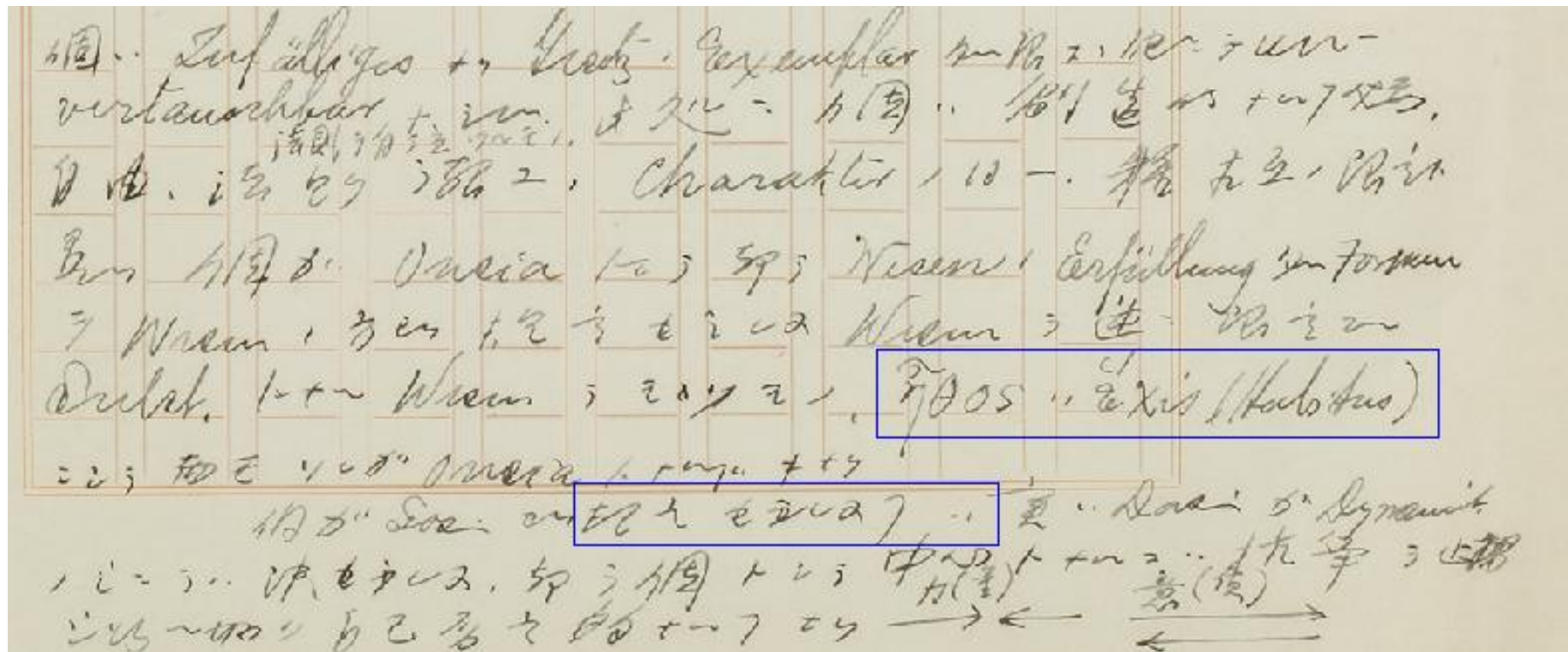
編集の記録の保持

SMART-GSではすべてのリソース(マークアップ、テキスト、画像など)に一意的なURIが割り当てられており、そこには作成した日時、作成したユーザなどの情報が含まれる



このことは史料研究の過程を詳細に記録することを可能にし、さらに複数のユーザによる共同研究を促進する

共同翻刻の力：田辺元史料研究会での例



日本語のくずし字、
ドイツ語、英語、
古代ギリシャ語の混在

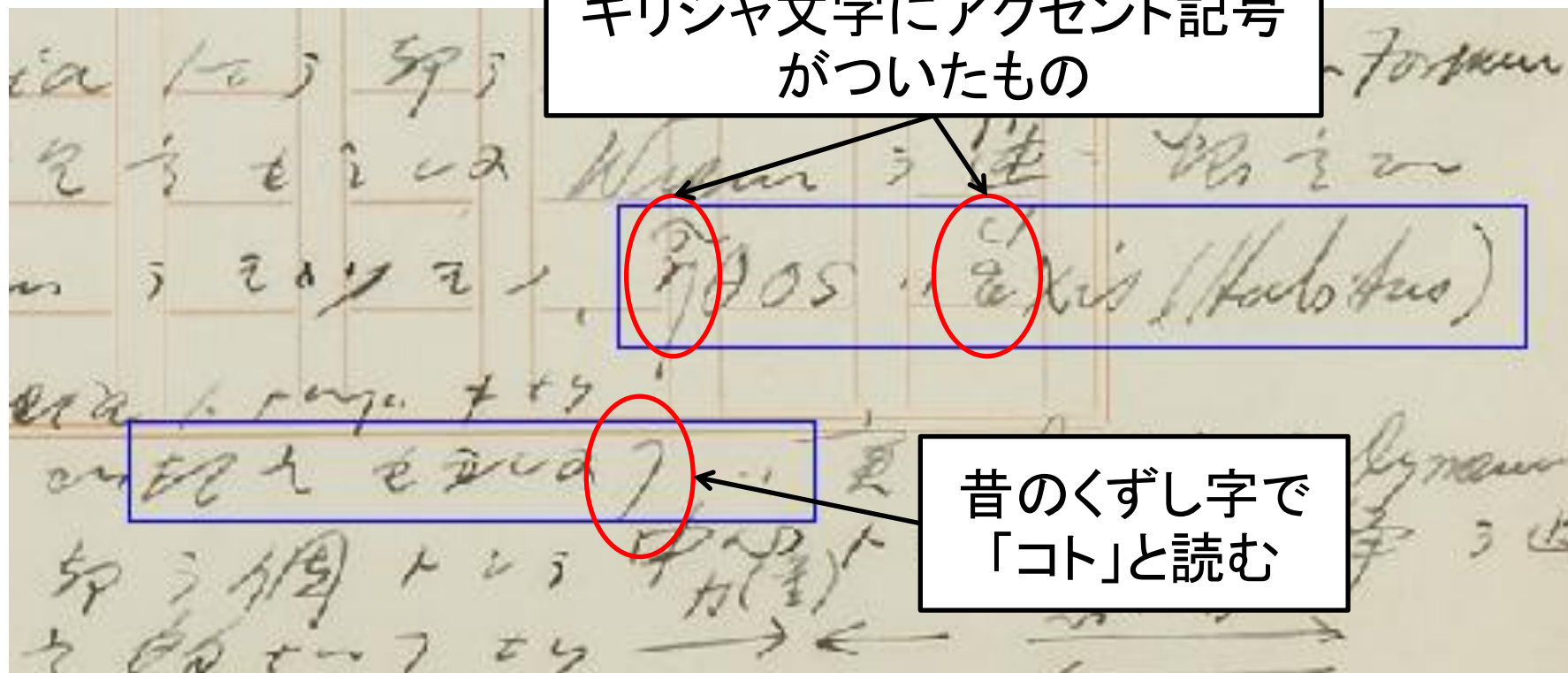
新カント派、現象学、
数学、物理学
等の専門用語

極めつけの
悪筆

一人の人間で読むのは非常に困難

共同翻刻の力：田辺元史料研究会での例

ギリシャ文字にアクセント記号
がついたもの



昔のくずし字で
「コト」と読む

これらは授業に出席していた
学生によって指摘されて判明したこと

専門家でもさじを投げた田辺の手稿の解読が
非専門家による共同作業によって着実に進んでいる

A free, open source tool enabling
community trans
document and m

Scripto brings the power of
collections. Designed to allo
to transcribe a range of diff

共同翻刻のためのプロジェクトと
ツール開発が現在
世界中で進められている

The Center for Digital Theology at Saint Louis University

Creating Digital Tools and Resources for Humanities Scholarship

Saturday, November 10, 2012

Collaborative Manuscript Transcription: What does it mean to "support TEI" for manuscript transcription?

Blog Archive

- ▼ 2012 (10)
- ▼ November
- Collaborative Manuscript Transcription: What does it mean to "support TEI" for manuscript transcription?
- Seeking a Manuscript
- September
- June (2)
- May (2)
- February
- January (3)
- 2011 (17)



UCL Home » Transcribe Bentham » Transcription Desk Log In / Create Account

Navigation

- Transcription Desk
- Transcription Guidelines
- Select a Manuscript
- Blog
- Discussion Forum
- Recent changes
- Random page

★ Help improve Transcribe Bentham by contributing material to transcribe: Bentham's [South Wales](#); Transcribe Bentham's [3D version of Bentham's Autobiography](#)

Wikipediaのように
世界中のユーザが
自由に参加・協力

<http://scripto.org/>
<http://digital-editor.blogspot.jp/>
http://www.transcribe-bentham.da.ulcc.ac.uk/td/Transcribe_Bentham

