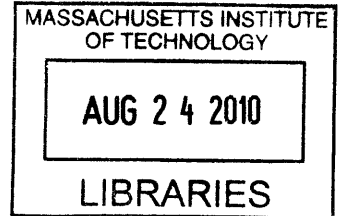


**The Zero Touch Experience: Intent Based  
Contextual Morphing on Mobile Devices using  
Localized Keyword Distributions**

by

Jong-Moon Kim

B.S., Massachusetts Institute of Technology (2009)



Submitted to the Department of Electrical Engineering and Computer  
Science

in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

**ARCHIVES**

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

© Massachusetts Institute of Technology 2010. All rights reserved.

Author .....

Department of Electrical Engineering and Computer Science

May 17, 2010

Certified by .....

Professor Glen Urban

David Austin Professor of Marketing

Chairman, MIT Center for Digital Business

Thesis Supervisor

Accepted by .....

Dr. Christopher J. Terman

Chairman, Department Committee on Graduate Theses



# The Zero Touch Experience: Intent Based Contextual Morphing on Mobile Devices using Localized Keyword Distributions

by

Jong-Moon Kim

Submitted to the Department of Electrical Engineering and Computer Science  
on May 17, 2010, in partial fulfillment of the  
requirements for the degree of  
Master of Engineering in Electrical Engineering and Computer Science

## Abstract

With the rise of smart phones integrated with GPS devices, we have seen the advent of location-based contextual services. "Contextual" in this sense, refers to a simple query of nearby points of interest based on the current location. This type of filtering is but the most rudimentary of what can be done with such information. Depending on if the user is at the location because of recreation or work, the phone should provide varying services appropriate for that purpose. In this thesis, I present a technique to perform inference on user purpose and an implementation of that technique in a demonstration application called Concierge. Concierge showcases how purpose can be used to provide a compelling, personal mobile experience. The application uses a Bayesian inference system with Gittins index utilizing location, past behavior, search queries, as well as other data present on the phone to make an assessment about the user's purpose. Using this data, the application assembles the most relevant applications, offers deals and discounts for appropriate nearby businesses, and shows information about the user's friends and their statuses. With Concierge, the most interesting content simply appears without any user input; hence the Zero Touch Experience.

I discuss how such an inference system is designed and how it was implemented in a first-stage demonstration for France Telecom/Orange and then explore the implications as it pertains to mobile applications, mobile advertisement, and social interaction.

Thesis Supervisor: Professor Glen Urban  
Title: David Austin Professor of Marketing  
Chairman, MIT Center for Digital Business



## Acknowledgments

To begin, I would like to thank Professor Glen Urban for being an incredible mentor, advisor, and friend. The days I have spent in his lab are amongst the most memorable I have had at MIT. It was there that I was exposed to a vast realm of disciplines from marketing science to computer science, from managerial communication to statistical analysis. I sincerely believe it was through his guidance I entered into one of my most creative periods of my life. I consider myself very lucky to have had the opportunity to be working with someone so innovative, wise, and generous. His patience and warmth in dealing with people was especially an inspiration that will not be a lesson forgotten soon.

I want to thank Ted Tomlison and Chris Perciballi for being fantastic lab mates who made it a pleasure to work with for various projects who have been instrumental in ensuring the success of our projects. I would like to especially thank Kevin Wang who first introduced me to Glen and has been a great mentor and friend. I would also like to thank Clarence Lee who introduced me to deeper statistical methods and inspired me to pursue on deeper studies in marketing science.

I would like to thank Professor Gui Liberali for his friendship and time offered to help me fully understand the mathematical mechanics of morphing. I also would like to thank Professor Erin MacDonald for her invaluable work with the General Motors Morphing project and assistance even from far away.

In addition, I would like to thank Dorothee Bergin and Diego Panama whose innovation and diligence helped bring the FT/Orange project into fruition, which is the subject of this thesis.

I would like to thank Professor John Hauser and Professor John Little for encouraging me to further pursue an academic career continuing at MIT after graduation.

Furthermore, I would like to thank Professor Renee Gosline, Professor Juanjuan Zhang, Oliver Baecker, and Tal Snir, for their excellent contribution with our Gen Y project and Mobile Marketing project.

This research is conducted at the Center for Digital Business and the Marketing

Department at the Sloan School of Management at MIT with the financial support of FT Orange and Suruga Bank. I would like to thank our sponsors for making our research possible.

Finally, I would like to thank my parents who have been instrumental in supporting me anyway that they can, always making sure I had everything I needed to grow. I would like to thank Daniel, my little brother, who inspires me with his incredible diligence and commitment to excellence. Lastly, I would like to thank my baby sister, Angela, who has always brought a smile to my face regardless of the circumstances.

# Contents

<b>1</b>	<b>Introduction</b>	<b>13</b>
1.1	Vision . . . . .	13
1.2	Thesis Scope and Overview . . . . .	14
<b>2</b>	<b>Localized Keyword Distributions</b>	<b>15</b>
2.1	Intuition . . . . .	15
2.2	Location-Based Design Parameters . . . . .	19
2.2.1	Data Structure . . . . .	19
<b>3</b>	<b>Content Morphing</b>	<b>25</b>
3.1	Motivation . . . . .	25
3.2	Architectural Design . . . . .	26
3.2.1	Population Segmentation using Cognitive Style . . . . .	26
<b>4</b>	<b>Concierge</b>	<b>31</b>
4.1	The Concierge Application . . . . .	31
4.2	Mobile Applications . . . . .	34
4.3	Local Search . . . . .	34
4.4	Offers and Deals . . . . .	37
4.5	Social Statuses . . . . .	37
4.6	Incorporating Personal Preferences . . . . .	40
<b>5</b>	<b>Contributions and Future Research</b>	<b>43</b>
5.0.1	Localized Applications . . . . .	44

