An overview with pictures: Automated Scoring With a
Semantic Grammar Rubric Expression Language (SGREL) Approach

16 December 2018

This demo will step through each capability of the SGREL Scoring Web Application platform and, along the way, show how to: create an institution and users, establish a rubric and an item type to score; define in detail a scoring rubric; and load, score and report on item responses.

In these examples the item will be a critical thinking question and the rubric will be a two-factor scoring metric based on the recognition of a proper claim and the citation of relevant evidence (as defined by the item author), In another scenario documented separately, the institution and group is an automobile dealership group, the items to score are inbound sales calls automatically recorded by their infrastructure, and the calls will be scored by a simple metric used to train and coach sales representatives.

This demo will be operational at http://CogWrite.ai; the car dealer / sales call scoring demo is available at http://CharmScience.ai and is available for review upon request and with the permission of the client.

Note also that full size, full resolution of all the figures in this document can be viewed at http://www.cogwrite2.com/docs/sgrel_ug_images/.

Prepared to support the review of the SGREL "rubric expression language" scoring platform / project….

Harry A. Layman
63 Mandria
Newport Coast, CA 92657

(949) 945-3373

© 2018 Harry A. Layman - All Rights Reserved

Table of Contents

I. Overview: Plan for this Demo *6*

A. Users and Groups: Organization of Data on this Platform **6**

B. Administration **6**

C. Items and Rubrics **6**

D. Target Response Elements **6**

1. TRE Groups 6
2. TRE sub-expressions 6

E. Item – Response Associations **6**

F. Upload and Process Documents, Calls, and Item Responses **6**

1. Single item loading, scoring and reporting 6
2. Batch loading / importing 6
3. Batch Scoring and Reporting 6
4. Item Response, Document & Call Record Review 6
5. Scoring single documents, call records or item responses 6

G. Score Reporting Review **6**

II. Users and Groups: Organization of Data on this Platform *7*

A. Log In **7**

B. Groups **8**

C. Users **9**

1. Add User 10
2. Manage Users 11
3. Delete User 11
4. Edit User Information 12

III. Administration *12*

A. Server Status page **13**

B. Log File page **14**

IV. Items and Rubrics *14*

A. Items **14**

1. Manage Items 15
2. Add Item 15
3. Edit Item 16

B. Rubrics **17**

1. Rubric Metadata 17
2. Add / Edit Rubric Popover 18
3. Rubric Target Response Elements 20
4. Rubric Scoring Formula 22

V. Target Response Elements *23*

A. TRE Groups **23**

B. TRE sub-expressions **24**

VI. Item – Response Associations *25*

VII. Upload and Process Documents, Calls, or Item Responses *27*

A. Single item loading, score and reporting **27**
1. Single item loading 27
2. Single Item Scoring 30
3. Single Item Reporting 33
4. The sa71 report: Scoring Detail Report 36
5. The sa61 report: End User Scoring Detail Report 39
6. The sa72 and sa62 Reports - Differences from sa71 and sa61 41

B. Batch Scoring **45**

C. Batch loading / importing **46**

VIII. Score Reporting ***46***

Table of Figures

Figure II-1 SGREL Scoring Web App Landing Page 7
Figure II-2 Log-In Page 8
Figure II-3 Manage Groups Page 8
Figure II-4 Add User Group Page 9
Figure II-5 Add User Page 10
Figure II-6 Manage Users Page 11
Figure II-7 Delete User Page 11
Figure II-8 Manage Users Page 12
Figure III-1 Server Status Page 13
Figure III-2 Log File View Page 14
Figure IV-1 Manage Items Page 15
Figure IV-2 Add Item Page 16
Figure IV-3 Edit Item Page 16
Figure IV-4 Manage Rubrics Menu Page 17
Figure IV-5 Manage Rubrics Page 18
Figure IV-6 Add / Edit Rubric Page 19
Figure IV-7 Edit Rubric Final Scale Score Descriptors 20
Figure IV-8 Manage TRE Page Summary 21
Figure IV-9 Edit TRE Page Summary 22
Figure V-1 Manage TRE Subexpression Menu 24
Figure V-2 Manage TRE Subexpression - Pending Changes 25
Figure VI-1 Manage Item-Rubric Associations 26
Figure VI-2 Add Item-Rubric Associations 27
Figure VII-1 Load and Process (a Single) Document Page 28
Figure VII-2 - Text field Entry on Upload Page 29
Figure VII-3 Upload File / Text Processing Complete 29
Figure VII-4 Navigate to Review Item Response Data 30
Figure VII-5 Navigate to Final Row to See Item Response 31
Figure VII-6 Show *IR Score* Page for a Single *Item Response* 32
Figure VII-7 Human Summary Scoring Popover Window 33
Figure VII-8 Item Response Scoring Complete Message 33
Figure VII-9 Get Reports for IRA Page (Modal Window) 34
Figure VII-10 Report Choice Dropdown for IRA 1 (i3r1) 35
Figure VII-11 Detail Score Report: sa6x and sa7x Options 36
Figure VII-12 sa71 Score Report Pt 1 IR Metadata & FSSs 37
Figure VII-13 sa71 Score Report Pt 2 Match Details 38
Figure VII-14 sa71 Score Report Pt 3 Sentence Details 39
Figure VII-15 sa71 Score Report Pt 3 Sentence Details - Zoomed 39
Figure VII-16 sa61 Score Report - All 3 Sections 40
Figure VII-17 sa72 New Section 3 (Manual Scores) 42
Figure VII-18 sa62 New Section 3 (Manual Scores) 43
Figure VII-19 Batch Scoring in Process 45
Figure VIII-1 Sample One - Score and Feedback 46
Figure VIII-2 Sample Two - Score and Feedback 47

I. Overview: Plan for this Demo

A. Users and Groups: Organization of Data on this Platform

B. Administration

C. Items and Rubrics

D. Target Response Elements

1. TRE Groups

2. TRE sub-expressions

E. Item – Response Associations

F. Upload and Process Documents, Calls, and Item Responses

1. Single item loading, scoring and reporting

2. Batch loading / importing

3. Batch Scoring and Reporting

4. Item Response, Document & Call Record Review

5. Scoring single documents, call records or item responses

G. Score Reporting Review

II. Users and Groups: Organization of Data on this Platform

As a platform solution, the SGREL Scoring Web Application (SSWAP) expects and supports different institutional groups of users to share use of the platform while isolating each user group's data from the other user groups.

Within each user group organization, the SSWAP allows any number of separate user accounts, each of which can be assigned one or more roles within the applications. Available roles for institutional users include Regular User, who can upload artifacts to score, score them, and generate score reports of various kinds. In addition, regular platform users can create and edit

The SGREL scoring platform organizes **users** and data by customer or **group**, and isolates data from access by **users** of one group to data of another **group**.

To begin our demonstration (hereinafter "demo"), we will create a new group and add a user to that group. This user account will be used for the remainder of this demo.

A. Log In

Logging into the platform a platform level, not only a group level admin, the menu below is shown with the choice to create a new group.

The home page for the platform at CogWrite.ai is here:

Figure II-1 SGREL Scoring Web App Landing Page

The log-in page is shown below.

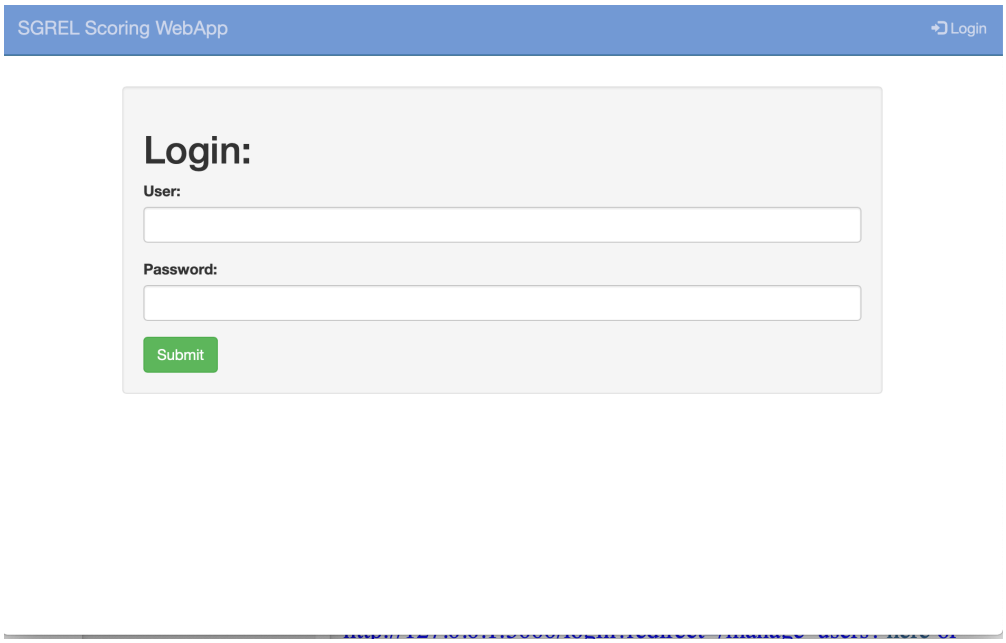


Figure II-2 Log-In Page

B. Groups

Logging in and selecting Groups --> Manage Groups shows the manage **groups** page:

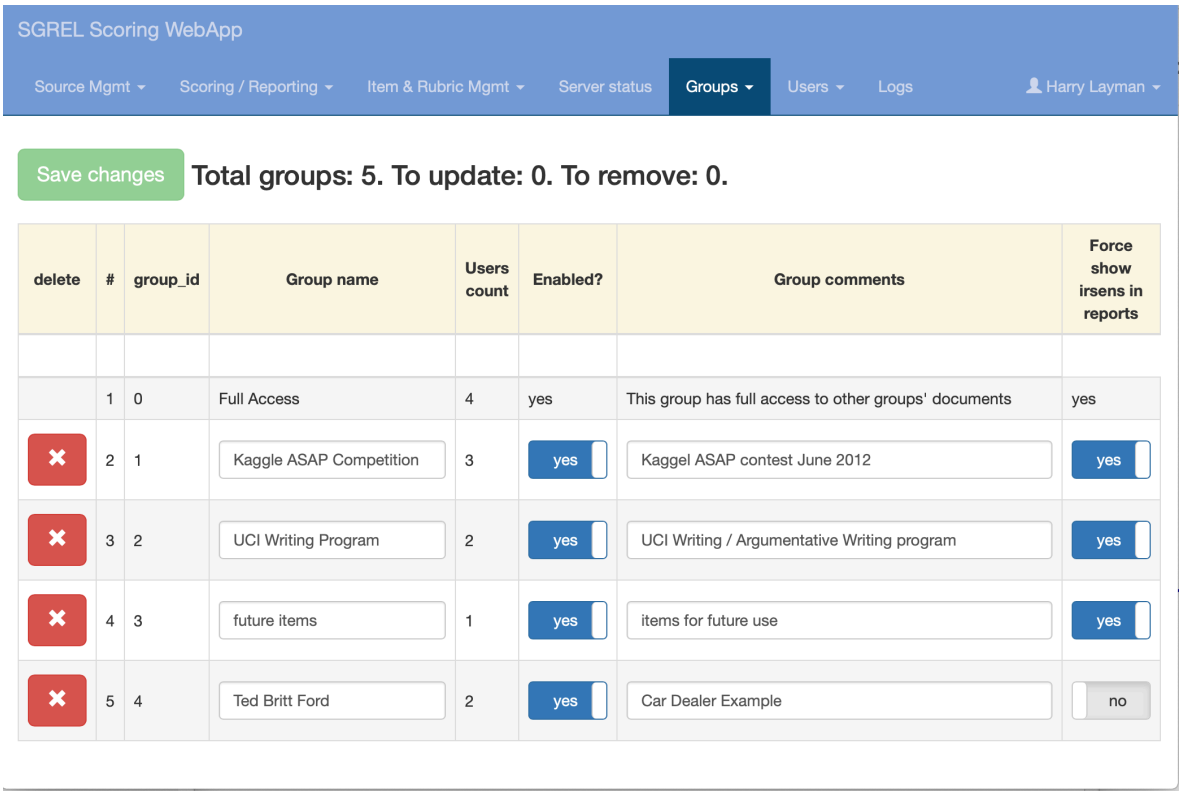


Figure II-3 Manage Groups Page

This page allows groups to be edited, disabled or deleted. To add a new group, we use the Group -> Add Group menu choice, which presents the add **group** page, as shown below.

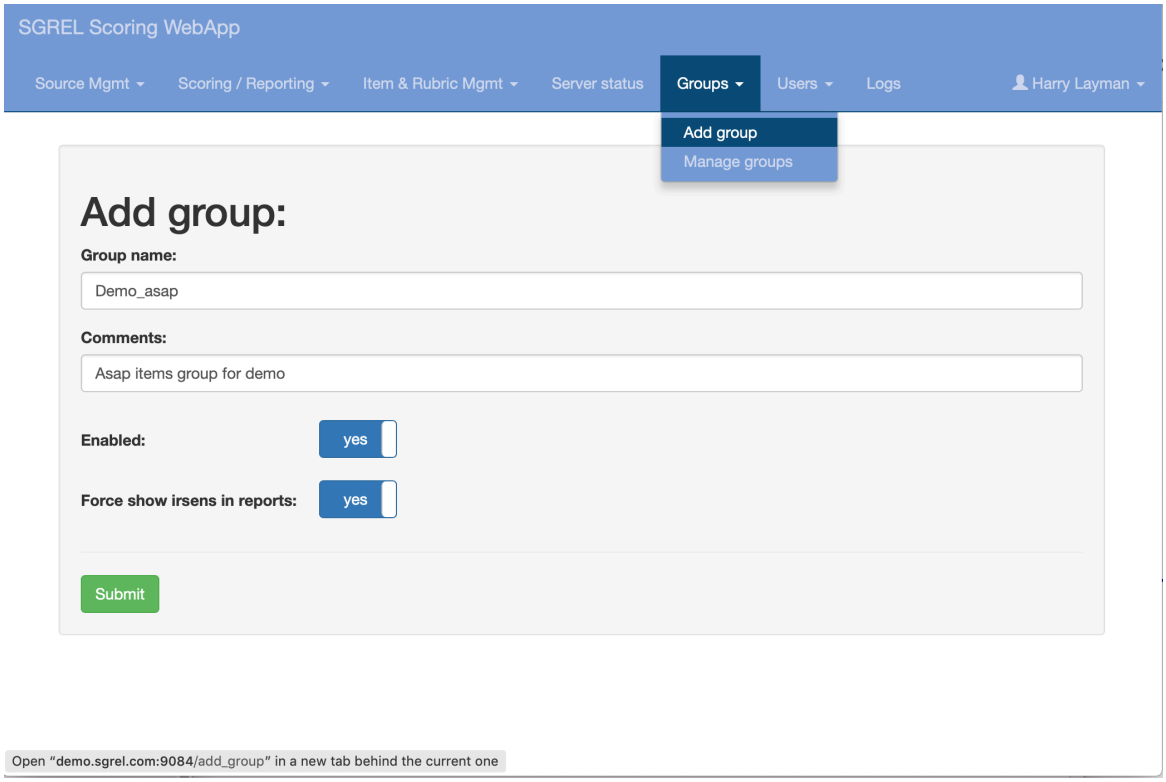


Figure II-4 Add User Group Page

The first group attribute choice will allow the new group to be enabled / disabled (disable will prevent log-ins from group users); the can be changed as by the platform admin as shown on the page for managing groups above. The second option insures that score reports list all item response lines below each item response score information on every scoring report. If not set, some reports may only list portions of answers for which credit was awarded.

C. Users

Next we will add a user to this group, as shown the add user page below. This page is used to establish the following user attributes:

- **User** name. No spaces, start with a letter. Make them unique.
- Password. Any value allowed.
- Repeat. To insure password entered correctly.
- User full name. Documentation.
- **Group** id:name: This is a drop down list that contains the available groups.
- Enabled: toggle switch that can disable or enable log-ins for users in the group.
- Demo: This **role** prevents this user from any update activity; other roles can limit which menus are available.

- Group Admin: This **role** gives this user special access to some functions -- editing common TRE sub-expressions for use in rubrics, for example. In future could be used to allow managing the groups own' users.
- Admin: This **role** allows this user to check the server status or see server log files.
- Super Admin: This **role** enables platform administration functions, like creating new groups.

1. Add User

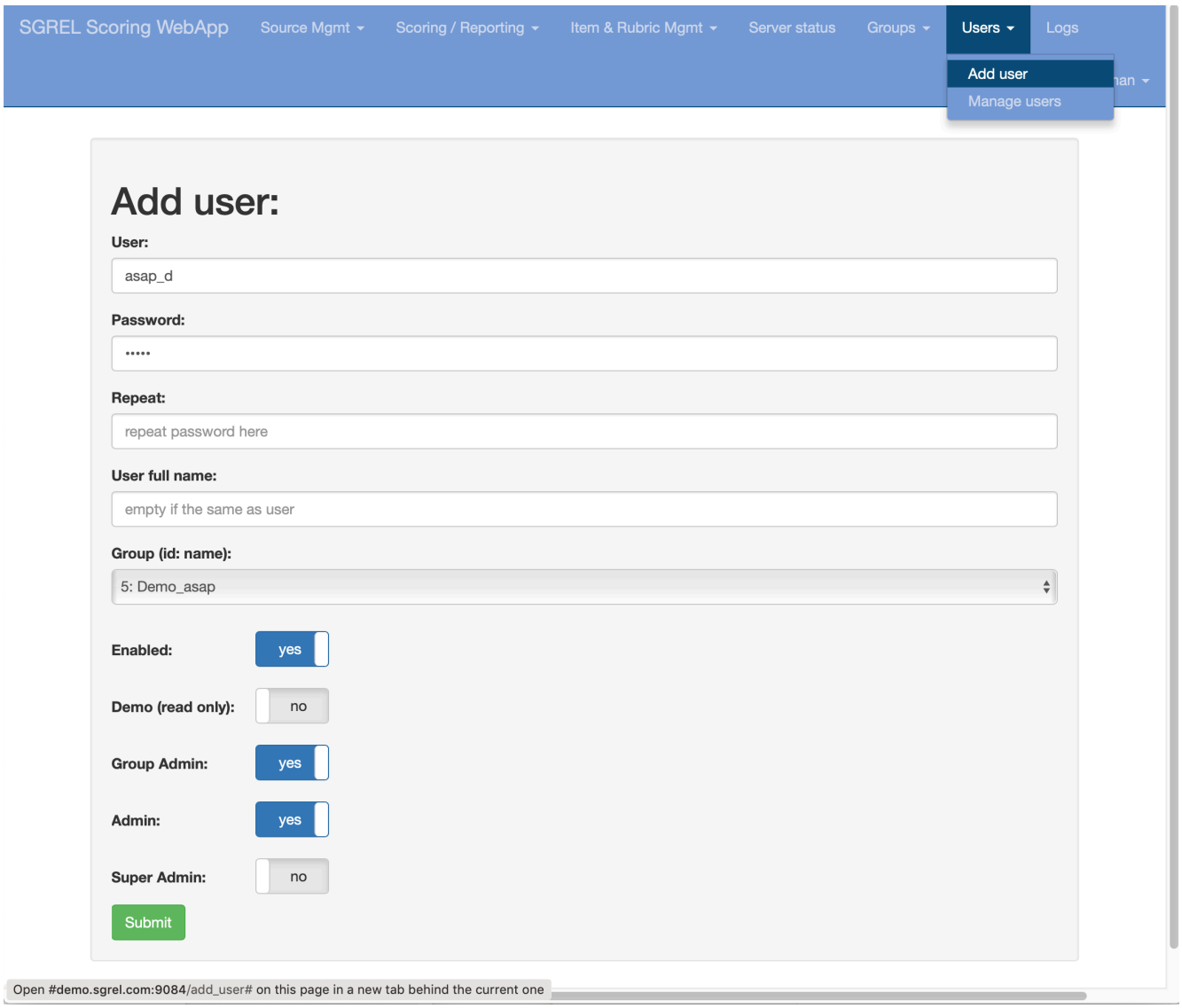


Figure II-5 Add User Page

Web application pages for a) managing users, b) deleting users and c) editing user information are shown in the following three pages. Note pending information is highlighted -- red for delete, blue for changes. No changes are actually made until the "save changes" button is clicked; reloading the page will throw away a pending change. These UI conventions are used across all of the SGREL Web App pages.

