

# Habits of Online Creatives

## Using an on-line social networking website to identify trends in the young urban professional creative community in the Calgary city region

BEN LI, University of Calgary (bali@ucalgary.ca)

*On-line social networking websites provide interesting opportunities to collect moderately detailed primary data about social and professional linkages among individuals between the ages of 18-25 within specified geographic communities. On the popular Facebook social networking website, over 11,620 individuals self-identify as being from the University of Calgary, the Calgary area, or both. This is a biased and self-selected, but statistically significant sample of the highly talented youth population in the Calgary area. Since participants in this community are actively interested or participating in largely informal, but documented social, professional and common interest networks, and because the Facebook network for University of Calgary students is fairly representative of the actual University of Calgary population, the diversity and quantities of participants' identified linkages may be viewed as a proxy for non-participant members of the young urban professional community in Calgary.*

*Drawing from self-declared data made public and searchable on Facebook, a number of quantitative analyses are revealing:*

- 1) Based on employment information (including firm and position) a number of different significant relationships between creative young people and firms are examined.*
- 2) Based on interest, activities, etc., various descriptions used by creative individuals to describe their own activities are evaluated.*
- 3) Based on employment and education information (including institution and level of study), relationships between educational experience and creative professional activities.*

*Although this self-selected dataset undoubtedly under-represents overall networking and creative activity in Calgary (and also of the age group sampled), the relatively simple analyses described above should provide useful information and trends with which to effectively target highly-educated and creative young individuals for further study.*

# Introduction

Online social networking websites allow individual Internet users to exchange information and form geographically delocalized communities. When an individual registers with a social networking website, they must complete a brief profile providing basic personal information. Any registered user may search profiles to identify other users with shared interests or other desired characteristics. User profiles contain information such as age, gender, geographic location, employer name and occupation, as well as more detailed information about education, marital status, hobbies, activities, favourite books, etc. Any subset of this information may be made publicly available to other users at the profile owner's discretion, and any information made public in that manner is viewable by any Internet user. In this study, we leverage the public visibility of these profiles to collect aggregate information about creative individuals who use Facebook ("Facebook Creatives") to identify and study their characteristics.

To that end, this study aims for two goals throughout: First it seeks to provide a variety of basic insights, confirmations and rebuttals into several aspects of the local, young creative community. Second, it proposes a number of questions in different areas that may be addressable by the dataset, which could inspire future research efforts.

## What is Facebook?

"Facebook is a social utility that connects people with friends and others who work, study and live around them," according to the Facebook website (2007). Established as an online version of traditional hard-copy compilations of photos of people used by departments and schools to help members identify each other, its primary audience at inception was university students, faculty and staff. Initially, a valid university ("college" in the Facebook parlance) e-mail address was required to register on the Facebook website, but registration has been open to all Internet users since September 2006.

When joining Facebook, members may select one or more networks to join. A network is based on a real-life community, such as a university, high school, or company, or on a city or geographic region, such as Calgary, Alberta. Users must possess a valid e-mail address for the network they wish to join, except for geographic networks which may be joined freely.

Individuals join Facebook to support and extend their

real-life relationships with colleagues from their current networks. Through Facebook's network directory and search features, members are able to find other members, organize events, contact potential participants, and re-connect with individuals from previous educational and work experiences. Members are encouraged to identify other found members as friends as appropriate, and to solicit friends outside of the Facebook community to join Facebook. Any individual registered on the website may also form a sub-community or special interest group based on any common interest.

Members are also able to build new relationships with individuals they meet through Facebook via network, event and group discussion boards. In order to find other members, or to be found by former colleagues, members are incentivised to accurately self-identify with interests, occupations and descriptors when completing and updating their profiles. Each dimension of the profile is searchable (see Figure 1a). Finding a friend, or being found, potentially expands each member's social network and expertise with their interests. For some individuals, listing occupations, interests, and affiliations may also be a matter of pride since the site provides an outlet for publicly displaying their accomplishments.

At any time, a user may elect to make none, some, or all of their profile information visible to other users, and they may edit or update the information at any time.

As a social networking site that grows by referral, Facebook membership grows on a daily basis (as do some sub-populations such as groups within Facebook). That feature results in new data set has the advantage of growing by integral percentage points on a daily basis, but also the disadvantage of not being viewable in its entirety at any moment in time.

## Why Facebook?

Facebook is an interesting data source because it covers a number of demographic, personal and professional dimensions of individuals and communities. In addition, the data points are not tainted by third party interpretation since they are provided directly by the users they characterize. The data is also highly regular because standard forms and input validation at the time of user data entry restrict the input of illegitimate or exceptional responses. Facebook, and other social networking websites, represent a new form of user-supplied but independently verified data not previously available, and especially not in such a clean form.

Figure 1a. Key features of a profile.

Basic demographic information

Personal preferences

Friends and networks

Education and employment

Posted items, group memberships, wall and notes

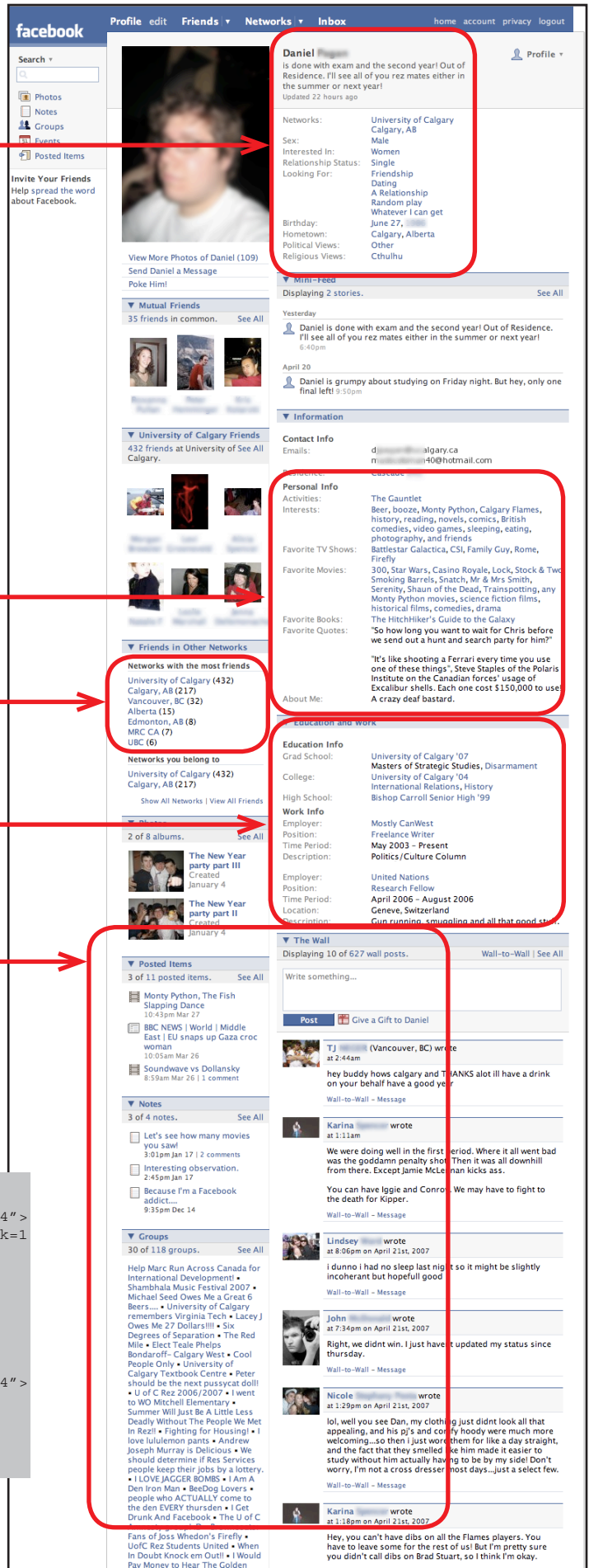


Figure 1b. HTML source code excerpt containing some fields and values.

```
<td class="label">Hometown:</td>
<td class="data"><div class="datawrap"><a href="http://ucalgary.
facebook.com/s.php?adv&k=10010&n=-1&c1=Calgary&o=4">
Calgary</a>, <a href="http://ucalgary.facebook.com/s.php?adv&k=1
0010&n=-1&r1=100001&o=4">Alberta</a></div></td></tr>
<tr id="Political Views"><td class="label">Political Views:</td>
<td class="data"><div class="datawrap"><a href="http://ucalgary.
facebook.com/b.php?k=10010&n=-1&p1=8&o=4">Other</a></
div></td></tr>
<tr id="Religious Views"><td class="label">Religious Views:</td>
<td class="data"><div class="datawrap"><a href="http://ucalgary.
facebook.com/s.php?adv&k=10010&n=-1&re=Cthulhu&o=4">
Cthulhu</a></div></td>
```

## Motivations

Drawing from self-declared data made public and searchable by Facebook, this study focuses on three broad but revealing quantitative analyses.

- 1) Based on employment information (including firm and position) a number of different significant relationships between creative young people and firms are examined.
- 2) Based on interest, activities, etc., various descriptions used by creative individuals to describe their own activities are evaluated.
- 3) Based on employment and education information (including institution and level of study), relationships between educational experience and creative professional activities.

Together, these factors contribute to our understanding of the role of creative people in the local innovation system. Specifically, understanding how to identify young creatives, their patterns, as well as how their creativity is manifest through personal and professional pursuits enables the city region to potentially better identify, develop and allocate creative human capital. Understanding creative roles among firms and the young people who occupy them may also result in innovative and competitive advantages.

## Methodology

### Data Gathering

An account was created on the Facebook site by supplying the minimum articles of information necessary to create a profile in order to gain access to the website. The account was affiliated with both the “Calgary, AB” and the “University of Calgary” networks. The “Calgary, AB” network is one based on geographic affiliation, denoting an affiliated user’s primary city from whichever perspective, while the “University of Calgary” network is one of institutional affiliation, denoting current or previous membership in the University of Calgary community (affiliation with this network requires a valid University of Calgary e-mail address).