5.0.2 Receptivity and Privacy Concerns . . . . . 44



# List of Figures

2-1	Google Characterizes Who I Am Dependent on Keyword Distribution	16
2-2	Google Characterizes Location Dependent on Keyword Distribution .	16
2-3	Localized Keyword Distributions Overlayed over Map . . . . .	17
2-4	Top 5 Locations with Highest Incidence of the Keyword <i>Taxi</i> . . . . .	18
2-5	Hyperdimensional Graph Representation of Localized Keyword Distri- butions . . . . .	20
2-6	Extracting Keyword Distribution Given Current Location in Hyperspace	21
3-1	Scenario without Latent Variable Segmentation . . . . .	27
3-2	Scenario without Latent Variable Segmentation . . . . .	29
4-1	Concierge Application Process Flow . . . . .	33
4-2	Multiple Dots Issue . . . . .	34
4-3	Concierge Mobile Application Screen . . . . .	35
4-4	Mobile Application Morphed to Different Purposes . . . . .	36
4-5	All Businesses in a Section of Downtown Boston . . . . .	36
4-6	Local Search Morphed to Different Purposes . . . . .	37
4-7	Offers and Deals Screen . . . . .	38
4-8	Various Deals for Different Purposes . . . . .	39
4-9	Social Status Screen . . . . .	41
4-10	Purpose Correction Screen . . . . .	42



# List of Tables

2.1	Sample Keyword Distribution Extracted from Current Point in Hyper-space . . . . .	22
-----	---	----



# Chapter 1

## Introduction

### 1.1 Vision

Imagine taking out your phone today. You must flick through multiple pages of applications before finding the one you want. With hundreds of thousands of applications currently on the app store and hundreds of new applications being introduced everyday, you find it difficult to know which applications to install and stay updated. You find this process a nuisance, or worse, you simply don't bother.

Now imagine it's the year 2013. When you pull out your mobile device, it immediately understands why you pulled it out. While waiting for the subway, the device has listed three of your favorite games and has preloaded three recommended games ready for you to play. When you pull out your device at 12:00 pm on a weekday, you are greeted with lunch applications as well as local lunch coupons. When you walk into a movie theater, the movie ticket application is already preloaded with which you immediately buy tickets from your mobile device without visiting the box office. You walk into a hospital, the hospital application is already preloaded that tells you check-in procedure and estimated wait time before the doctor can see you.

There is no "application downloading" or "navigating to a website". It makes intuitive sense: I know what I want, this device provides it for me in the easiest way possible.

## 1.2 Thesis Scope and Overview

The contribution of this thesis is enumerating the concept of Localized Keyword Distributions and documenting Concierge, a demonstration application utilizing Intent Based Contextual Morphing. Intent Based Contextual Morphing is important because it minimizes friction in delivering relevant content to the user. As mobile devices become more and more connected, users will be inundated with too much content that will make it difficult, if not impossible, to utilize optimally. By automatically loading the most relevant content based on the user's purpose, the cognitive load required to organize and navigate through drops down to zero. It is only with improving the user's bandwidth to handle more content can we hope to bring more to the already saturated mobile experience.

In Chapter 2, I introduce the idea of localized keyword distributions. I discuss how keywords from a variety of sources can be utilized construct a conceptual map of a region. I will enumerate a strategy utilizing machine learning techniques from which we can conduct inference about the user's purpose.

In Chapter 3, I will explain the concept and mathematical model of Content Morphing. I will illustrate the concept with a discussion of previous work with Suruga Bank utilizing cognitive styles and explain the mathematics behind using the Bayesian inference engine and Gittins index.

In Chapter 4, I introduce Concierge, the mobile assistant application built for demonstration. I discuss how this application explores the three spaces for possible implementation: applications, local business search and deals, and social statuses.

In Chapter 5, I will describe the contribution of this thesis, and conclude with a discussion on ideas for future research as they relate to Marketing Science and Computer Science.

## Chapter 2

# Localized Keyword Distributions

In this chapter, I will describe the concept of Localized Keyword Distributions. This is the idea of having a collection of keywords describe a particular combination of time, geographical space, and personal attributes. I will describe an inference design that utilizes Support Vector Machines using kernels that can be implemented to derive estimations.

### 2.1 Intuition

When I conduct a Google search, Google chronicles all my search phrases into my personal keyword distribution. A keyword distribution is simply a set of all the different words I have searched on Google in the past. It is this keyword distribution that enables them to target advertisements relevant to me. When we take a look at the search query itself, we realize the query has one defining attribute: who made that query. What we realize is that now the query describes that attribute; that is, Google characterizes who I am based on my past search keywords. I search for "Canon" and "MIT academic calendar", Google can conclude I am probably interested in cameras and am a current student at MIT.

With so many smart phones out in the world, we can imagine there is a volume of mobile-based Google search queries. For these search queries, we realize they have two defining attributes: who made that query and *where that query was made*. Similarly