2. Manage Users

Note red button for delete, editable fields can be effected directly (change text, switch button position, select drop-down menu choice).

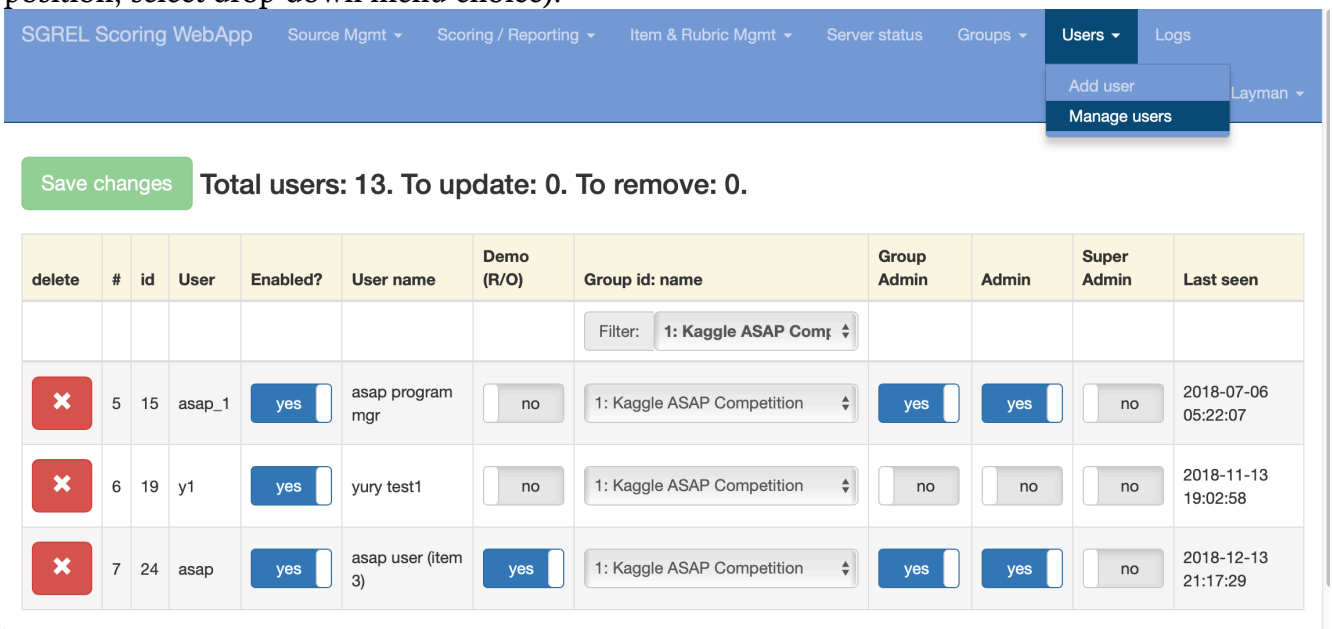


Figure II-6 Manage Users Page

Note "save changes" not yet enabled and no pending changes (zero to update or remove in heading).

3. Delete User

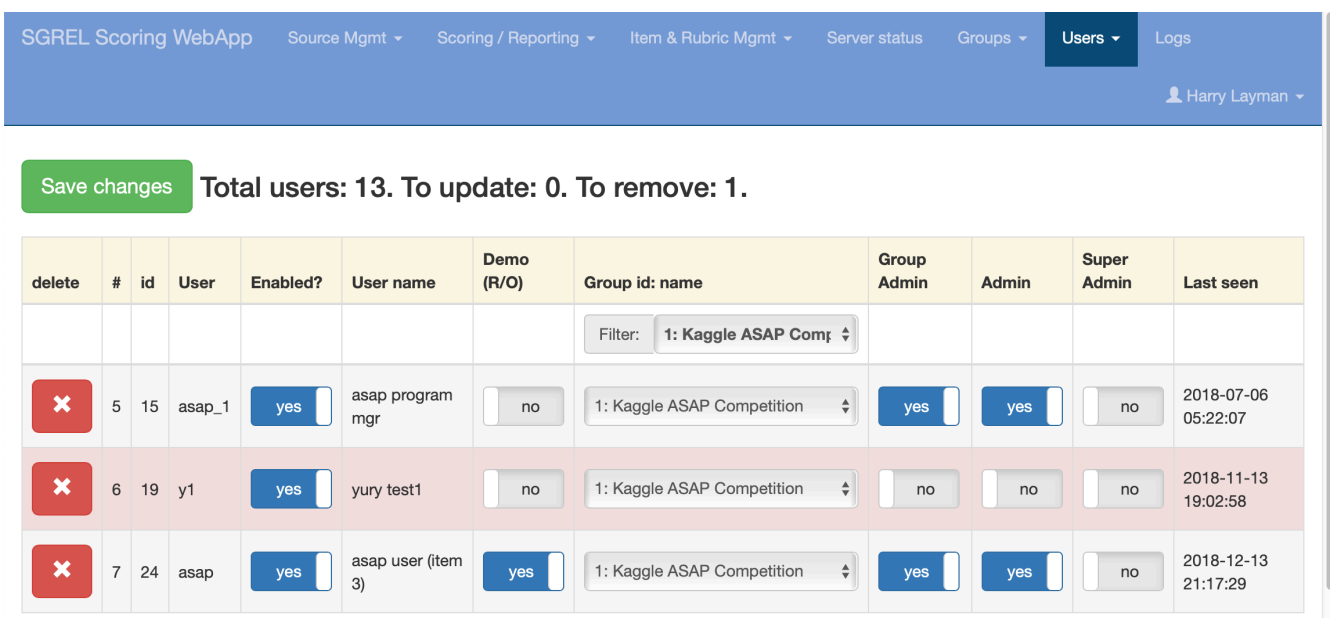


Figure II-7 Delete User Page

Clicking red button highlights **pending delete** row; message at top of page shows 1 pending removal. Save Changes button is enabled.

4. Edit User Information

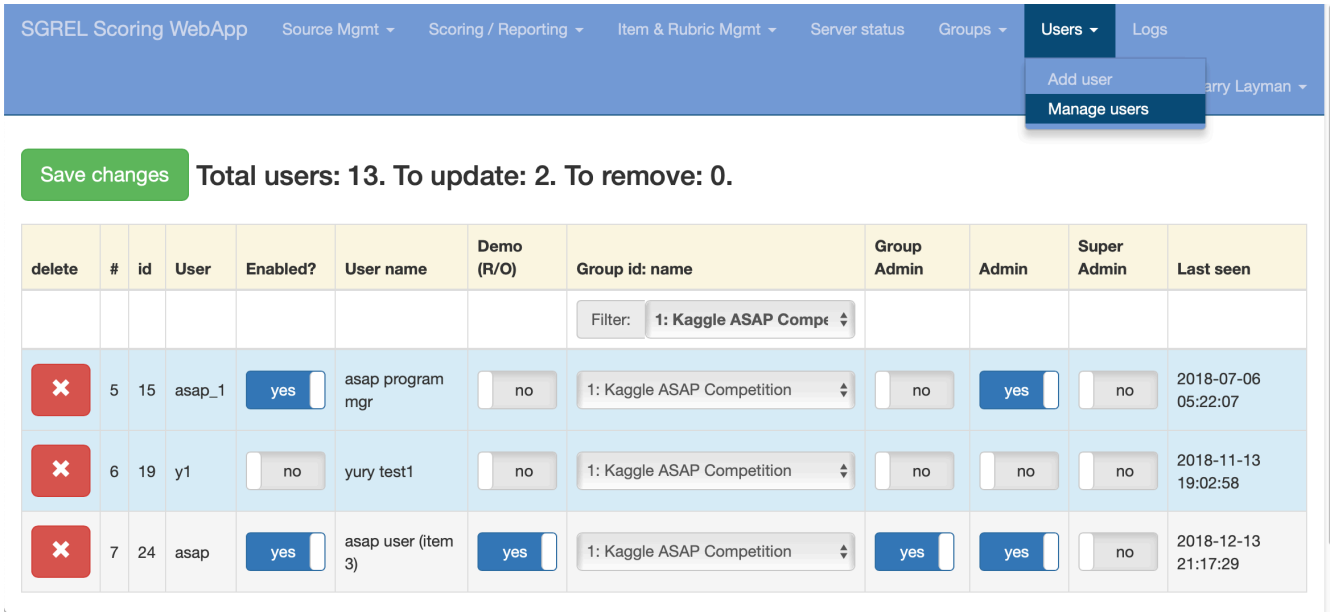


Figure II-8 Manage Users Page

Note pending update rows are blue. Count of updated rows in heading, **save changes** button enabled. Reload page to discard changes.

III. Administration

All users with the Admin role have access to the Server status page, which shows the current state of the SGREL web app platform, and whether scoring services are available. During periods of significant editing for scoring rules or other administrative activity, a user organization may want to prevent scoring by regular users until scoring rule or other updates are complete.

This page also provides information about the current state of the platform. The server status page is shown below. The switch next to "Processing status" can be used to disable scoring.

The following page shows log entries also available to admin users.

A. Server Status page

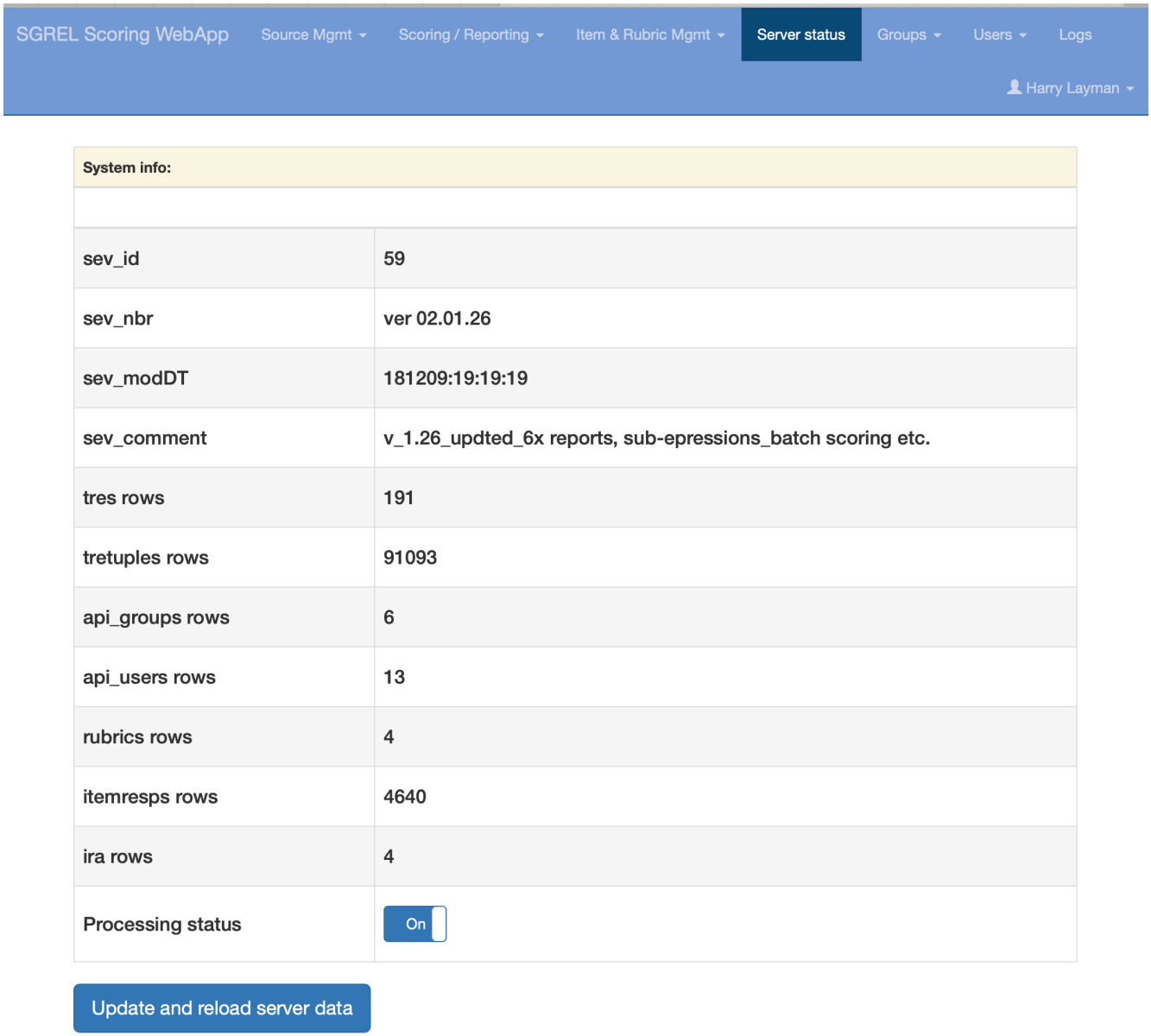


Figure III-1 Server Status Page

B. Log File page

SGREL Scoring WebApp | Source Mgmt | Scoring / Reporting | Item & Rubric Mgmt | Server status | Groups | Users | Logs | Harry Layman

Page scroll: Max rows: 20

log_id	log_date	log_dt	log_level_int	log_level_str	log_msg	log_user	log_request_method	log_remote_addr	log_request_url	log_url_name	log_api_function_name	log_reque
1181	Thu Dec 13 2018	12/13/2018, 2:05:25 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	get_server_status	200
1180	Thu Dec 13 2018	12/13/2018, 2:05:25 PM	20	info	Request	harry	GET	68.231.212.160	http://demo.sgrel.com:9084/server_status	server_status	null	200
1179	Thu Dec 13 2018	12/13/2018, 2:05:21 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	list_all_groups	200
1178	Thu Dec 13 2018	12/13/2018, 2:05:21 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	list_all_users	200
1177	Thu Dec 13 2018	12/13/2018, 2:05:20 PM	20	info	Request	harry	GET	68.231.212.160	http://demo.sgrel.com:9084/manage_users	manage_users	null	200
1176	Thu Dec 13 2018	12/13/2018, 1:54:23 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	list_all_groups	200
1175	Thu Dec 13 2018	12/13/2018, 1:54:23 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	list_all_users	200
1174	Thu Dec 13 2018	12/13/2018, 1:54:23 PM	20	info	Request	harry	GET	68.231.212.160	http://demo.sgrel.com:9084/manage_users	manage_users	null	200
1173	Thu Dec 13 2018	12/13/2018, 1:52:43 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	list_all_groups	200
1172	Thu Dec 13 2018	12/13/2018, 1:52:43 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	list_all_users	200
1171	Thu Dec 13 2018	12/13/2018, 1:52:43 PM	20	info	Request	harry	GET	68.231.212.160	http://demo.sgrel.com:9084/manage_users	manage_users	null	200
1170	Thu Dec 13 2018	12/13/2018, 1:50:07 PM	20	info	Request	harry	POST	68.231.212.160	http://demo.sgrel.com:9084/api	api_endpoint	add_user	200
1169	Thu Dec 13 2018	12/13/2018, 1:43:27 PM	20	info	Request	harry	GET	68.231.212.160	http://demo.sgrel.com:9084/add_user	add_user	null	200
1168	Thu Dec 13 2018	12/13/2018, 1:43:24 PM	20	info	Request	harry	GET	68.231.212.160	http://demo.sgrel.com:9084/	home_page	null	200
					User							

Figure III-2 Log File View Page

IV. Items and Rubrics

Scoring on the SGREL scoring platform involves scoring "**item responses**" to **items** with **rubrics**, both of which must be defined by platform **users**. Note that for some users, an "item response" could be a constructed response to an assessment item, while for others it could be a sales call, a customer service inquiry or an e-commerce transaction.

A. Items

Each kind of thing that is to be scored on the platform might be considered an "item type", and each specific instance of that item type is defined on the SGREL scoring platform as an "Item". Items refer to a single question for which responses are to be scored. To facilitate interchange of data with other systems (e.g. test delivery systems, test response analytics, etc.), a number of attributes are included for items which could identify them on other systems (e.g. located at URL, or long and short names); such fields are optional. Such fields are optional.

Item management, including adding of new items, takes place on the item management page shown below. On the item management page, the orange-yellow + button can be used to add a new item; the red x button can be used to delete an item, and the blue pencil button can be used to edit an item. IF a URL is entered for an item, the presence of a URL is noted by a green URL download button and can be used to download a remote PDF or other asset at the universal record locator location.

1. Manage Items

The basic item management page illustrating all these controls and the primary item attributes is shown below.

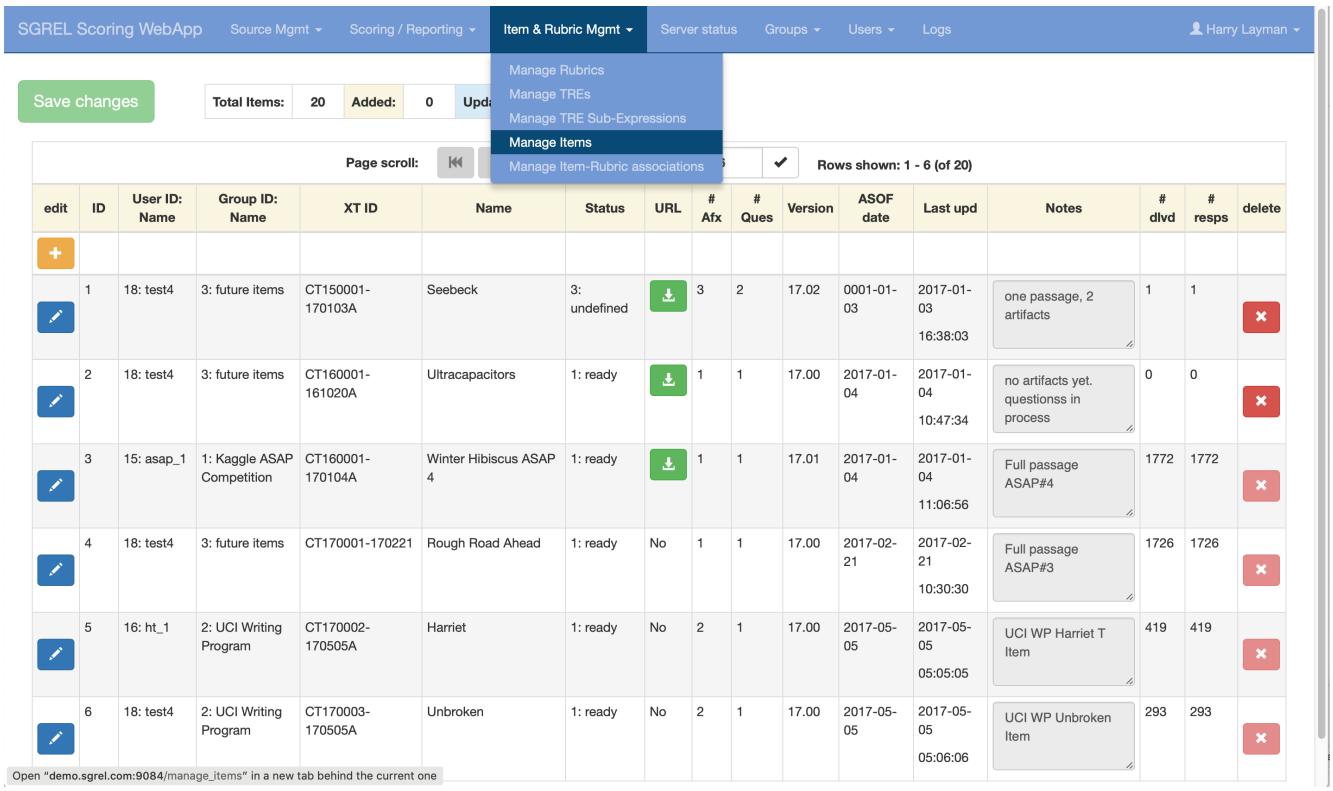


Figure IV-1 Manage Items Page

Clicking on the yellow + button brings up a modal "add item" window as shown below.

