

(14:21:28) dr ttol: to the profiles  
(14:22:05) dr ttol: the emails should be saved in a database  
(14:22:14) David Gucwa: but the user doesn't need to wait for that, right?  
there's nothing in the crawled friends' profiles that the user needs for his  
own?  
(14:22:14) dr ttol: so before the sendmail(), it checks to make sure the email  
isnt in the database  
(14:22:31) dr ttol: because there will be common friends, and we only want one  
email  
(14:22:40) dr ttol: right  
(14:22:46) David Gucwa: ok  
(14:22:50) dr ttol: well  
(14:22:54) dr ttol: except the friend requests  
(14:23:17) David Gucwa: okay but that means the friend already has a profile  
with connectu  
(14:23:20) dr ttol: that part of the software needs to crawl through their  
friends list and check to see if they're already on connectu  
(14:23:30) dr ttol: if they are, send a friend request through connectu software  
(14:23:31) David Gucwa: so we don't need to crawl their facebook account  
(14:23:46) dr ttol: if they're not, add email to database of emails sent, and  
then also do a sendmail()  
(14:23:58) dr ttol: if email is already in database, bypass sendmail  
(14:24:02) David Gucwa: as long as we store their facebook id on their connectu  
account, we can look them up on connectu without crawling their profile on  
thefacebook  
(14:24:23) dr ttol: ok  
(14:25:05) dr ttol: the point is, if im a user, i will import my facebook  
account, and i expect the software to look for my friends automatically and send  
requests out if they're on connectu, and send invite emails out if they're not  
(14:25:38) David Gucwa: does the user get an option to send out invites or not  
(14:26:10) dr ttol: no  
(14:26:13) David Gucwa: good  
(14:26:14) dr ttol: its part of the first checkbox  
(14:26:28) dr ttol: if they want auto friends, auto invites is part of the  
functionality  
(14:26:35) David Gucwa: that makes things a lot simpler  
(14:26:57) dr ttol: but you still need to gather the email addresses from the  
profiles  
(14:27:43) David Gucwa: right  
(14:28:55) dr ttol: im just thikning that if we crawl through the entire  
facebook school by school now  
(14:29:09) dr ttol: then we essentially dont need to hit their site during  
impotr time  
(14:29:15) dr ttol: because we'll have all the info locally  
(14:29:25) dr ttol: we can just verify their login info with facebook  
(14:29:36) dr ttol: and if succesful, the user will claim ownership to a certain  
profile we stored locally  
(14:29:47) dr ttol: and all the friend requests will be sent out  
(14:29:52) dr ttol: (whether through connectu or via email)  
(14:29:54) David Gucwa: I don't think it's feasible to crawl their entire site  
before ours goes live  
(14:30:38) dr ttol: ok, but if you had a friend, Adam Wood  
(14:30:41) dr ttol: and I had the same friend  
(14:30:47) dr ttol: that means the importer will hit thefacebook twice  
(14:30:50) dr ttol: same profile  
(14:30:59) dr ttol: and if theres a third friend that also had Adam Wood listed  
(14:31:00) David Gucwa: the importer will cache any profile it crawls through

(14:31:01) dr ttol: it hits it again  
(14:31:09) dr ttol: okay  
(14:31:13) dr ttol: so only one hit per profile  
(14:31:15) dr ttol: right?  
(14:31:16) David Gucwa: yes  
(14:31:19) dr ttol: ok good  
(14:31:25) David Gucwa: I just don't think we can expect to do them all in advance  
(14:31:31) dr ttol: ok  
(14:33:21) dr ttol: but it may be good to have, at some point, software that caches, or do you expect the importer to do all the caching we need  
(14:34:28) David Gucwa: I was expecting to write the caching for just the importer  
(14:34:48) David Gucwa: the info will be there for any other reasons we need it  
(14:35:08) dr ttol: ok  
(14:35:10) dr ttol: eta  
(14:38:03) David Gucwa: for spidering/caching on import, the rest of the day probably  
(14:40:03) dr ttol: ok, what about friends  
(14:40:21) dr ttol: checking to see if friends are there, and triggering a friend request event  
(14:40:36) dr ttol: and db of emails that we already sent invites to  
(14:40:59) David Gucwa: do you know how easy or hard it is to add a friend to a connectu profile  
(14:41:05) dr ttol: ask joel  
(14:41:12) David Gucwa: It would be pretty easy to just pass some information to an existing script if one exists  
(14:41:22) dr ttol: im not sure which is better  
(14:41:25) dr ttol: to restrict one invite per email  
(14:41:27) dr ttol: or no restriction  
(14:41:41) dr ttol: "One of our members, David Gucwa, has just invited you to join him in ConnectU"  
(14:41:50) dr ttol: "One of our members, Wayne Chang, has just invited you to join him in ConnectU"  
(14:41:59) dr ttol: someone may get 20+ emails  
(14:42:06) dr ttol: that might not be bad  
(14:42:08) David Gucwa: that could get pretty obnoxious  
(14:42:20) dr ttol: ok  
(14:42:22) dr ttol: so one email per user?  
(14:42:30) David Gucwa: well I don't know if obnoxious is what you want  
(14:42:36) dr ttol: no  
(14:42:39) David Gucwa: it would certainly be a good incentive to join connectu and make the invites stop  
(14:42:39) dr ttol: we dont want to be evil  
(14:42:41) David Gucwa: heh, okay  
(14:42:58) David Gucwa: I can remember each email then  
(14:43:03) dr ttol: ok  
(14:43:26) dr ttol: where is your home town?  
(14:43:38) David Gucwa: sandwich, on the cape  
(14:44:17) dr ttol: ok  
(14:44:41) dr ttol: there is going to be a cambridge office opening soon. i want amherst to be the better one  
(14:45:46) David Gucwa: ok  
(14:46:07) dr ttol: so everything will be finished tonight?  
(14:46:45) David Gucwa: everything for the spidering part  
(14:46:51) David Gucwa: I still have a bunch of other work to do  
(14:46:59) David Gucwa: I have to make all the other importers