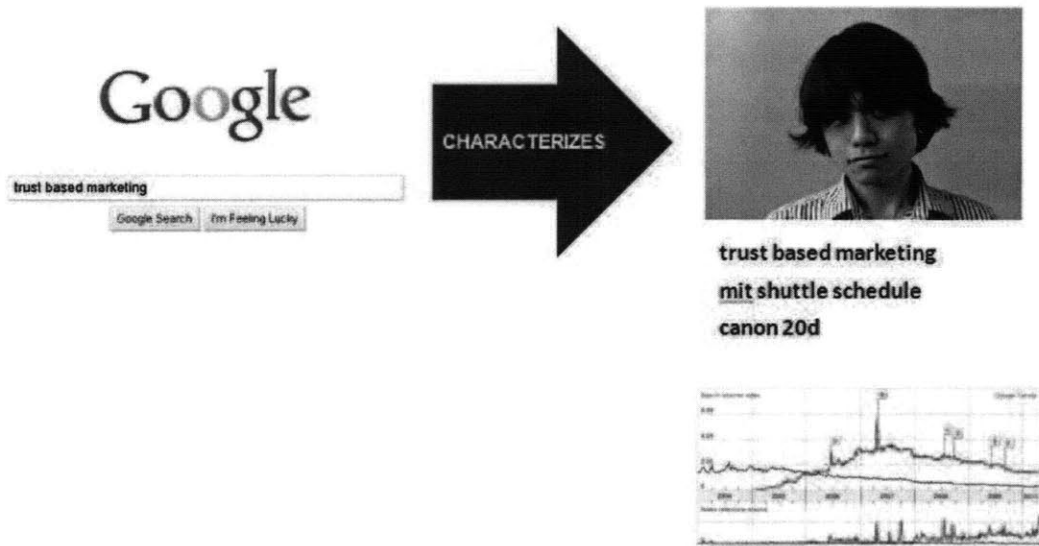


Figure 2-1: Google Characterizes Who I Am Dependent on Keyword Distribution

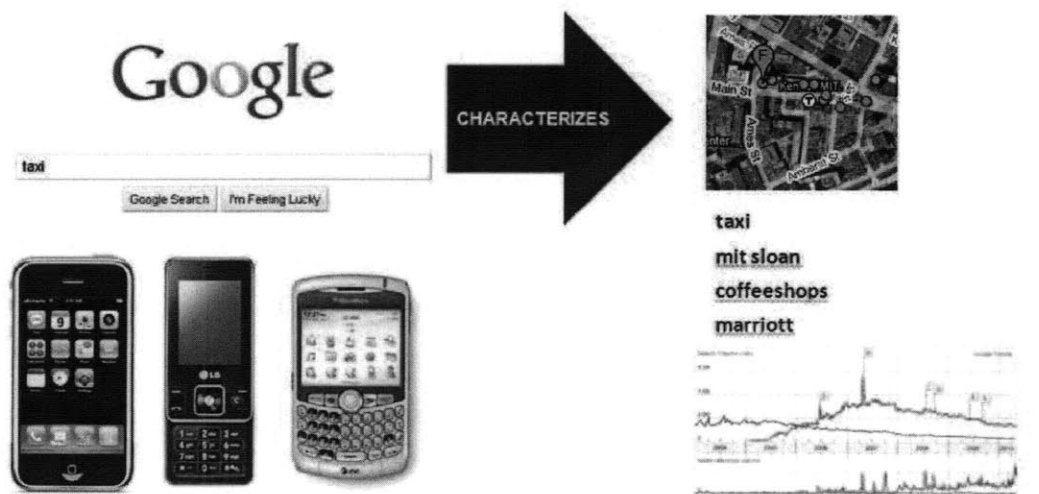


Figure 2-2: Google Characterizes Location Dependent on Keyword Distribution



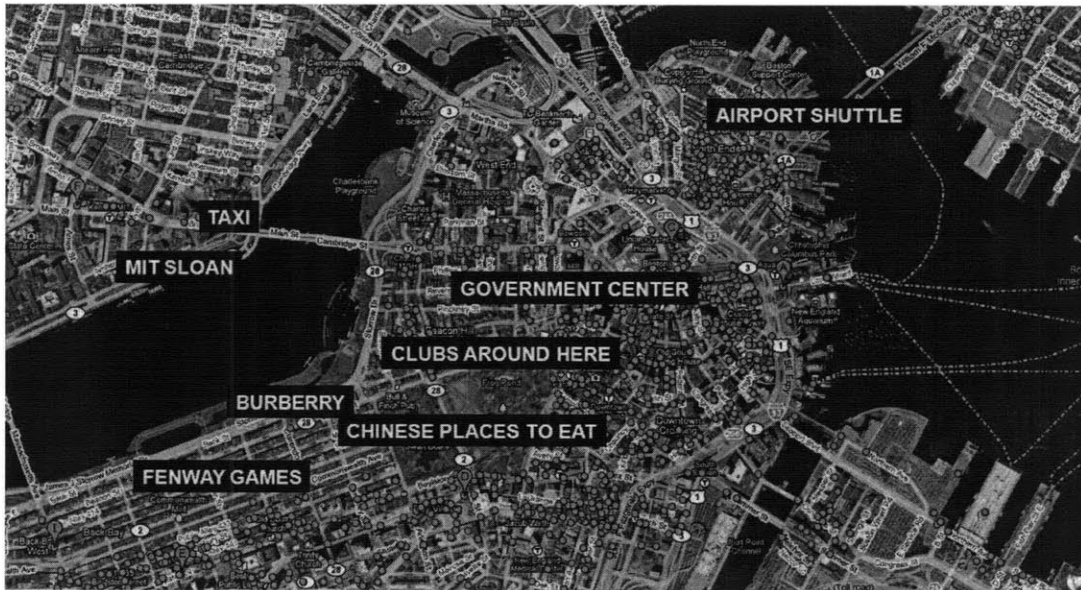


Figure 2-3: Localized Keyword Distributions Overlaid over Map

as before, the query now describes who made that query, but also, it now describes where that query was made. It may seem like a subtle point, but this is where the crux is based. If that location were a person, we could try to get a judge of his personality based on different keywords this location has "searched" for. Google could then, presumably serve up targeted advertisements catered to this location-person. If this location-person searches for "Boston Clubs and Bars", we realize we can classify it as favoring nightlife. We arrive at the powerful notion that we have a way to extract out the *ideas and interests* of that location.