In the initial data collection, profiles of 6,346 registered users were downloaded using the Firefox web browser and the “DownloadThemAll” extension (Parodi and Verna, 2006) between 8 p.m. and 10:30 p.m. on April 7, 2007.

The profiles were collected in batches using the social

networking site’s profile browser tool:

A number of batches were chosen, divided by age, gender, and network. The choices for age were each year of age from 19 and 27, inclusive. The choices for gender were “male” and “female”. The choices for network were “Calgary, AB” and “University of Calgary”. Profiles of interest could be registered with either network, or both networks simultaneously, but they cannot be registered as both male and female simultaneously, nor in multiple single year age groups simultaneously.

Each batch consisted of profiles matching a particular combination of age group, gender and network (for example, all 27-year-old males in the “Calgary, AB” network). Facebook’s profile browser tool was instructed to return a list of all members in alphabetical order by name whose profiles matched the criteria for the batch. DownloadThemAll was instructed to download and store matching profiles as HTML files on a local computer workstation.

The Facebook profile browser was limited to displaying a maximum of 550 profiles for any particular query. Where the number of profiles returned for a batch exceeded 550, the profile browser was then used to obtain a random list of individuals matching the criteria for the batch, not sorted in any way. Those profiles were then downloaded

Two additional batches were also subsequently collected, repeating the procedure above for both the “Calgary, AB” and “University of Calgary” networks, but with no age or gender specified, and results in random order. 1,405 profiles and 1,499 profiles were retrieved for the two networks, respectively. In conjunction with the publicly disclosed total size of those two networks (90,222 and 11,620 members, respectively), these samples enable an approximate age and gender distribution of the social networking website’s user population to be calculated.

A subsequent data collection, using the same profile browser tools as above, but with only the place of employment specified, was used to collect all available profiles for individuals working at several particular firms of interest. Firms were deemed interesting if they appeared multiple times in the previous batches.

### Data Extraction and Recording

Matched field-value pairs from each downloaded profile were extracted using a custom text parser programmed by the author in the PHP scripting language. The data extraction method was based loosely on web data parsing and extraction techniques developed by Camacho *et*

al. (2004), which were enhanced by the author to handle the specific data format used by the Facebook website. The enhancements consisted primarily of using POSIX-style regular expressions as the matching token instead of fixed strings (as in Camacho), and were required to distinguish between multiple similar semantic HTML data structures which the Camacho technique did not directly handle. (Source code for the parser is available from the author's website.)

An extracted field-value pair has the form *field-name: field-value*, such as "Sex: Male", "Groups: 13", or "Activities: Writing, photography, coding, design".

Once extracted, the matched field-value pairs associated with each user profile were automatically inputted by the text parser into a relational database running the MySQL software, with each profile being assigned a unique "UserID" which is attached in all instances to data originating from that user's profile. The social networking website's unique user identifier was also recorded in the database for deduplication purposes. Duplicate profiles (obtained, for example, when a user is a member of both the "Calgary, AB" and the "University of Calgary" networks and appeared in the random samples of each network) were detected by comparing the unique user identification number of each new profile to be entered into the database, with all the unique user identifiers currently known to the database. If a match occurred, the profile would not be recorded to the database again.

The accuracy of the extraction and recording process was verified by testing the process on several cases covering extreme and edge conditions for data variety, and then comparing the imported results in the database with the actual contents of the profile. For example, the test cases included at least one profile listing multiple graduate schools, another listing one, and a third not listing any graduate school information at all.

Not all possible information was imported from each profile. For example, relationship status and gift box contents serve no immediate purpose for this study, and so were not parsed, recorded or verified, although extending the parser and database to include those fields at a future time would be trivially simple.

## Confidentiality

The data used in this study is publicly available and disclosed with the explicit consent of the users on the social networking site. Users may choose to disclose none, some, or all details of their profiles, such as work,

education, birthdate, address, preferences etc. to other registered users. In the data extraction process, the name of each individual is stored but not used in analysis, while, as stated above, the unique identification number given by the social networking website to each user is stored and used to identify duplicate records describing the same individual occurring in multiple batches of data (as may occur when an individual sampled in the "Calgary, AB" and "University of Calgary" networks, and also has an employment history with one or more firms of interest).

Although it may be practically possible to backtrack to particular individuals based on combinations group affiliation, education and employment, etc., the search tools website itself offers more efficient and direct methods to find particular individuals based on any desired set of parameters for whatever benevolent or malicious intents.

## Post-Processing

After the data was imported, a semi-automated rule-based process was used to normalize data in free-form fields where multiple different textual strings were used by different users to refer to the same thing. For example, the institutions "Southern Alberta Institute of Technology", "SAIT", "S.A.I.T." and "S. Alberta Tech" are used to denote the same organization, and were normalized to "Southern Alberta Institute of Technology" (which was the preferred denotation by the majority of users). The pattern used for each normalization step was recorded in a separate file. This normalization was performed by identifying common mis-spellings and synonyms by visual inspection of data sorted alphabetically by the field to be normalized, and defining regular expressions which would match (or near-match) and replace text in the matched fields with a normalized version. Firm names and religious affiliation were the only data on which these normalizations were performed. Depending on the item in question, anywhere from a single entry to almost half the entries referring to the same item deviated in spelling from the most common form.

A number of other post-processing functions involving subjective input were also performed on the data, with the results recorded in the same database using the same keys but different columns as the parsed data. Post-processing steps included: a) classifying and recording key parent firms by visual inspection and regular expression matching, b) classifying and recording key group and religious affiliation by visual inspection and

regular expression matching, c) automated counting of the number of the number of activities and interests listed, d) automated counting of the number of creative attributes in the personal and employment aspects of each user's profile, by regular expression matching with lists of key word roots as described next.

## Defining the creatives

Richard Florida (2005, pg. 27–28) offers the following list of creative occupations: architects, aesthetic workers, engineers, scientists, artists, writers, high-end managers, planners, analysts, health care, finance, law. That list is nearly sufficient for the purposes of this analysis, and forms the basis of a dictionary used to automatically identify individuals engaged in creative occupations.

Listed work information (that is, the combination of a position and an optional description in a user's free-text profile) matching one of Florida's occupations directly (e.g. lawyers, engineers), or indirectly (e.g. paralegals practice law; designers work with some combination of aesthetics, engineering, planning) was used to categorize the individual's employment as being creative.

In addition, two categories were added—knowledge workers and people who create things with knowledge—since their work is creative and uses information as a starting material in assembly where the product creation algorithm allows for variation based on professional judgment.

Any profile of an individual matching the expanded definition of creative in either the employment position or job description was identified using regular expression matching and marked as describing an individual employed in a creative capacity. The number of occurrences of those words was also recorded.

Separately, two other creative identifications were performed using the same technique, matching contents of the "Activities", "Interests" or "About Me" sections of profiles instead of matching work information.

First, individuals whose "Interests", "Activities" or "About Me" sections were classified as self-identified creatives if they contained any of the following words (or their roots): "artistic", "clever", "cool", "demiurgic", "deviceful", "fertile", "formative", "gifted", "hep", "hip", "ingenious", "innovational", "innovative", "innovatory", "inspired", "inventive", "original", "originative", "productive", "prolific", "stimulating", "visionary", and, "way out". The list of words was drawn from the entry for "creative" in Roget's New Millennium Thesaurus, First Edition (2007). The number of occur-

rences of those words was also recorded.

Second, individuals who included one of the following words or their roots (drawn from the same thesaurus for the entries for "artistic" and "original") in their "Interests", "Activities" and "About Me" sections were classified as externally identified creatives. For "artistic", the words were: "aesthetic", "creative", "cultivated", "cultured", "decorative", "dramatic", "elegant", "exquisite", "fine", "graceful", "grand", "harmonious", "ideal", "imaginative", "musical", "ornamental", "pictorial", "picturesque", "pleasing", "poetic", "refined", "rhythmical", "satisfying", "sensitive", "stimulating", "stylish", "sublime", and, "tasteful". For "original", the words were: "avant-garde", "causal", "causative", "cherry", "conceiving", "creative", "demiurgic", "devising", "envisioning", "fertile", "formative", "fresh", "generative", "imaginative", "ingenious", "innovational", "innovative", "innovatory", "inspiring", "inventive", "new, novel", "originative", "productive", "quick", "ready", "resourceful", "seminal", "sensitive", "unconventional", "unprecedented", "untried", "unusual".

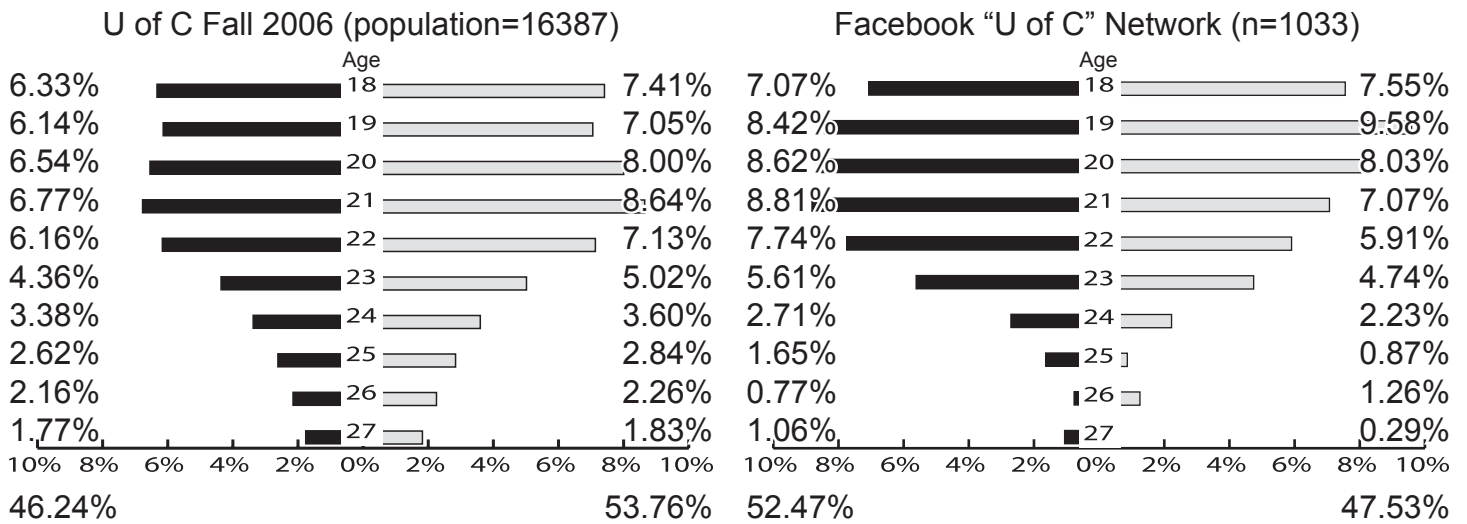
To prevent multiple-counting, words appearing in the 'creative' list of words and either or both of the 'artistic' or 'original' lists of words were retained in the 'creative' list, and removed from the latter two.