(14:47:13) David Gucwa: speaking of which I need logins for the other sites we're importing from

(14:48:50) dr ttol: right

(14:49:06) dr ttol: ill get that to you, but we need to finish thefacebook importer first

(14:49:12) David Gucwa: ok

Conversation with drttol at 2005-01-27 16:44:12 on David Gucwa (aim)

(16:44:12) dr ttol: status?  
(16:44:23) David Gucwa: caching umich  
(16:44:36) dr ttol: how fast?  
(16:44:37) David Gucwa: I'm going to put the script on the server and see if it goes any faster  
(16:44:43) David Gucwa: it's been going for like 10 minutes  
(16:44:49) dr ttol: how many profiles  
(16:45:11) dr ttol: tfb is run by a techie, so becareful  
(16:45:25) David Gucwa: there are two steps, it retrieves all the id's first, then it gets the profile for each id  
(16:45:30) David Gucwa: it's still on the first step, up to 6000 ids  
(16:45:51) dr ttol: ok, does it get the pictures as well  
(16:46:29) David Gucwa: it gets the urls of the pictures  
(16:46:44) dr ttol: what about for the importer  
(16:48:23) David Gucwa: same  
(16:48:42) David Gucwa: it displays the pictures when you're choosing which one you want, but it doesn't store the picture anywhere locally yet  
(16:49:02) dr ttol: when they pick the picture, does it store it locally  
(16:49:05) dr ttol: to the cu database  
(16:49:24) David Gucwa: no  
(16:49:30) David Gucwa: I don't know where they're supposed to go  
(16:49:37) dr ttol: ask joel  
(16:50:04) David Gucwa: yeah that's something I'll need to take care of eventually  
(16:50:13) David Gucwa: but I'm trying to get the crawler finished up for now  
(16:50:27) dr ttol: ok  
(17:27:05) dr ttol: status  
(17:27:11) David Gucwa: importing 4 schools  
(17:27:16) David Gucwa: it's taking a while  
(17:27:17) dr ttol: all profiles?  
(17:27:26) David Gucwa: I think it's actually going slower on the server  
(17:27:28) David Gucwa: than on my computer  
(17:27:35) David Gucwa: still just grabbing id's  
(17:28:00) dr ttol: hm  
(17:28:03) dr ttol: which server are you using  
(17:28:19) David Gucwa: 69.44.59.182  
(17:28:27) dr ttol: it could e that the schoools have more ids than you think  
(17:28:40) dr ttol: so the crawler works?  
(17:28:47) David Gucwa: yeah  
(17:29:22) dr ttol: which universities  
(17:29:24) dr ttol: do we have umass  
(17:29:29) David Gucwa: no I don't have a umass login  
(17:29:38) dr ttol: rdegutis@student.umass.edu  
(17:29:40) dr ttol: ruth1783  
(17:29:49) David Gucwa: I'm getting umich, cornell, yale and harvard  
(17:30:06) dr ttol: ok  
(17:30:07) dr ttol: get umass next  
(17:30:40) David Gucwa: k  
(17:31:00) dr ttol: we want to import the top i2hub schools first  
(17:31:48) David Gucwa: ok  
(17:31:54) David Gucwa: well I've used all the logins you gave me  
(17:32:09) David Gucwa: so if you give me more logins I can start importing  
(17:32:13) dr ttol: ok  
(17:32:14) dr ttol: you mean  
(17:32:17) David Gucwa: also I'm waiting for wellesley to start showing emails again