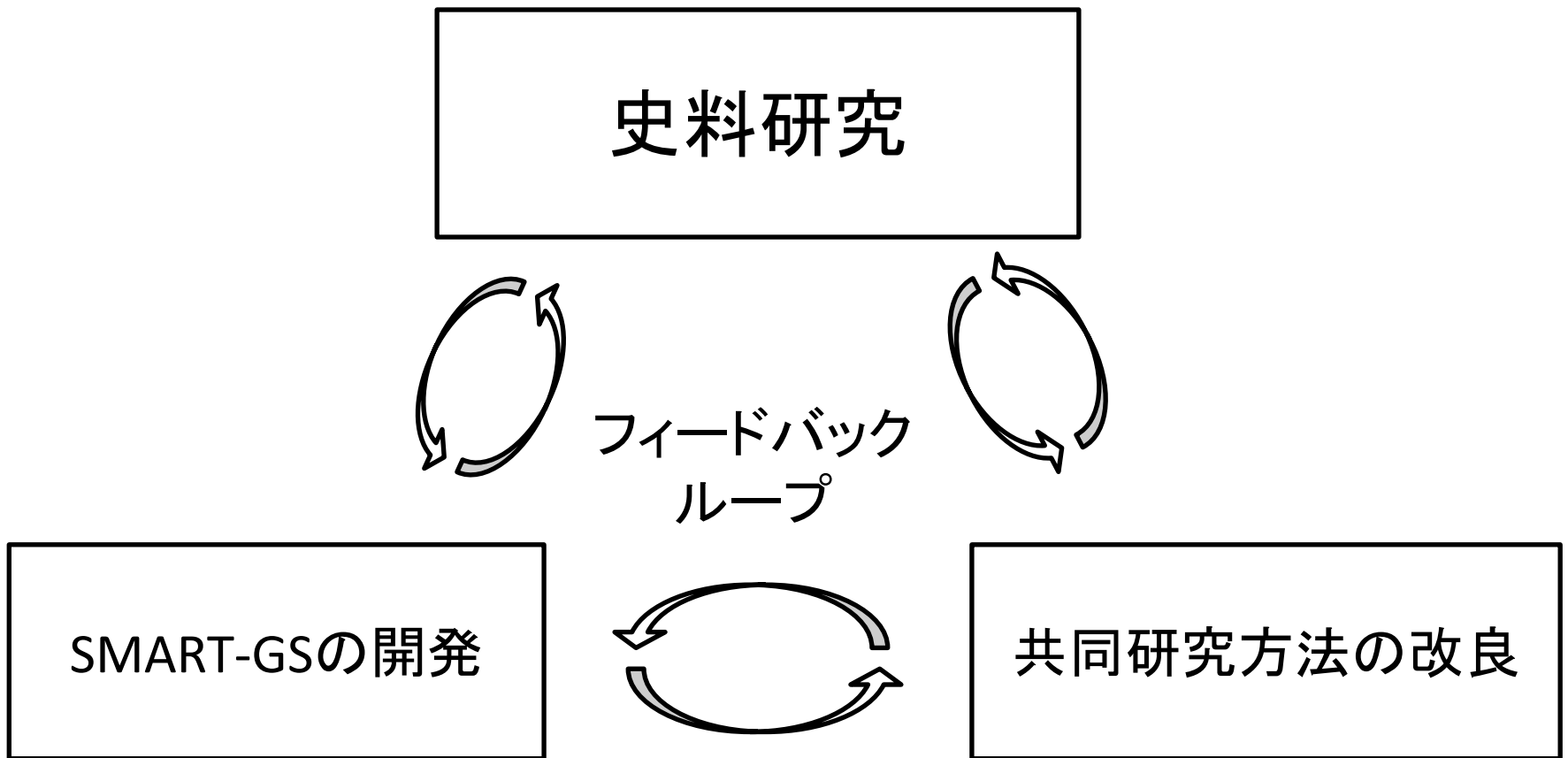
Transcribe Bentham

SMART-GSの独自性

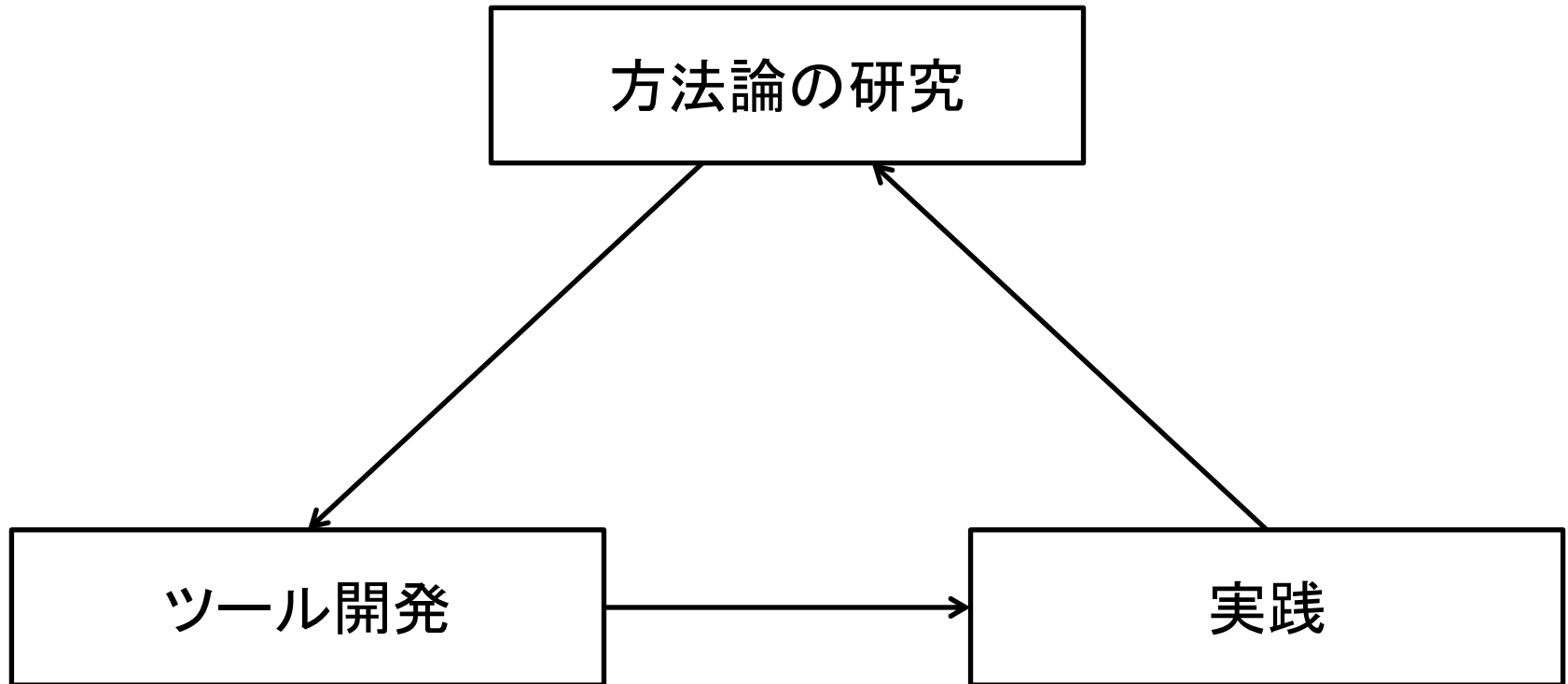
実際にツールを使用する歴史学者が
主要な開発者であるということ

プロの歴史家にとって何が必要なのか
現場で何が本当に有用なのか
ということがすぐに(on the hoof)反映される

Reflexiveな研究と開発

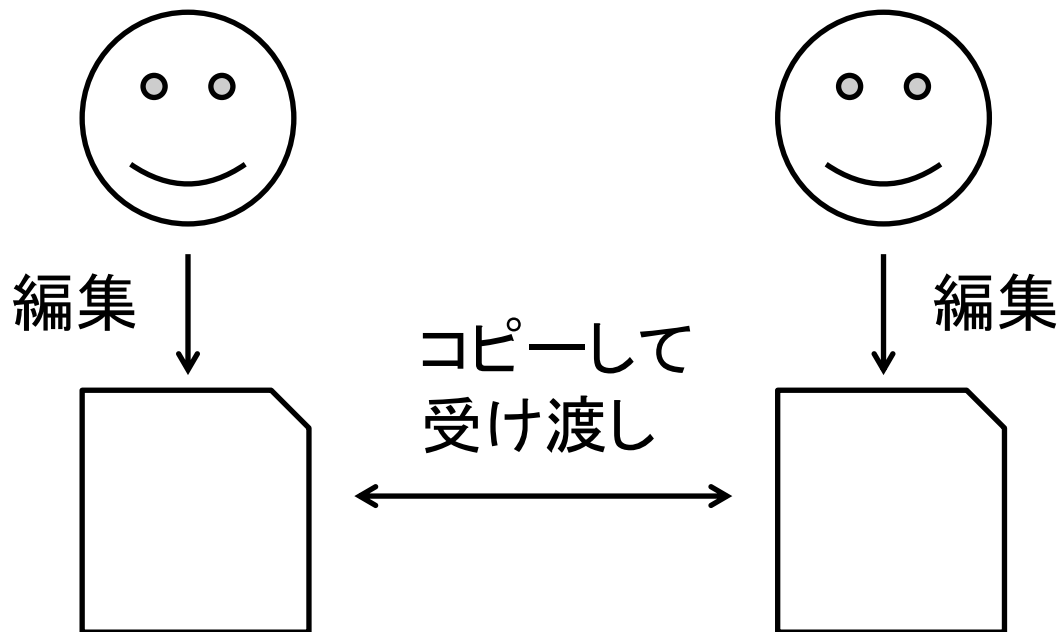


Cf. Digital Humanities の基本図式



Cf. 永崎研宣、「人文学はデジタル技術に何を期待しうるか」

ファイルの共有における難点



異なるバージョンが存在することになり
それらの違いを管理するための
プロトコルを作る必要がある

田辺史料研究会での例

毎週異なる担当者が翻刻を進め、研究会の前に担当者は20121106pre.gsx等の名前をつけたファイルを他のメンバーに送信。



研究会で検討した後に、20121106post.gsxという名前に変更したファイルを再びメンバーに送信。



次の担当者が翻刻を行い、名前を20121113pre.gsxと変更して研究会の前に他のメンバーに送信。

様々なファイル共有システム, サービス



github



Google Drive

SugarSync 

SMART-GSネットワーク版

- Subversion的なファイルの共有とバージョン管理システムを採用
- HCPサーバというウェブ上のレポジトリでSMART-GSを使って行う史料研究(プロジェクト)を管理