Together, it was hoped that these definitions of members belonging to the creative class encompassed "those creative humans using creativity to benefit the economy" (Florida 2005, pg. 35).

## Subsequent data exploration

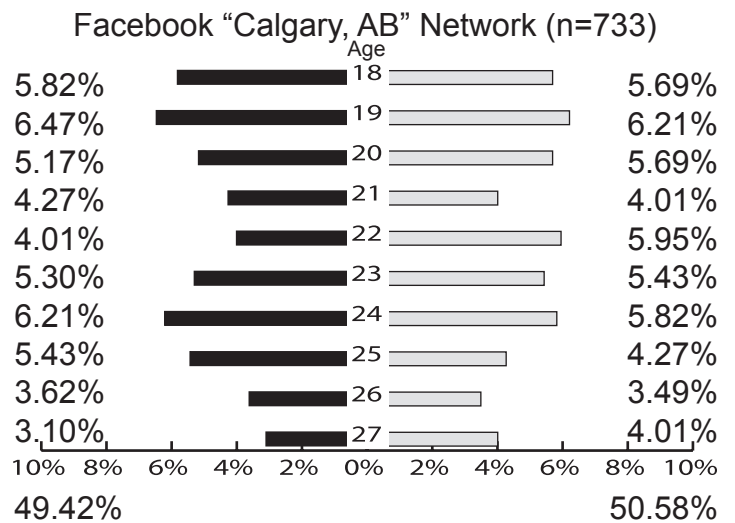
After the data was gathered, a number of subsequent data manipulations were performed to count things of interest in the original data. For example, the number of group memberships and activities, and the major religious affiliation were calculated and recorded into the database, into new columns after the initial data entry process.

A number of measures were created based on counts of data in the database. Some of these measures were paired and their correlations automatically calculated (in a manner similar to that suggested by Geng and Hamilton, 2007) and graphed and then manually reviewed to ensure that expectable correlations were present, and to ensure that nothing unexpected distinguished the general population or creative subpopulations. Findings from these correlations are discussed in the next section.



Source: University of Calgary Office of Institutional Analysis.

**Figure 2.** Age and gender distributions of University of Calgary population, and random Facebook samples of the “Calgary, AB” and “University of Calgary” networks.



## Results

Note: Users of Facebook can limit the amount of profile information displayed to individuals not designated as “Friends”. Therefore, not all users returned by the profile browser had accessible profiles, and not all parts of all profiles were publicly visible or populated. Therefore, in this section, we consider only those profiles which provide data in the fields under analysis.

### What the indicators tell us

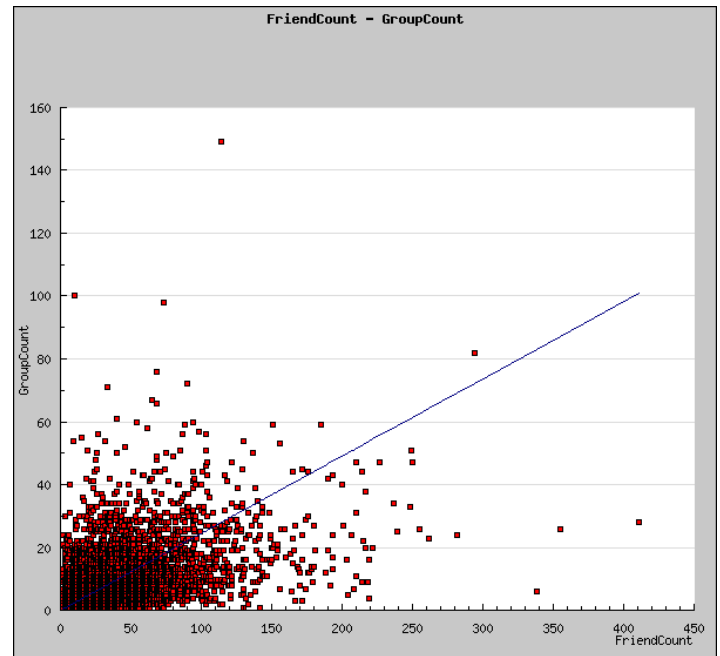
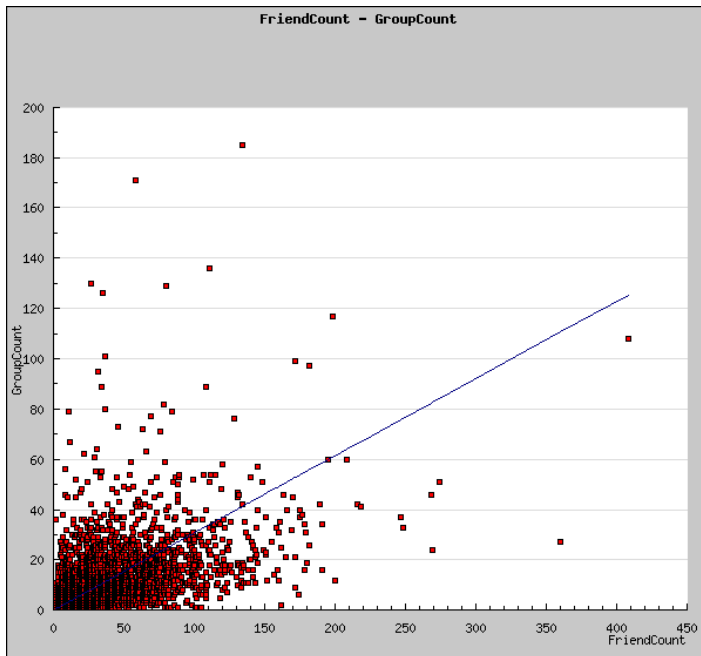
A number of objective indicators contained within the Facebook profiles can tell us things about their owners:

**Wall** – Each user’s “Wall” is a place where friends can communicate information to the user in a public way (only individuals designated as friends on Facebook can write to other friends’ walls). Wall activity indicates active friendship and trust-building activities.

**Groups** – Facebook allows users to create and join

any number of voluntary groups which may pertain to almost any topic. Joining a group means that, minimally, the user sought out a group, or noticed that a friend had joined it, and is an indicator of community participation.

**Activities** – The “Activities” section of each user’s profile is a place where the user may supply free-form lists and descriptions of things they do. Activities are similar to “**Interests**”, another free-form field where users can describe their interests. Both are similar to the “**About Me**” section, which is intended to allow other free-form input from the user. Users sometimes include overlapping information in those three sections, or refer to one section from within another. In this analysis, we treat all three, Activities, Interests, and About Me, as being sources of information of approximately equal strength. We search Activities, Interests and About Me for words that relate to creativity, as well as count the elements in those lists to determine to what magnitude a user participates in



Creatives:

$[r_2] \Rightarrow 0.17492936014223$

n=2807

Non-Creatives:

$[r_2] \Rightarrow 0.1960712981371$

n=6135

**Figure 3.** Number of Friends vs Number of Groups for Facebook Creatives (left) and non-creatives (right).

creative activities.

**Work** – The work and “**Education**” sections of a profile give organizational, locational, and temporal information about a profile owner’s work and education history. They both allow positions/programs to be described using free-form text.

**Books, Movies, TV** – Users may put any text in these categories. Most users include lists of favourite works and artists. In future, books might be used as a proxy for gaining new personal knowledge, but for this study, the sections are given equal weight. The lengths of these lists, rather than the qualitative nature of their contents are counted here.

### Parting the Facebook creatives

Initial analysis divided the population into two classes: “Facebook Creatives”, and non-creatives. Facebook Creatives were defined as anyone having at least one of the following characteristics: a) defined themselves as “creative” (or one of the synonyms) through the Activities, Interests or About Me sections of their profiles; b) listed “artistic” or “original” (or one of their synonyms) in the Activities, Interests or About Me sections of their profiles; c) listed one or more creative activities as listed in the Activities, Interests or About Me sections of their profiles; d) had a job title or description which matched

the expanded Florida definition of creative occupations given previously. Non-creatives are Facebook users having none of the qualities of the Facebook Creatives defined above.

## General observations

### Scope of coverage

According to Facebook, the number of users in the “Calgary, AB” network has 90,222 members as of April 2007, while the “University of Calgary” network has 11,620 members. The two random sample sizes of 1,405 members of the “Calgary, AB” network and 1,499 members of the “University of Calgary” network account for 1.557% and 12.90% of the populations using Facebook, respectively. In total, 10,962 profiles were obtained from the “University of Calgary” and “Calgary, AB” networks.

### Population distributions

The population distribution recorded by the University of Calgary showed some differences with the randomly sampled Facebook population belonging to the “University of Calgary” network, although they are in general agreement (see Figure N). First, the University of Calgary population has more females than males overall,



by 7.53%, whereas in the randomly sampled Facebook “University of Calgary” population, there were 4.53% fewer females than males. Second, the actual University of Calgary population peaks doubly at 18 and 21 years of age for both males and females, while the randomly sampled Facebook population peaks only once at 19 for the female population, and only once at 20 for the male population. In the randomly sampled Facebook population, the number of users reporting their age drops dramatically at years of age greater than 24, while the trend is more gradual for the actual University of Calgary population.

These differences in the two distributions are explained at least in part by the fact that participation in Facebook by U of C students is incomplete, and that not all participants list their age on Facebook.