2. Add Item

The ad item popover is shown in the figure below. The attributes required include:

- **Group** ID and **User** ID are specified for the logged user adding the new operation; they cannot be set by the user explicitly.
- **Item** name should be descriptive; XT ID can specify an external identifier for the item.
- Status for the item can be set to: start, in review, ready and retired.
- # artifacts can indicate the number of separate files required to deliver the item
- # questions reflect the number of distinct questions ask. This is normally 1.

- Total delivered and total responses are used for statistical purposes and not required. The # delivered can tally counts for payment of copyright fees if applicable.
- At URL swtich -- if set to yes, can provide a URL that links to a source document for the item.
- Version number is information only. Same for "as of date" (default set to today).
- Last update data and time are also defaulted to "now" and optional.
- Notes is descriptive only.

As you can see, items can be defined with almost no data This primary use for this data in the system at present is for inclusion on score reports (e.g. to identify what item is being scored).

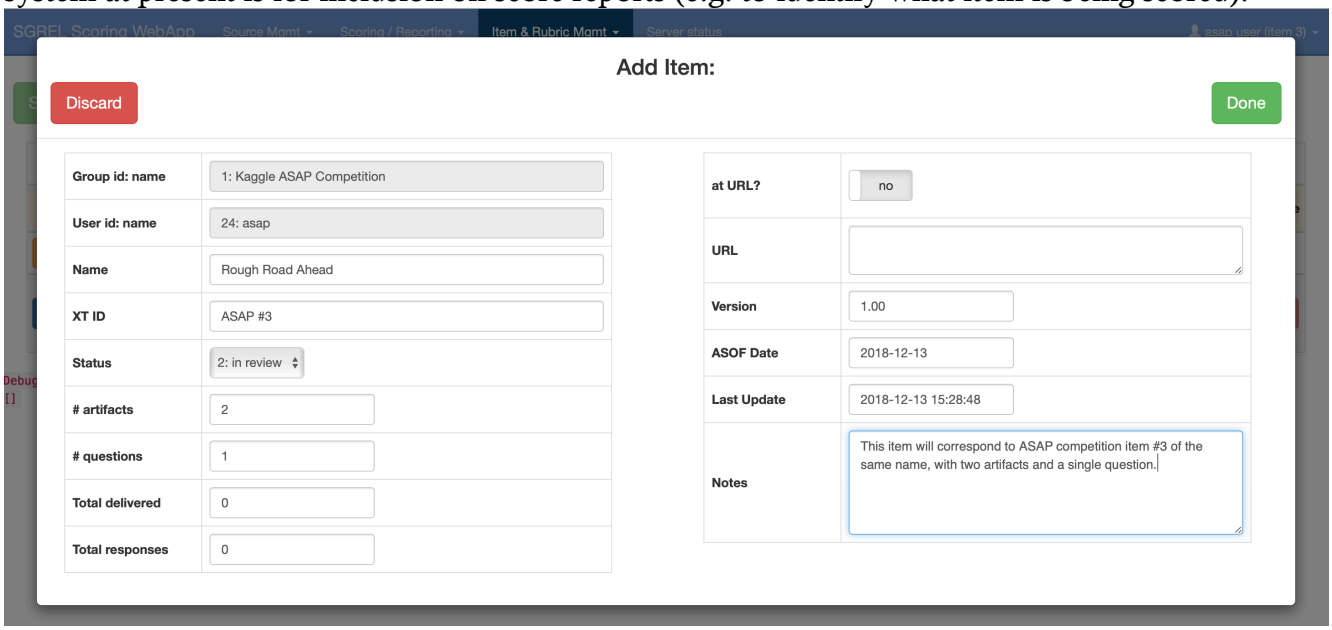


Figure IV-2 Add Item Page

3. Edit Item

As shown below, the item edit modal window is the same as the add item window.

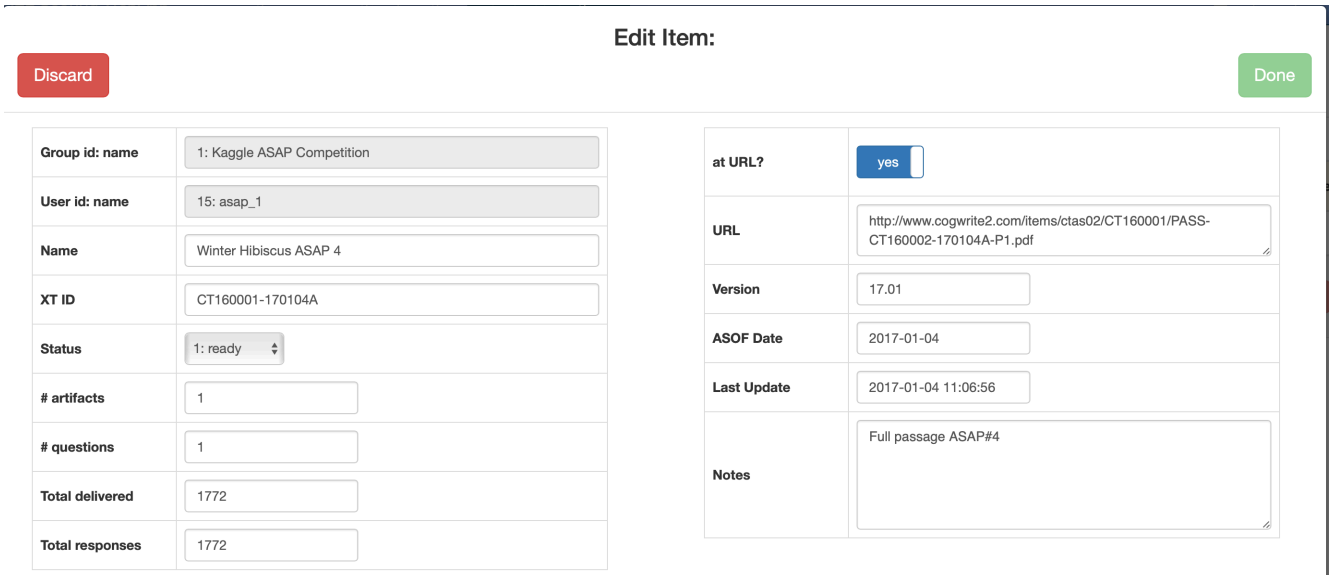


Figure IV-3 Edit Item Page

B. Rubrics

Rubrics specify the scoring rules for items scored on the SGREL scoring platform by the automated scoring agent. Rubrics have three Parts: the rubric metadata; the target response elements and the scoring formula. Each of these are described below. The second component of the rubric, the TREs, drives the heart of the automated scoring process and is covered more fully in the following chapter.

1. Rubric Metadata

Access to the manage rubrics menu is from the Item & Rubric Mgmt page as shown here:

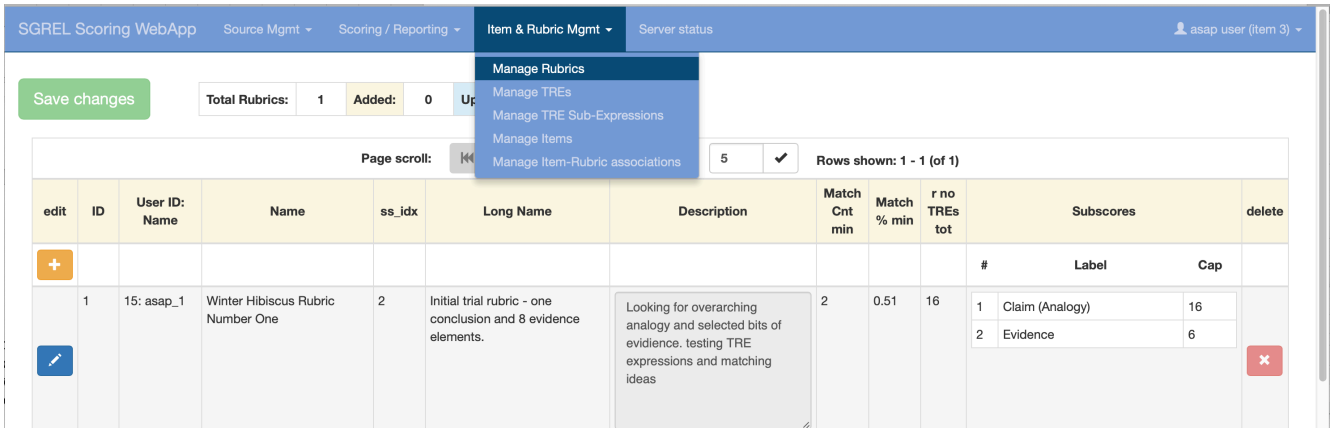


Figure IV-4 Manage Rubrics Menu Page

The full "manage rubrics" capability -- to **add**, **edit** or **delete** rubrics -- is provided from this page. As across the system, new rubrics are added with the yellow + button; they are edited with a similar popover window from the blue-pencil button, and they are deleted by using the red x button. As before, changes are signaled by highlighting, and not made until after the enabled "save changes" green button is active and then clicked.

All primary **rubric** metadata are shown on this display. In addition to the **rubric** name and ID number, there is a rubric description and up to seven pairs of sub-score Label values (text) and sub-score Cap values (small integers). In the example above we see this rubric addresses two subscores described as "Claim (analogy)" and "Evidence". These two subscores are shown to have maximum values of 16 and 6.

The image below shows seven rubrics, including rubrics defined with two, three and seven sub-scores.

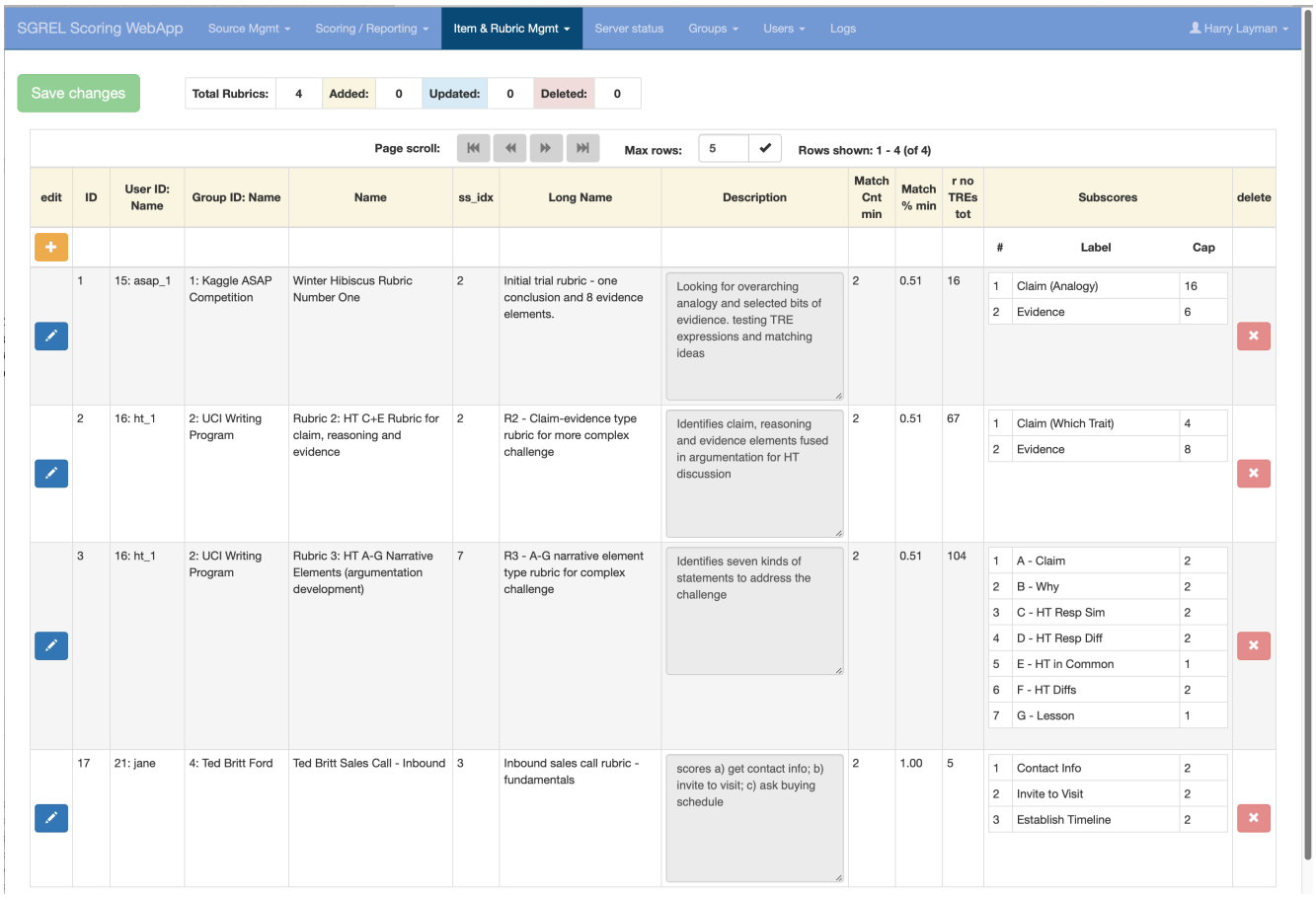


Figure IV-5 Manage Rubrics Page

The number of sub-scores is indicated in the center column labeled "ss_idx". Note row four is 3 and row 3 is 7. The columns for Match Count Min and Match % Min are not used at this time. The column just before the "Subscores" column on the right side of the page shows the computed number of TREs defined for a rubric; this topic is covered in the TRE chapter which follows this one.

2. Add / Edit Rubric Popover

Adding a new rubric and specifying rubric metadata and the final scaled score formula (if any), are done using the modal popover window that opens when the yellow + button is checked on the manage rubrics page.

Editing an existing **rubric** and its metadata and **final scaled score** formula is done on a modal popover window with the same layout and fields as used when adding a new rubric.

These identical pages are described below.

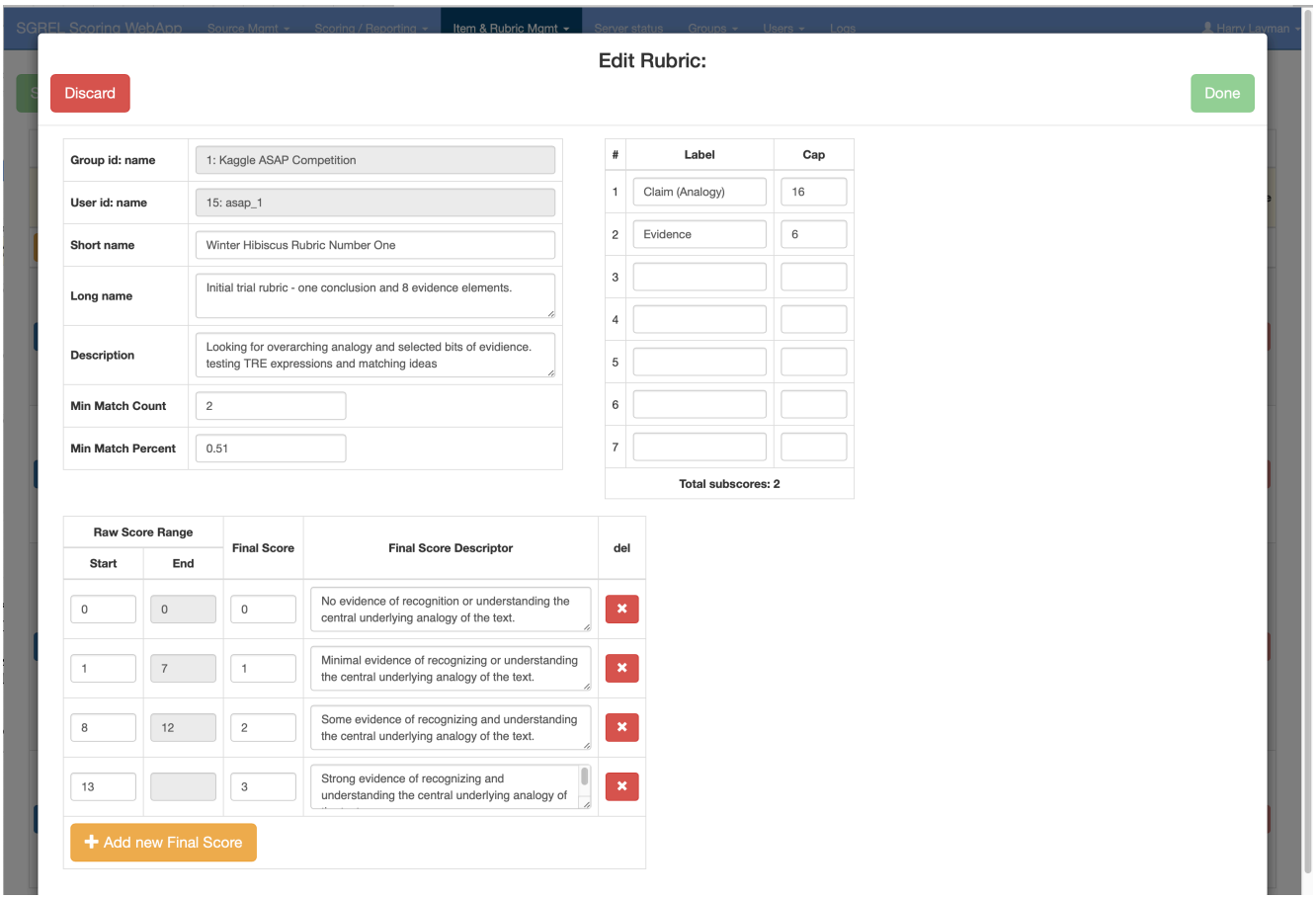


Figure IV-6 Add / Edit Rubric Page

The primary metadata here that can be entered or edited includes the same short and long name fields seen on the manage rubrics page, in addition to the up-to-7 labels and cap values for sub-scores. What is new here is a set of optional entries for a final score value and final score descriptor based on a "raw score range".

The lower quadrant of fields for the "final score descriptor" information (which in this response web app design could be on the far right of a very wide display) includes one or more rows that must define a sequential, increasing set of "bands" or score ranges and provide some text for what a score in that band signifies. [See following figure.] This could be as simple as "Overall score of 1 on a 3 point scale." Or something more complex as in the example above, which provide an explanation of score zero as "No evidence of recognition or understanding the central underlying analogy of the text." on up to, for a value of 3, "Strong evidence of recognizing and understanding the central underlying analogy of the text."

As covered in more depth in the next chapter, scoring on this platform uses rules, constructed using a rubric expression language, that assign score points (small integer values) for concepts found in the text that are specified in TREs. Evidence, claim statements, reasoning or other "target elements" are identified via flexible expressions, and if they are found, the scoring engine is told to assign a specific number of points for these "match candidates". Once all match candidates have been identified, TRE attributes such as group numbers and sub-score index values dictate how to combine and filter these match candidates across TRE groups and sub-

scores according to well defined rules, creating a set of "best matches" which are then used to calculate the final "sub-scores" and final "raw score" for an item response.
The final score descriptor is selected based on the where the final raw score falls in the bands defined by the set of raw score ranges, and the descriptor associated with that final raw score number is used on score reports (as can be sub-score values and indeed every TRE match included in the final score.)

A close up view of the portion of this window used to specify and edit the final score ranges and descriptors for a rubric is shown in the figure below.

