

# Online Appendix to: Mining and Quality Assessment of Mashup Model Patterns with the Crowd: A Feasibility Study

CARLOS RODRÍGUEZ, University of Trento  
FLORIAN DANIEL, Politecnico di Milano  
FABIO CASATI, University of Trento

---

## A. DESIGN OF TASK PAGES FOR PATTERN MINING

Figure 12 is a screen shot of the *questionnaire* used to assess workers' knowledge of Yahoo! Pipes and to decide which worker to reward. The same questionnaire is used inside CrowdFlower to assess workers in each of the crowd task designs used in the pattern mining and assessment experiments described in this article.

Figures 13 and 14 illustrate the screen shots of the *Random3* and *ChooseN* task designs for pattern identification described in Section 4.1. Like the task design shown in Figure 4, these two design are implemented as external web pages executed on our own web server and linked from within CrowdFlower.

We acknowledge one limitation pointed out by one reviewer regarding two questions included in the task ("have you ever seen this pattern?" and "have you ever used this pattern?"). For the scale we used for these questions (5-point Likert scale), it would be more correct to rephrase these questions as "how often have you seen this pattern?" and "how often have you used this pattern?"

## B. DESIGN OF TASK PAGES FOR QUALITY ASSESSMENT

Figure 15 illustrates the details of how we assessed the quality of identified mashup model patterns. The form shows one model pattern and asks the user (both workers in the crowd experiment and us in the expert assessment) to rate the pattern in terms of understandability, usefulness, reusability, and novelty.

Figure 16 explains the design of the pairwise pattern quality assessment task. The core idea is to provide the worker with two patterns and to ask him or her to choose which pattern is better in terms of understandability, usefulness, reusability, and novelty.

## C. EXAMPLES OF MINED MODEL PATTERNS

Figure 17 discusses, for each of the four assessment criteria, examples of good and bad patterns as identified by the crowd in the *Naive* pattern mining experiment.

1) Have you used Yahoo Pipes before?  
☐ Yes ← **Correct answer**  
☐ No

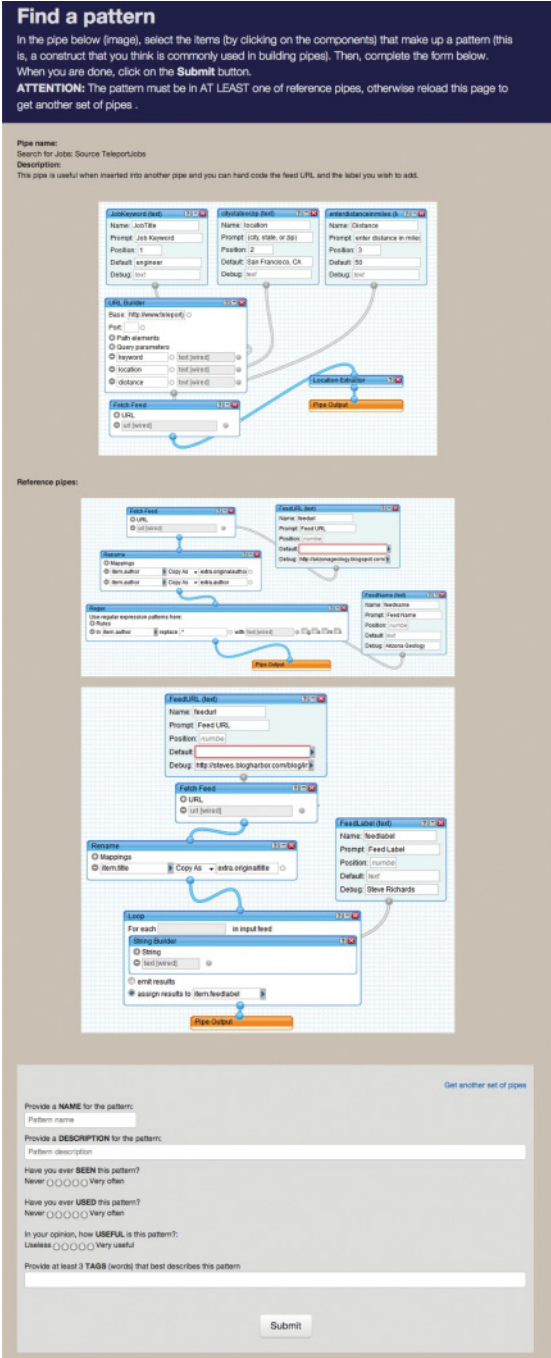
2) What is the maximum number of Pipe Output components you can have in a pipe?  
☐ Exactly 1 ← **Correct answer**  
☐ Exactly 2  
☐ As many as you want