Overall, the randomly sampled Facebook population is younger than the actual University of Calgary population by six months, with mean ages of 20.81 and 21.36 years, respectively. The University of Calgary population and the “University of Calgary” Facebook sample are satisfactorily comparable for ages 18 to 23 or 24.

Compared with the general “Calgary, AB” population as sampled on Facebook, which is bimodal around 19 and 24 years of age, both University of Calgary samples are less lumpy and younger. The average age of the “Calgary, AB” sample was 22.14 years. With respect to gender, the “Calgary, AB” population is more balanced in aggregate than either the University of Calgary or Facebook “University of Calgary” populations, although clear swings of the data in both directions are visible at different age groups in the “Calgary, AB” population.

### Counting creatives

From the entire data set, 3,213 Facebook creatives were identified, 7,749 non-creatives were identified. Of those, 2,761 have been employed.

From the random samples of the “Calgary, AB” and “University of Calgary” Facebook networks only, 22.4 of the “Calgary, AB” network members were classified as creative while 27.9% of “University of Calgary” network members were classified as creative.

### Obvious correlations

A number of correlations come as no surprise: Profiles in either the Facebook Creative or non-creative classes which listed many friends also participated in many groups, and interacted with friends via their messaging walls ( $R^2$  values of 0.17 to 0.31,  $n > 2800$ , with non-creatives having slightly higher  $R^2$  values, see Figure N

for an example: “Number of Friends” plotted against “Number of Groups”). Individuals in either class who listed many movies of interest also listed many TV shows of interest ( $R^2 = 0.186$  for non-creatives,  $n = 2,518$ ;  $R^2 = 0.144$ ,  $n = 2,627$  for creatives), and also, but to a lesser correlation, books of interest ( $R^2 = 0.133$ ,  $n = 2607$  for non-creatives,  $R^2 = 0.135$ ,  $n = 2327$  for creatives). These observations do not say much about Facebook Creatives in particular, or how they interact with the local innovation environment.

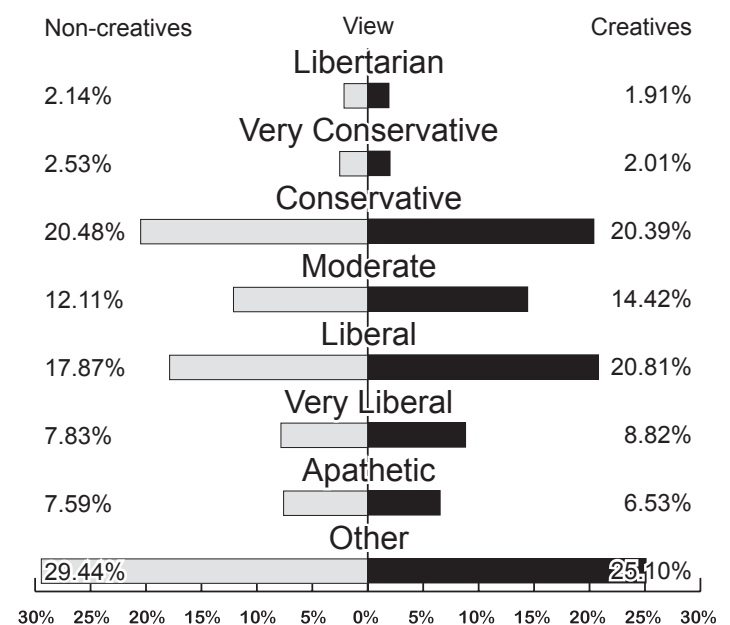
### An a-political affair

As a group, Facebook Creatives express slightly more moderate to liberal views than their non-creative counterparts. However, their political views are not significantly different from each other (see Figure 4).

Non-creatives make up 25.16% of the libertarian-very conservative part of the political spectrum, while Facebook Creatives compose slightly less at 24.31%. By contrast, 27.34% of Facebook Creatives declared liberal or very liberal views, while only 25.70% of non-creatives chose liberal or very liberal for their views. Also, more non-creatives consider themselves apathetic or of another political persuasion than Facebook Creatives.

Although it may be tempting to speculate about conservative creatives and liberal dullards, the lack of spe-

**Figure 4.** Political views of non-creatives and creatives.  $n = 4711$  (2568 non-creatives, 2143 creatives).



cificity in the labels obscures any distinction that users may wish to make with respect to sharing views about various political spectra on their profiles. For example, the concept that a person may be socially liberal and financially conservative is not expressible on the social networking website's political views measure, yet many Albertans would consider themselves to be "Red Tories" fitting that description.

Also of cautionary note are the large and unbalanced portion of both creatives and non who would consider themselves to be of "Other" political leaning, with no clear indication of what exactly that category encompasses.

### More meaningful correlations

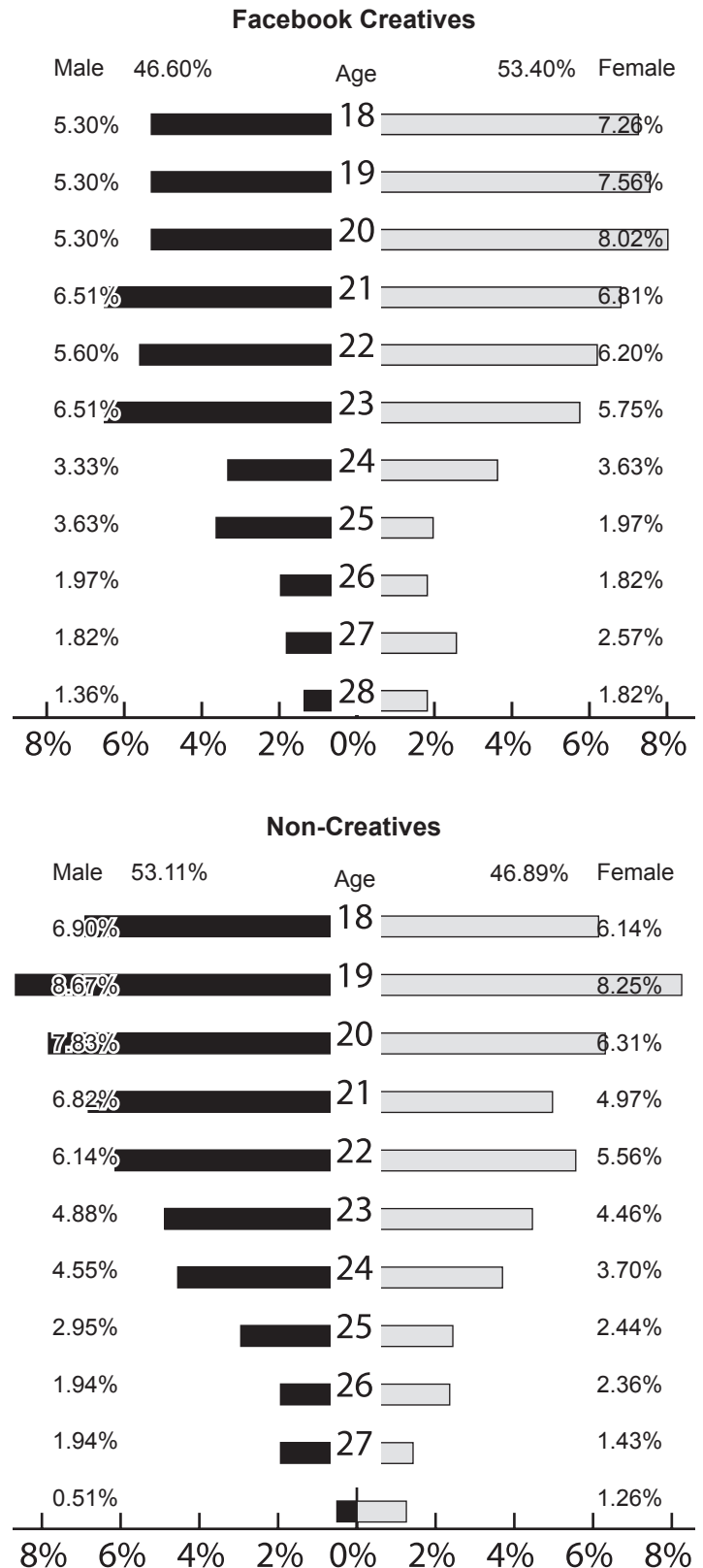
A much more meaningful number is obtained by examining the relationship between the number of times a person has been employed to the total number of creative descriptors in the positions or descriptions of those jobs. Those two measures give us a positive relationship with an  $R^2$  value of 0.156 for non-creatives ( $n=1,268$ ), and 0.168 for Facebook Creatives ( $n=2,035$ ).

Creatives, as measured by creative activities and creative work, also have interests in larger quantity than non-creatives ( $R^2=0.079$  for non-creatives vs  $R^2=0.112$  for Facebook Creatives).

Perhaps more importantly, the lack of meaningful relationships between most of the basic pairings of measures shows that they are sufficiently independent as potential indicators of creativity.

Potentially interesting non-correlates for both Facebook Creatives and non-creatives included lengths of activity lists and the number of colleges attended (it's plausible to hypothesize that exposure to new colleges would also expose the individual to new activities), and the number of musical interests and number of friends (apparently, one doesn't make friends on Facebook by virtue of the number of different bands one enjoys).

The number of activities a person participated in did not relate at all to the number of posts on their wall ( $R^2=0.000$ ,  $n=2,112$  for non-creatives;  $R^2=0.010$ ,  $n=2,322$  for creatives). (Note that more Facebook Creatives than non-creatives both listed activities and interacted with friends, but that is a side-effect of using activities to help identify creativity.) This was somewhat unexpected in that social capital theories generally agree that participation in group activities generally follows, or is followed by building new relationships with friends. Examining the data in more detail in a future study, with



**Figure 5.** Gender distributions among randomly selected creatives (top,  $n=661$ ) and non-creatives (bottom,  $n=1,188$ ) in the "Calgary, AB" and "University of Calgary" networks.

respect to categorizing the individualistic or group nature of groups listed, may reveal some relationship with social capital.