HCPサーバによる管理

HCPサーバ

レポジトリ

プロジェクト

イメージ・サーバ

画像の名前と
URLのデータ
ベース

アーカイブ

画像

Web

同期

プロジェクト

ダウンロード

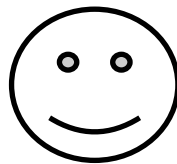
ローカル

翻刻・ノートなど

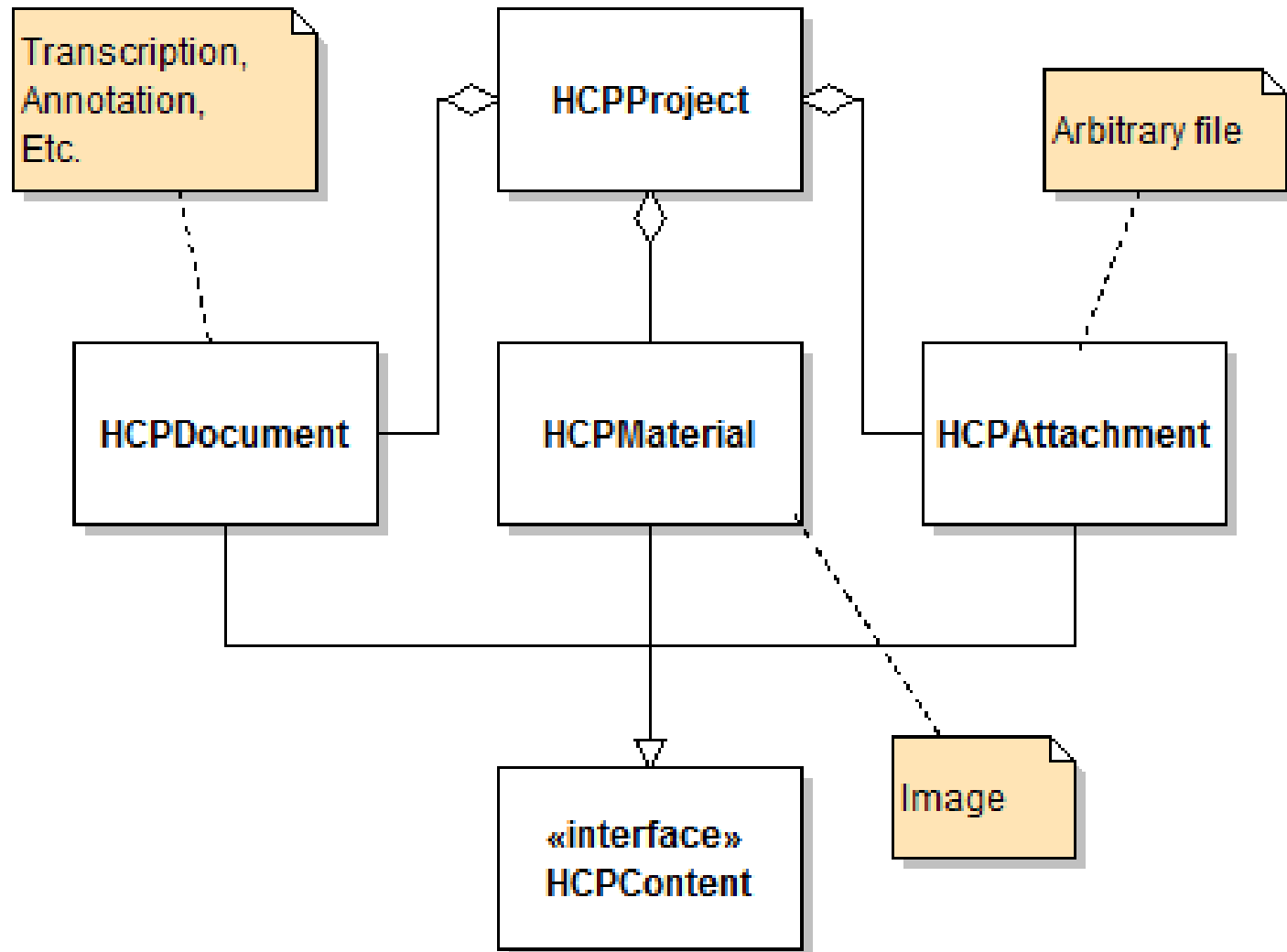
画像

その他のファイル

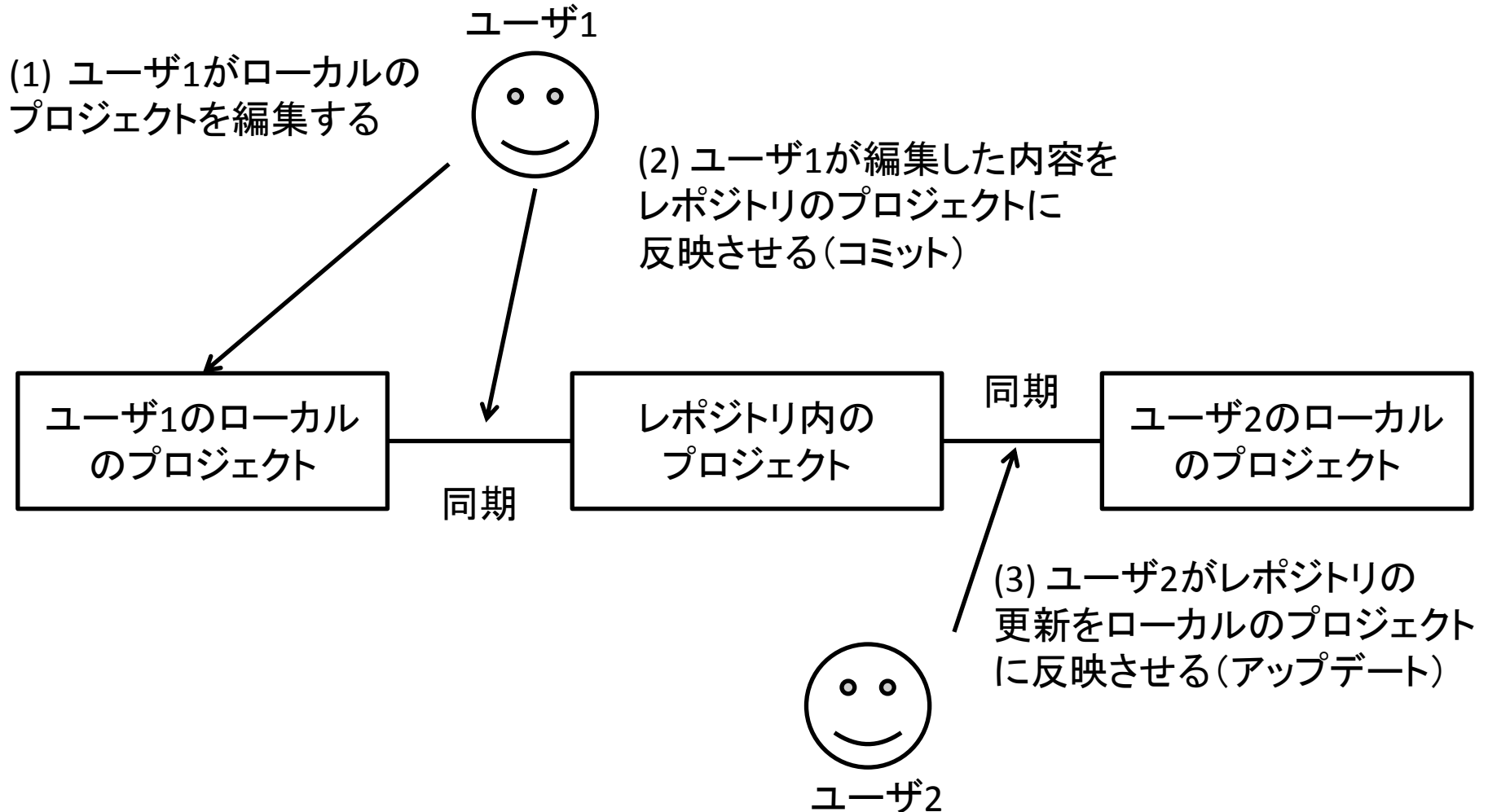
作成・編集・閲覧



プロジェクトの構成

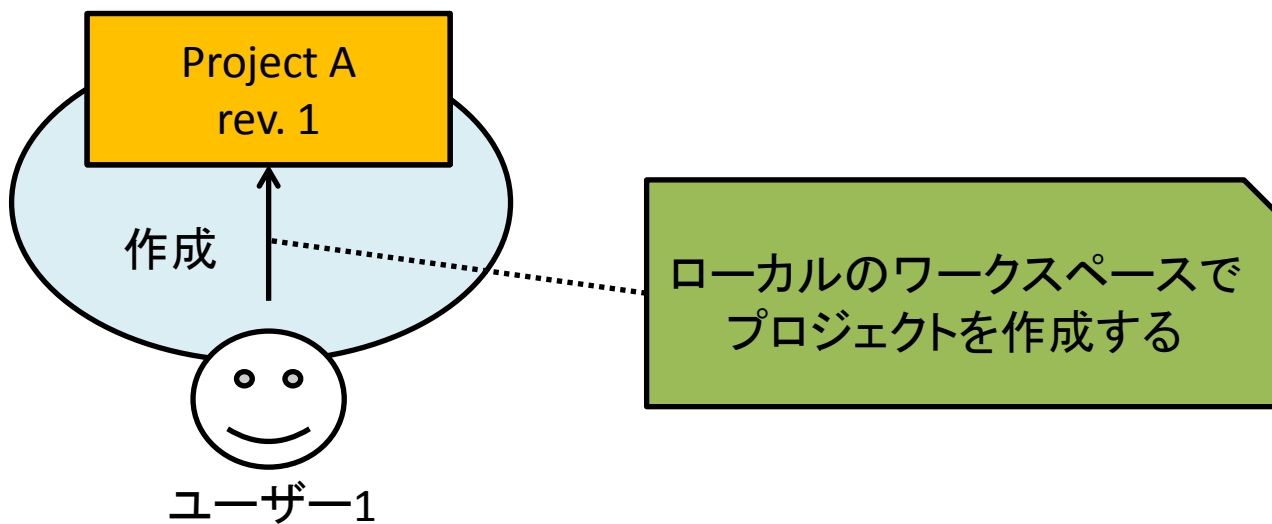
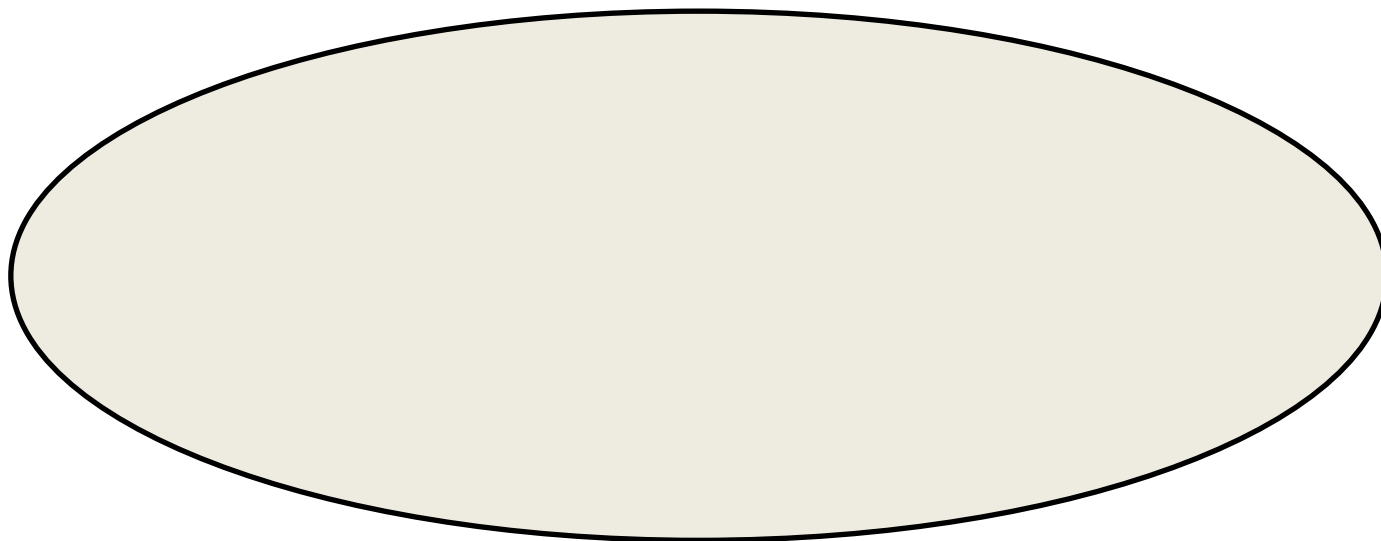