3) To build a working pipe, one must have in his/her pipe what of the following?  
☐ Only modules  
☐ Only connectors  
☐ Both modules and connectors  
☐ Both modules and a pipe output module  
☐ Modules, connectors and a pipe output module ← **Correct answer**  
☐ None of the above

4) Assume you would like to fetch news from an online newspaper in RSS format. Which of the following components is the most appropriate for this task?  
☐ Fetch CSV  
☐ Fetch Data  
☐ Fetch Feed ← **Correct answer**  
☐ RSS item builder  
☐ All of them

5) Which of the following components can be embedded into a Loop component?  
☐ Item builder  
☐ Fetch data  
☐ Find First Site Feed  
☐ URL builder  
☐ All of them ← **Correct answer**

Fig. 12. Screen shot of the questionnaire used to assess workers' acquaintance with Yahoo! Pipes. The figure also highlights the correct answer of each question.



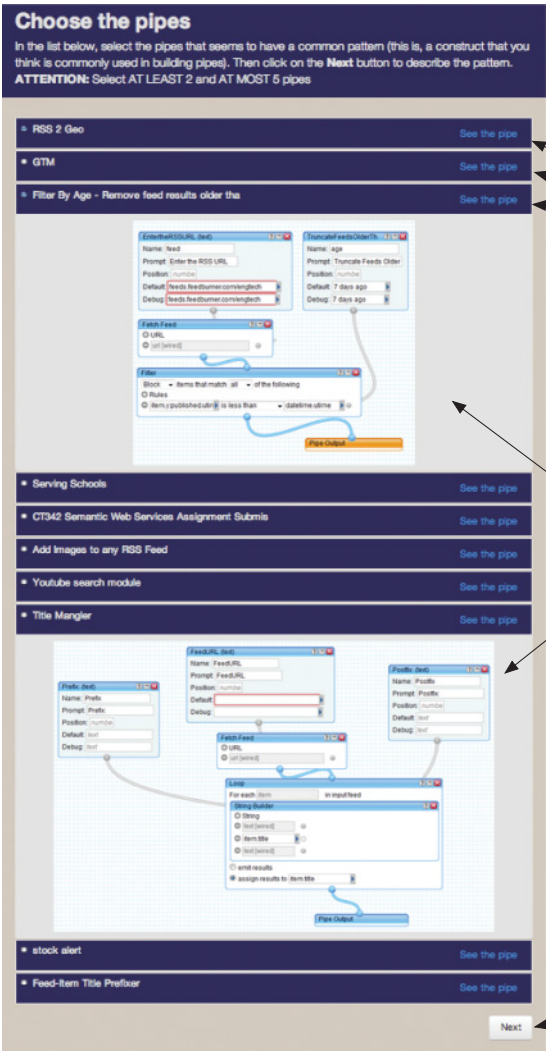
Short description of task with **instructions** for the worker

Main pipe of the task. Here workers can **select the components** of the identified pattern (if any).

Two randomly chosen pipes complementing the main pipe of the task. Workers are asked to compare the main pipe with these two to **identify similarities**.

Input form asking for additional **metadata**.

Fig. 13. Screen shot of the *Random3* task UI implemented for the identification of mashup model patterns from a set of three different pipes models.



Short description of task with instructions for the worker

The worker can select n pipes from 10 randomly chosen pipes. By clicking on the tabs, the worker can open (select) and close (unselect) pipes.