### Gender balance of creatives

An interesting divergence was observed with respect to the Facebook Creative and non-creative populations distributed by age and gender (see Figure 5). Among Facebook Creatives, females make up a 6.80% majority for all ages except 23-25, while among non-creatives, males have a 6.22% majority over females, but are less well represented at ages 21, 26 and 27. Overall, the Facebook Creative sample appears to have a less balanced population distribution than does the non-creative population.

Some of the variance may be explainable by the fact that the Facebook Creative population is in part defined

by employment, where older individuals tend to have more employment roles and hence more opportunities to have had a creative role. However, removing creative employment from the definition of Facebook Creatives increased the overall lead of females to 12.02% over males among Facebook Creatives, and increased the gap to 7.41% of males over females among non-creatives. Employment appears to act as a dampening factor of the gender difference.

The relationship between gender and creativity may be the topic of future research.

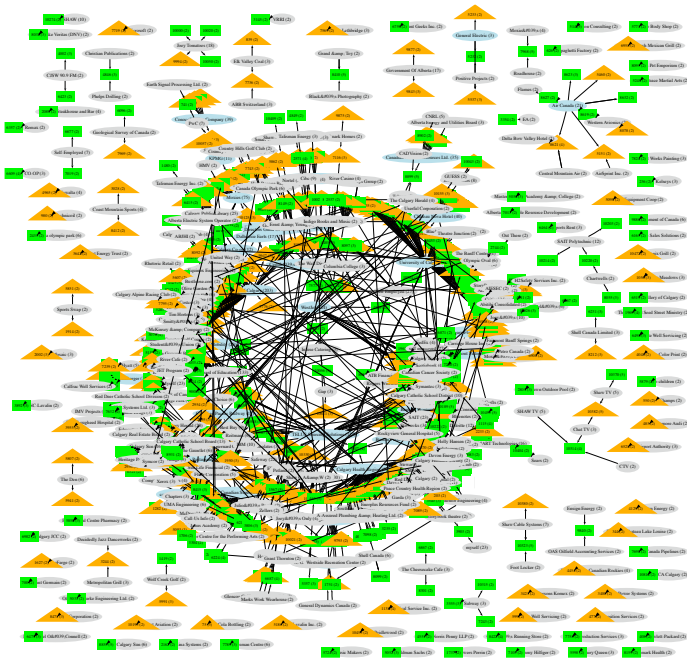
### Firms, creatives, and networks

Firms in this study were more easily connected to each other through creatives than non-creatives. As Figure 6 shows, when considering all employees who have had

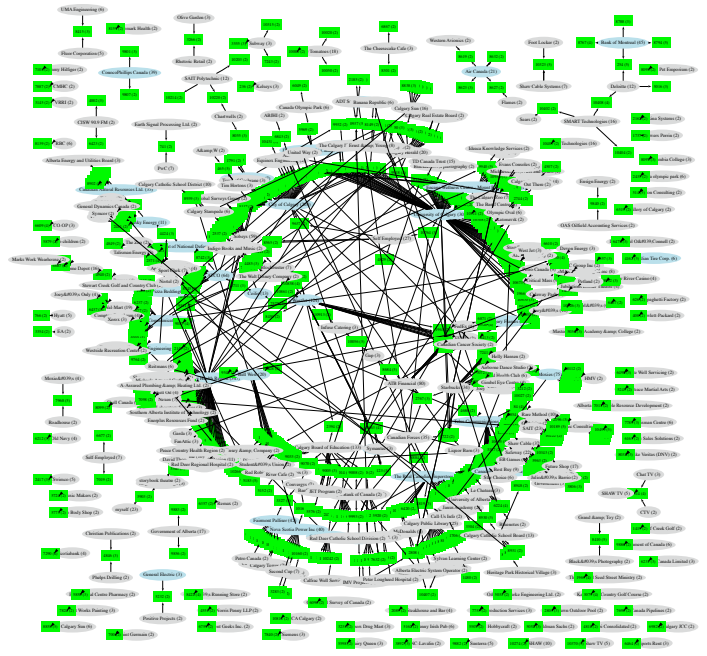
**Table 1.** Network linkages vs. firm size for firms employing at least two people. Creative definition includes individuals who work at a creative position.

Min. # of employees	Creative links	Non-cr. links	Links ratio C/NC	Firm links	Creative workers	Non-cr. workers	Workers ratio C/NC	Firms
2	673	792	0.849	1465	418	492	0.849	395
3	562	655	0.858	1217	385	439	0.876	211
4	518	597	0.867	1115	365	420	0.869	150
5	483	577	0.837	1060	354	412	0.859	127
6	474	560	0.846	1034	349	406	0.859	115
7	449	524	0.856	973	338	390	0.866	94
8	422	505	0.835	927	323	382	0.845	80
9	411	491	0.837	902	318	377	0.843	75
10	404	463	0.872	867	315	365	0.863	66
12	390	439	0.888	829	306	352	0.869	56
14	366	411	0.890	777	296	334	0.886	48
15	361	406	0.889	767	293	329	0.890	46
16	358	401	0.892	759	292	326	0.895	44
18	334	381	0.876	715	277	316	0.876	37
20	330	371	0.889	701	273	310	0.880	35
22	309	348	0.887	657	259	293	0.883	31
24	292	334	0.874	626	246	287	0.857	27
26	283	328	0.862	611	238	282	0.843	26
28	274	319	0.858	593	230	276	0.833	23
30	274	319	0.858	593	230	276	0.833	23
35	258	308	0.837	566	219	269	0.814	21
40	245	277	0.884	522	211	246	0.857	16
45	223	257	0.867	480	190	229	0.829	13
50	221	248	0.891	469	188	222	0.846	12
60	210	245	0.857	455	179	219	0.817	11
70	196	224	0.875	420	166	202	0.821	9
80	189	212	0.891	401	159	191	0.832	8
90	160	195	0.820	355	134	174	0.770	6
125	142	182	0.780	324	119	164	0.725	5
140	118	145	0.813	263	98	132	0.742	3
210	93	108	0.861	201	77	99	0.777	2
310	35	61	0.573	96	31	55	0.563	1

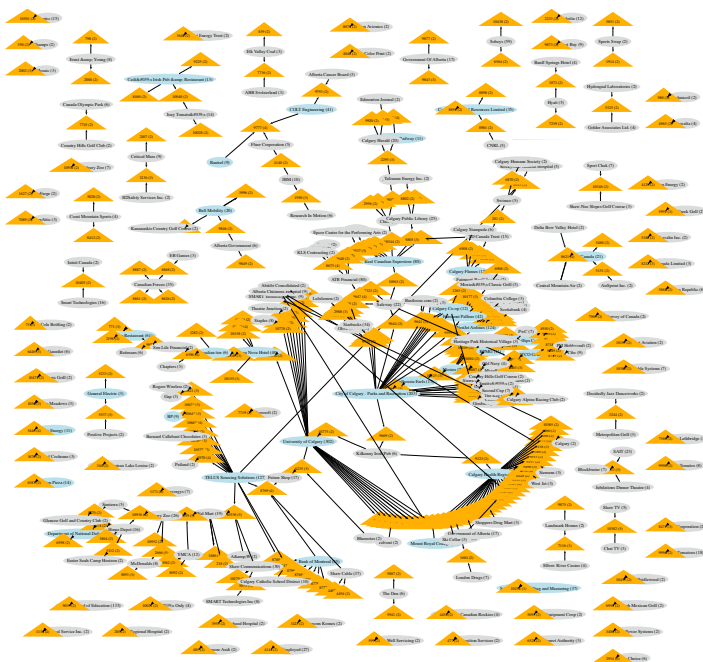
\* Successive rows with no changes in links or workers omitted



**Combined:** Creative links: 1009 Non-Creative links: 456  
 Firm links: 1465 Creatives represented: 623 Non-Creatives  
 represented: 287 Firms represented: 403



**Creatives:** Creative links: 1009 Firm links: 1009 Creatives  
 represented: 623 Firms represented: 323



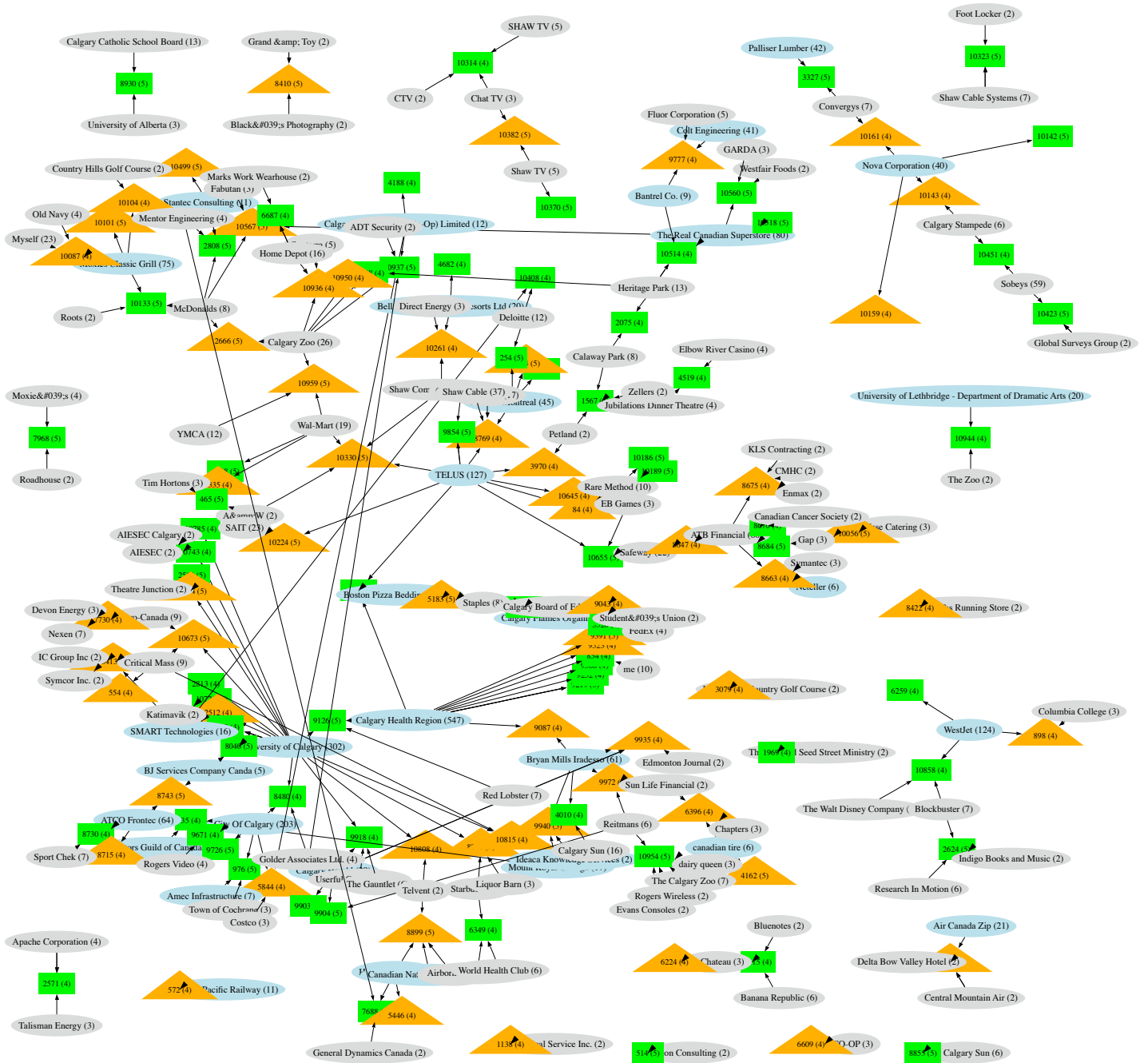
**Non-Creatives:** Non-Creative links: 456 Firm links: 456  
 Non-Creatives represented: 287 Firms represented: 208