Raw Score Range		Final Score	Final Score Descriptor	del
Start	End			
0	0	0	No evidence of recognition or understanding the central underlying analogy of the text.	✖
1	7	1	Minimal evidence of recognizing or understanding the central underlying analogy of the text.	✖
8	12	2	Some evidence of recognizing and understanding the central underlying analogy of the text.	✖
13		3	Strong evidence of recognizing and understanding the central underlying analogy of the text.	✖

✚ Add new Final Score

Figure IV-7 Edit Rubric Final Scale Score Descriptors

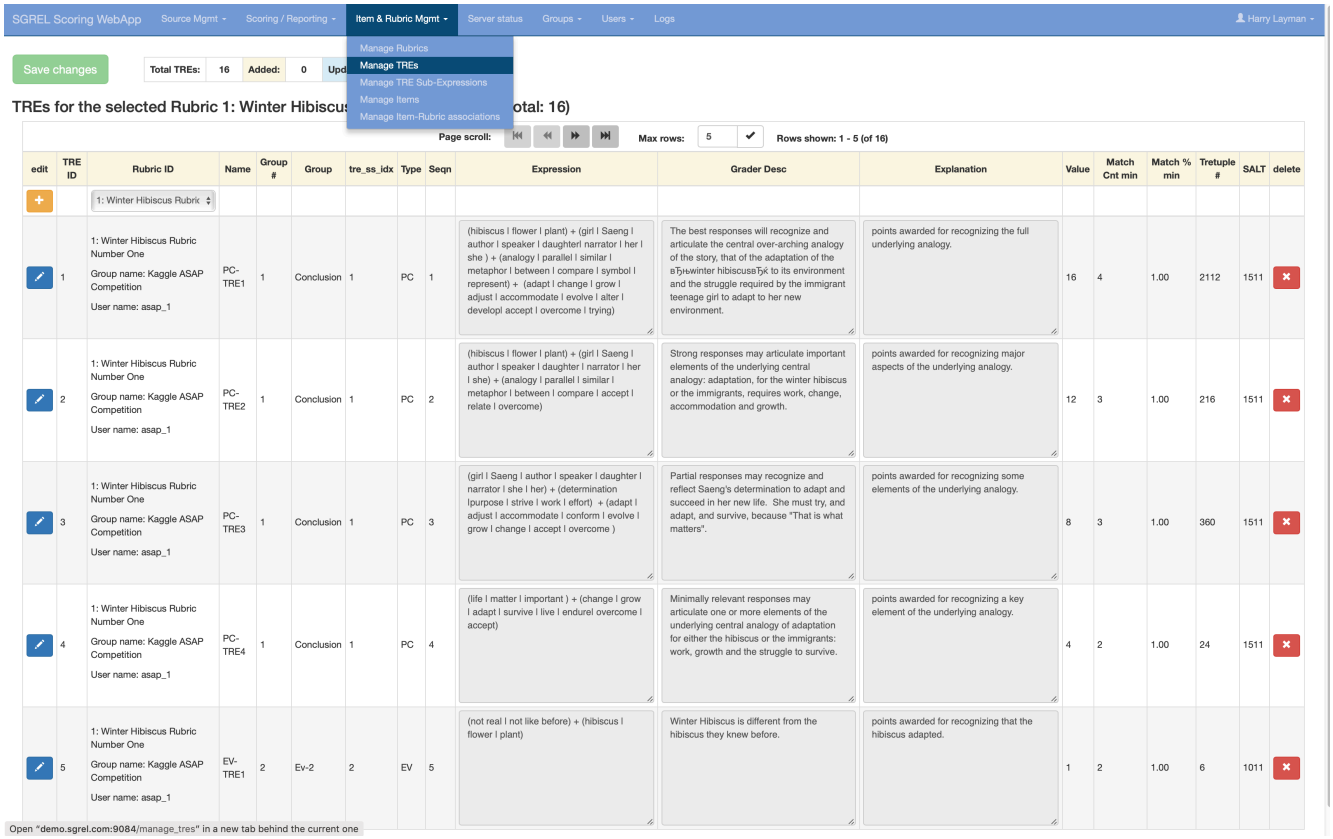
As before, the red x buttons allow range elements to be deleted and the orange + button allows new ranges to be defined.

Note that different standard report formats can report sub-scores and raw total scores, or can focus on the scaled final score values provided by these ranges, depending on the report selected for reporting for a given set of item responses to a particular item-rubric combination.

3. Rubric Target Response Elements

Target response elements are discussed in detail in the next chapter. Examples of a "manage TRE" page and an "add TRE" window are included below to provide an overview of the variety and complexity of the data used to specify a TRE.

Manage TREs presents a TREs for a selected Rubric (or all rubrics) in a scrollable grid, where individual TREs can be deleted (red x), selected for edit (blue pencil), or new TREs can be added (yellow x) with buttons as for other entities.



edit	TRE ID	Rubric ID	Name	Group #	Group	tre_ss_idx	Type	Seqn	Expression	Grader Desc	Explanation	Value	Match Cnt min	Match % min	Tretuple #	SALT	delete
+		1: Winter Hibiscus Rubric															
	1	1: Winter Hibiscus Rubric Number One Group name: Kaggle ASAP Competition User name: asap_1	PC-TRE1	1	Conclusion	1	PC	1	(hibiscus I flower I plant) + (girl I Saeng I author I speaker I daughterI narrator I her I she) + (analogy I parallel I similar I metaphor I between I compare I symbol I represent) + (adapt I change I grow I adjust I accommodate I evolve I alter I developl accept I overcome I trying)	The best responses will recognize and articulate the central over-arching analogy of the story, that of the adaptation of the вЂ¦winter hibiscusвЂ¦ to its environment and the struggle required by the immigrant teenage girl to adapt to her new environment.	points awarded for recognizing the full underlying analogy.	16	4	1.00	2112	1511	
	2	1: Winter Hibiscus Rubric Number One Group name: Kaggle ASAP Competition User name: asap_1	PC-TRE2	1	Conclusion	1	PC	2	(hibiscus I flower I plant) + (girl I Saeng I author I speaker I daughter I narrator I her I she) + (analogy I parallel I similar I metaphor I between I compare I accept I relate I overcome)	Strong responses may articulate important elements of the underlying central analogy: adaptation, for the winter hibiscus or the immigrants, requires work, change, accommodation and growth.	points awarded for recognizing major aspects of the underlying analogy.	12	3	1.00	216	1511	
	3	1: Winter Hibiscus Rubric Number One Group name: Kaggle ASAP Competition User name: asap_1	PC-TRE3	1	Conclusion	1	PC	3	(girl I Saeng I author I speaker I daughter I narrator I she I her) + (determination Ipurpose I strive I work I effort) + (adapt I adjust I accommodate I conform I evolve I grow I change I accept I overcome)	Partial responses may recognize and reflect Saeng's determination to adapt and succeed in her new life. She must try, and adapt, and survive, because "That is what matters".	points awarded for recognizing some elements of the underlying analogy.	8	3	1.00	360	1511	
	4	1: Winter Hibiscus Rubric Number One Group name: Kaggle ASAP Competition User name: asap_1	PC-TRE4	1	Conclusion	1	PC	4	(life I matter I important) + (change I grow I adapt I survive I live I endurel overcome I accept)	Minimally relevant responses may articulate one or more elements of the underlying central analogy of adaptation for either the hibiscus or the immigrants: work, growth and the struggle to survive.	points awarded for recognizing a key element of the underlying analogy.	4	2	1.00	24	1511	
	5	1: Winter Hibiscus Rubric Number One Group name: Kaggle ASAP Competition User name: asap_1	EV-TRE1	2	Ev-2	2	EV	5	(not real I not like before) + (hibiscus I flower I plant)	Winter Hibiscus is different from the hibiscus they knew before.	points awarded for recognizing that the hibiscus adapted.	1	2	1.00	6	1011	

Figure IV-8 Manage TRE Page Summary

The manage TREs page presents existing TREs based on the drop-down selection list for a specific rubric or for all rubrics. These first five TREs are alternative expressions designed to match all or some of the elements of an underlying analogy that a proper claim statement in response to this item would cite. The attributes on each row include the TRE expression (in the left-most grey column), followed by a description for the grader and an explanation for the student, appropriate for the presence or absence of the target concept this scoring element is designed to represent. Other attributes include computed elements such as the number of tretuples generated by an existing TRE expression, and attributes set by the rubric author (e.g. tre group number, used to group together TREs from which only the best, highest-valued match should be selected. Point value (immediately right of the gray columns), and match count minimum, the following column, to the right of "point value", which specifies the number of words in the given expression that much be matched for a TRE to be considered as a potential "match candidate" during scoring.

The elements presented below are collected for new TREs and are editable by existing TREs in a popover window with the layout presented. Some notes on usage follow -- but again, this topic is covered in step-by-step detail in the TRE chapter which follows.

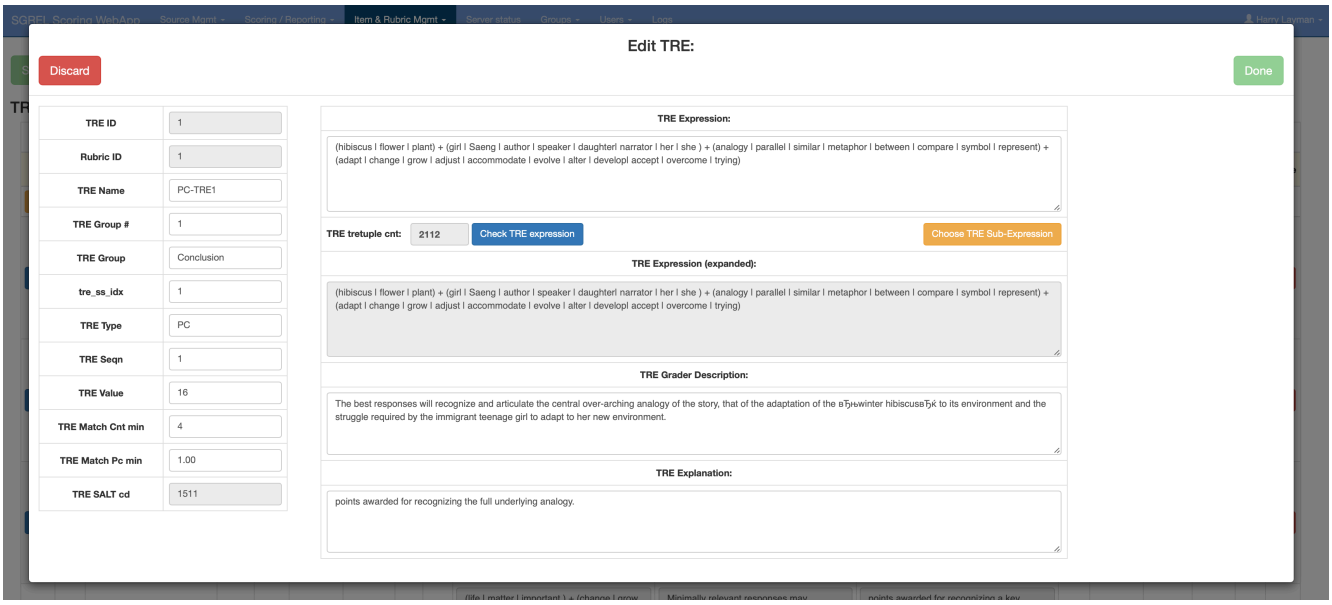


Figure IV-9 Edit TRE Page Summary

This "Edit TRE" modal window includes a **blue** button that checks the syntax of the TRE expression after it has been entered or edited, and the TRE cannot be saved until the edit for proper syntax is successful. When a TRE expression is edited, the blue check button is enabled. When successfully edited, the TRE tretuple count value is updated (giving an idea of the complexity of the expression) and the "Done" button is enabled.

The yellow button on this phrase allows the TRE author to substitute into the TRE expression and TRE Subexpression, something that can be defined elsewhere (again, see next chapter). As an example, the entire phrase "<Saeng>" could be defined as a TRE subexpression "(Saeng | author | speaker | daughter | narrator)" simplifying its use in multiple TREs.

4. Rubric Scoring Formula

a) Implicit Scoring Formula

The implicit scoring formula for the SGREL scoring platform is to a) sum all the point values of all the "best match" TREs that survived the filter and aggregation processing based upon "TRE group number" designations for each TRE; based on meeting match count minimums; and based on prioritizing on point value (firstly) and then on "match percent" (only after match count minimums have been met), within each sub-score category (ss_idx value specified in the TREs) and subject to sub-score cap or maximum values.

This process allows a "raw total score" to be calculated based on the TRE-assigned point values. Score reports may value these "sub-score" values by themselves, or on conjunction with (or instead using) the final score value described below.

b) Final Scoring Formula

The final score value is simply a user-specified set of point values defined by a set of score bands assigned to the final raw score value from the initial pass at SGREL scoring. In this way, detailed evidence and claim values can be assigned for different sorts of concepts included in an answer, and scales can be set for TRE point values in different subscores to achieve whatever relative weight might be desired in a final raw score calculation. But in addition to this flexibility, the final score value ranges can then be portioned as needed into useful performance bans based on actual experience with an assessment, relative difficulty compared with other assessments, relative difficulty of individual sub-score constructs, and other factors.

V. Target Response Elements

A. TRE Groups

In some instances, a scoring formula may want to award different point scores for different degrees of correctness -- for example, to give partial credit for a partially correct response, or more credit for the best response and somewhat less credit for a lesser response.

In the case of more evidence gets more points, say, or more examples get more credit, a rubric may want to simply combine all the credits found into a sub-score total. But where there is a desire to distinguish between partial and full credit, a "TRE group" in the form of a single "TRE group number" can be assigned to two or more TREs, each with their own TRE expression. Matches by such a group of TREs are first compared to each other, and multiple qualifying matches (number of words matched and % match exceed thresholds) from multiple TREs are rank ordered, first by point value, then by strength-of match as expressed in the % match value, and the "best" of the matches is selected to assign the point value to their common sub-score.

An example:

For the rubric with rubric id 1 in the demo database for the asap item, item id 3, there are six TREs with tre_ss_idx = 1 (sub score 1, for Claim) and tre_group_no = 1, meaning that only one of those six TREs will be counted toward the sub-score for strength of claim. Further, there point values for the six TREs are 16, 12, 8 and 4 (with two having 8 and two having 16 point values). The higher point values have longer expressions and higher match count minimums that require a more robust statement of the underlying analogy that is central the best answer. These at the most robust require reference to a) Saeng; b) the Hibiscus; c) comparison or relation between the two; and d) adaptation. The lowest scores accept observations about either Saeng or the hibiscus, and some notion of adaptation, growth or struggle.

This rubric has six different evidence TRE group numbers; and some phrases -- like ways in which the flower in the new country is like the one back home -- were designed to accept any of three examples of similarity expressed in the short story. All have the same TRE group number, so at most one point is awarded for citing evidence of similarity; other evidentiary points are similarly supported by multiple TREs but limit the credit for a specific kind of evidence to a single point as stated in the English language rubric for the item.

Note well: future versions of the UI (manage_tres page) should allow not just selection of TRE by rubric, but sorting by tre_group_no (or other attributes), and filtering + sorting by sub-score (tre_ss_idx). Near term, changing the sort to use the tre_seqn field, not the tre_id, will allow user more control and ability to see related TREs on the same page. Also, we need a tre report -- or simple "export tre to xlsx" (currently done with SQL).

B. TRE sub-expressions

In developing the TRE expressions for identifying concepts required to describe the analogy in the story for the Winter Hibiscus item, item 3 in the demo database, many TREs require common sub-expressions -- for identifying Saeng, or the Hibiscus, or the idea of an analogy, to cite three examples.

This is facilitated by the ability to define a subexpression -- say "Saeng", as "(Saeng | author | speaker | daughter | narrator)", and then use <Saeng> in the TRE expression composition window.

TRE subexpressions, accessible by fellow group members, are composed and edited in the manage TRE Subexpressions page, under Item & Rubric Management of the main menu.

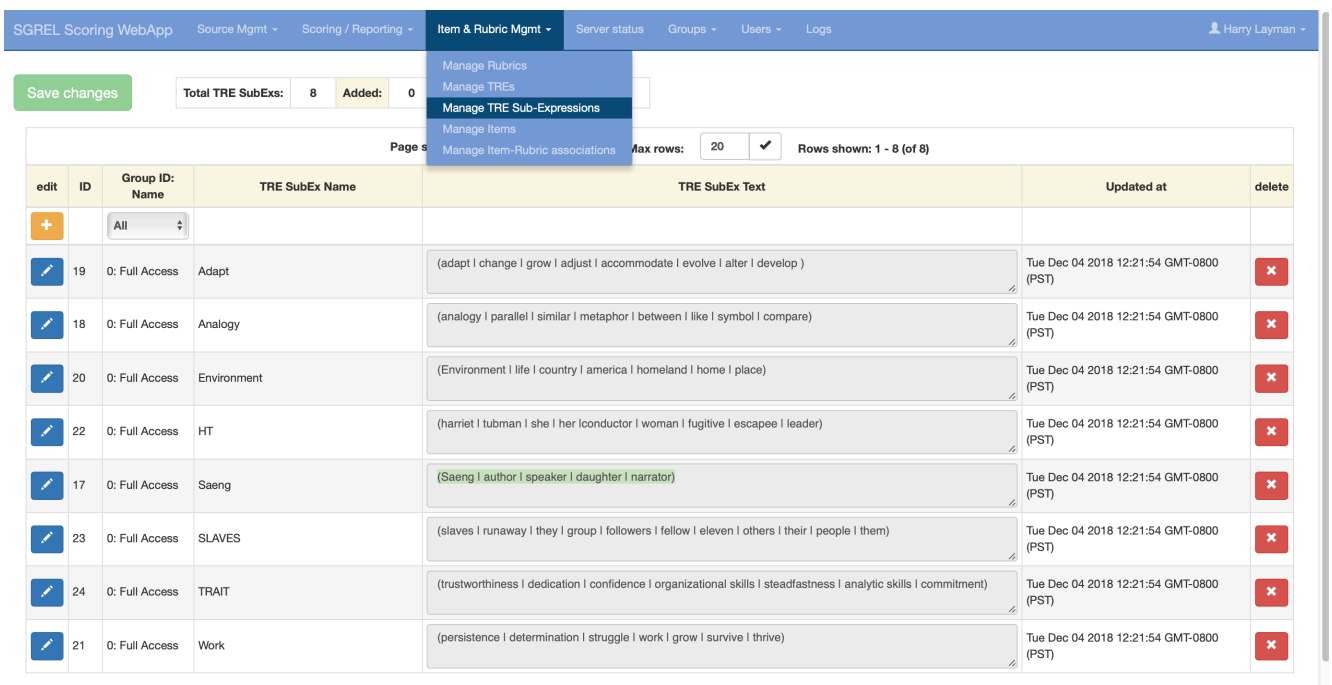


Figure V-1 Manage TRE Subexpression Menu

As with other object management pages, this page allows TRE Subexpressions to be added (with the yellow + button), edited (with the blue pencil buttons) and deleted with the red x buttons. When edit changes are pending, edited elements are highlighted in blue and pending deletes in red. New rows are also highlighted, and the Save Changes button is enabled, as shown below.

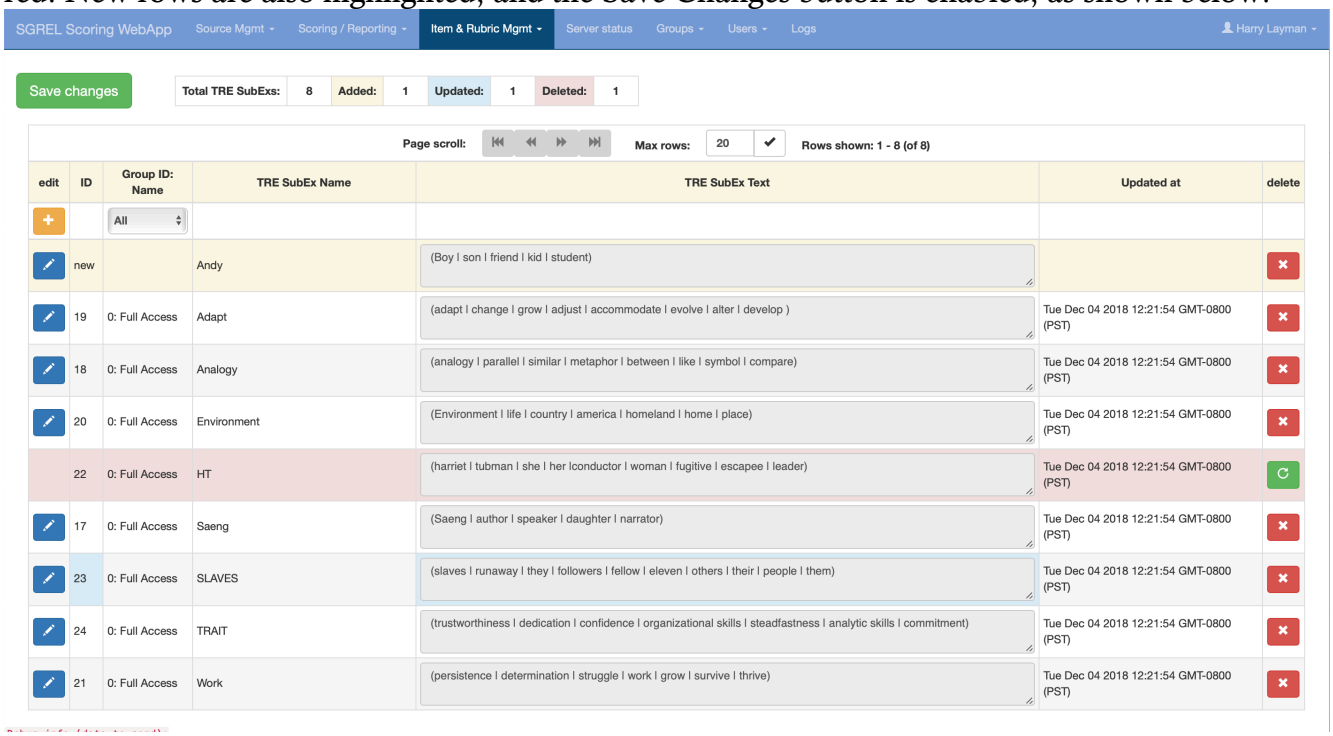


Figure V-2 Manage TRE Subexpression - Pending Changes

In the figure above there are above, there is a delete pending for the row mostly highlighted in red, and a change pending for row 23 where the row ID column is tinted blue. The row at the top, highlighted with yellow, and with no value (yet) in the "updated at" column, is a new addition. The top left "Save Changes" button will save all three kinds of changes and refresh the list of tre subexpressions.

