

## CUNY Academic Commons - Feature #11941

### CSV export for Reckoning data

2019-10-03 12:01 PM - Boone Gorges

<b>Status:</b> Resolved	<b>Start date:</b> 2019-10-03
<b>Priority name:</b> Normal	<b>Due date:</b>
<b>Assignee:</b> Raymond Hoh	<b>% Done:</b> 0%
<b>Category name:</b> Reckoning	<b>Estimated time:</b> 0.00 hour
<b>Target version:</b> 1.18.0	
<b>Description</b>	
From <a href="#">#11937</a> :	
2. Can we build in an option to export the Reckoning data to CSV? We probably only need to export info from the user summary overview, not the second page with full length posts.	
Laurie, can you please provide some more specifics on where you'd expect the interface to appear and what it should include? - I'd assume, for example, that we only need the ability to export Reckoning data for an entire site, <b>not</b> on an individual user basis - but it could be that I have this the wrong way round. - What info would be included? User names, display names? Would the export just have total counts for each user? Or would it be more specific, with links and/or content to specific items? If the latter, it would probably be necessary to have a separate CSV for counts, since the export columns would be along the lines of 'User name', 'post URL', 'comment URL', and the export would contain a row for each item instead of each user. Or maybe we need three exports: counts, posts, comments? To answer this, it would be helpful to describe the gradebook workflow you envision. (Presumably we'd also include WP Grade Comments info if available, but this would also be on a per-post/comment basis.)	
<b>Related issues:</b>	
Related to CUNY Academic Commons - Feature #11937: Integration between WP Gra...	<b>Assigned</b> 2019-10-03
Related to CUNY Academic Commons - Feature #12447: Tallying Forum Participation	<b>Resolved</b> 2020-02-19

### History

#### #1 - 2019-10-03 12:01 PM - Boone Gorges

- Related to Feature #11937: Integration between WP Grade Comments and Reckoning added

#### #2 - 2019-10-03 02:29 PM - Laurie Hurson

- File screen 1.png added

- File Screen 2 Extended User Summary.png added

This export function could appear in the Users dashboard menu as Users > Export User Summary

This could take admins to a page to determine what content will be exported. Like you suggested Boone, I think it would make sense to provide multiple export options.

(1) User Summary Overview Export: First option could include information from the user summary overview (based on page 1). Export CSV would include the following information for each user

- User name
- Post name
- Post category (if added)
- Post URL (if possible but not absolutely necessary here)
- Date posted
- Total post count
- Total Comment count

(2) Extended User Summary Export: Second option would be to export extended single user data (based on page 2. This CSV would contain:

- User name
- Post name
- Post category
- Post URL (if possible)
- **Post content**
- Date posted
- Total post count
- Total Comment count
- **Comment content**

I'm not sure if full content export is possible only on a user by user basis (so you can only do one at time) or if we could allow export of all extended user summaries at one time AKA exporting all (or most) content organized by user (this is outlined below in CSV 2)

If possible maybe on the export selection screen we could create the option to select or export ALL user data that produces two csvs:

1. User Summary Overview Export
2. Extended User Summary Export (all users)

### #3 - 2019-10-03 03:04 PM - Boone Gorges

Thanks, Laurie. I'm not sure how to reconcile the columns you've suggested. Think of it this way: Each row in a CSV represents a "record" or an "item". Things like "total post count" only make sense if the record stands for a **user**. If the record is a **post**, then "total post count" is not really a meaningful number for that record. Likewise, if a record is a user, then it doesn't make sense to talk about "post date", since a user may represent many posts (or none at all).

It's possible to divide this content up into exports in many different ways, but I don't want to create extra work for the dev team, or conceptual overhead for instructors. So I think the important question is: what workflow(s) are we enabling by allowing CSV export? For example, if I'm an instructor who builds his own Excel spreadsheet for grades, I might find it helpful to get an export of all posts and comments related to my class - in this case, records would be posts or posts+comments. I could then (say) use Excel's sorting and counting tools to organize the records by date, so that I can transfer items over to the corresponding assignment columns in my gradebook spreadsheet. In this case, overall user counts wouldn't be interesting to me, since I need to fill in my gradebook on a per-assignment basis. In contrast, I can imagine someone wanting more "summary" type information, or wanting information that can easily be massaged for import into Blackboard (? not sure if this exists). But I'm not sure which of these scenarios is most likely to be useful to actual instructors.

In the absence of other thoughts, my inclination is to provide data on as fine-grained a basis as possible. This means individual records for each piece of content (posts + comments). Instructors interested in more general summary-type data can use the specific data in order to calculate what they need.