Imagine we are standing outside of Penn Station, one of the busiest bus and rail stations in the world. As travelers exit the station, a large fraction will be looking for taxis for transportation. Many of these travelers looking for taxis would have smart phones searching for the keyword "Taxi" at that location. Imagine we are Google, trying to serve a mobile advertisement, but we had absolutely no clue what the traveler's interests are. We can imagine we would do very well simply by taking the most popular localized keyword, "Taxi", and serving a taxi ad for that traveler. Imagine we are a Taxi company who wants to put up 5 physical billboards on the island of Manhattan. We have reduced this task to a triviality as we simply take the

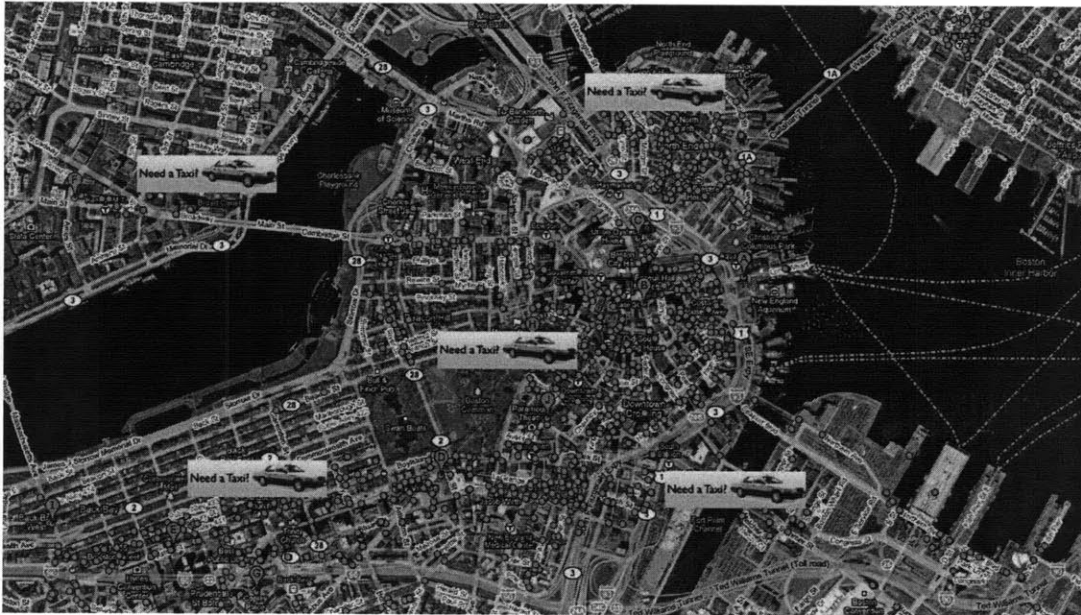


Figure 2-4: Top 5 Locations with Highest Incidence of the Keyword *Taxi*

5 locations with the highest incidence of the "Taxi" keyword. What is neat is the fact even people without smart phones now benefit.

What we must realize is that this is far separated from the current proximity-based matching utilized by location-based advertisers today. These advertisers would perhaps post sales in the department store across the street from Penn Station or book sales in the neighboring Barnes and Noble. Some of the smarter advertisers may even have an employee hard-code this location to serve taxi-based information since it is a transportation hub. The key is the fact that our system is entirely automated, robust, and dynamic.

An example might be standing at the finish line of the Boston Marathon and being given information about the pizza parlor down the street. If it were any other day, this might have been acceptable content. With our system with dynamic measuring of the most interesting keywords queried at that location, we would accurately infer we should probably give the user information about the marathon, and better yet, live standings of the current runners.

What is most exciting is the fact we now have extremely rich sources of localized

keyword data including: foursquare, Twitter, mobile SMS, amongst others. If we were to aggregate these data sources into a single distribution, we would have a high density data cluster ripe to be utilized for next generation applications.

## 2.2 Location-Based Design Parameters

Now that we have established our intuition, I will explain the current form of the incoming data and the challenges that arise when we attempt to extract user purpose from this data.

### 2.2.1 Data Structure

For each keyword datapoint, we record at minimum the following attributes:

- Geographical Location - Longitudinal and Latitudinal
- Day of the Week - Sunday through Saturday
- Time of Day - Hours and minutes

The reason why we list out Day of the Week and Time of Day as separate variables is because of the fact Day of the Week is an orthogonal measure to Time of Day. As we have the concept of "weekly" events and activities, in order for our system to be cognizant of such patterns, we have it as an axis. For completion's sake, we would also have to have "Time of Month" and "Time of Year" to provide for monthly and yearly activities. For the purposes of this demonstration, we will restrict our discussion to the above four axes.

As our chief interest is with repeating keywords, we have little need to record the actual timestamp (combination of year, month, day, time) of when the keyword was recorded. This data is meant to give us a predictor of future instances of keywords. If that particular keyword has been repeated only once in the past, it will naturally be overpowered by surrounding, stronger keywords. We may later choose to add a