VI. Item – Response Associations

Scoring items requires creating items and rubrics, and then associating them with each other. Such an association is an "item-response association" or IRA. By creating the ira as a relationship, we can score item responses with different rubrics that measuring different things. For the original research project, I scored a set of responses with both a "claim + evidence" rubric for argumentation, and then could scored the same responses with a rubric designed to identify seven kinds of narrative elements requested by the item instructions -- things like "lessons we can learn from …" and "how is HT like the followers… and how is she different", etc. called a "seven narrative elements" rubric.

Once an item - rubric association is defined, when scoring a response to an item, the user is given a choice of available rubrics to use or scoring -- that is, if more than one is available. Otherwise, the only available rubric available is assumed.

Scoring is the first stage at which rubric - item associations come into play. When reviewing scoring and reporting capabilities, keep in mind the underlying IRA required for scoring to be done.

IRAs are managed with a “Manage Item - Response Associations” menu choice with the usual grid organization allowing new IRAs to be created or existing IRAs to be edited or deleted, as shown below.

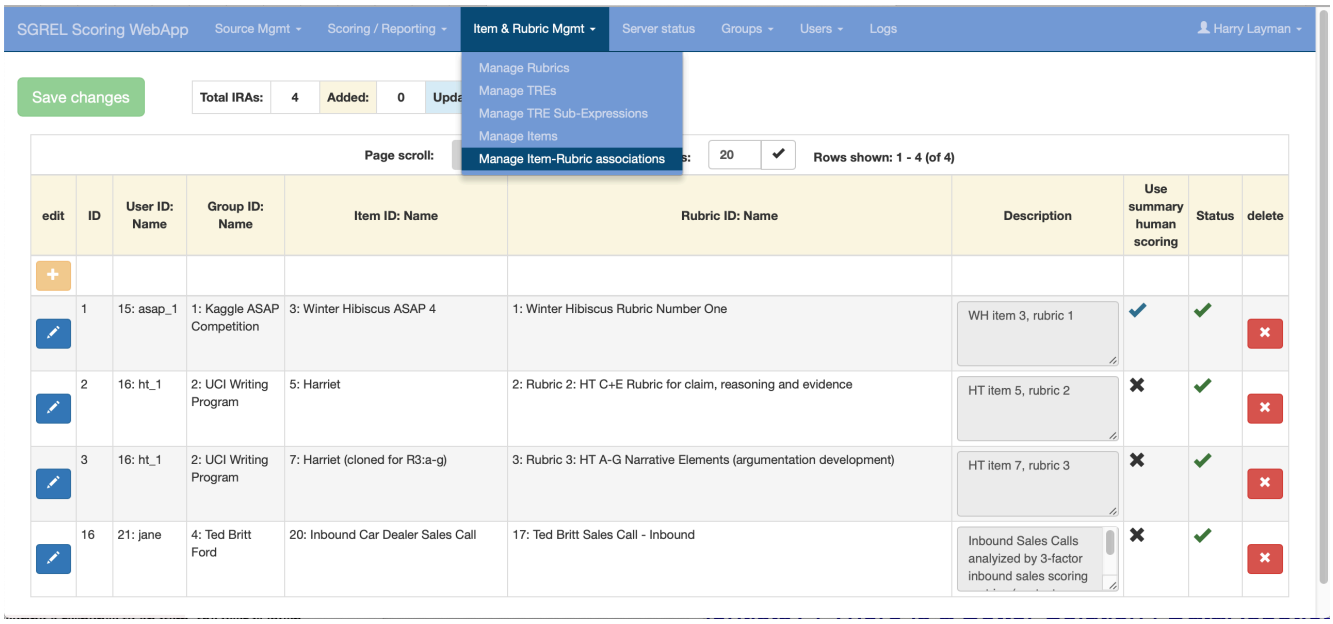


Figure VI-1 Manage Item-Rubric Associations

Creating a new IRA allows the user to select the item from one dropdown list and the rubric from another, as shown below (after pressing the yellow x button).

To add new IRAs, a user must have available, un-associated items and rubrics to assign to each other. In the figure below there is one available item and two available rubrics. Both rubrics are listed on the right; to add an association, select one of the two rubrics and click “Done”. (the second is selected below). Back at the manage iras page, you must press “Save Changes” or nothing will be changed.

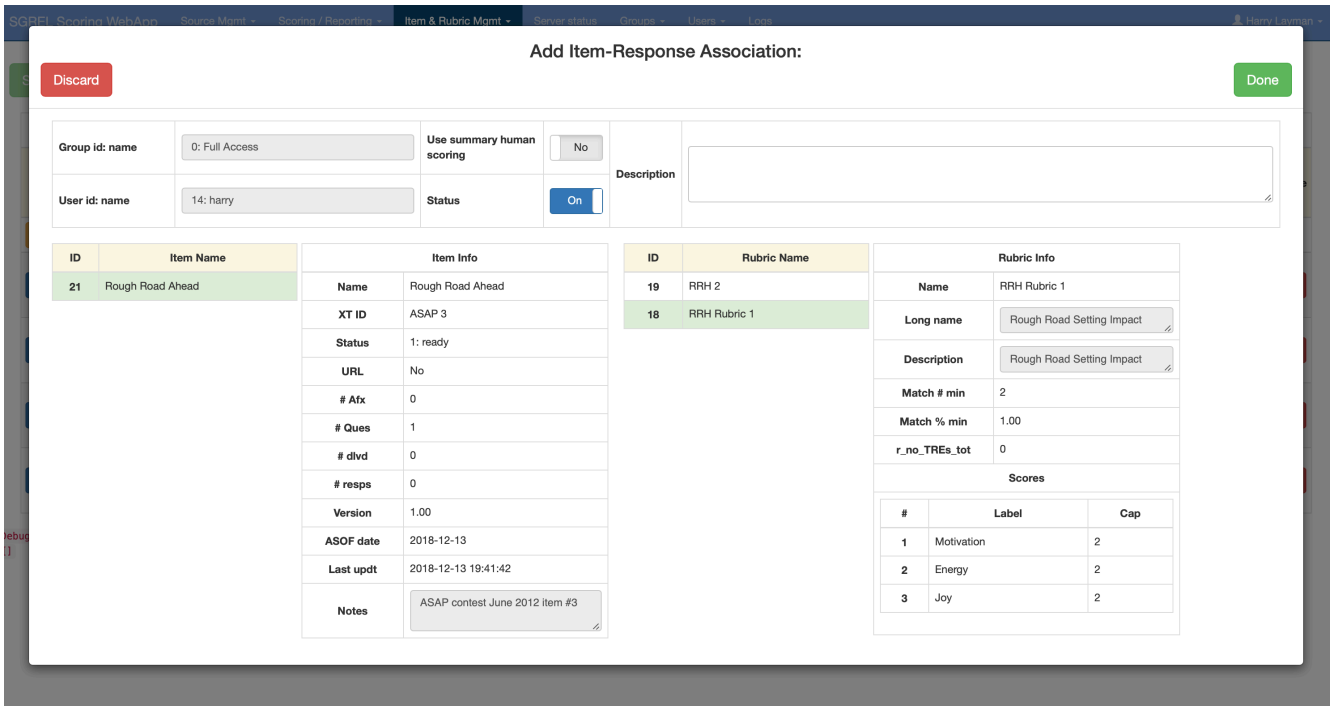


Figure VI-2 Add Item-Rubric Associations

VII. Upload and Process Documents, Calls, or Item Responses

Once items are defined and rubrics have been created, item responses can be scored.

First, we will walk through pasting in and scoring an item response right on the web page. Next we will load a document from the disk to score.

A. Single item loading, score and reporting

1. Single item loading

First we will navigate to the upload and process a document page. Second we will directly enter and score an item response -- or two -- to demonstrate the scoring system and one of the existing rubrics in the demo database.

From the main menu, navigate to “Source Mgmt -> Upload and process a document” as shown in the next figure. In the text file name, enter something like “Demo Response 181212 01 Harrys free form text.txt”.

In the enter text here box, enter “The author concluded the story as she did to underscore the analogy between the winter hibiscus in the title and Saeng, the immigrant girl would also adapt and survive in her new found home. The flower in her new country had petals and the fragrance of the flower from her home. Despite not being as pretty as the Hibiscus in Vietnam, the smell told her it was the same flower. Somehow it could survive here in the cold winters, and she would find a way to survive as well.”

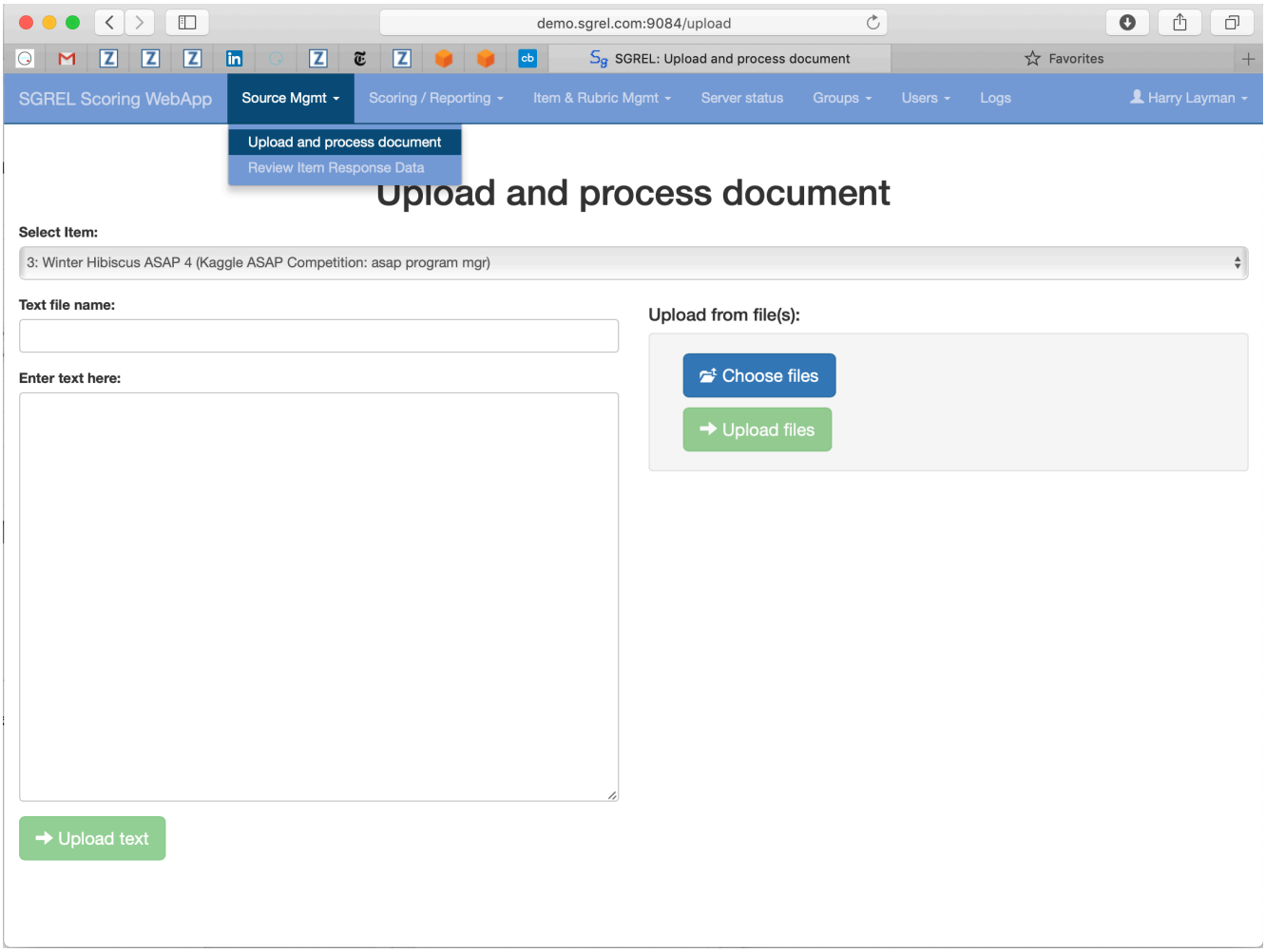


Figure VII-1 Load and Process (a Single) Document Page

With the page now containing your answer such as shown below and remember - select item 3 (in this example, or some item) from the drop down! As of version 1.28, you can also associate a loaded document with a specific item response author / creator (irac), which might be a student, a CSR or whoever generated the actual speech or text you are scoring (this is an optional association). In academic applications, item responses are associated with students / candidates, whereas in training / coaching and call center environments the calls being scored are from CSRs or other employee types.

Keep in mind that in production systems, artifacts to score will be loaded in a bulky, typically in a batch mode, perhaps pulling data from an IVR, CRM or other sales force management / call center related technology component.

The "upload" page, or the page which allows an individual artifact to be uploaded into the system for scoring, is shown in the next figure. Processing of that artifact (whether auto-transcribing audio to text and then parsing that into a conversation, or uploading a docx or text file directly -- or even pasting in text from a web form or other source) results in the creation of an scorable

artifact (or "ir" -- item response, in the parlance of this platform, or call or whatever) and is assigned a unique id and other attributes.

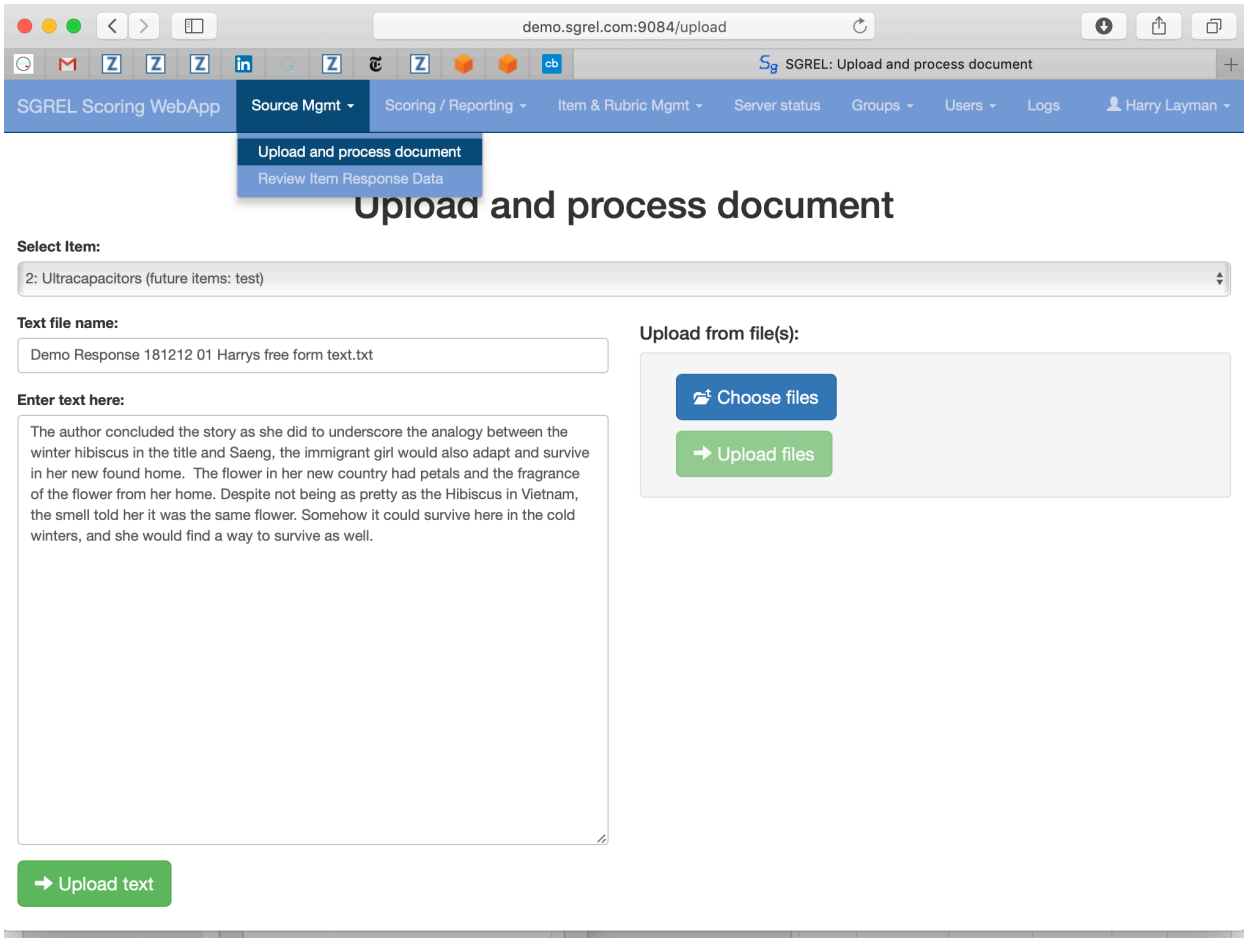


Figure VII-2 - Text field Entry on Upload Page

The result should be something like this on completion:

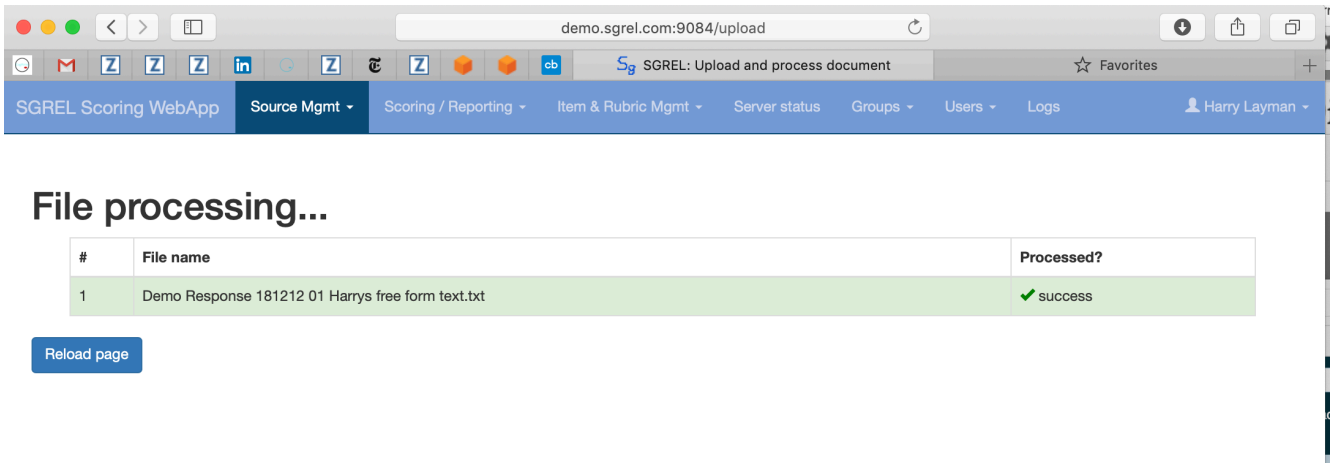


Figure VII-3 Upload File / Text Processing Complete

2. Single Item Scoring

Reloading the page will allow loading or entering another item to score. Instead of using the blue botton, navigate to the item response review (ir_review) page under Source Mgmt ->Review Item Response Data, and select "3:Winter Hibiscus ASAP" from the selection box under the column headed with "Item id" as shown in the next figure (if you are following the example; point is you selected the "item" that described the kind of document you wanted to score, and on the review page you again want to select it from documents that are an instance of that document you defined earlier.