**Figure 6.** Firm connectivities via employees in the Calgary area as represented on Facebook. Only firms employing more than one person on Facebook and users who have had at least two employment experiences are shown. Blue ovals are firms for which all profiles of employees were obtained from Facebook. Grey ovals are all other firms. Green squares are creatives. Orange triangles are non-creatives. The number in the parenthesis next to a firm indicates total number of people employed and found in FaceBook, the number in parenthesis next to a person indicates total employments. (Readers of the PDF version of this document can find exquisite detail by zooming in.)

more than one employment experience and firms employing more than one person, each Facebook Creative provide more than twice as many links between firms studied than do non-creatives (recall that the default definition of creative includes counts of job creativity). Excluding job creativity from the definition of creative, the number of interfirm links and firm representations generated by the Facebook Creative and non-creative groups becomes exactly proportional to the size of

the groups. Of those studied, 251 firms were linked by 418 creatives (0.6005 firms per non-creative), while 295 firms were linked by 492 creatives (0.5995 firms per creative). The result is similar if we increase the minimum number of employments to three (0.476 firms per creative, 0.500 firms per non-creative), four (0.627 firms per creative, 0.666 firms per non-creative).

The most striking feature to note about this comparison is that people employed in creative jobs are



**Figure 7.** A much more readable maps of 151 Calgary firms and workers (creative workers include the creative work definition) with employment instances less than 4 removed. Most of the combined network as in Figure 6 remains connected, but the majority of connections (166 out of 298) are made by a 71 creatives (60 non-creatives are also represented). Nodes are labelled in the same manner as in the previous figure.

**Table 2.** Network linkages vs. number of employments for firms employing at least two people. Creative definition includes individuals who work at a creative position.

Min. # of employments	Creative links	Non-cr. links	Links ratio C/NC	Firm links	Creative workers	Non-cr. workers	Workers ratio C/NC	Firms
1	1560	2672	0.583	4232	1305	2372	0.550	554
2	673	792	0.849	1465	418	492	0.849	395
3	317	324	0.978	641	157	163	0.963	242
4	166	132	1.257	298	71	60	1.183	151
5	83	49	1.693	132	34	21	1.619	88

responsible for such a large fraction of inter-firm linkages (68.8% in the example in Figure N). This indicates that a firm's connectivity to other firms via workers is largely related to how it defines its job roles (or rather, how employees reported about the creativity of their jobs in the information provided to Facebook). The comparison also indicates that creative activities and interests of employees do not on their own translate into more inter-firm connectivities since varying those attributes of employees did not affect inter-firm linkages.

It bears repeating that self-identification of creativity, or creativity identified through non-vocational activities have little effect on inter-firm linkages via staff mobility within the network studied. This is surprising not just because it implies that personal creativity and professional creativity are distinct and separable characteristics, but because externally-identified creativity is positively linked to worker mobility (number of employment experiences), which should help to build inter-firm linkages.

Table 1 lists the number of Facebook Creatives (defined by including creative work). The most obvious feature is that the ratio between Facebook Creatives to non-creatives falls steadily with an increasing number of employees, but peaks at 0.895 and 0.846 creatives per non-creative when a firm has 16 employees listed on Facebook, and again when the number of employees is 50, respectively. That is despite the fact that within the population, the ratio between Facebook Creatives and non-creatives is 0.629. At firm with more than 50 listed employees on Facebook, the ratio slowly declines, reaching a ratio closer to that predicted by population alone: 0.563. Lesser peaks are also observed at 80 and 210 employees listed. A similar pattern holds for the number of linkages originating from creatives to firms, and the first peak also holds for links into firms.

Such local islands of peak creative employment and linkage are difficult to explain within the current data

set and warrant further investigation.

Turning our attention to Figure 7 of the network of 151 Calgary area firms and employees, we find that removing employees who have worked at fewer employment opportunities from the network reduces the number of non-creatives than the number of Facebook Creatives. Examining Table 2, from which the data for Figure 7 is drawn, it is readily visible that non-creatives dominate both in numbers and linkages when considering individuals with single employment experiences, whereas at two employment experiences and above, the Facebook Creatives clearly contribute proportionally more numbers and linkages than non-creatives. This reaffirms the previous observation that externally observable personal creativity is positively linked to the number of employment opportunities experienced, although it does not reveal the direction of causation.

**Table 3.** Calgary Firms with employees on Facebook with five employment instances.

Firm	Number of employees	Number of Creatives
Shaw Communications	5	3
McDonalds	4	1
Calgary Zoo	3	1
Wal-Mart	3	0
A&W	2	0
Ajilon Consulting	2	0
Calgary Board of Education	2	1
Critical Mass	2	0
Gap	2	1
Global Surveys Group	2	2
Grand & Toy	2	0
Rare Method	2	2
Reitmans	2	2
Staples	2	1
Student's Union	2	0
Tim Hortons	2	0
(59 other firms with a single employee not shown)		

Table 3 lists several of the 88 firms in the last row of Table 2 having at least two employees each with five previous employment experiences.

In it, we see an uneven distribution of Facebook Creatives and non-creatives among firms. This is to be expected since the number of firms having at any point employed an individual with five employment experiences far exceeds the number of Facebook Creatives who have had that much previous employment experience. This again suggests that the pool of Facebook Creatives are highly mobile and circulate to multiple firms.

### Firm connectivities and employment

It is clear from Table 4 that the majority of interfirm connectivities for firms represented by Facebook users are through Facebook Creatives, as opposed to non-creatives. Indeed, when the links are tallied, we find that a minority of the working population, consisting of those individuals labelled as Facebook Creatives, are responsible for the majority of inter-firm links. Combined with the previous observation that creative work is also positively linked to multiple employment

instances, we can draw the conclusion that having creative positions at firms are a mixed blessing.

Firms offering the most creative positions can benefit from new tacit information gained through the circulation of creative individuals within the workforce. However, the mobility of the workforce itself, and particularly of creatives, results in significant staff turnover.

In the survey, 19.7% of individuals were found to have multiple employment experiences (out of 5,573 individuals with at least one employment experience), with 9.1% of them belonging to Facebook Creatives (defined without considering previous employment in a creative capacity).

Also interesting to note is that in the all-inclusive dataset, creative occupations made up 28% of all jobs. Indeed, even though non-creatives outnumber creatives by an approximate ratio of 2.6 to 1, at least twice as many creatives are multiply employed as creatives in all multiple employment categories (see Table 4). Recalling Table 2, there is approximately one Facebook Creative for every non-creative in firms among employees with two or more employment experiences.

**Table 4.** Employment instances vs creativity

Creatives			Non-Creatives		
Instances	#	percentage	Instances	#	percentage
5	51	1.21%	5	8	0.0%
4	58	1.37%	4	24	0.36%
3	168	3.97%	3	64	0.95%
2	486	11.49%	2	238	3.54%
1	2212	52.28%	1	2264	33.64%
0	1256	29.69%	0	4133	61.40%
Total	4231	100.00%	Total	6731	100.00%

Creatives excluding work			Non-Creatives excluding work		
Instances	#	percentage	Instances	#	percentage
5	37	1.15%	5	22	0.28%
4	39	1.21%	4	43	0.55%
3	103	3.21%	3	129	1.66%
2	333	10.36%	2	391	5.05%
1	1445	44.97%	1	3031	39.11%
0	1256	39.09%	0	4133	53.34%
Total	3213	100.00%	Total	7749	100.00%

## Creative employment implications for education

Table 5 compares Facebook users attending the University of Calgary and the Southern Alberta Institute of Technology (SAIT) with respect to the number of creative job indicators in their derived from their attendees profiles. (A creative job indicator is simply one match in either a Facebook user's job title or description against the expanded list of creative occupations described previously. Thus, a person titled "Industrial Design Manager" with a description containing "engineering samples" would gain three creative job indicators from that employment, while another user whose job title was simply "Data Analyst" would gain only one indicator. Indicators are cumulative across jobs.)

The table shows that while SAIT attendees on Facebook have fewer non-creative jobs (total number of creative job indicators for a person equals 0), almost 6% more SAIT attendees on Facebook than University of Calgary attendees have been employed in a job with a single creative indicator, and slightly more SAIT attendees have held a job with two creative indicators than those attending University of Calgary. However, University of Calgary attendees outnumber SAIT attendees in having experienced jobs contributing a total of more than two creative job indicators.