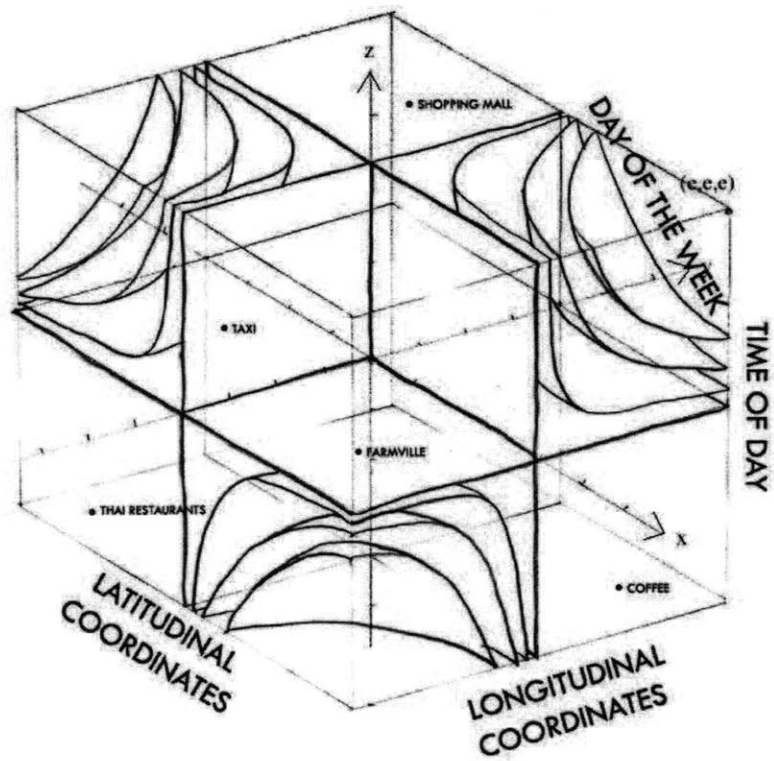


Figure 2-5: Hyperdimensional Graph Representation of Localized Keyword Distributions

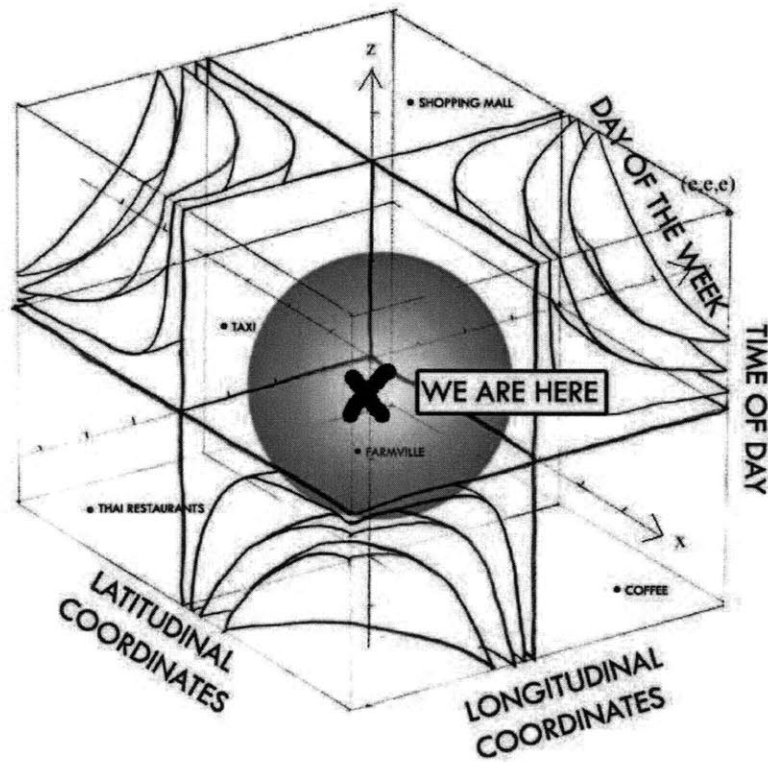


Figure 2-6: Extracting Keyword Distribution Given Current Location in Hyperspace

temporal decay to give heavy weighing to events and keywords happening now, but that constitutes a minor modification and is for later discussion.

We can think of this data existing in a hyperdimensional space where each of these attributes represent an axis. Each keyword searched at a particular place and a particular time is represented by a point in hyperspace. If we are now thinking about how to get the most interesting keyword from that location, we first need to realize that at any time we send in a query, we ourselves are currently at a particular place in hyperspace. That is, we currently have a location, time of day, and day of the week and are represented by a point on that graph. We can then do a straightforward extraction of nearby keywords within a certain distance from our current point. We can choose to weigh these keywords according to distance from our current point with

keywords that are close by getting more weight.

What does a point close by mean? It means that point is either geographically closer, temporally closer, or closer in terms of the day of the week. If we take the keyword with the highest weighted incidence from our query point, we now have a keyword that describes our current point in hyperspace. A nuance is the fact distance between the axes will never be equally scaled, where the Days of the Week axis will have a much shorter range compared to Time of Day. The *Figure 2-6* that depicts the sphere as range from our current point should more resemble a thin cylinder.

One obvious purported misstep might be the case where a particular keyword happens all throughout the week instead of as weekly functions. For that particular keyword, we would see a dense strip in the Day of the Week axis, much different in shape to the sphere inference shape we utilize. The answer to this is the fact the way we collect data, any particular day of the week that we happen to query, the sphere will still see a mass incidence of that keyword, correctly picking it out.

Keyword	Distance
coffee	2.35
taxi	5.32
marriott	5.36
taxi	7.32
taxi	7.53
starbucks	7.98

Table 2.1: Sample Keyword Distribution Extracted from Current Point in Hyperspace

From *Table 2.1*, we realize "coffee" is the closest keyword but "taxi" is by far the most frequent. The Localized Keyword Distribution is designed for flexibility in usage where we can weigh proximity and frequency to varying degrees. For our current usage, we will favor frequency greatly.

In addition to the base data, if the keyword is accompanied by additional metadata or if we can cross reference the sender with another data source, we record that as well:

- Age of User
- Sex of User
- Income of User
- Interests of User
- Friends of User
- User's Home Address - Used to determine relative distance from query location
- User's Work Address - Used to determine relative distance from query location
- Is Part of Routine Sequence - Determines if user frequents location or if it is a one time or rare occurrence.
- Time of Day

We must realize that with additional axes, the data will become more and more sparse.