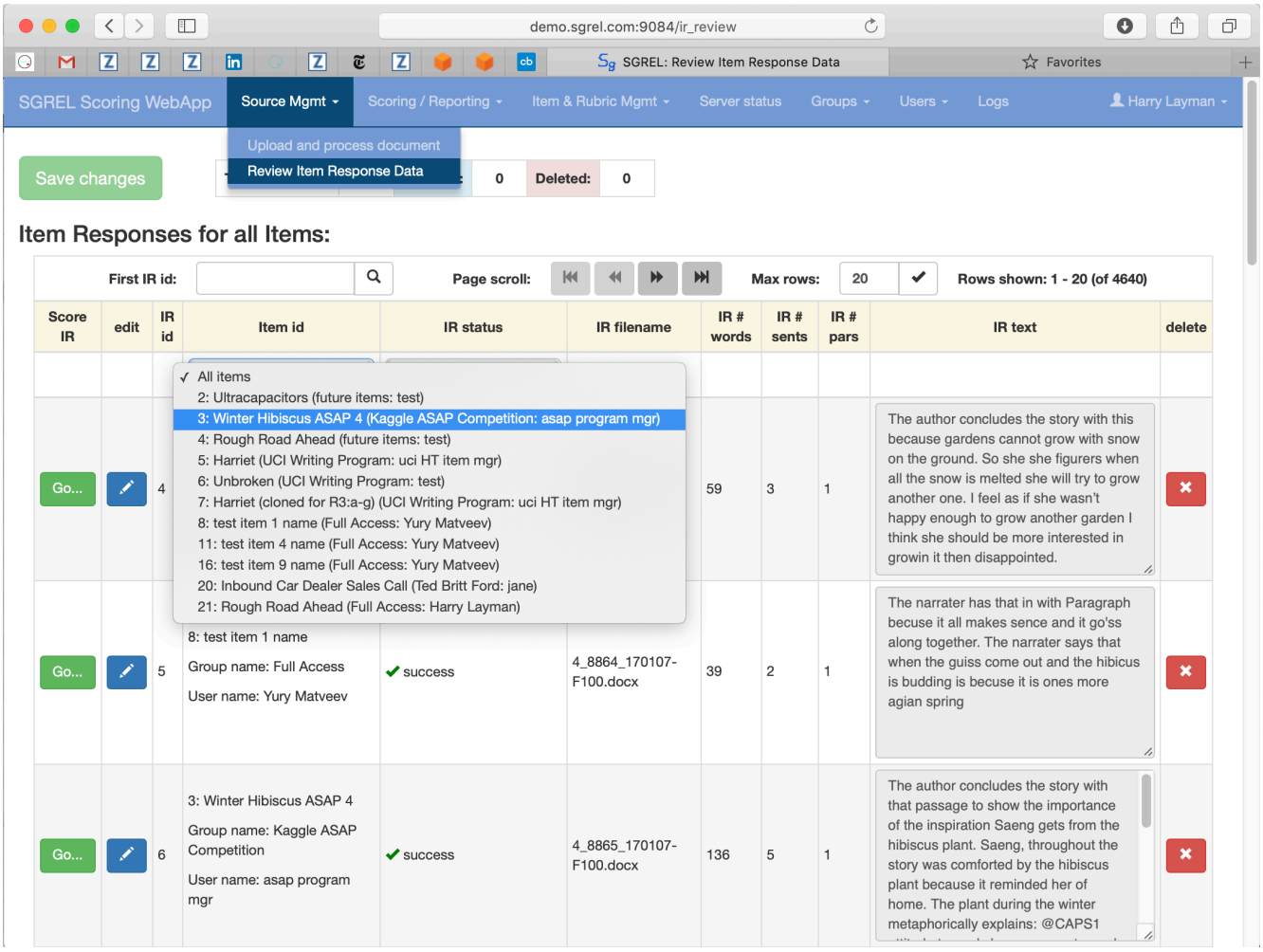


Figure VII-4 Navigate to Review Item Response Data

With item 3: Winter Hibiscus ASAP showing in the drop down selection list, the page should show a view of the 1700+ item 3 responses in the demo database. Use the ">>|" skip to end button and you should have the final 20 (or 5 or whatever you choose in the "max rows" control on the top) in the item response data set for item3.

Scrolling to the last item should reveal the item response just processed. In the page below, you can see that the item response was loaded as it was entered, assigned a "file name" as entered, and shows a word count of 87, a sentence count of 4 and a paragraph count of 1. The itemresp_id assigned is (in this example) 14225 per the figure below.

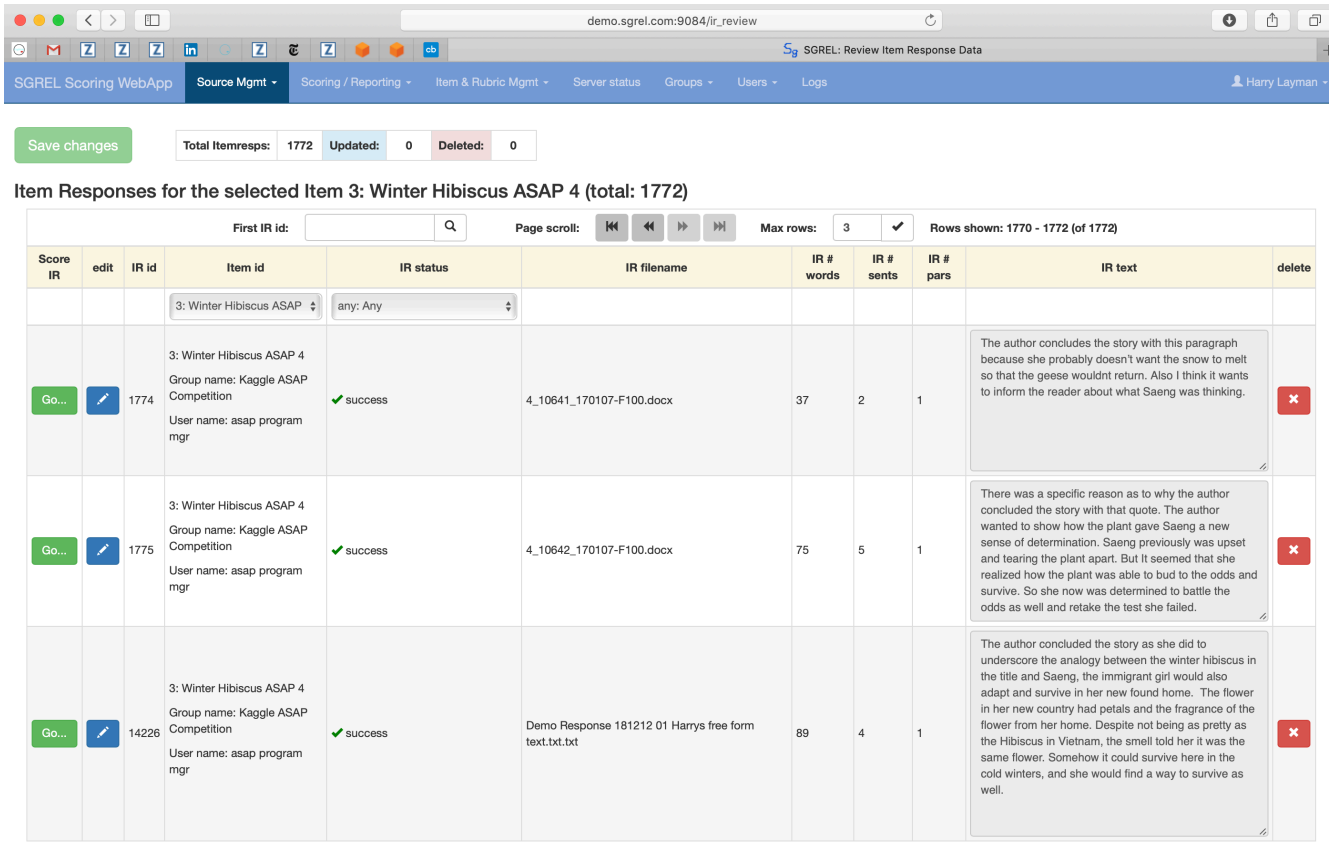


Figure VII-5 Navigate to Final Row to See Item Response

The final row in the table shown above contains the item response just entered. Note that by scrolling and expanding the text in the far right "IR text" column you can see and verify the upload of the entire response.

The next step is to click the "go" button on the left hand end of the bottom row, which brings you to an individual item response scoring page for this item response. The IRA selection box across the top of the **item-rubric association info** area shows item 3, rubric 1, as this is the only rubric associated with this item id -- otherwise you'd be able to switch to a different rubric. Note the URL here also shows the item response id and "ir_score" as the page name (see below).

Also note you can adjust the width of the screen to rearrange the controls for easier viewing.

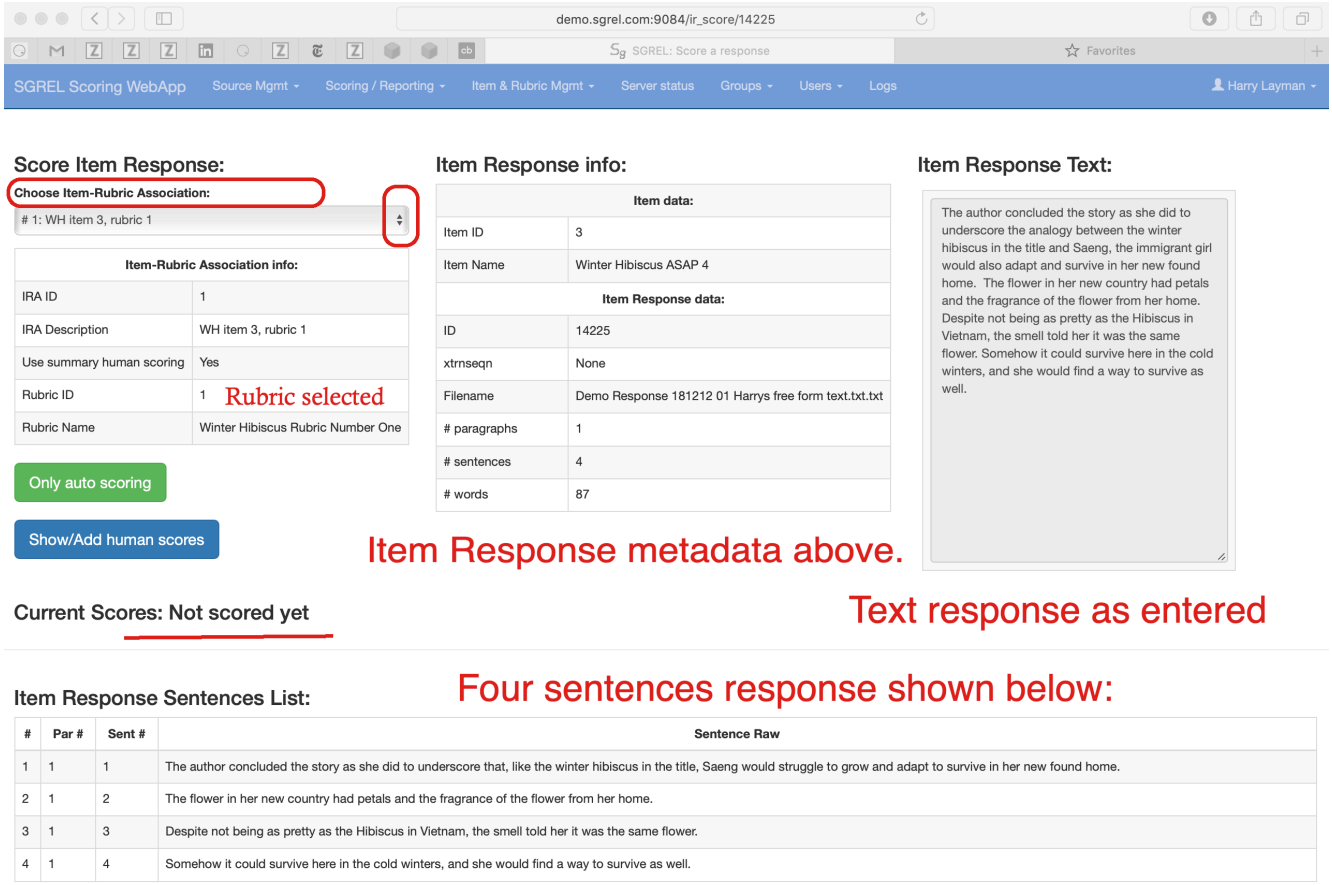


Figure VII-6 Show *IR Score* Page for a Single *Item Response*

This item is not yet scored, but the rubric is selected and there are two choices: enter human scores (blue button) or auto score only (green button). When re-scoring there will be more choices as a result of existing automated or manual scores: "show previous scores", "show / add manual scores", and below the "current score" section, a red "delete auto/human scores" button to erase all existing scores for the selected item response / ira (item-rubric association) combination.

This IRA specifies "summary scoring only"" so let's enter a human score -- 16 for a complete and correct for the claim, and 3 for evidence (citing evidence that flower is the same flower (fragrance and petals), but that flower is different not as pretty and could survive in the winters). As the initial claim statement and final statement make reference to work and struggle, lets score that 3 points of evidence for a total score of 19, and compare that to the automated score.

Clicking the blue button brings up the human scoring popover window shown in the next figure. I have entered the 16 and 3 claim and evidence scores; the next choice is to simply record the human scores (blue button) or record them and also auto-score the result (green button). We will click the green button.

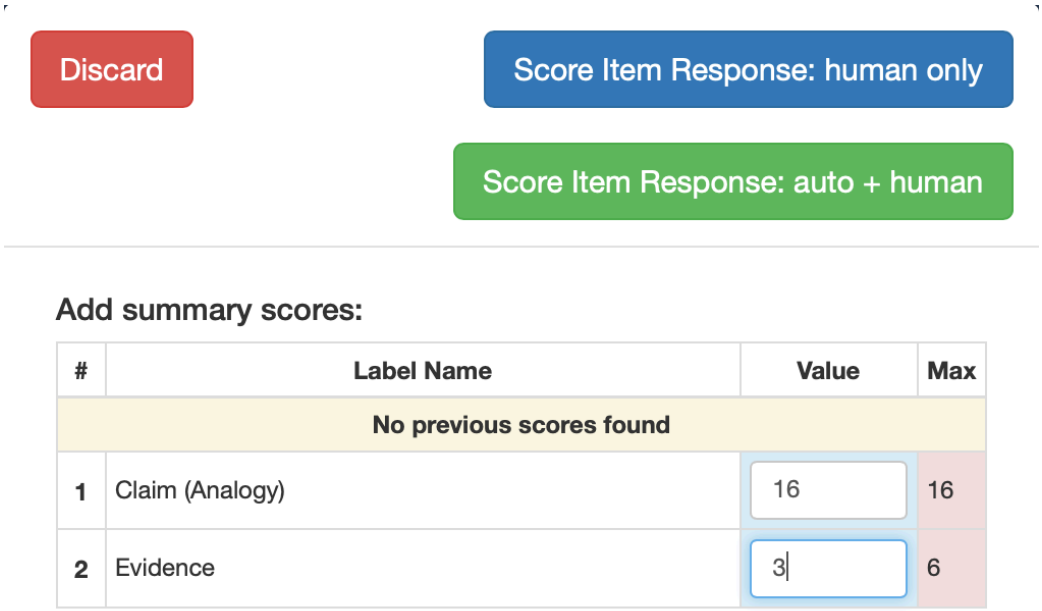


Figure VII-7 Human Summary Scoring Popover Window

The scoring complete message then appears:

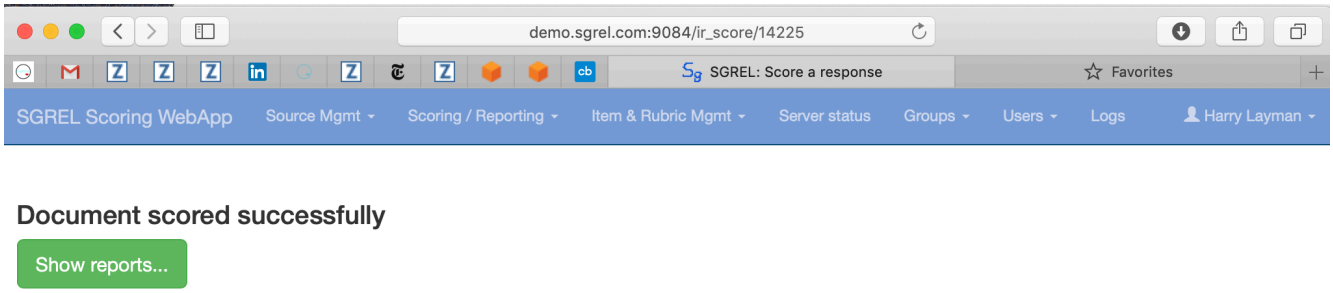


Figure VII-8 Item Response Scoring Complete Message

3. Single Item Reporting

The Green "Score reports…" button is the next logical choice. This button opens a modal window for generating reports as shown below.

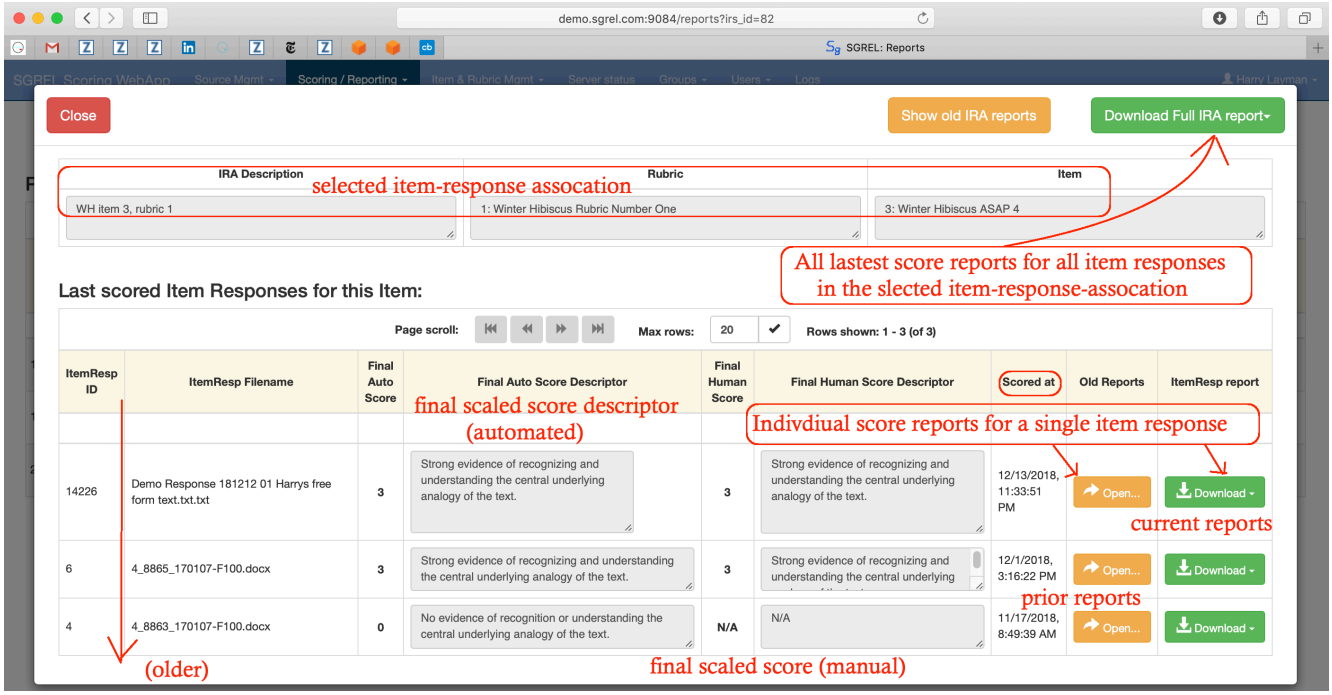


Figure VII-9 Get Reports for IRA Page (Modal Window)

This popover window heading is "Last scored Item Response for this Item" (see above).