複数ユーザによる 共同翻刻



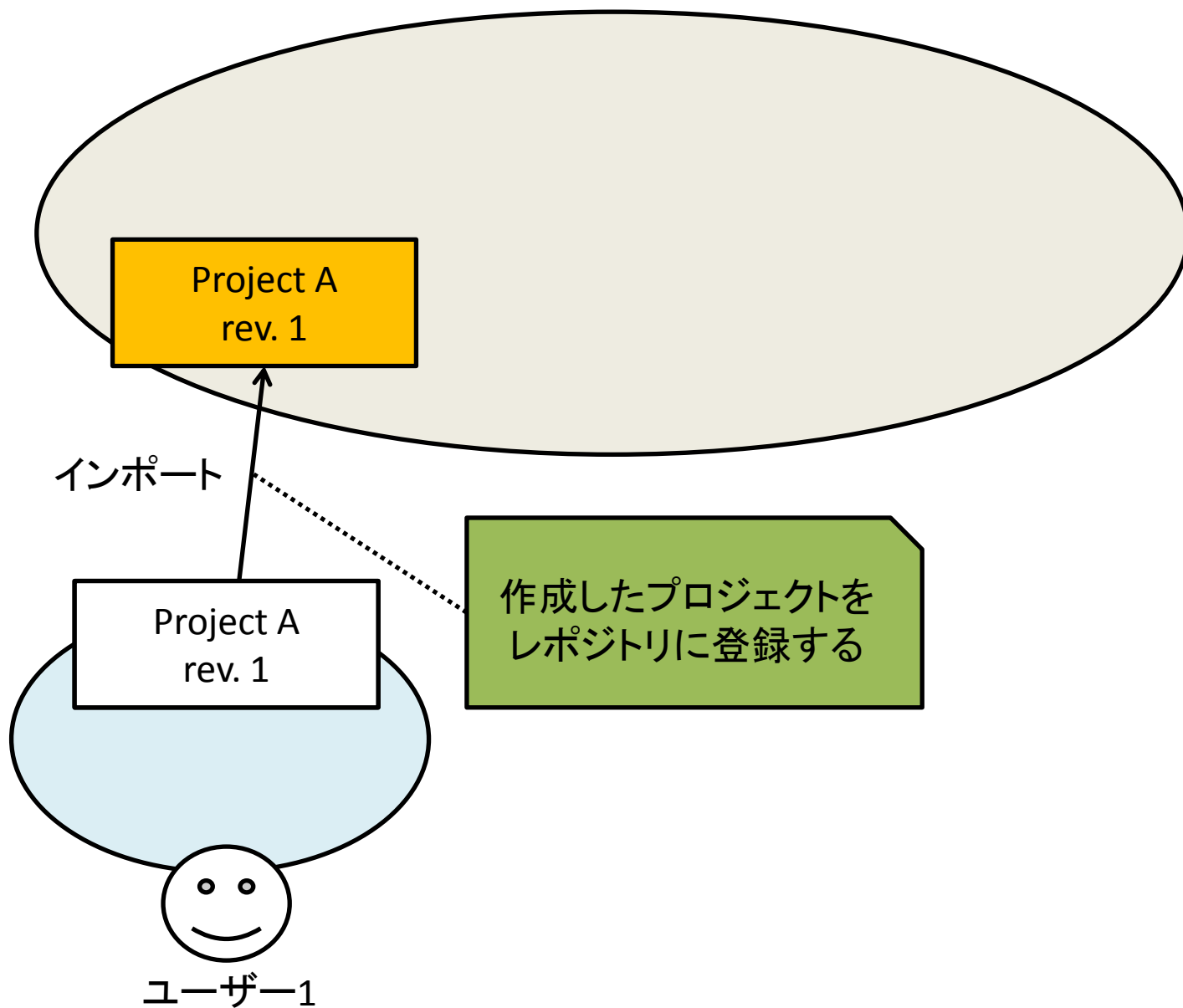
バージョン管理

レポジトリ



バージョン管理

レポジトリ



バージョン管理

レポジトリ

Project A
rev. 1

チェックアウト

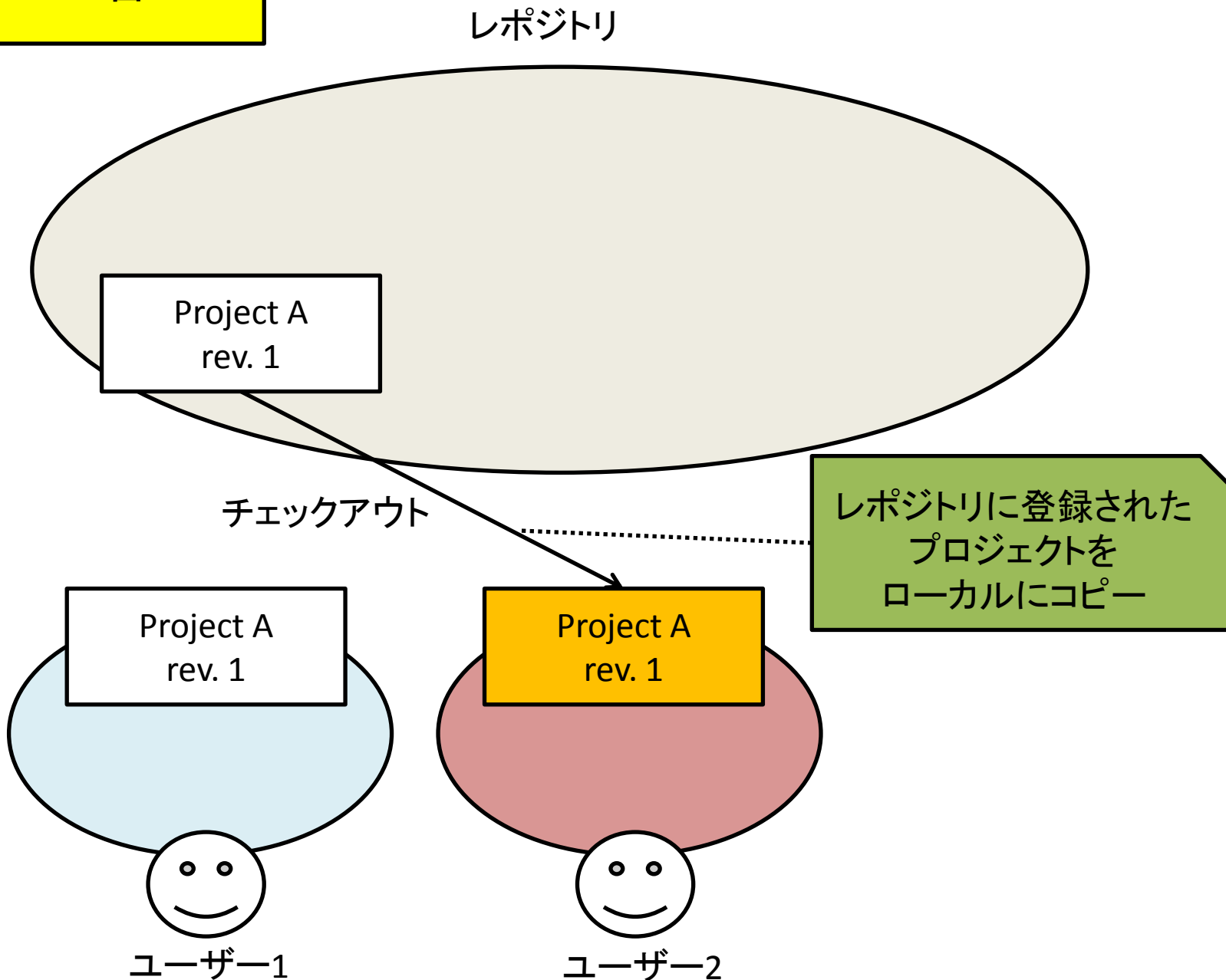
Project A
rev. 1

Project A
rev. 1

レポジトリに登録された
プロジェクトを
ローカルにコピー

ユーザー1

ユーザー2



バージョン管理

レポジトリ

Project A
rev. 1

Project A
rev. 1

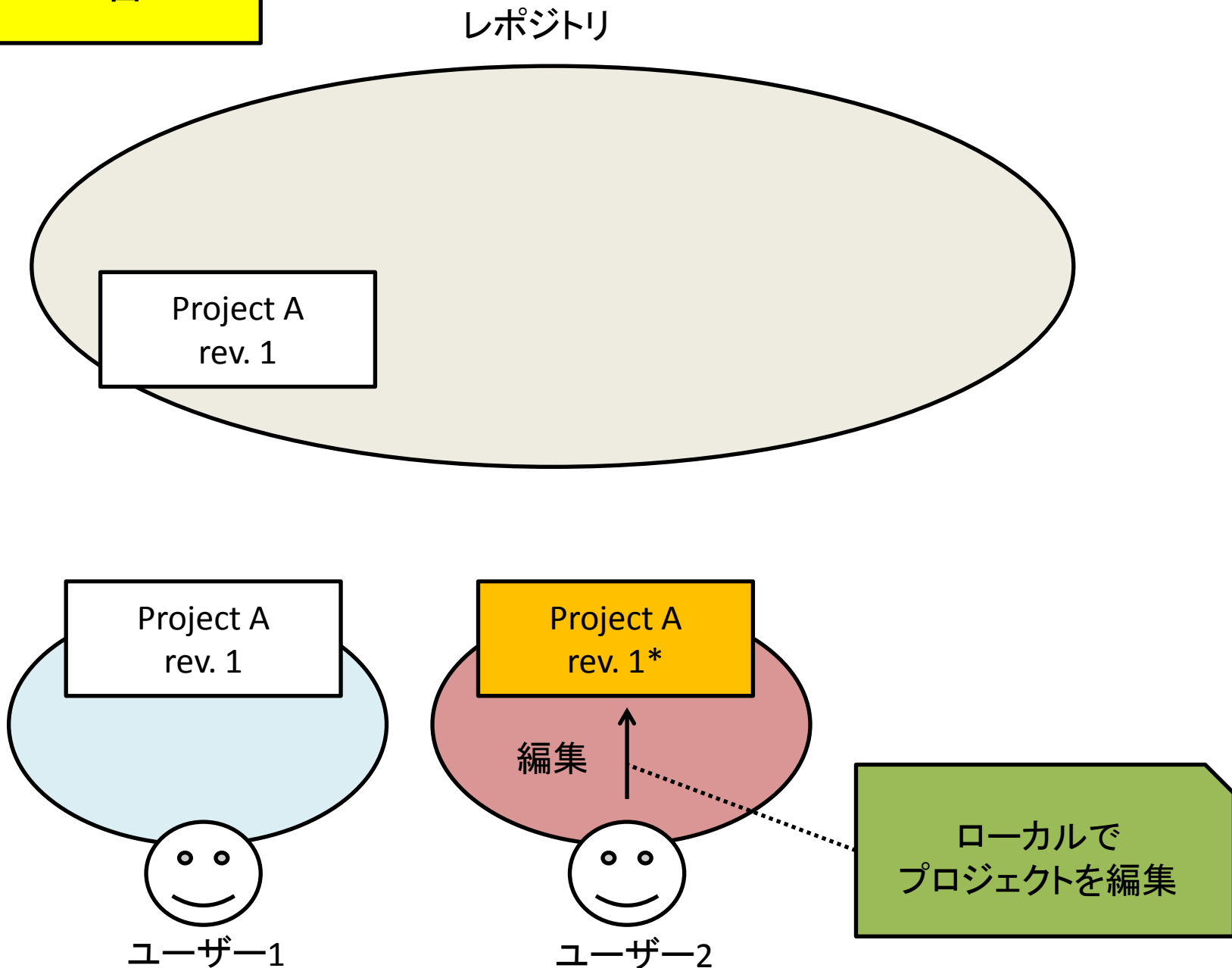
Project A
rev. 1*

編集

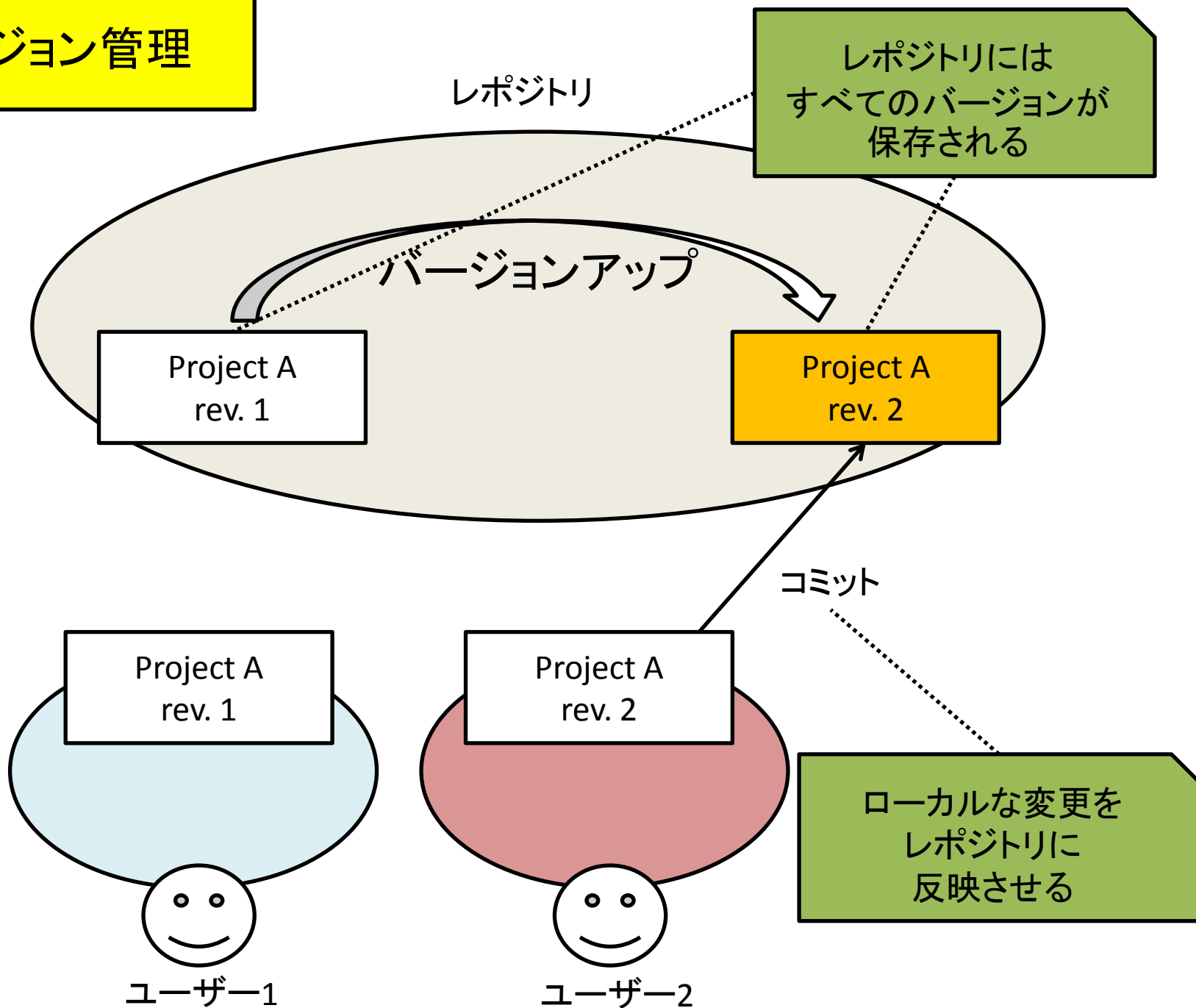
ローカルで
プロジェクトを編集

ユーザー1

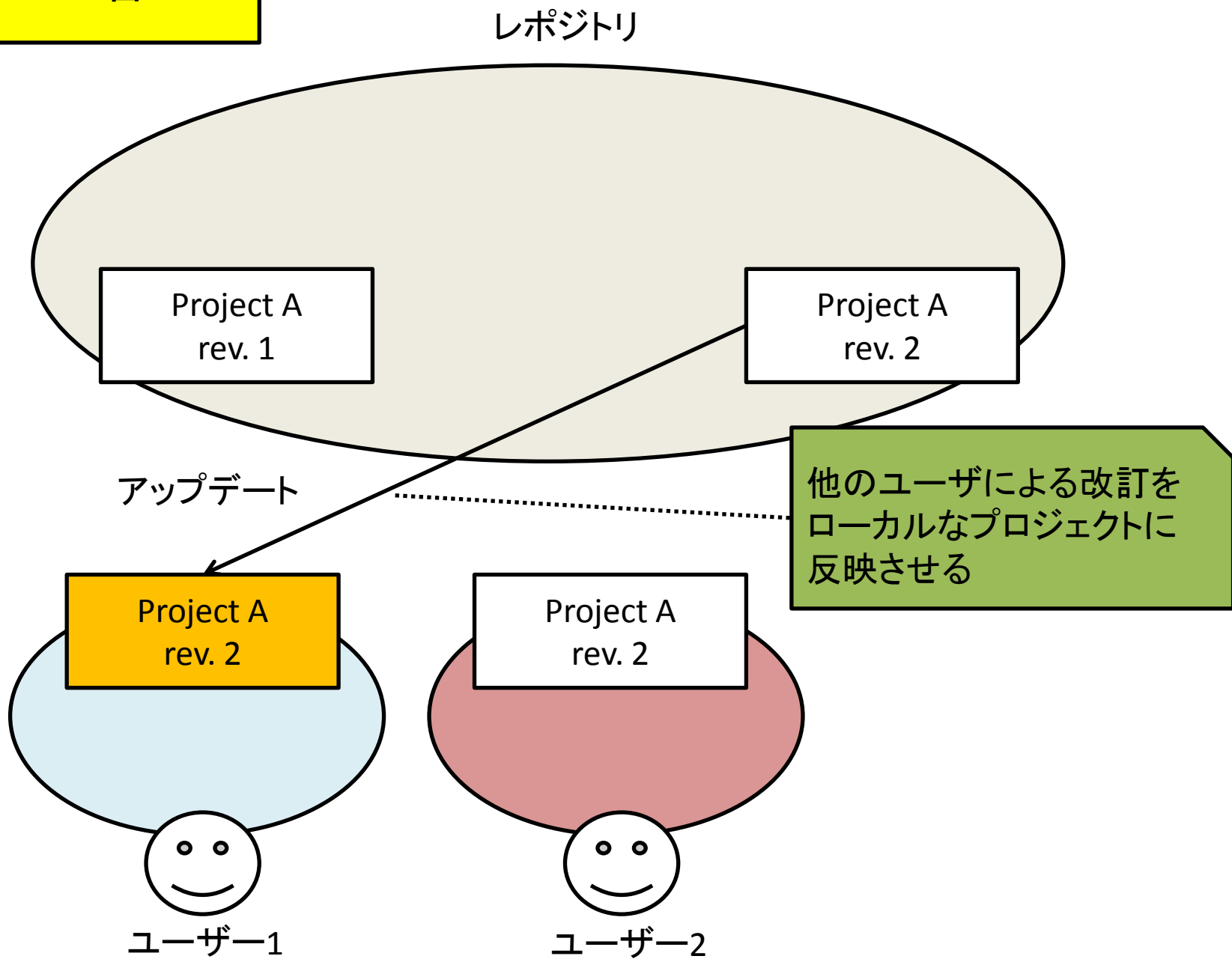
ユーザー2



バージョン管理

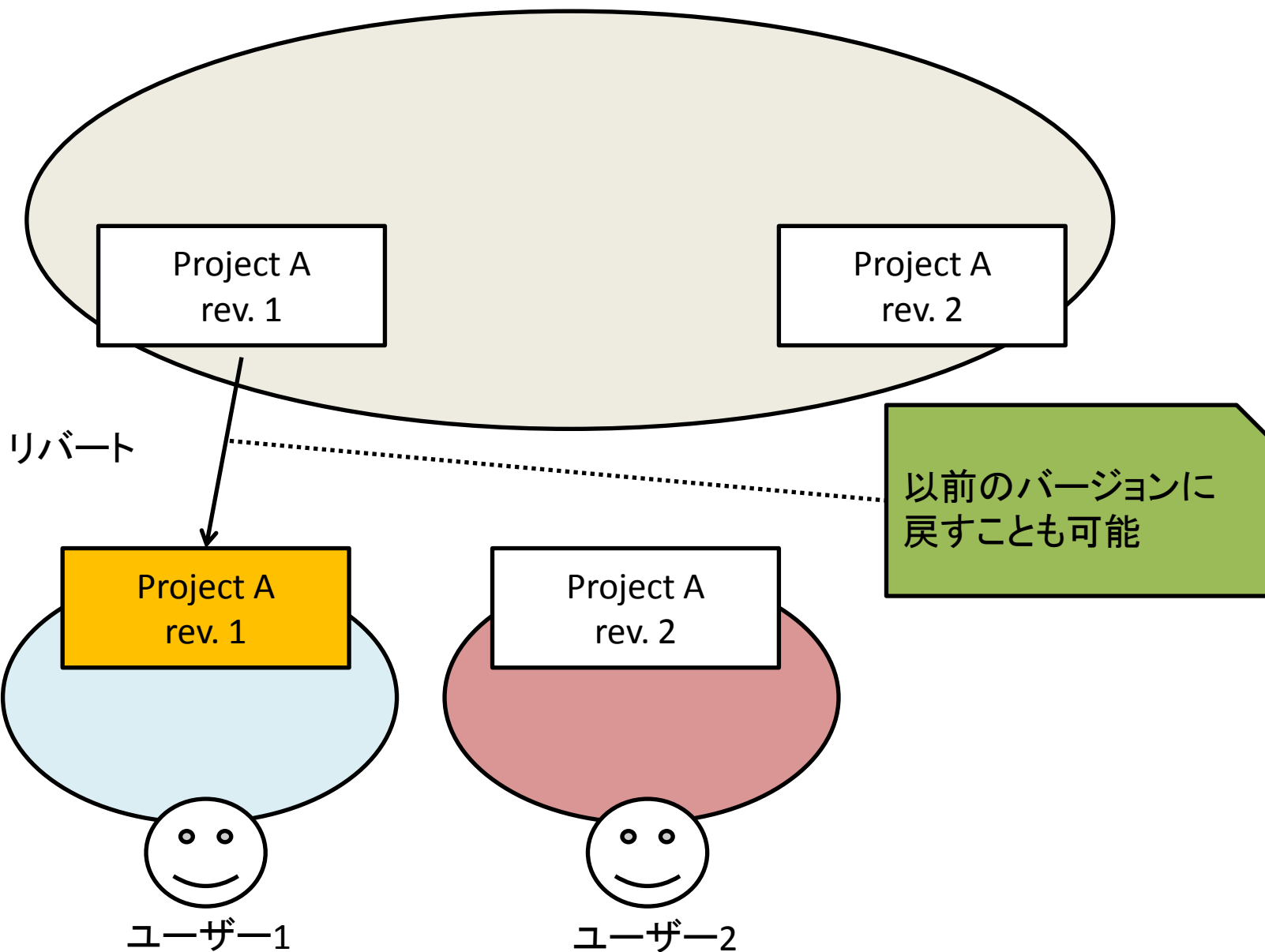


バージョン管理



バージョン管理

レポジトリ



編集の衝突

レポジトリ

Project A
rev. 1

同じバージョンを
別々に編集

Project A
rev. 1*

編集



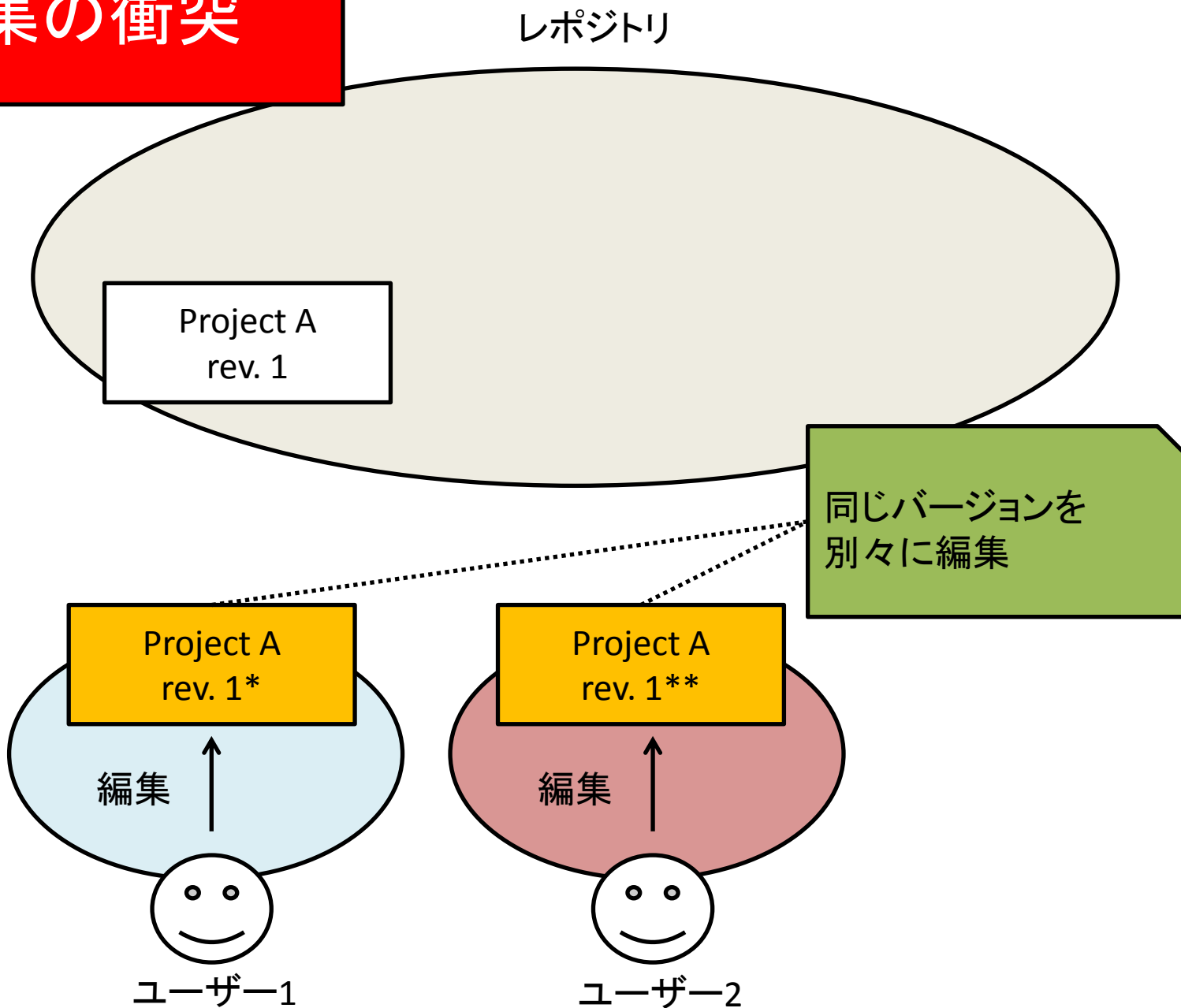
ユーザー1

Project A
rev. 1**

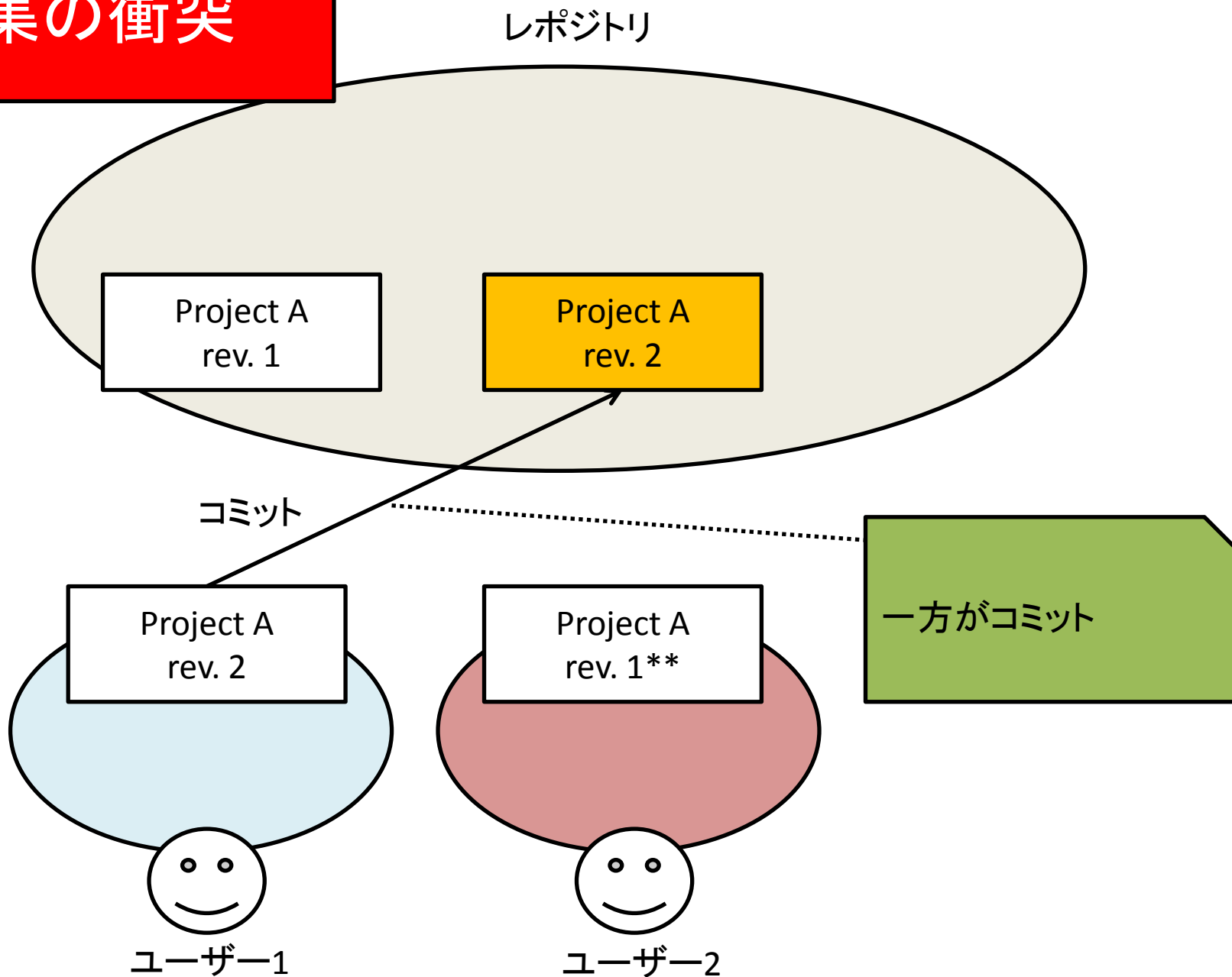
編集



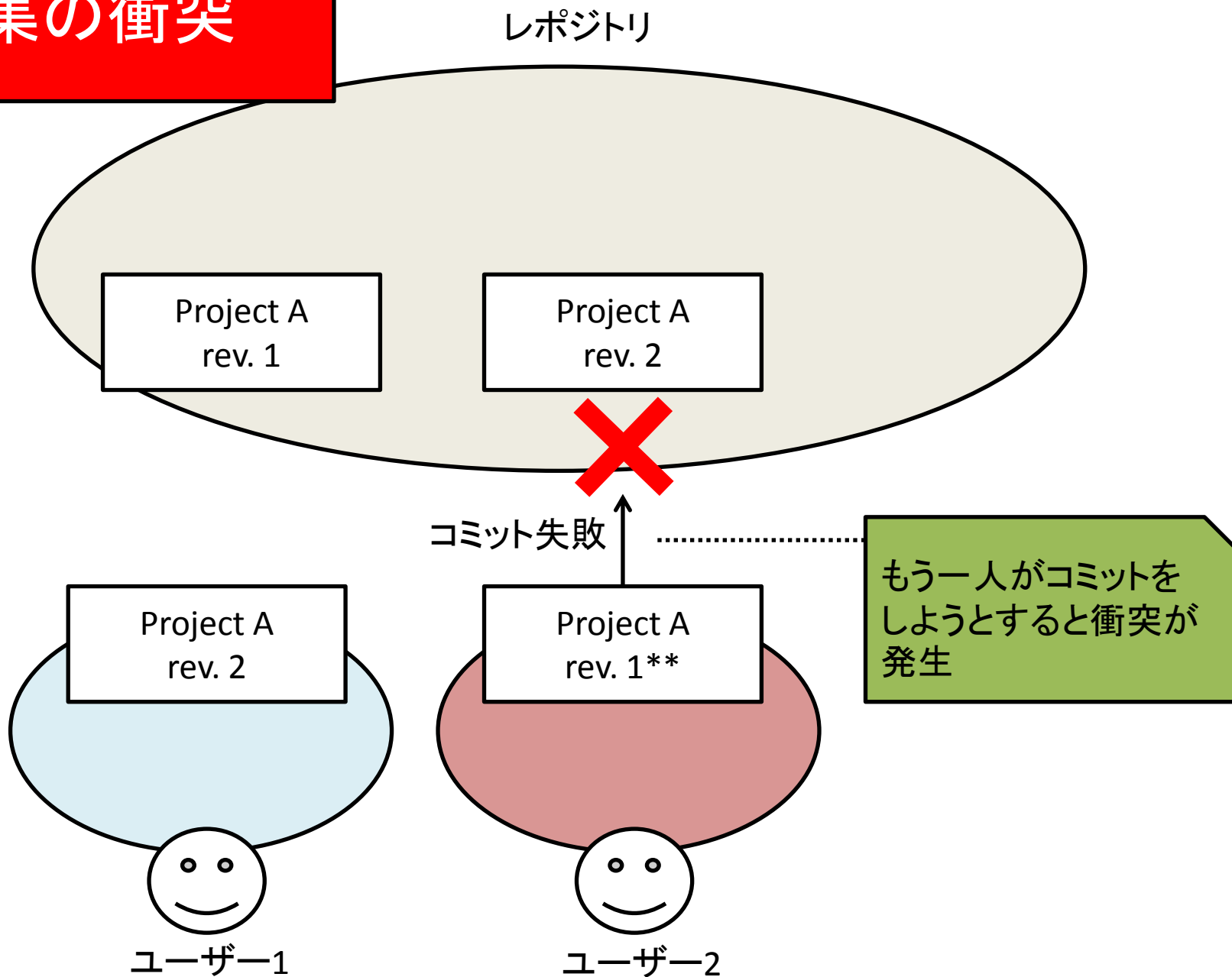
