

Extraction process for Sakai Gradebook and Sakai Events Data sets from the CMS

The extraction of CMS Sakai Gradebook and CMS Usage Data (Sakai Events) is extracted from Sakai using the following list of queries.

Currently, we are running queries on Cerritos, Redwoods and Savannah. We also run queries on Marist for building the analytics model. We are responsible for running the Savannah queries and Marist queries. rSmart is responsible for running the Cerritos and Redwoods queries.

For convenience these are the main queries used. The queries used with some of the variations of the following queries depending on the university and dataset we are extracting.

Queries to extract Sakai Events:

Query A (Context Not Null)

```
INSERT INTO temp_sakai_event_table select ss.TITLE, sn.SESSION_USER, sn.SESSION_START,
sn.SESSION_END, se.* from SAKAI_EVENT se join SAKAI_SESSION sn on se.SESSION_ID =
sn.SESSION_ID join SAKAI_SITE ss on (se.CONTEXT = ss.SITE_ID and CONTEXT IS NOT NULL) where
ss.TITLE like '%10F';
```

Notes:

- 1) The course TITLE in this example is '10F'. This will retrieve all Fall of 2010 courses. The field will need to be modified for other semesters. Some schools may not use the same naming convention and the queries will need to be modified appropriately.

Query B (Context Null)

```
INSERT INTO temp_null_sakai_event_table
select null, sn.SESSION_USER, sn.SESSION_START, sn.SESSION_END, se.EVENT_ID, se.EVENT_DATE,
se.EVENT, SUBSTR(se.REF, LOCATE('/',se.REF, 2) + 1, INSTR(se.REF, '-p') - LOCATE('/',se.REF, 2) - 1),
se.SESSION_ID, se.EVENT_CODE, se.CONTEXT from SAKAI_EVENT se join SAKAI_SESSION sn
on se.SESSION_ID=sn.SESSION_ID where se.EVENT_DATE > DATE('2010-08-02') and se.EVENT_DATE <
DATE('2011-02-25') and se.CONTEXT IS NULL and se.EVENT in('pres.begin', 'pres.end');
```

Notes:

- 1) The Date fields used in the query need to match the date range for the semester of interest. Having a date range greater than the actual range is fine as the ERP enrollment would be used to match and filter out the subset of interest.

Query C (merges Context Null results into Context Not Null Event table)

```
INSERT INTO temp_sakai_event_table
select ss.TITLE, tse.SESSION_USER, tse.SESSION_START, tse.SESSION_END, tse.EVENT_ID,
tse.EVENT_DATE, tse.EVENT, tse.REF, tse.SESSION_ID, tse.EVENT_CODE, tse.CONTEXT
from temp_null_sakai_event_table tse
join SAKAI_SITE ss on (ss.SITE_ID=tse.REF) where ss.TITLE like '%10F';
```

Notes:

- 1) Title needs to be modified to correspond to the semester of interest.

Query D (output specific fields to csv file)

```
select title, session_user, event, event_date from temp_sakai_event_table into outfile
'/tmp/10F_ALL_SAKAI_EVENTS.csv' fields terminated by ',';
```

Notes:

- 1) See OAAI document for naming conventions used to name files.

Queries to extract Gradebook data

```
select s.title worksite, g.id gradebook, r.student_id username, c.weight, c.name category, o.name
object, o.points_possible, r.points_earned, o.due_date, o.ungraded, o.version
from ( GB_GRADABLE_OBJECT_T o left join GB_CATEGORY_T c on o.CATEGORY_ID = c.ID
join GB_GRADE_RECORD_T r on o.id = r.gradable_object_id )
join GB_GRADEBOOK_T g on g.id=o.gradebook_id
join SAKAI_SITE s on g.gradebook_uid=s.site_id
where s.TITLE like '%10F' and o.object_type_id = 1 and o.not_counted <> 1
and (c.removed is null or c.removed <> 1) and (o.removed is null or o.removed <> 1)
order by due_date, object, username into OUTFILE '/tmp/10F_GRADES.csv' FIELDS TERMINATED by ',';
```

Notes:

- 1) Title pattern for the semester needs to be changed appropriately.

Other queries used:**Query Full Id Map based on semester**

```
select distinct suim.user_id, suim.eid from sakai_user_id_map suim, sakai_site ss, sakai_site_user ssu
where ss.title like "%11F" and ssu.site_id = ss.site_id and ssu.user_id = suim.user_id into outfile
'/tmp/marist_id_map_11F_2012_02_03.csv' fields terminated by ',';
```

Query for all students enrolled in a Course

```
select ssu.USER_ID, EID from sakai_site_user ssu, sakai_user_id_map suim where  
ssu.site_id='478ec4da-f534-40d6-8c10-eaadbe286b38' and ssu.USER_ID = suim.USER_ID into outfile  
'/tmp/savannah_ACCT_2101_01_12S.csv' fields terminated by ',';
```

To Query for a subset of events for specific Courses

```
select title, session_user, event, event_date from temp_sakai_event_12S_table where (title like  
'ENGL_1102_%_12S' or title like 'ACCT_2101_%_12S' or title like 'MATH_1111_%_12S' or title like  
'BUSA_2182_%_12S' or title like 'MATH_1113_%_12S' or title like 'ARTS_1101_%_12S') into outfile  
'/tmp/savannah_sakai_events_2012_01_27.csv' fields terminated by ',';
```