Implicit in this statement is the identity of the selected IRA, rubric and item id / names listed across the page just above these words. This modal window puts the most recently scored item response first, with older item responses listed below in the order scored, from newest to oldest. Date / time scored is shown in the "Scored at" column. Item responses that are associated with this IRA that have not been scored (either manually or by the software) are listed among these entries but have "N/A" in the respective (manual or automated) final score descriptor column.

The top row indicates that most recently scored artifact (for this IRA) for which a report can be requested. Notice that on this listing all the choices are for item responses for a specific item id and scored with a specific rubric. In this case it is for a specific Item-Rubric Association between **item** 3 **responses** and scored by **Rubric** 1. For this item response association, there are three responses listed in this figure that have been scored so far. Of the three, the top row is the most recent and the item response we just scored.

From this page it is possible to generate any report for the scoring of any item response for the item id and rubric id that are grouped into this **ira** (item-response association).

Not only can all the different reports be requested here for any response in the group, but the top right button for "Full IRA report" will generate a report using any available report format for every scored item in a single xlsx file.

On an individual item response row, you can see and compare the "final scaled scores" for both human and automated scoring if available, plus you can see the date / time scored and can request any individual report (report for an individual item response) or any prior version of the report generated under prior versions of the rubric (unless the group admin has chosen to remove them). Old reports are selectable by date / time generated and report format name.

Selecting an individual report for the new item response is shown in the figure below. Two report formats are recommended: SA61 and SA71. The sa62/sa72 report formats show detail human scores associated to individual sentences in the response, but since these item responses for this item were "summary score" by humans, no such human scoring detail is available, so the report formats offer more blank space but no additional data than on the sa61/sa71 which are more compact. The automated scoring is built from detail that is always tied into the item response at specific sentences, so there is no difference there between the sa6X and sa7X reports.

The report format sa6X is more a user-oriented, results and meaning report whereas the sa7X report contains full detail on the underlying scoring decisions -- particularly useful in debugging TREs or documenting and justifying scoring decisions. The sa 7X reports are more of a "scoring data dump" and are not yet easily digestible for the untrained / students, CSRs and others being scored.

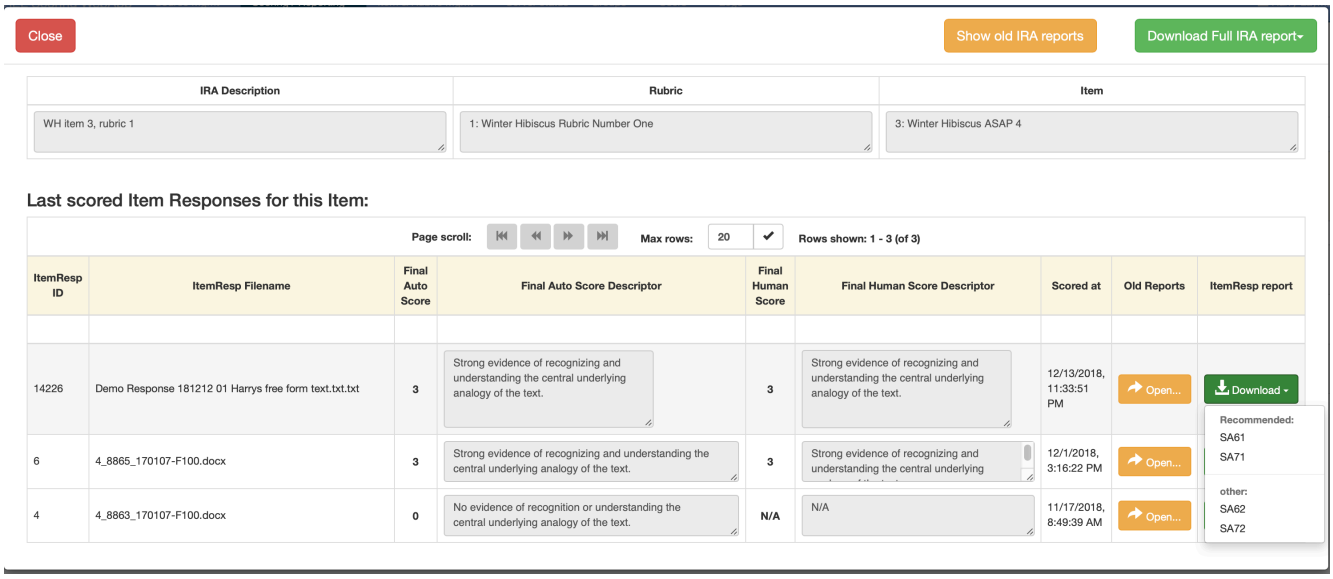


Figure VII-10 Report Choice Dropdown for IRA 1 (i3r1)

Reporting options are as follows: format (with content focused on how the details of the matching decisions -- what tretuples were the best matches, who normalized word tokens were matched, versus what did the matches for particular TREs mean, in terms of point values and scoring feedback) and support for human scores that are either entered at a summary level or for those associated with specific sentences. A two-by-two matrix shown in the figure below indicates how the report names map to these two formats and data content expectations.

Report Use	Debug Scoring: sa7x	User Feedback: sa6x	Notes
Summary Human Scoring: saX1	sa71	sa61	These are designed for when there is NO detailed (sentence level) human scoring, (Usually short-ish responses)
Detailed Human Scoring: saX2	sa72	sa62	These provide full detail for human and automated score reporting. Every scored match is identified
Report Content	Shows details of scoring decisions in data rich (overload!) fashion.	Show implications of scoring decisions -- rationale and actionable feedback.	

Figure VII-11 Detail Score Report: sa6x and sa7x Options

Given the sample item response just loaded was for the demo **ira** (item-rubric-association) for item 3 using rubric 1, this **ira** was set up for "summary human scoring". With an average of only five sentences per response, the location of a specific claim or evidentiary point was less a focus on the feedback (and is expected to be largely self-evident). This indicates that the two reports most useful for this sample item response are the sa61 and sa71 reports. These two report examples are shown in the paragraphs which follow.

4. The sa71 report: Scoring Detail Report

The results of scoring the sample item response with the demo rubric with rubric_id of 1 using the sa71 report format is described section by section and shown in the figures below. The complete xlsx file for this score report is named

sa71-ItemID_3-ItemRespID_14226-IraID_1-2018-12-14_07-51-33.xlsx

and located in the associated file repository identified on the first page of this report. This report name includes the format and support for summary manual scoring (sa71) which indicates this will have all the details of the automated scoring included in the report body.

The sa71 report has three sections:

- Item response information,
- Automated scoring information, and
- Item response sentences.

a) Sa71 Item Response Information Content and Layout

The sa71 Item Response Information Section begins the report for a given response with basic item response metadata and attributes:

- In a blue area, item response attributes including status, word count, sentence count, paragraph count, original score, file name and location;
- in some cases, in the green area, there may be (if available) historical scoring information for the response by other rubrics, systems or processes (identified as sh1, sh2, sh3, ss1 and sm1);
- final scaled score scores and descriptors, if available, for both human (manual) and automated scores if such scoring has been completed and recorded. Human score data is on a brown background and automated score data is on a purple background.
- Four vertical sections underneath the summary scoring data:
  - Subscore Labels (light orange, under "scoring summary" heading)
  - Manual scoring, subscore totals and max, for each category / label (light brown)
  - Automated scoring, subscore totals and max, for each category / label (purple)
  - Three rows showing the item, rubric and ira id's and names.

This first section of the sa71 score report is shown in the figure below.

SGREL SA71 Score Report for Itemresps 14226 to item 3, Winter Hibiscus ASAP 4 , using rubric 1, Winter Hibiscus Rubric Number One.
Notes: Rubric 1 item response scores include human scores (if available) at the sub-score level only, not at the sentence level.

1	Item Response Information	itemresp_id	xtrnseqn	status	wordcnt	sentcnt	paracnt	origscore	itemresp_filename
		14226		3	89	4	1		Demo Response 181212 01 Harrys free form text.txt.txt
		sh1	sh2	sh3	ss1	sm1			itemresp_location
		h final score	3	h desc	Strong evidence of recognizing and understanding the central underlying analogy of the text.				
		a final score	3	a desc	Strong evidence of recognizing and understanding the central underlying analogy of the text.				

Scoring Summary	Manual Scoring		Automated Scoring		Item ID	3	Winter Hibiscus ASAP 4
Label	Value	Max	Value	Max	Rubric ID	1	Winter Hibiscus Rubric Number One
Claim (Analogy)	16	16	16	16	ira_id	1	WH item 3, rubric 1
Evidence	3	6	3	6			
Total Score	19		19				

Figure VII-12 sa71 Score Report Pt 1 IR Metadata & FSSs

b) Sa71 Item Response Automated Scoring Match Detail

The sa71 Item Automated Scoring Match Detail Section is wide and shown in two parts below. There is one line in this section for each "best match" that the automated scoring system assigned and used during scoring. The two images below appear in the report side-by-side as a single wide array. The data in this section includes:

- A blue shaded section heading area;
- A pink shaded row of column headings;
- For each row:
  - A sequence number counting all the best matches;
  - The best match table key (irbm_id)
  - The associated sentence id (irsen_id)
  - The Associated TRE id (tre_id)
  - The TRE's tretuple id for the specific tretuple expansion that made the match (tretuple_id)
  - Other TRE attributes for the TRE that matched, including group number, match count attained, match count minimum, match count percent, tre_group name, tre value, and
  - The actual cleaned version of the tretuple text -- sequence of word tokens -- that were matched to the subject sentence; and
  - The TRE's explanation indicating the target concept and rationale

Of all these fields, the key items are a) the tret_clean, or tretuple text (token string) to show what matched to the target sentence; the m_cnt or match count and the mcm or match count minimum, and how these compare, and how they compare to the number of tokens in the tret_clean and (in the next section) the words in the sentence matched by the irsen_id value on the same row as this match.

Automated Scoring:

#	irbm_id	irsen_id	tre_id	tretuple_id	grp_no	m_cnt	mcm	mpct	tre_group	value
1	202	138959	1	237698	1	8	4	0.28	Conclusion	16
2	204	138959	14	1426	5	4	2	0.15	Ev-5	1
3	205	138960	12	1475	8	4	2	0.33	Ev-8	1
4	203	138961	6	1339	3	6	3	0.33	Ev-3	1
										19

tre_group	value	tret_clean	tre_explanation
Conclusion	16	hibiscu;girl;analog;adapt	points awarded for recognizing the full underlying analogy.
Ev-5	1	surviv;girl	points for recognizing importance of persistence or determination.
Ev-8	1	petal;flower	points for noting way(s) that the winter hibiscus is the same plant as in home country.
Ev-3	1	hibiscu;not;pretti	points for recognizing local hibiscus less pretty.
	19		

Figure VII-13 sa71 Score Report Pt 2 Match Details

c) Sa71 Item Response Sentence Detail

The sa71 Item Response Sentence Detail Section shows the item response sentence-by-sentence, as parsed by the scoring engine, and with "original text" and "normalized" or "cleaned" text for

each sentence. This complete listing of the item response data by sentence, with associated sentence_id values, allows the tretuple matches identified by TRE (tre_id) and item response sentence (irsen_id) in the prior section, to be examined in detail by comparing the "cleaned" tretuple string with the "cleaned" sentence string of the matching sentence identified, to assess if the match is working (in terms of grouping, filtering and match count minimums and such, as desired.

Seen in it's full width, this third and final section of the sa71 report is shown below.

Item Response Sentences:

irsens:

irsen_id	irpara _seqn	irsen _seqn	irsen_text	irsen_cleaned
138959	1	1	The author concluded the story as she did to underscore the analogy between the winter hibiscus in the title and Saeng, the immigrant girl would also adapt and survive in her new found home.	author conclud stori a she did to underscor analog between winter hibiscu in titl saeng immigr girl would also adapt surviv in her new found home
138960	1	2	The flower in her new country had petals and the fragrance of the flower from her home.	flower in her new countri had petal fragranc flower from her home
138961	1	3	Despite not being as pretty as the Hibiscus in Vietnam, the smell told her it was the same flower.	despit not be a pretti a hibiscu in vietnam smell told her it wa same flower
138962	1	4	Somehow it could survive here in the cold winters, and she would find a way to survive as well.	somehow it could surviv here in cold winter she would find way to surviv a well

Figure VII-14 sa71 Score Report Pt 3 Sentence Details

However, to appreciate the reveal of how the free form text entered by the examinee might compare with the "normalized" version of the text (by being compared to the normalized version of the tretuples shown in the prior sa71 report section), a larger versio of the left and right sides shown above are included below.

irsen_text
The author concluded the story as she did to underscore the analogy between the winter hibiscus in the title and Saeng, the immigrant girl would also adapt and survive in her new found home.
The flower in her new country had petals and the fragrance of the flower from her home.
Despite not being as pretty as the Hibiscus in Vietnam, the smell told her it was the same flower.
Somehow it could survive here in the cold winters, and she would find a way to survive as well.

irsen_cleaned
author conclud stori a she did to underscor analog between winter hibiscu in titl saeng immigr girl would also adapt surviv in her new found home
flower in her new countri had petal fragranc flower from her home
despit not be a pretti a hibiscu in vietnam smell told her it wa same flower
somehow it could surviv here in cold winter she would find way to surviv a well

Figure VII-15 sa71 Score Report Pt 3 Sentence Details - Zoomed

5. The sa61 report: End User Scoring Detail Report

The sa61report is similar to the sa71 report but is more focused on diagnostic information and does not reveal, or have the clutter, of the details about how the automation decided to score specific sentences.

Rather, the sa61report contains three sections:

- Item response information (the same as on sa71);
- Automated scoring information (without "tretuple" details but showing the tre explanation / feedback); and
- The item response presented as free-form text for easy reference.

This report provides the minimum essential information to document the automated score results, and includes the summary human scoring information (if any has been applied to this item response).

The first two sections of the report are quite similar, with the second section removing tretuple-specific data from the automated scoring data, but leaving the explanations for why each "best match" tre ended up contributing to the score and how. Further the specific way the scoring engine parsed the response into sentences is not presented, but rather the entire response is included as the third section of the report.

The entirety of the sa61 report for this item response created during this demo is shown in the figure below.

SGREL SA61 End User Score Report for Itemresps 14226 to item 3, Winter Hibiscus ASAP 4 , using rubric 1, Winter Hibiscus Rubric Number One.
Notes: Rubric 1 item response scores include human scores (if available) at the sub-score level only, not at the sentence level.

1	Item Response Information	itemresp_id	xtrnseqn	status	wordcnt	sentcnt	paracnt	origscore	itemresp_filename
		14226		3	89	4	1		Demo Response 181212 01 Harrys free form text.txt.txt
		sh1	sh2	sh3	ss1	sm1			itemresp_location
		h final score	3	h desc	Strong evidence of recognizing and understanding the central underlying analogy of the text.				
		a final score	3	a desc	Strong evidence of recognizing and understanding the central underlying analogy of the text.				

Scoring Summary	Manual Scoring		Automated Scoring		Item ID	3	Winter Hibiscus ASAP 4
Label	Value	Max	Value	Max	Rubric ID	1	Winter Hibiscus Rubric Number One
Claim (Analogy)	16	16	16	16	ira_id	1	WH item 3, rubric 1
Evidence	3	6	3	6			
Total Score	19		19				

Automated Scoring:

#	irsen_id	tre_id	value	grp_no	ss_idx	tre_type	tre_group	tre_explanation
1	138959	1	16	1	1	PC	Conclusion	points awarded for recognizing the full underlying analogy.
2	138959	14	1	5	2	EV	Ev-5	points for recognizing importance of persistence or determination.
3	138960	12	1	8	2	EV	Ev-8	points for noting way(s) that the winter hibiscus is the same plant as in home country.
4	138961	6	1	3	2	EV	Ev-3	points for recognizing local hibiscus less pretty.
			19					

Item Response:

irsens count:	4	First Sentence:	138959	Last Sentence:	138962

The author concluded the story as she did to underscore the analogy between the winter hibiscus in the title and Saeng, the immigrant girl would also adapt and survive in her new found home. The flower in her new country had petals and the fragrance of the flower from her home. Despite not being as pretty as the Hibiscus in Vietnam, the smell told her it was the same flower. Somehow it could survive here in the cold winters, and she would find a way to survive as well.

Figure VII-16 sa61 Score Report - All 3 Sections

The report itself can be found in the file:

sa61-ItemID_3-ItemRespID_14226-IraID_1-2018-12-14_07-51-30.xlsx

6. The sa72 and sa62 Reports - Differences from sa71 and sa61

The sa72 and sa62 reports differ from their sa71 and sa61 counterparts only by having an additional third section added between section 2, which contains automated scoring information, and what was section 3. Section 3 in the case of sa72 was a listing of the item response that was scored in a sentence-by-sentence, raw and cleaned versions side-by-side format. In the case of sa71, section three contained a free-form version of the original scored text for reference.

A view of the new section 3 for sa71 and sa61 is presented in the next two figures. Again, full forms of the sample reports can be found on the image repository for this report or generated by the user accounts on the demo site at CogWrite.ai.

Also note in, the final example, that if an sa62 report is generated for an item scored with summary human scoring, the new "human scoring" section that lists all the sentences is suppressed as superfluous since ira's with this attribute (summary manual scoring only) set have no human score data associated with specific sentences.

a) Sa72 New Section 3 - Manual Scoring Detail

The figure below shows content from report
sa72-ItemID_5-ItemRespID_3503-IraID_2-2018-12-01_23-30-12.xlsx.