Regarding user-specific vs entire-class, it's possible to do both, but I have a feeling - again, based on my guess on how instructors would use this, though this is largely conjecture - that per-user export would be tedious. A likely workflow is that the instructor will download all data at the end of the semester, or maybe a few times per term, and then use it to enter into a gradebook en masse. It's far easier to do this with a single spreadsheet, which can then be sorted/filtered/searched if I'm looking for content written by specific users. As in the foregoing paragraph, I'd err on the side of maximal info here; instructors who want less data (eg only the data corresponding to a single student) can use tools in their spreadsheet program to narrow it down. This also simplifies things for me, since I'll only have to build a single exporter.

IMHO it's not good to put full comment/post content into the spreadsheet. It takes it out of context (especially relevant for comments) and clutters up the spreadsheet interface. Instead, we provide the URLs, and instructors can click through if they want to see the content. Happy to be corrected about this, though.

Given the above, here's what I'm suggesting:

- A single CSV export, which will aggregate all content from across the entire site, by all users
- Each row in the CSV will be either a post or a comment. There'll be a column along the lines of 'type' that will contain either 'post' or 'comment' so that the instructor can sort/filter as needed
- Columns would be as follows:
  - User display name
  - User WP user name
  - Type (post or comment)
  - Post name
  - Post category (if added)
  - Post URL
  - Comment URL (if type=comment)
  - Date posted

Does this seem like an OK way forward? (Sorry for the massive braindump.)

### #4 - 2019-10-03 03:30 PM - Laurie Hurson

Sorry for making you do all the heavy lifting! I understand the organizational conundrum and requirements better now, though.

I agree that a whole class csv is most help and that full text comments and posts are too much info (and out of context).

I think the most likely scenario is that the prof wants to organize the information by student to see if that person has completed all the posts, comments, etc, which could be done by filtering by user name. Or they want to organize by post, in which case they can filter by post category. So, yes, this sounds like it will work.

### #5 - 2019-10-03 03:31 PM - Laurie Hurson

By filtering I mean, once the professor exports the csv and opens the file with excel or another program.

### #6 - 2019-10-04 01:28 PM - Boone Gorges

- Status changed from Reporter Feedback to Assigned

Excellent - sounds like we have a rough plan for this. I've tentatively placed it in the spring semester, but if we clear the fall docket sooner than expected, we'll see about moving back.

**#7 - 2020-07-15 11:26 AM - Boone Gorges**

- Target version changed from 1.17.0 to 1.18.0

**#8 - 2020-09-23 11:06 AM - Boone Gorges**

- Related to Feature #12447: Tallying Forum Participation added

**#9 - 2020-09-23 11:12 AM - Boone Gorges**

- Assignee changed from Boone Gorges to Raymond Hoh

Ray, let's couple this one with [#12447](#). Feel free to modify the Reckoning plugin directly if necessary, though I imagine the CSV assembler could be pretty much standalone. And the process of looping through and collecting content probably doesn't need Reckoning at all - just query for all posts and comments, and plop them into a CSV. Put the Download button somewhere in the Dashboard where it makes sense to you, and we can deal with placement/styling down the road if necessary.

**#10 - 2020-10-27 06:05 PM - Raymond Hoh**

I'm just about to finish this up.

For the CSV columns, these are what I'm using:

- User display name (If comment is anonymous, this uses the comment's Name field)
- Username (this uses user\_login. If comment is anonymous, this will default to 'Anonymous')
- Type (post or comment)
- Content URL (post URL if post, comment URL if comment)
- Post title (only applicable for posts. Should this use the post title or the post slug? I used the post title.)
- Post category (only applicable for posts. If posted in multiple categories, will be a comma-delimited list.)
- Date posted (I'm currently using GMT, but should I use the local timezone instead? Or have another column for local timezone?)

Also let me know if there should be any additional columns in the CSV that we might be forgetting.

For comments that are from non-logged-in users, I'm currently adding those comments as well, but let me know if I should omit non-logged-in users from the CSV.

**#11 - 2020-10-28 02:12 PM - Boone Gorges**

Thanks, Ray! This list looks good to me, but Laurie should weigh in.

**#12 - 2020-10-28 03:00 PM - Laurie Hurson**

Hi Ray and Boone

Thanks for this! It looks good to me too.

The only column addition I would suggest (if possible) would be to add a column that indicates which post a comment was made on. I know this could be gleaned by clicking the comment URL but would it be possible to add a column that surfaces the post info that a comment was made on? I think this would be helpful to know since professors often ask students to comment on certain posts.

I can also imagine that in the future professors might want the CSV to include post or comment content. For now this export structure makes sense but if there are use cases down the line where professors want post content in the export would we be able to amend the export file to include post content?

**#13 - 2020-10-28 03:48 PM - Raymond Hoh**

Thanks for the feedback, Laurie.

I know this could be gleaned by clicking the comment URL but would it be possible to add a column that surfaces the post info that a comment was made on? I think this would be helpful to know since professors often ask students to comment on certain posts.