# Chapter 3

## Content Morphing

In this chapter, I will describe the concept of Content Morphing. This is the idea that there are latent attributes for the user for which we can provide appropriately differentiated content. One such example is cognitive style morphing where we differentiate content based on the way the user comprehends information. This concept provides us the construct to transform inferences on Localized Keyword Distributions into discovering the user's purpose.

### 3.1 Motivation

Everybody processes information in their own way. While some people tend to be more visual, other people tend to be more verbal. While some people like going into detail, other people like to focus on the big picture. Cognitive style, simply put, is the way an individual thinks, perceives, and utilizes information. We pose the question: if we knew about the cognitive style of a person beforehand, would it make it easier to provide information in a much more meaningful and interesting way?

Imagine you are a car salesman trying to sell a car. You see your first customer approaching and you happen to know he likes to follow his gut instinct and is very visual. You might choose to take him through the showroom of the latest vehicle and take him on a test drive where you can pose the question, "Does it feel right?". A little while later, you see your second customer approaching where you know she is deeply

analytical and enjoys drilling down into nuanced details. You might provide her with detailed comparison charts against competitor cars and expound on the various features on why your car has much better value. If you had taken her through the showroom instead, we can imagine she might have been much less receptive.

What you have just done is Content Morphing. Based on the customer's cognitive style, you chose the appropriate experience that that customer would find most natural. Imagine going to a website that automatically modifies itself to make it as meaningful and interesting as possible for you. It is towards that vision we take a look how Content Morphing was implemented using cognitive style and how it might extend towards intent based morphing.

## **3.2 Architectural Design**

The Content Morphing system functions by classifying individuals as particular cognitive styles and then providing the content variant optimized for each cognitive style based on active learning.

### **3.2.1 Population Segmentation using Cognitive Style**

In order to perform effective Content Morphing, we must employ appropriate latent variables to segment the population. Cognitive style is an enduring latent variable. That is, we infer the cognitive style based on external observations because it is not directly observable (latent) and that it generally stays the same for that person (enduring). To examine why such segmentation is necessary, let us walk through a scenario where we do not utilize cognitive style.

#### **Scenario without Latent Variable Segmentation**

Imagine we seek to provide personalized web content to each visitor. In order to provide personalization, each individual is directly linked up to a content variant. This would be akin to say, if Janet shows up, show her variant A, or if Tom shows up,

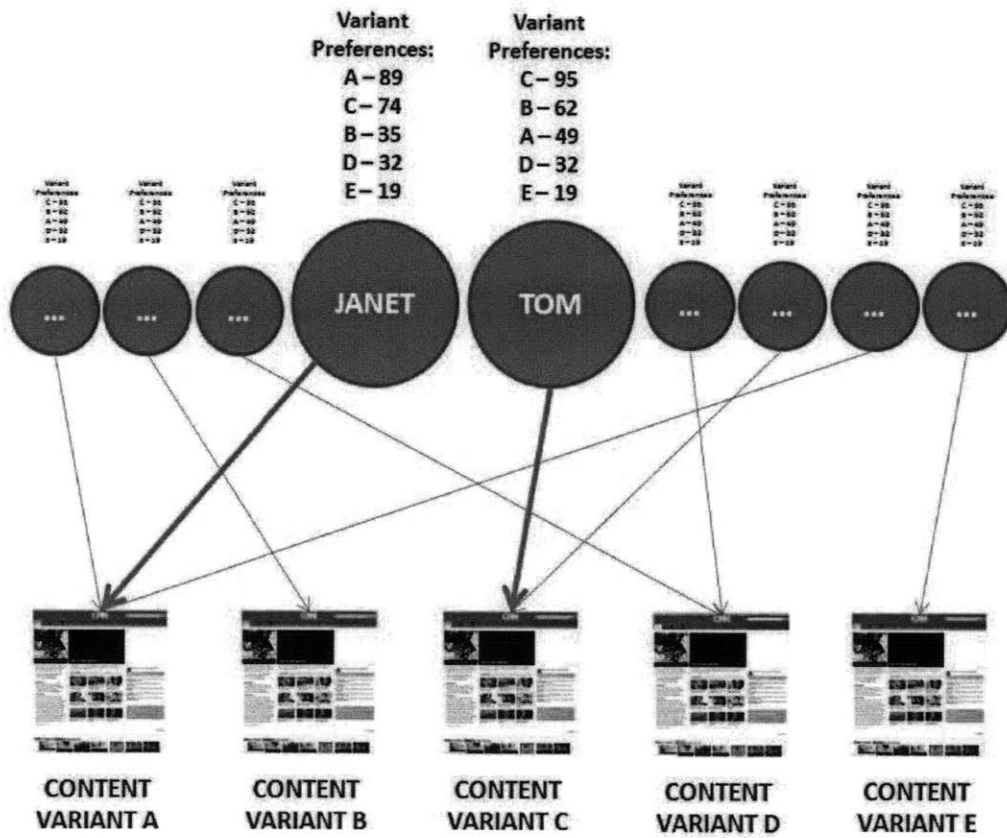


Figure 3-1: Scenario without Latent Variable Segmentation

show him variant B. This raises two serious issues where first, we need a comprehensive system to completely describe each individual and second, we need a individual scoring metric to determine which content variant is best for that user.

For the first issue, in order to get an accurate description of each individual, we would require a multitude of variables such as demographics, areas of interest, personality amongst others. It is a well-known issue that getting web users to provide to extensive personal data is notoriously difficult. In the car salesman example, we have a definite advantage where we can take in facial expressions, voice intonation, style of dress and body language as strong priors for assessing what the individual is like. Across the web, users are near anonymous save for the keyboard and mouse interactions.