The two pipes chosen by the worker.

Leads the worker to the actual pattern identification page (similar to the one of the Random3 task design).

Fig. 14. Screen shot of the *ChooseN* task UI implemented for the identification of mashup model patterns from a set of three different pipes models freely chosen out of 10 available models.

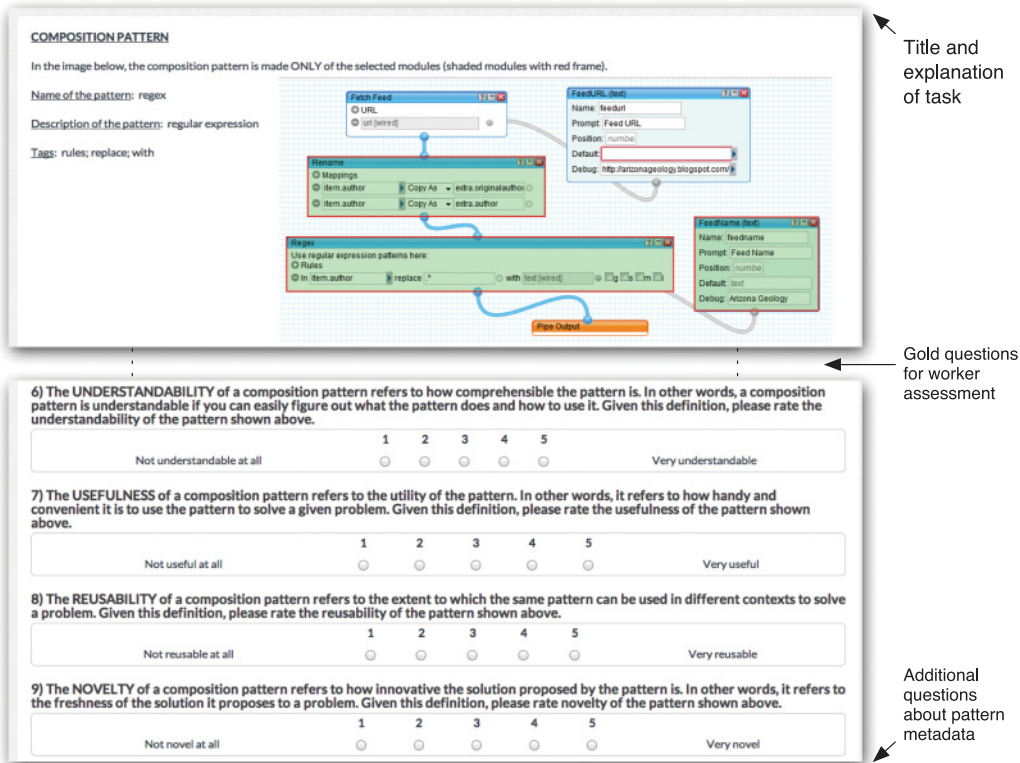


Fig. 15. Screen shot of the task implemented for the assessment of pattern quality. Both the experts and the crowd in the *Individual* setting use the same task design to perform their evaluation.

Assessment of composition patterns from Yahoo! Pipes

Introduction

In this task, you need to assess the quality of composition patterns from Yahoo! Pipes. A composition pattern is a fragment of modules and connections that altogether accomplish specific functionality by solving a given problem. In order to perform this task, you need to read the task instructions described below.

12 READ carefully ALL the instructions provided for this task. ONLY AFTER you read all the instructions proceed with the execution of them.

13 Inspect the two images (composition patterns) and try to decipher these before and try to understand what each pattern does.

14 Answer questions 1 to 12. The 12 questions concern these questions refer to the images below. You can answer these questions based on your own application of the knowledge acquired.

15 Answer the 16 remaining questions (questions 13 to 28). These questions refer to general pipes modeling knowledge.

Assessment


PIPEPATTERN

Introduction

Name of the pattern: KTV-Tube

Description of the pattern: Change Provider's feeds to a program to feed that can be used by a Personal Tube or any feed reader.