ユーザー2



編集の衝突



編集の衝突



編集の衝突

レポジトリ

Project A
rev. 1

Project A
rev. 2

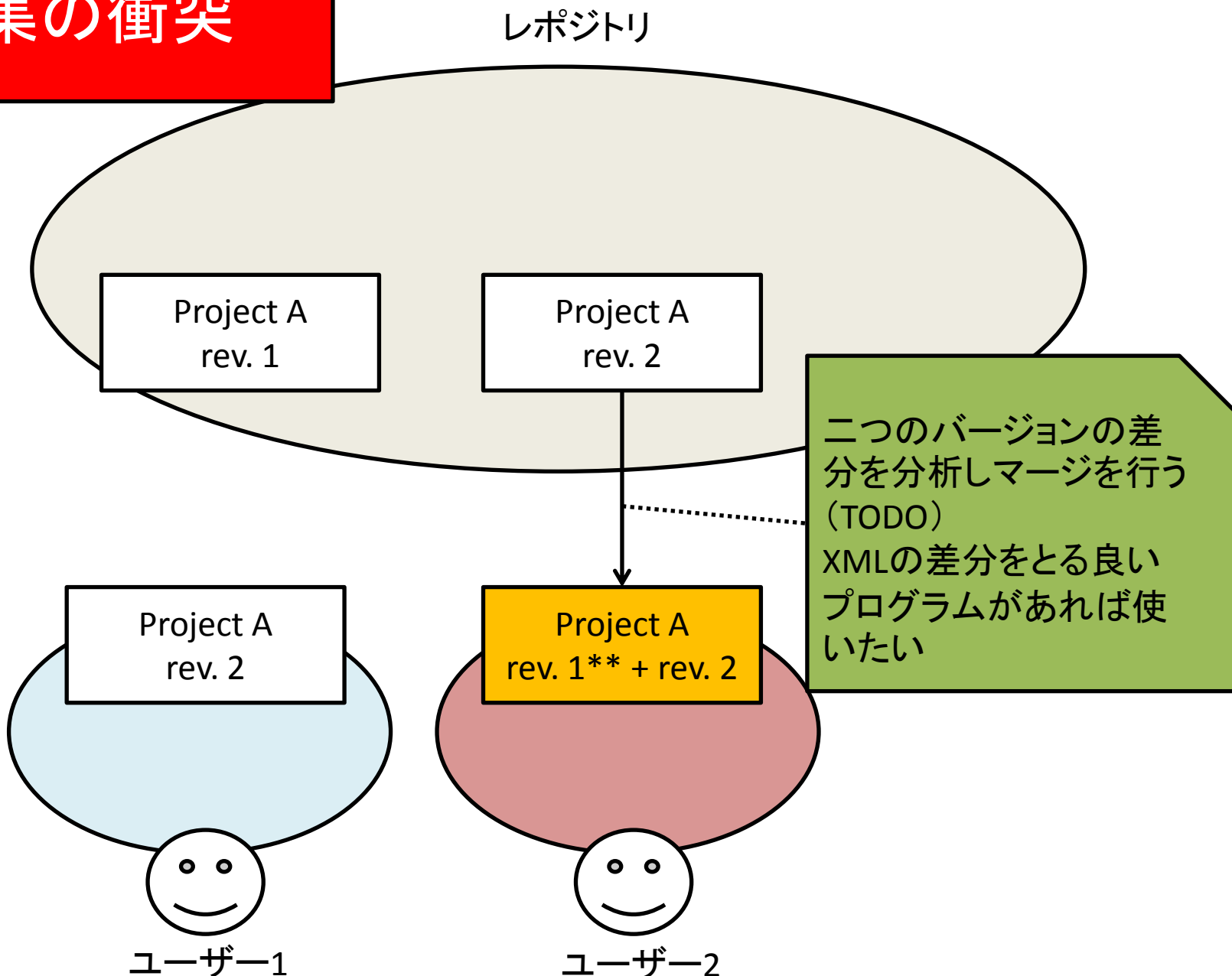
Project A
rev. 2

Project A
rev. 1** + rev. 2

二つのバージョンの差分を分析しマージを行う
(TODO)
XMLの差分をとる良い
プログラムがあれば使いたい

ユーザー1

ユーザー2



SMART-GSのUser Interface Project Explorer

プロジェクトに対する
アクション

The screenshot shows the SMART-GS Project Explorer window. The top menu bar includes: Create, Open, Authenticate, Update, Commit, Import, Checkout, Export, Compare, Remove, and Fetch Contents. The interface is divided into three main panes: Workspace, Repository, and WORKSPACE. The Workspace pane on the left lists local projects, with 'test-project' selected. The Repository pane in the middle lists projects from the web, also with 'test-project' selected. The WORKSPACE