For the second issue, the only way we can be sure we are providing the best content variants right for the user is by directly asking the user to evaluate all different variants available. This polling process would be detrimental to the user experience. We need a method to passively optimize the right variants based on user behavior.

The content variant design is largely left arbitrary as the designer will have to play guesswork coming up with different designs he feels like a grouping of users might like. In terms of scalability, if we were to add on a new variant, we would have to query all visitors for their refreshed variant preferences. We are severely limited in the number of variants we can provision for the engine.

### **Scenario with Latent Variable Segmentation**

Now let us take a look how incorporating the cognitive style latent variable helps us alleviate the two issues. Cognitive style distills down the identity of an individual to the core of what matters most: how information will be received. Other variables such as demographics, areas of interest and personality are mere proxies to help us determine the reaction to the content format. If we have cognitive style, we arguably do not need any other data to make an informed decision. The advantage of using a latent variable such as cognitive style is that we can utilize almost any user behavior for classification. As long as we have a sufficient training set of user behaviors and

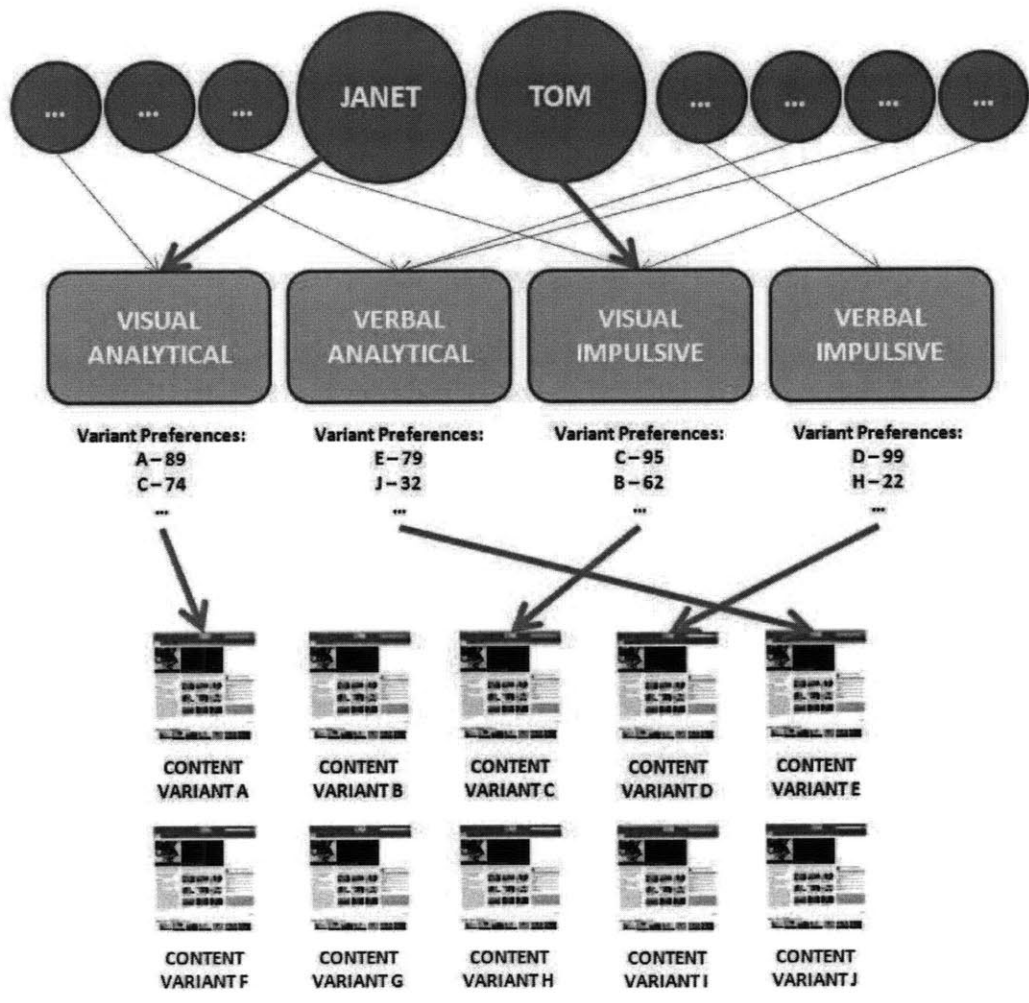


Figure 3-2: Scenario without Latent Variable Segmentation

known cognitive styles, we can utilize Bayesian inference based on user action.

For answering the second issue, we utilize a Bayesian inference engine with what is known as the Gittins Index which will be covered in the section following. What it does is that it actively searches for the best possible content variant per cognitive style through a particular combination of exploration and optimization. In terms of scalability, we can very easily add a new content variant and have the inference engine optimize over the latest collection of variants. In fact, it is favorable to have as many variants as possible that span the space of all possible variants as that will increase the probability an extremely good variant / cognitive style fit will occur. It is important to note that we are no longer designing these variants with a particular person or cognitive style in mind. As long as the designer makes designs with a large distance in between, the engine will automatically find the best suitable.

# Chapter 4

## Concierge

In this chapter, I will describe Concierge, the demonstration application built on the iPhone in cooperation with France Telecom / Orange. The goal of this demonstration is to visualize what an implementation of intent based mobile morphing might look like had it been available today. This application provides the stepping stones towards a functional prototype that would utilize Content Morphing in conjunction to Localized Keyword Distributions to impute user purpose. France Telecom is interested in providing such an application to improve the mobile experience for all carrier customers.

The application includes four different components, Apps, Search, Deals, and Friends, each representing different spaces intent-based morphing could be utilized. The application was designed to span a vast variety of use cases to stimulate discussion for possible product offerings. Currently, this project has culminated in demonstrative form and is currently being circulated around FT as well as other network providers.