Human Scoring:

Sent Seqn #	Claim (Which Trait)	Evidence	irsen_id	irsen_cleaned
1	0	0	118200	everyon ha their own qualiti on be good leader
2	0	0	118201	some are differ come are same
3	0	0	118202	in
4	0	0	118203	railroad run to canada
5	0	0	118204	from harriet tubman conductor on underground railroad by ann petri is about woman that wa onc slave led eleven ru
6	0	0	118205	she runaway slave walk all night got few rest but with help peopl they manag to surviv by be fed given shelter
7	0	0	118206	harriet tubman becam great leader to runaway slave risk her life to do so
8	0	0	118207	accord to barbara white an expert in educ leadership presid beyond better develop author
9	0	0	118208	seven qualiti good leader
10	4	0	118209	state that there are seven person qualiti good leader white identifi trustworthi is most essenti in enabl harriet to guid sl
11	0	0	118210	barba white note that
12	0	1	118211	leader need to be trust be known to live higher life with honesti integr
13	0	1	118212	harriet promis slave to get them food shelter she did
14	0	0	118213	she kep her word for that she got trust slave
15	0	0	118214	they had all been warm safe well
16	0	0	118215	thi quot is import becaus it tell me that harriet did not give up on them wa determin to help them
17	0	0	118216	clearli thi qualiti good leadership wa most essenti in enabl harriet to guid slave to north becaus all slave trust what she
18	0	0	118217	they sat on ground near her wait patient until she awaken
19	0	0	118218	they had to come to trust her implicit total
20	0	0	118219	petri word suggest that throughout their journey to freedom slave have come to trust her
21	0	1	118220	thi help her surviv becaus runaway did not harm her turn their back against her
22	0	0	118221	if harriet had no been trustworthi she might have not been abl to guid slave to north becaus moment she fell asleep sl
23	0	0	118222	they would have also disclos stop place hide place
24	0	0	118223	trustworthi is veri import becaus they need to trust each other in order to move on
25	0	1	118224	runaway rag dirti hungri cold did not steal gun a they might have set off by themselv turn back
26	0	0	118225	thi show that slave wait patient for harriet to awak
27	0	0	118226	she had gun to protect herself slave
28	0	0	118227	slave had no attent to use it on her becaus they trust harriet to give them their freedom so that they could no longer b
29	0	0	118228	throughout stori harriet slave were alik but differ to
30	0	0	118229	they were both afraid get caught were exhaust hungri
31	0	0	118230	harriet wa veri determin to get slave their freedom
32	0	0	118231	slave howev were have doubt
33	0	0	118232	from harriet stori we can learn that we should take risk for what we believ is right
34	0	0	118233	she believ that slaveri wa just wrong horribl thing so she risk her life to save slave led them to freedom
35	0	0	118234	without trustworthi their could be no relat between anyon
36	0	0	118235	good leader must be trust worthi to help other
	4	4		

Figure VII-17 sa72 New Section 3 (Manual Scores)
Note human scoring decision for each sub-score by sentence are highlighted in the report above.

b) Sa62 New Section 3 - Manual Scoring Detail

The figure below shows content from report
sa62-ItemID_5-ItemRespID_3503-IraID_2-2018-12-09_19-47-20.xlsx

Human Scoring:

Sent Seqn #	Claim (Which Trait)	Evidence	irsen_id	irsen_text
1	0	0	118200	Everyone has their own qualities on being a good leader.
2	0	0	118201	Some are different and come are the same.
3	0	0	118202	In.
4	0	0	118203	The Railroad Runs to Canada.
5	0	0	118204	from Harriet Tubman: Conductor on the Underground Railroad by Ann Petry is about a woman that was once a slaved and led eleven runaway slaves all the way to C
6	0	0	118205	She and the runaway slaves walked all night and got few rest but with the help of people they manage to survive by being fed and given shelter.
7	0	0	118206	Harriet Tubman became a great leader to the runaway slaves and risk her life to do so.
8	0	0	118207	According to Barbara White, an expert in educational leadership and president of Beyond Better development and author of.
9	0	0	118208	Seven Qualities of a Good Leader.
10	4	0	118209	states that there are seven personal qualities of a good leader White identifies, Trustworthiness is the most essential in enabling Harriet to guide the slaves to North becau
11	0	0	118210	Barba White notes that.
12	0	1	118211	A leader needs to be trusted and be known to live his/her life with honesty and integrity.
13	0	1	118212	Harriet promised the slaves to get them food and shelter and she did.
14	0	0	118213	She kepted her word and for that she got the trust of the slaves.
15	0	0	118214	They had all been warm and safe and well-fed.
16	0	0	118215	This quote is important because it tells me that Harriet did not give up on them, and was determine to help them.
17	0	0	118216	Clearly, this qualities of a good leadership was most essential in enabling Harriet to guide the slaves to the North because all the slaves trusted what she did and said.
18	0	0	118217	They sat on the ground near her and waited patiently until she awaken.
19	0	0	118218	They had to come to trust her implicit, totally.
20	0	0	118219	Petry words suggest that throughout their journey to freedom the slaves have come to trust her.
21	0	1	118220	This helped her survive because the runaways did not harm her or turn their back against her.
22	0	0	118221	If Harriet had no been trustworthiness, she might have not been able to guide the slaves to the North because the moment she fell asleep the slaves would of have betray
23	0	0	118222	They would of have also disclose the stopping places and the hiding places.
24	0	0	118223	Trustworthiness is very important because they need to trust each other in order to move on.
25	0	1	118224	The runaways, ragged, dirty, hungry, cold, did not steal the gun as they might have and set off by themselves or turn back.
26	0	0	118225	This shows that the slaves waited patiently for Harriet to awake.
27	0	0	118226	She had a gun to protect herself and the slaves.
28	0	0	118227	The slaves had no attention to use it on her because they trusted Harriet to give them their freedom, so that they could no longer be slaves.
29	0	0	118228	Throughout the story Harriet and the slaves were alike but different to.
30	0	0	118229	They were both afraid of getting caught and were exhausted and hungry.
31	0	0	118230	Harriet was very determined to get the slaves their freedom.
32	0	0	118231	The slaves however were having doubts.
33	0	0	118232	From Harriet storys we can learn that we should take risk for what we believe is right.
34	0	0	118233	She believed that slavery was just a wrong and horrible thing, so she risked her life to save slaves and led them to freedom.
35	0	0	118234	Without trustworthiness, their could be no relation between anyone.
36	0	0	118235	A good leader must be trust worthy to help others.
	4	4		

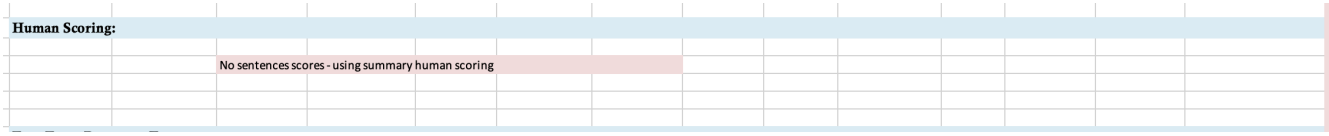
Figure VII-18 sa62 New Section 3 (Manual Scores)

.

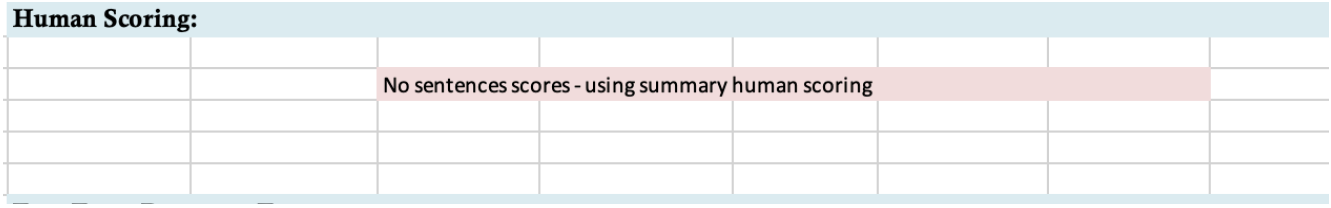
c) Sa62 Section 3 for Manual Scoring Details When N/A

The figure below shows content from report
sa62-ItemID_3-ItemRespID_14226-IraID_1-2018-12-15_07-23-28.xlsx

Note that this report was for an item response that is part of an **ira** group with manual summary scoring only, and the report includes section 3 to show human scoring detail. But Since human scoring details being absent contrast directly with the function of this software to print the detailed human scores (which are not recorded), in lieu of the new section three a message is printed to indicate that this detail human scoring is not available for this item response / ira.



Or more clearly:



B. Batch Scoring

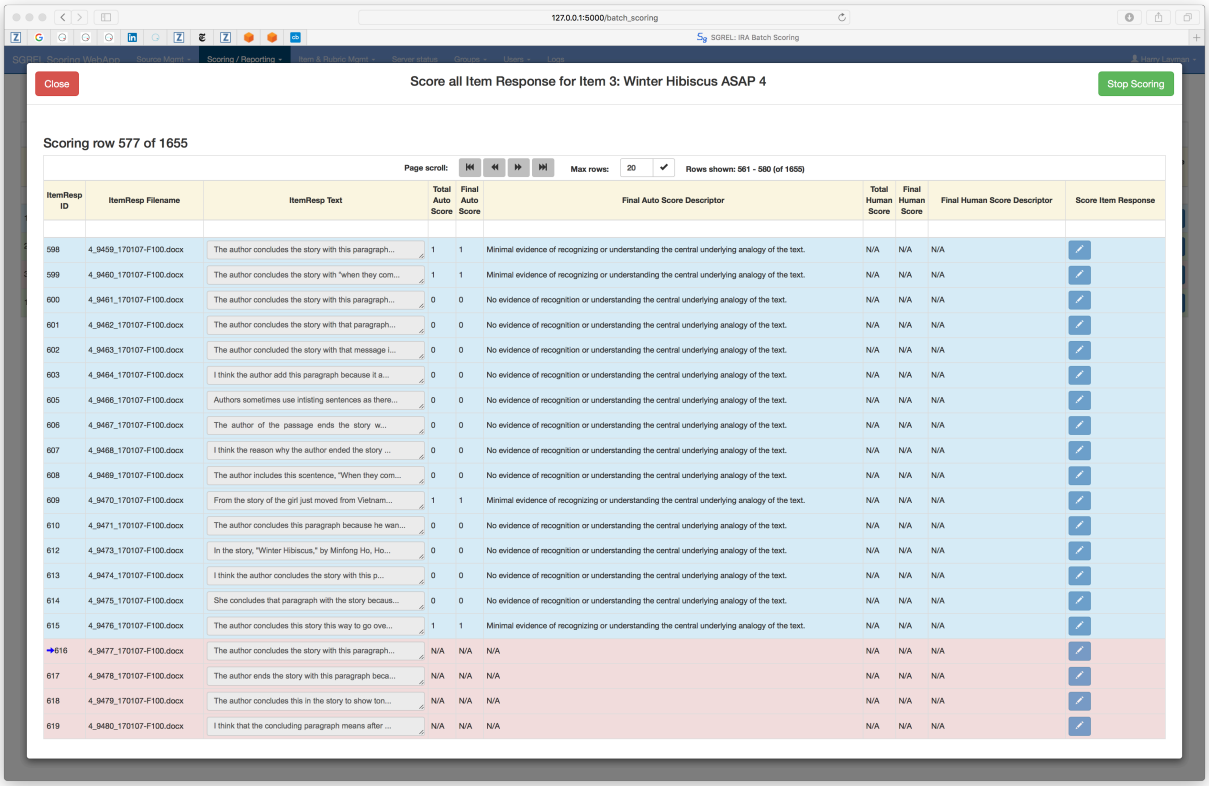
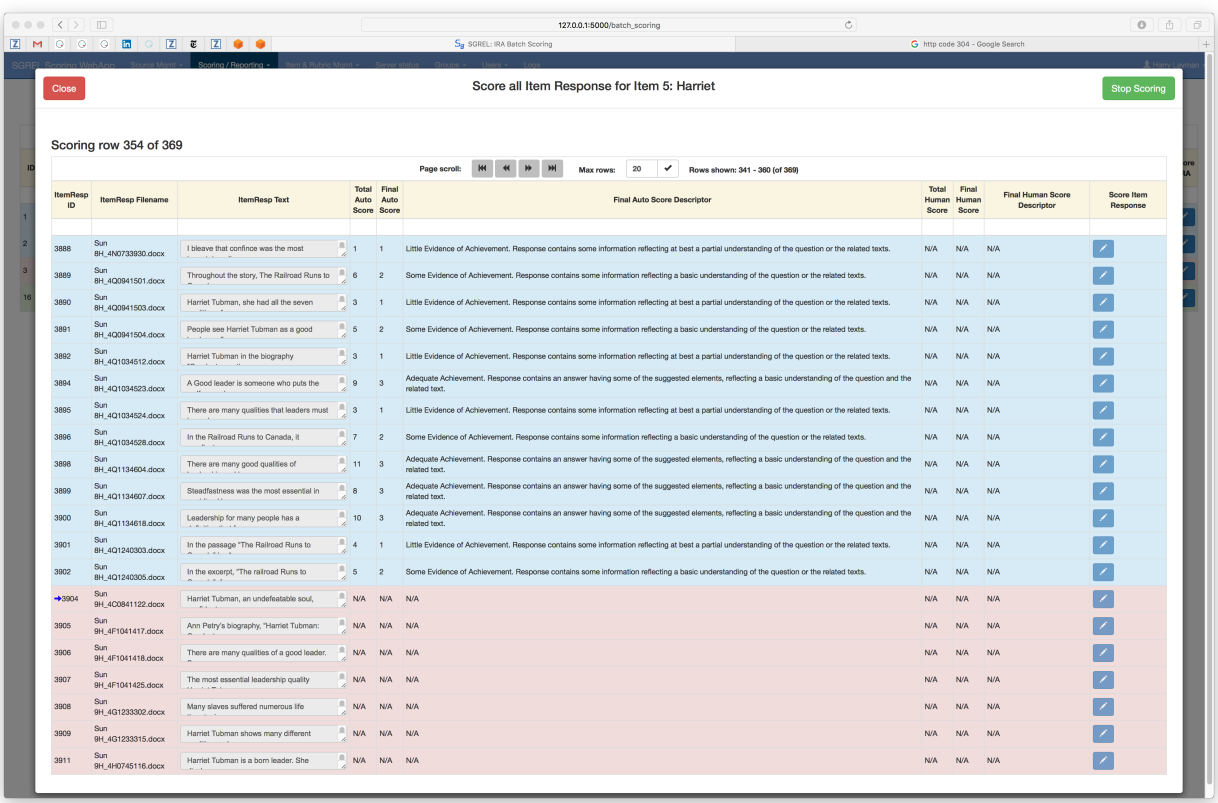


Figure VII-19 Batch Scoring in Process



C. Batch loading / importing

TBD

VIII. Score Reporting

Prototype versions of the desired scoring reports are presented in the figures below. Existing scoring capabilities store all the data necessary to format and build these and many other report formats that could contain valuable scoring details and diagnostic feedback on scored assessment items.

3705 - Moon 10H_4G1031806

Leaders all around the world and in history had great qualities in order for them to become a leader. One great leader in history is the abolitionist, Harriet Tubman. She helped free slaves for many years because they all deserved to be free. Although, without the cooperation with the slaves, Harriet Tubman wouldn't have been able to help free them. One of Harriet's most essential quality was that she was trustworthy. (1) The slaves entrusted her with their safety and is why they were able to escape slavery. They entrusted their lives to Harriet Tubman and shows that Harriet Tubman was a trustworthy person.

When Harriet gathered up a group of slaves, she promised them that they be able to get food and shelter. Even though they had very rough conditions, Harriet knew that they would be able to provide the resources they need. "She had told them about the place where they would stay, promising warmth and good food, holding these things out to the as an incentive to keep going." (The Railroad Runs to Canada). Harriet makes a promise to them in order for them to keep going and following her. The slaves believed Harriet when she promised them food and shelter. They all had to believe that they will find the resources. Harriet wasn't sure about keeping her promise, but she had to believe in order for the slaves to trust her.

As the leader, Harriet had to make sure that all of her followers trust her. Unfortunately, one slave spoke up and said that they wanted to go back. Although, Harriet wouldn't allow that to happen. She knows what could possibly happen if a slave returned. "If a runaway returned, he would turn traitor, the master and the overseer would force him to turn traitor." (The Railroad Runs to Canada). She explains to the group the possibilities of returning and they all believed her because she is the more experienced out of them all. She knows what she is doing and she needs the slaves to believe in her. (4) That is why Harriet is trusted as leader because she has more knowledge of what is going on unlike the slaves.

Throughout the journey, Harriet would get tired just like the rest of the slaves. (3) She would sleep whenever she wasn't able to walk anymore. While she sleeps, the slaves wait for her to wake up since she is the leader. "She was leading them into freedom, and so they waited until she was ready to go on." (The Railroad Runs to Canada). Her followers know that she is the only one who can guide them to freedom. The believe that she would take them to a better place so that they could all be free. They trust her with their lives so much that they respect her resting time. "A good leader "walks the talk" and in doing so earns the right to have responsibility for others." (Seven Qualities of a Good Leader). Harriet earned their respect, therefore, the slaves will wait for her to wake up. They would only allow Harriet to guide them and to tell them what to do. She's placed a huge impact on their lives and entrusts her with their freedom.

As a leader, Harriet proved how great she was by leading the slaves to freedom. They all trusted her and cooperated with her in order to be free. (2) Through her actions of freeing the slaves, Harriet shows that one of her most essential qualities is that she is trustworthy.

Scoring Results Key:
Claim score: 4/4 Evidence Score: 3/8 **Total score: 4 + 3 = 7 / 12**

Detailed Score Report and Feedback
HT Item Response - Claim + Evidence Score
3705 - Jane Doe 18 August 2016

tre_group	Loc	Sentence begins with...	tre_explanation	point value
claim	1	One of Harriet's most essential …	Strong claim, identifies a trait as highly important.	4
evidence	2	They all trusted her and coop …	Notes that HT was trusted by the slaves.	1
evidence	3	Throughout the journey, Harriet …	HT, at times, was tired, hungrey and afraid, just like the slaves.	1
evidence	4	She knows what she is doing …	Notes HT will do anything to help her group succeed.	1

Category	Max Points	Basis
Claim	4	Statement of one most essential trait for maximum credit; fewer points for partial response.
Evidence and Reasoning	8	Identifying attitudes, thoughts and actions that demonstrate the trait or connect it to leadership.

Raw Score Range	Final Scaled Score	Final Score Descriptor
12	4	Exceptional Achievement. Response contains all suggested elements and reflects a thorough understanding of the question and related texts.
8 - 12	3	Commendable / Adequate Achievement. Response contains many or all of the suggested elements reflecting a good understanding of the question and the related texts.
5 - 7	2	Some Achievement. Response contains an answer having some of the suggested elements, reflecting a basic understanding of the question and the related text.
1 - 4	1	Minimal / Little Evidence of Achievement. Response contains minimally relevant information reflecting at best a partial understanding of the question or of the texts.
0	0	No evidence of understanding the question or the related texts.

Figure VIII-1 Sample One - Score and Feedback

Moon 10H_4A1030202_13503

[Automated scoring, sub-scores A to G, tagged sentences alternate colours; 6 sentences tagged.]

Everyone has their own qualities on being a good leader. Some are different and come are the same. In "The Railroad Runs to Canada," from Harriet Tubman: Conductor on the Underground Railroad by Ann Petry is about a woman that was once a slaved and led eleven runaway slaves all the way to Canada. She and the runaway slaves walked all night and got few rest but with the help of people they manage to survive by being fed and given shelter. Harriet Tubman became a great leader to the runaway slaves and risk her life to do so. According to Barbara White, an expert in educational leadership and president of Beyond Better development and author of "Seven Qualities of a Good Leader," states that there are seven personal qualities of a good leader White identifies. Trustworthiness is the most essential in enabling Harriet to guide the slaves to North because the slaves needed to trust her to get them their freedom.

Barba White notes that, "A leader needs to be trusted and be known to live his/her life with honesty and integrity." Harriet promised the slaves to get them food and shelter and she did. She kepted her word and for that she got the trust of the slaves. "They had all been warm and safe and well-fed." This quote is important because it tells me that Harriet did not give up on them, and was determine to help them. Clearly, this qualities of a good leadership was most essential in enabling Harriet to guide the slaves to the North because all the slaves trusted what she did and said.

"They sat on the ground near her and waited patiently until she awaken. They had to come to trust her implicit, totally." Petry words suggest that throughout their journey to freedom the slaves have come to trust her. This helped her survive because the runaways did not harm her or turn their back against her. If Harriet had no been trustworthiness, she might have not been able to guide the slaves to the North because the moment she fell asleep the slaves would of have betrayed her and leave her all alone in the woods. They would of have also disclose the stopping places and the hiding places. Trustworthiness is very important because they need to trust each other in order to move on.

"The runaways, ragged, dirty, hungry, cold, did not steal the gun as they might have and set off by themselves or turn back." This shows that the slaves waited patiently for Harriet to awake. She had a gun to protect herself and the slaves. The slaves had no attention to use it on her because they trusted Harriet to give them their freedom, so that they could no longer be slaves.

Throughout the story Harriet and the slaves were alike but different to. They were both afraid of getting caught and were exhausted and hungry. Harriet was very determined to get the slaves their freedom. The slaves however were having doubts. From Harriet storys we can learn that we should take risk for what we believe is right. She believed that slavery was just a wrong and horrible thing, so she risked her life to save slaves and led them to freedom. Without trustworthiness, their could be no relation between anyone. A good leader must be trust worthy to help others.

Automated Score:

Narrative Elements:	A	B	C	D	E	F	G	Total
Best Match Scores	2	2	1	4	1	4	1	
Rubric Sub-Scores	**2**	**2**	**1**	**2**	**1**	**2**	**1**	**11**

Figure VIII-2 Sample Two - Score and Feedback