pane on the right shows the content of the selected project, including a directory tree with files like '0002', '0001', 'Attachments', and 'Images'. Below the directory tree is a preview of a document page with handwritten Japanese text. Red circles and arrows highlight these key elements.

Workspace

- share20121116pre
- 20121110(1)
- ConflictTest120831
- share20121109
- ConflictTest120816
- share20121109(2)
- 20121116(1)
- test-project
- 20121116(3)
- share20121109pre
- 20121116(2)
- 20120831(5)

Repository

- test-project2
- test-project-20120524(1)
- test-project-20120524(2)
- test-project-20120524(3)
- aaa
- ProjectC
- ProjectB
- ConflictTest120816
- ConflictTest120831
- 20120831(1)
- 20120831(2)
- 20120831(3)
- 20120831(4)
- 20120831(5)
- ProjectA
- share20121109
- share20121109(2)
- test-project
- 20121110(1)
- 20121116(1)
- 20121116(2)

WORKSPACE

- test-project
 - /テキスト/0002
 - /テキスト/0001
 - Attachments
 - Images
 - 画像/S09_004
 - 画像/S09_003
 - 画像/S09_002
 - 01

ローカルに
持っている
プロジェクト

Web上にある
プロジェクト

選択されたプロジェクト
のコンテンツ一覧

選択されたコンテンツ
のプレビュー

新しいプロジェクトを作成する

HCPサーバに関連するURLを指定

ローカルのワークスペースのディレクトリを指定

プロジェクトの名前を入力

The image shows a 'Create New Project' dialog box with the following fields and buttons:

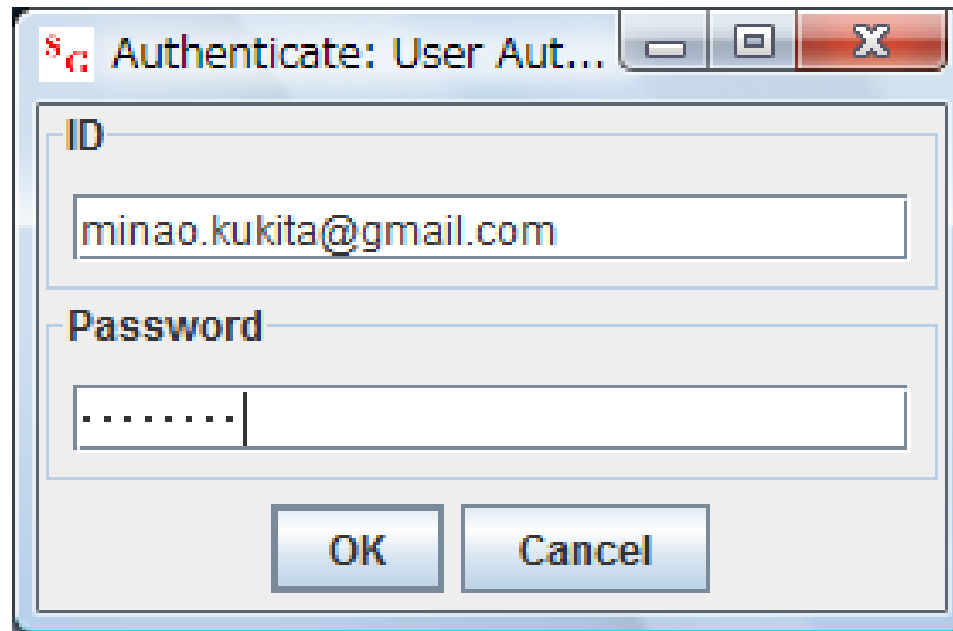
- Repository API URL:**
- Authorization URL:**
- Image Server URL:**
- Workspace Directory:**
- Project Name:**
-
-

Annotations (green boxes with arrows pointing to the dialog fields):

- Annotation 1 points to the Repository API URL, Authorization URL, and Image Server URL fields.