For comments, I'm currently not populating the "Post title" column, but I can add the post title for comments here if it helps tie the comment to a post more easily. If the "Post title" column is too busy, perhaps we could also add a new "Post ID" column for easier identification?

One question I have is whether we should have anonymous comments in the CSV. Currently, anonymous users are not displayed in Reckoning's "User Summary" page, so to be more in line with Reckoning, maybe we shouldn't add anonymous comments to the CSV. One advantage to having anonymous comments in the CSV though is sometimes a student might not log in to post a comment and might just post anonymously with their email address tied to the comment. I think if we do add anonymous comments, we should also add a "User email" column so a professor can identify if an anonymous user is actually a student.

I can also imagine that in the future professors might want the CSV to include post or comment content. For now this export structure makes sense but if there are use cases down the line where professors want post content in the export would we be able to amend the export file to include post content?

It would be easy to add the post / comment content to the CSV. I think if I were a professor, I might want to have this included into the CSV especially if I wanted to do some work offline. Reading the content through the CSV would require some sorting gymnastics though. If we did want to add the content, should this be the raw value of the content or the rendered content? Raw value would mean things like shortcodes and embed URLs would not be transformed into HTML. I'm guessing we would want rendered content.

**#14 - 2020-10-28 07:00 PM - Raymond Hoh**

- *File reckoning-csv.png added*

To those that want to test, I've added a first pass on cdev.

To test, ensure that your site has the Reckoning plugin activated. Next, on the "Users > User Summary" admin page, you will notice a new "Download CSV" button:

reckoning-csv.png

At the moment, I've chosen not to add this functionality directly into the Reckoning plugin, but I can if we deem that the main Reckoning plugin can benefit from this feature.

The "User email", "Post ID" and "Content" columns are not added to the CSV in this iteration, but let me know if we want to add them or not. Anonymous comments are still in this iteration as well.

(Commit - <https://github.com/cuny-academic-commons/cac/commit/4e70d8d8dd1daaf49628d1332e731994516b901b>)

**#15 - 2020-11-04 11:36 AM - Laurie Hurson**

Hi Ray and Boone,

I just tested this and it works great! A few thoughts:

- I don't think we need "User email" or "Post ID" columns.

- I am torn about adding a content column. I can see the value of having content in order to work offline but it will also make the sheet much messier. If we do add content, I think rendered content would be ideal but if unrendered is much easier to provide I think that would be fine. (might also format better to a csv output?) On the other hand, it might make sense to wait until we have a few use cases to determine if there is a demand for content download.

- I was confused about the default to GMT timezone. I downloaded once with site's default timezone settings and downloaded after setting timezone to UTC-5 and the CSV defaults to GMT in both scenarios. I think most profs typically do indicate their timezone in the settings of the site so they can accurately see when posts were published. Is there a way to reflect the time/date settings from the site in the CSV? If not, perhaps we choose EST as default time zone in the CSV to reflect that most profs are teaching in NYC timezone?

- Can we change the language that describes the download. Now it says: "You can also download a CSV of all user data below" but I think this is misleading because without content it's not really all data. Perhaps it should read: "You can also download a CSV of summarized user activity data below" or something along those lines (which mirrors the title of the page- summary vs. more detailed data).

Thanks all!

**#16 - 2020-11-04 04:57 PM - Raymond Hoh**

Thanks for the feedback, Laurie.

I've changed the Date posted column to use whatever timezone WordPress has set on the "Settings > General" admin page. And I've also adjusted the blurb above the "Download CSV" button. This is available for testing on cdev. (Commit - <https://github.com/cuny-academic-commons/cac/commit/d3eb0be18b16706de1abc6b3a231ee612284e600>)

Do you have any further thoughts about anonymous comments in the CSV? I've written some reasons why it might be useful to add them [here](#).

**#17 - 2020-11-05 11:11 AM - Laurie Hurson**

Thanks Ray.

I think it makes sense to include anon comments with a column for user email. I think if students get confused and do not sign in but comment as anon the email will help sort out who commented.

**#18 - 2020-11-05 06:07 PM - Raymond Hoh**

- *Status changed from Assigned to Testing Required*

Thanks Laurie. I've added a new column for user email to the CSV.

This can be tested on cdev. (Commit - <https://github.com/cuny-academic-commons/cac/commit/2cb5e078c04d5042903f6f24d1662d0fe747d61b>)

**#19 - 2020-11-06 09:54 AM - Laurie Hurson**

Hi Ray,

Tested this on Cdev - updates to CSV fields look good and time column mirrors site settings for timezone.

**#20 - 2020-12-08 11:31 AM - Boone Gorges**

- *Status changed from Testing Required to Resolved*

**Files**

---

screen 1.png	298 KB	2019-10-03	Laurie Hurson
Screen 2 Extended User Summary.png	624 KB	2019-10-03	Laurie Hurson
reckoning-csv.png	22.6 KB	2020-10-28	Raymond Hoh