This implies that, on average, University of Calgary attendees are either employed in a larger number of creative roles than their SAIT counterparts, or each of their employments had more creative indicators than individual employments by SAIT attendees.

### Grad school vs creative/employment

Table 6 compares Facebook users attending a graduate

Creative job Indicators	U of C		SAIT		Total
0	4161	85.07%	526	77.24%	4687
1	508	10.39%	109	16.01%	617
2	153	3.13%	27	3.96%	180
3	43	0.88%	12	1.76%	55
4	16	0.33%	3	0.44%	19
5	6	0.12%	2	0.29%	8
6	3	0.06%	2	0.29%	5
7	1	0.02%	0	0.00%	1
<b>Total</b>	<b>4891</b>		<b>681</b>		<b>5572</b>

**Table 5.** Number of creative job indicators per attendee for University of Calgary and Southern Alberta Institute of Technology.

school program and users attending an undergraduate program with respect to the number of creative job indicators in their derived from their attendees profiles. In general, its features resemble those of the previous table comparing the creative employment of University of Calgary and SAIT attendees. Comparing graduate degree program attendees with undergraduate degree program attendees reveals that proportionately more graduate school attendees have a higher number of creative job indicators than undergraduate attendees, with graduate school attendees holding a proportionally larger percentage of jobs with three or more creative indicators than undergraduate attendees.

Data extracted from collected profiles includes a person job title and post-secondary education program focus, which could provide a valuable platform for further studies of the relationship between level and focus of education and creative employment in the professional world.

### Words matter

Some words used to describe creative occupations and activities were more effective than others, as shown in the heat maps in Tables 7 and 8. Florida's list of creative occupations had a high success rate matching positions in Facebook Creatives' profiles, particularly among manager and planner occupation categories, but relatively less success with respect to matching words used in job descriptions. Choosing one of the alternative occupation categories for some words (for example, categorizing programmer or marketing occupations as analysts) does not significantly change the overall hot spots on the map.

Creative job Indicators	Graduate		Undergraduate		Total
0	746	75.97%	6421	77.32%	7167
1	153	15.58%	1252	15.08%	1405
2	42	4.28%	390	4.70%	432
3	15	1.53%	118	1.42%	133
4	8	0.81%	51	0.61%	59
5	10	1.02%	37	0.45%	47
6	5	0.51%	29	0.35%	34
7	0	0.00%	1	0.01%	1
8	0	0.00%	3	0.04%	3
9	1	0.10%	0	0.00%	1
10	2	0.20%	2	0.02%	4
<b>Total</b>	<b>982</b>	<b>100.00%</b>	<b>8304</b>	<b>100.00%</b>	<b>9286</b>

**Table 6.** Number of creative job indicators per attendee for graduate school attendees versus undergraduate attendees.



The appended list of knowledge worker words tended to match more profiles than Florida's list, and did so more consistently in both the job positions and descriptions. The teaching and instructing roles tended to produce the most matches in the appended list, possibly due to the significant number of current and former University of Calgary education students within the "University of Calgary" and "Calgary, Alberta" Facebook networks.

Both Florida's list and the knowledge worker list, however, matched job positions better than job descriptions overall.

With respect to the personal profile heat map, creative activities drew the most hits in the "Activities" and "Interests" categories, and to a lesser extent, in the "About Me" category. It is not surprising at all that Facebook Creatives would list indisputably creative activities such as writing and design within their activities and interests.

Some caution, however, is required in the interpretation of these heat maps. Extremely high numbers of matches indicates some loss of selectivity. "Manager", from the occupations list, which recorded 339 hits, could match a top level research manager, or the assistant manager at a fast food establishment, or the assistant to the one of the managers just described. While all three roles require some form of creativity, the natures of their roles in employment and the local innovation system differ significantly. Lack of specificity also exists in the personal profile heat map. "Cool", "new" and "music" are the top hits from the "About Me" list, yet each can have little to do with creativity in any sense important to the study of innovation. Someone who enjoys listening to "cool new music", for example, would register as a creative, yet such a data point provides no analytical value.

### Self identification of creatives

As a rule, Facebook Creatives do not self-identify as such. Only 328 individuals sampled listed themselves as creative by including the word "creative" or its synonyms in the relevant sections of their profiles. And even when such individuals self-identified as being creative provided virtually no predictive power about anything. Particularly, they did not show any significant correlation to number of colleges attended ( $R^2=0.000$ ,  $n=250$ ), group membership ( $R^2=0.000$ ,  $n=230$ ), interests ( $R^2=0.004$ ,  $n=222$ ), participation in activities ( $R^2=0.000$ ,  $n=203$ ), musical preference diversity ( $R^2=0.006$ ,  $n=232$ ) nor direct interactions with friends ( $R^2=0.000$ ,  $n=317$ ). Again, the small sample size of self-identified creatives should be recalled.

Self-identifying as a creative did slightly correlate with *Habits of Online Creatives* – April 25, 2007

Word	Position	Description	Category	Total
analyst	179	1	analyst	180
draft	19	6	architect	25
architect	16	3	architect	19
art	35	25	artist	60
chef	22	4	artist	26
stylist	20	1	artist	21
photo	18	27	artist	45
carpenter	14	2	artist	16
music	13	26	artist	39
dance	13	15	artist	28
piano	11	4	artist	15
DJ	10	3	artist	13
flair	1	0	artist	1
creat	10	23	creating	33
engineer	176	31	engineer	207
network	26	21	engineer	47
programmer	18	1	engineer	19
finance	7	5	finance	12
nurse	203	39	health care	242
therap	39	5	health care	44
doctor	7	5	health care	12
legal	17	10	law	27
lawyer	2	2	law	4
manag	339	94	manager	433
consult	86	16	manager	102
owner	68	10	manager	78
president	28	1	manager	29
everything	18	45	manager	63
c?o	17	1	manager	18
chief	12	3	manager	15
founder	5	1	manager	6
marketing	52	13	planner	65
plan	38	74	planner	112
researcher	14	2	scientist	16
scientist	5	1	scientist	6
writ	29	37	writer	66
freelance	11	2	writer	13
teach	217	132	knowledge worker	349
instructor	117	12	knowledge worker	129
develop	87	58	knowledge worker	145
design	64	55	knowledge worker	119
train	59	52	knowledge worker	111
communic	32	13	knowledge worker	45
build	2	68	knowledge worker	70

**Table 7.** Heat map and count of words matching creative occupations and job descriptions. Darker colours indicate a higher number of matches. Words are given in their root form, if applicable. "art" matches "art", "artist", "artisan", etc.

multiple employment when all the singly-employed and singly-self-creative members were removed from the sample ( $R^2=0.049$ ,  $n=222$ ). This very small sample size of creatives overall ( $n=328$ ) does not strongly dispute that self-identified creativity is a poor predictive measure. Self-identification with creativity related positively with external identification ( $R^2=0.040$ ,  $n=226$ ) but incompletely. Almost one-third of self-identified creatives would not be externally identified as creatives through their activities alone.

Word	About Me	Activities	Interests	List	Total
writ	88	155	185	Activity	428
danc	79	428	332	Activity	839
making	57	119	122	Activity	298
cook	41	143	270	Activity	454
comput	36	61	164	Activity	261
design	27	24	87	Activity	138
english	20	5	26	Activity	51
photog	18	76	277	Activity	371
film	15	12	67	Activity	94
draw	12	73	88	Activity	173
fashion	11	8	93	Activity	112
gadget	2	0	13	Activity	15
music	126	291	1122	Artistic	1539
pict	45	34	45	Artistic	124
styl	19	2	12	Artistic	33
grand	17	6	5	Artistic	28
drama	16	7	22	Artistic	45
culture	14	5	83	Artistic	102
fine	7	2	16	Artistic	25
ideal	6	0	2	Artistic	8
poet	5	9	50	Artistic	64
imaginat	4	0	2	Artistic	6
sensitive	4	0	0	Artistic	4
decorat	3	5	18	Artistic	26
rhythm	1	2	0	Artistic	3
satisfying	1	1	1	Artistic	3
aesthetic	1	0	0	Artistic	1
graceful	1	0	0	Artistic	1
harmonious	1	0	0	Artistic	1
refined	1	0	0	Artistic	1
tasteful	1	0	0	Artistic	1
cool	96	28	22	Creative	146
creat	33	18	39	Creative	90
origina	22	1	5	Creative	28
hip	14	13	18	Creative	45
inspir	8	1	5	Creative	14
invent	3	3	3	Creative	9
artistic	2	3	2	Creative	7
innovat	2	1	1	Creative	4
clever	2	0	1	Creative	3
gifted	2	0	0	Creative	2
productive	1	1	2	Creative	4
stimulat	1	1	5	Creative	7
visionary	1	0	0	Creative	1
way out	1	0	0	Creative	1
new	195	97	210	Original	502
quick	13	3	2	Original	18
ready	12	3	4	Original	19
fresh	10	5	8	Original	23
novel	6	5	14	Original	25
unusual	2	0	1	Original	3
cherry	1	0	0	Original	1
resourceful	1	0	0	Original	1
unconventional	1	0	0	Original	1
unprecedented	1	0	0	Original	1
devis	0	1	2	Original	3
envison	0	0	1	Original	1

**Table 8.** Heat map and count of words matching “About Me”, “Activities”, “Interests”.

Words (and roots) omitted from map above with no matches:

**Artistic:** cultivated, eleganc, exquisite, ornamental, pleasing, sublime.

**Creative:** demiurgic, deviceful, fertile, formative, hep, ingenious, prolific

**Original:** avant-garde, causa, conceiv, generative, seminal, untried

## External identification of creatives

In contrast, many more individuals were identified as being creative through their activities, and such identifications had greater value in distinguishing them from non-creatives. External identification of creatives had some correlation with interests ( $R^2=0.099$ ,  $n=2,601$  for non-creatives;  $R^2=0.121$ ,  $n=2811$  for creatives), but weak correlation with respect to musical preference diversity ( $R^2=0.038$ ,  $n=2,841$ ). It had a similar relationship for the non-creatives ( $R^2=0.032$ ,  $n=2,628$ ), and it had slightly better predictive power with respect to activities ( $R^2=0.063$ ,  $n=2325$  for non-creatives;  $R^2=0.067$ ,  $n=2,512$  for creatives). External identification of creativity had no relationship to college attendance ( $R^2=0.000$ ,  $n \geq 2,885$ ), direct interaction with friends via their own Wall ( $R^2=0.001$ ,  $n \geq 2,726$ ) and exactly no correlation with the number of friends a creative person had ( $R^2=0.000$ ,  $n \geq 2,789$ ). Externally-identified creatives didn't join groups ( $R^2 \leq 0.009$ ,  $n \geq 2,558$ ) or acquire friends ( $R^2=0.000$ ,  $n \geq 2,789$ ), nor interact with them on a one-to-one level ( $R^2 \leq 0.002$ ,  $n \geq 2,726$ ) with them through the wall.