- Annotation 2 points to the Workspace Directory field.
- Annotation 3 points to the Project Name field.

ユーザ登録とユーザ認証

レポジトリを利用するためには
HCPサーバにユーザ登録をした上で
パスワードによる認証を行う。

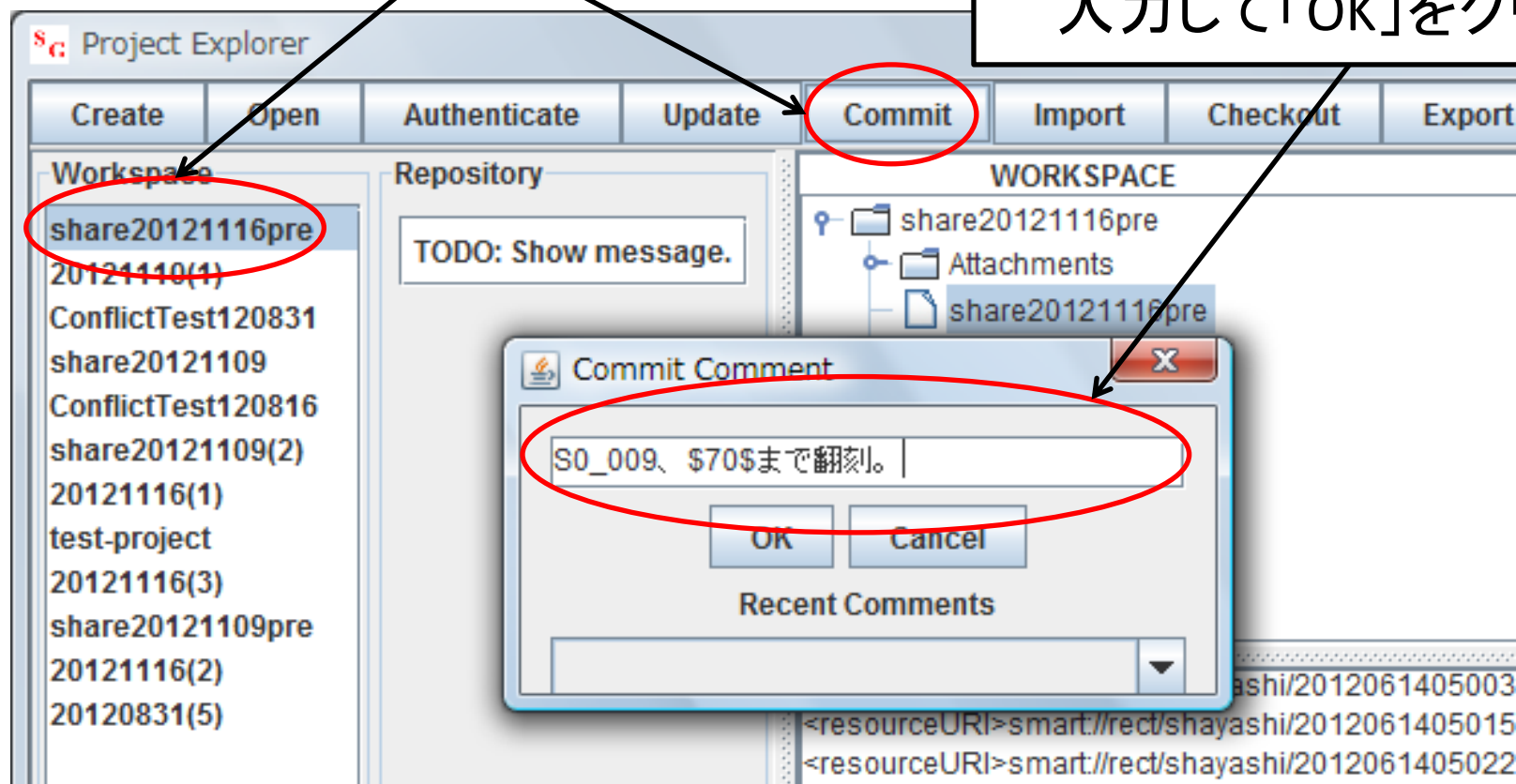


The image shows a standard Windows-style authentication dialog box. The title bar reads "Authenticate: User Aut...". Inside the dialog, there are two main sections. The first section is labeled "ID" and contains a text input field with the email address "minao.kukita@gmail.com". The second section is labeled "Password" and contains a text input field with masked characters ".....". At the bottom of the dialog, there are two buttons: "OK" and "Cancel".

コミット

ワークスペースのプロジェクトを選択して「コミット」ボタンをクリック

変更についての説明を入力して「OK」をクリック



今後の発展

バージョン管理のための
UIの洗練

サーチ結果の
自動翻刻

リアルタイムでの
オンライン共同作業

ご清聴ありがとうございました。