Tags for the pattern: RSS, KTV



Assessment

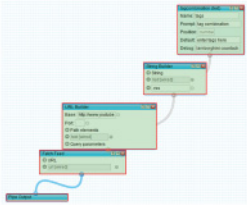
PIPEPATTERN

Introduction

Name of the pattern: URL-Feed

Description of the pattern: Retrieve feeds for feed

Tags for the pattern: RSS, Feed, URL



1) Which of the patterns (Pattern A or Pattern B) is more UNDERSTANDABLE? (The understandability of a composition pattern refers to how comprehensible the pattern is, in other words, a composition pattern is understandable if you can easily figure out what the pattern does and how to use it)

☐ Pattern A

☐ Pattern B

2) Which of the patterns above (Pattern A or Pattern B) is more USEFUL? (The usefulness of a composition pattern refers to the utility of the pattern, in other words, it refers to how handy and convenient it is to use the pattern to solve a given problem)

☐ Pattern A

☐ Pattern B

3) Which of the patterns above (Pattern A or Pattern B) is more REUSABLE? (The reusability of a composition pattern refers to the extent to which the same pattern can be used in different contexts to solve a given problem)

☐ Pattern A

☐ Pattern B

4) Which of the patterns above (Pattern A or Pattern B) is more NOVEL? (The novelty of a composition pattern refers to how original the solution proposed by the pattern is, in other words, it refers to the freshness of the solution, in particular to a problem)

☐ Pattern A

☐ Pattern B

5) Did you find the NAME of the patterns provided above helpful for understanding the patterns?

Not helpful at all    1    2    3    4    5    Very helpful

6) Did you find the DESCRIPTION of the patterns provided above helpful for understanding the patterns?

Not helpful at all    1    2    3    4    5    Very helpful

7) Did you find the TAGS of the patterns provided above helpful for understanding the patterns?

Not helpful at all    1    2    3    4    5    Very helpful

8) How easy was Yahoo! Pipes before?

☐ Very easy

☐ Easy

☐ Somewhat easy

☐ Somewhat difficult

☐ Difficult

☐ Very difficult

9) What is the maximum number of Pipe Output components you can have in a pipe?

☐ Unlimited

☐ Between 1 and 10

☐ Between 1 and 20

☐ Between 1 and 50

10) To build a working pipe, one must have in their pipe what of the following?

☐ Data module

☐ Data component

☐ Both modules and components

☐ Both modules and Pipe Output module

☐ Module, component and Pipe Output module

☐ None of the above

11) Assume you would like to fetch news from an online newspaper in RSS format. Which of the following components is the most appropriate for this task?

☐ Fetch RSS

☐ Fetch Data

☐ Fetch Feed

☐ RSS to HTML

☐ RSS to JSON

12) Which of the following components can be embedded into a Loop component?

☐ Fetch RSS

☐ Fetch Data

☐ Fetch Feed

☐ RSS to HTML

☐ RSS to JSON

Task title and description

The two patterns to compare

Options to choose best pattern regarding understandability, usefulness, reusability and novelty

Additional metadata questions

Gold questions for worker assessment

Fig. 16. Screen shot of the task UI implemented for the pairwise comparison and ranking of identified model patterns.

ACM Transactions on Internet Technology, Vol. 16, No. 3, Article 17, Publication date: June 2016.



	<p><b>Good example</b></p>	<p><b>Bad example</b></p>
Reusability	<p>For this criteria, the good example shown on the left performs a simple but concrete function, while the bad example on the right seems to perform many functionalities that, as a whole, may be hard to reuse into a different pipe to address a particular problem.</p>	
Novelty		
Usefulness	<p>The good example shows a pattern that seems useful because it allows the user to address a common need: provide authentication to access a web resource (feeds in this case). The bad example shows a combination of components that can be hardly useful in a different context of this particular pipe.</p>	
Understandability		

Fig. 17. Examples of mashup model patterns discovered by the crowd.