Creativity as identified through employment had intermediate to good correlative value overall. The degree to which a job is creative was strongly related to the number of times the person holding the job is employed ( $R^2=0.151$ ,  $n=1,268$  for non-creatives;  $R^2=0.180$ ,  $n=2,035$  for creatives), but this effect is less significant when non-creatives who only list a single employment instance and have only worked at one job are removed from the sample ( $R^2=0.040$ ,  $n=164$ ). Removing the same groups from the sample of Facebook Creatives results in the same type of effect, but to a lesser extent ( $R^2=0.101$ ,  $n=316$ ). This could indicate that creatives sample more jobs before settling on one, or that they are difficult to retain. Note, however that many creatives who are self-employed or employed on a contract basis may in reality go through several different jobs under the banner of a single employer or employment opportunity.

As with previous observations about the comparative predictive values of self-identification and external identification of creativity, the number of creative jobs held is only marginally predicted by self-identified creativity in Facebook Creatives ( $R^2=0.116$ ,  $n=305$  among Facebook Creatives;  $R^2=0.160$ ,  $n=449$  for non-creatives), and much more poorly by externally-identified creativity among Facebook creatives ( $R^2=0.021$ ,  $n=300$ ).

This is a key observation since it decouples personal creativity from professional creative employment.

# Further analysis and implications

## Validity and reliability

This dataset can be examined in two ways with respect to reliability: as a single entire dataset, or as two overlapping sets of data from which certain parts may be excluded.

First, because Facebook is an on-line representation of networks derived from prior real-life relationships, the opportunities for individuals to provide grossly false demographic information are limited (they will be discovered by peers in their network). Thus, we will assume without further justification that the externally-derived information contained in the dataset (education, employment) are subject to sufficient peer review so as to deter the display of false information. However, information about individuals who joined the social networking site independently of any existing real-world relationships can not be said to be as reliable because there is less review. Individuals having no friends and no wall posts (or choosing to suppress the display of such information to the public) make up less than 2% of the entire data set.

Second, however, activities, interests and favourites, are less verifiable through third parties in all cases. It may be the case that close friends who are familiar with each others personal preferences can detect obvious contradictions, but errors and lies within personal information can be given much more leeway due to incomplete information and the less critical nature of these data to external personal and professional relationships.

With respect to validity, original text from the gathered profiles was used as much as possible for analysis for objective measures. Counting and pattern matching were automated, resulting in high validity (Krippendorff, pg. 178). Data transformations were limited to normalizing variations on spelling and punctuation for firm names. Although the normalization rules, dictionaries, and field extraction rules are inputted by hand, data was never re-coded by hand.

## Precautions

Several precautions should be kept in mind when considering the results of this study. First, only self-selected individuals who sign up for Facebook are represented in the data set. Despite common-sense assumptions, factors such as Internet access may or may not be pervasive

within the young creative population.

Second, nearly 100% of males had accessible profiles, while approximately 90-95% of females did, varying with age. Therefore, more males were effectively sampled than females in the data obtained.

Third, employment and education information are under-representative of actual experiences because not all individuals list comprehensive career histories, and there appears to be little to no social penalty for omission of such data.

Finally, data points about religious affiliations, activities and employment may appear in the list of groups a user may be part of, but they are not necessarily detected in fields in which they would normally appear, or using terminology differing from local norms.

## “Creatives”

Self-assigned attributes of creativity were universally unusable in this study. Individuals' perceptions, or, more accurately, their descriptions of themselves with respect to being creative, were not detected using the firm-based definitions of creativity.

The vast majority individuals on Facebook do not identify themselves as being creative, or any of the other homonyms identified by Rogets. In the dozens of possible correlates examined, the population of Facebook Creatives were not clearly distinguishable from the Facebook population as a whole. External definitions of creativity—those defined by observable characteristics such as employment and activities—were far more useful in identifying creative individuals.

## Creative assistants?

In an age group where people are beginning their professional careers, many undoubtedly work in assistant roles as adjuncts to individuals whose professional careers the young people would like to emulate.

In the complete dataset, 358 assistant employment roles were identified in total using the same method by which creative roles were identified from the dataset. Of those, only 102 positions were not identified as creative positions, covering a diversity of positions and role descriptions.

By the definition of the role, any professional individual or organization requiring assisting on an ongoing basis is creative. Otherwise, the tasks performed while assisting would be broken out and codified into well-defined roles and given a proper position names. While assisting someone in the performance of a creative task

is not the same as performing that task, the activity can be a great opportunity for tacit knowledge transfer.

In addition, assistants can be gatekeepers for information to the creative, they might serve as creative troubleshooters or implementers of creative decisions, and they may be involved in creative planning, scheduling, and resource allocation tasks that at any given moment might fit any or all of Florida's descriptors of creative occupations.

However, as a direct intersection between current and future creative professionals, they have much potential as subjects of future study.

## Future work and Conclusions

According to Richard Florida, the creative class is a proportionally large source of wealth generation for proportion of the population (Florida 2005, pg. 29), and this study indicates that creative employees identified externally account for many interfirm connectivities and potentially knowledge flows via their own mobility. The mobility of creative youth, or youth in creative employment positions undoubtedly deserves further study as they will not only become our next generation of leaders, but also the next generation of creative infrastructure.

The generation of young Facebook Creatives entering the workforce has already broken away from some traditional definitions and concepts of creativity. A thesaurus reflects a language as it is in contemporary use. Many of the words by which we attempted to detect creativity utterly failed to do so with youth who have adopted the Facebook innovation at a relatively early stage.

When this paper was first proposed in early February 2007, there were perhaps 1,700 users of Facebook at the University of Calgary with a population of over 25,000 students. Within three months, the number of users at the University of Calgary has grown to seven times that number, but even at an impressive 11,620 members, it barely touches 1% of all University of Calgary alumni.

This paper described a single study of creative youth on Facebook in Calgary in terms of data collection, analysis and critical review of the available data. But like the Facebook social networking site itself, it has barely scratched the surface of its full potential.

At every step along the way, key observations were identified about the Facebook creatives, including their propensity to not describe themselves as creatives, their

apparent tendency to disconnect personal creativity from professional creativity, and their already significant role in the local network of firms.

Some observations and inferences were more startling than others. It's no surprise that skill level relates to creative employment, or that searching for creative activities would locate creative people, but the same might not be said about the fact that, as a whole, the Facebook creatives do not view themselves or their skills as anything special with respect to creativity. Yet they observably are, and that latent creativity has profound effects and implications yet to be explored.

As this was a first attempt at exploratory research with a new dataset and a combination of traditional and improvised tools, not all possible directions of research or analysis have been pursued, and many avenues of potential discovery have already been indicated within the current dataset alone:

- Since the dataset contains post-secondary educational and work history, we may be able to study the on-line attitudes and behaviour of students as they move from institution to institution to firm within an extended innovation network.
- Activities lists include extra-curricular and social capital-building activities which have not been characterised or examined in this study. How do they contribute to the development of creative professionals?
- It was noticed that women reveal less than men on some aspects of the profile. What other gender differences exist in the data set, particularly with respect to innovation and creativity?

The relative ease of gathering large data sets, along with the availability of the analytical tools developed for this study should enable anyone interested in any aspect of this data set and the creative youth it represents to easily replicate this type analysis for any other city or university, in any other set of dimensions.

In addition, there are several other significant social networking websites founded around communities other than educational institutions and cities. MySpace was founded on sharing user-generated multimedia content. Orkut, Google's social networking service popular among Internet-savvy users around the world. And Friendster which tends to connect people from geographically disperse regions.

These additional social networking websites provide ample opportunity to study specific dimensions of each broad topic addressed in this paper. Possibilities

include describing networks based on creative interests or occupation group, content analyses to identify and describe gender-, age- or geographically- based trends in self identification with occupation or interests, and the potential to track, through time, trends in any values or attitudes identified above. Further, this study may be easily repeated with minimal modification in any major city in Canada by selecting a different geographic focus.

## Reproducing this study in any city

Software source code for data parsing and exploration are available at the author's website at:

<http://www.ucalgary.ca/~bali/>

## Acknowledgements

The author acknowledges and appreciates the support and advice of colleagues at the University of Calgary, particularly the members of the MCRI II team.

## Literature Cited

David Camacho, Ricardo Aler, Juan Cuadrado, 2004.

“Rule-Based Parsing for Web Data Extraction” in *Intelligent Agents for Data Mining and Information Retrieval*. Masould Mohammadian, Ed. Idea Group Publishing.

Klaus Krippendorff, 1980. *Content Analysis – An Introduction to Its Methodology*. Sage.

Frederico Parodi, Stefano Verna, 2006. “DownloadThemAll” version 0.9.9.7. Obtained from:

<http://www.downloadthemall.net>

Anonymous. “About Facebook” <http://ucalgary.facebook.com/about.php>. Retrieved April 2, 2007.

Liqiang Geng, Howard J. Hamilton, 2007. “Choosing the Right Lens: Finding What is Interesting in Data Mining” in *Quality Measures in Data Mining*. Fabrice Guillet, Howard J. Hamilton, Eds., Springer.

Roget's New Millennium Thesaurus, First Edition v 1.3.1, 2007. Barbara Ann Kipfer, Ed. Lexico Publishing Group