### 4.1 The Concierge Application

Concierge is a mobile service that offers the most relevant applications, businesses, special offers, and social streams based on the user's inferred purpose. The application utilizes Bayesian inference on a several datasets available through the device as well as through the carrier. Specifically, Concierge provides relevant content in the following

incarnations:

- Mobile Applications - With over 200,000 applications available on the iPhone marketplace with hundreds of new applications getting introduced everyday, finding and managing mobile applications is becoming more and more of an arduous task. Concierge removes all hassle as it assembles a compilation of both already installed applications as well as recommended applications for the user's purpose.
- Local Search Results - Instead of simply returning all local business establishments near the user, Concierge automatically picks out the most useful and most likely destinations the with which the user might be interested.
- Offers and Deals - Taking local search results a step further, Concierge offers special offers and deals by localized establishments that would help fulfill the user's purpose.
- Social Status - Instead of traditional social status messages that become outdated within the hour, Concierge provides a semantically rich description that dynamically changes each time the user embarks on a new purpose.





# CONCIERGE

An app that infers your purpose and serves you relevant information. Simplify and leverage your life.

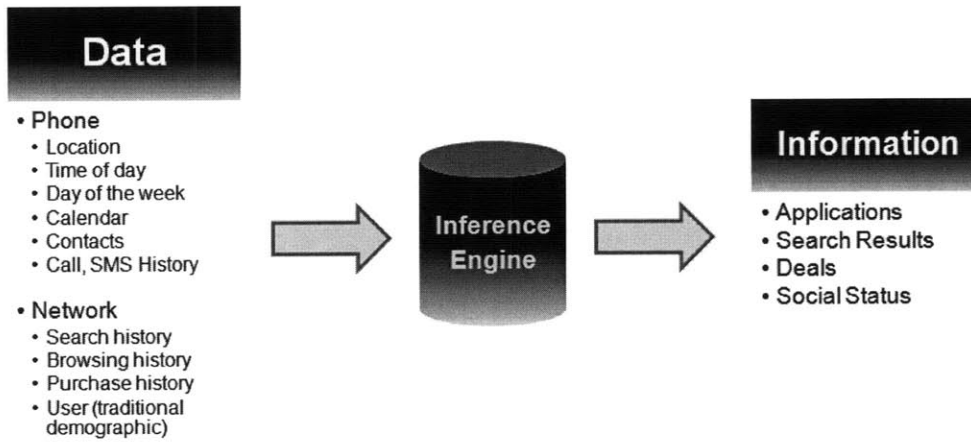


Figure 4-1: Concierge Application Process Flow



Figure 4-2: Multiple Dots Issue

## 4.2 Mobile Applications

As the Mobile Application market grew to maturity, we began witnessing what was known as the "Multiple Dots" issue on iPhones. That is, as users downloaded a large number of applications, it became increasingly difficult for them to open up appropriate apps when they needed them, having to flick through multiple pages visually scanning for the name. Having such friction discourages users from getting full value out of these applications which in turn, stalls mobile progress. In order to ensure rapid mobile innovation, we need a way to effectively utilize old applications while introducing new applications useful for the user.

The Apps page provides a listing of both already installed applications as well as other applications deemed useful for the inferred user purpose. From this screen, the user could in one click launch any of these applications.

At the top of the screen, we have sponsored applications that have paid advertisement fees to have their applications displayed. What makes this advertisement system different is the fact that these developers have bidded on user purpose. That is, much akin to how Google auctions off keywords, we auction off purpose. We can imagine why developers would want to purchase purpose at a premium; the people who see their apps will be actively thinking about the problem space the apps seek to solve. As they are provided right when necessary, chances are high the user will have a chance to appreciate the full value proposition.

## 4.3 Local Search

Simple proximity-based localization has serious limitations where in particular urbanized environments, the user is likely to be overwhelmed by all the nearby estab-



Figure 4-3: Concierge Mobile Application Screen



Figure 4-4: Mobile Application Morphed to Different Purposes



Figure 4-5: All Businesses in a Section of Downtown Boston



Figure 4-6: Local Search Morphed to Different Purposes

lishments and not be connected to one that answers her need. Concierge performs a powerful filtering where only businesses that are related to the current purpose is shown.

## 4.4 Offers and Deals

Taking local search a step further, Concierge also provides the ability for local businesses to provide exclusive deals through the system. These businesses would be willing to provide steeper discounts as they know that viewers of these ads will be extremely close by as well as thinking about their product category.

## 4.5 Social Statuses

Social statuses are currently ubiquitous online where we can find it everywhere from Facebook to Twitter, from Gmail to AIM. In parallel, location based social networks



Figure 4-7: Offers and Deals Screen



Figure 4-8: Various Deals for Different Purposes

are also on the rise, letting all of the user's friends know where she is currently at. What Concierge provides is a centralized way to refresh all status messages to the current user's purpose.

One of the most important value propositions is the fact that with current status systems, the text tend to become out of date and stale extremely quick. As soon as the user moves to a new location, as soon as he starts a new activity, the status message is no longer accurate. As Concierge constantly refreshes the status with the user's purpose, the text is guaranteed to stay fresh. This also provides a benefit over existing location based social networks where now, instead of merely seeing the location of friends, users read a semantically rich description of what their friends are doing at each of those locations.

## **4.6 Incorporating Personal Preferences**

Of course, with any probabilistic inference system, there will be the chance that the result is incorrect. The user will be able to correct the specified purpose with the actual purpose the user happens to be undergoing. This will update the system in two ways: 1. An additional weighing for that purpose in that particular location in hyperspace will be increased. 2. A much stronger personal preference bias will be placed for that location such that any future inferences made by the person nearby will result in a similar response.





Figure 4-9: Social Status Screen

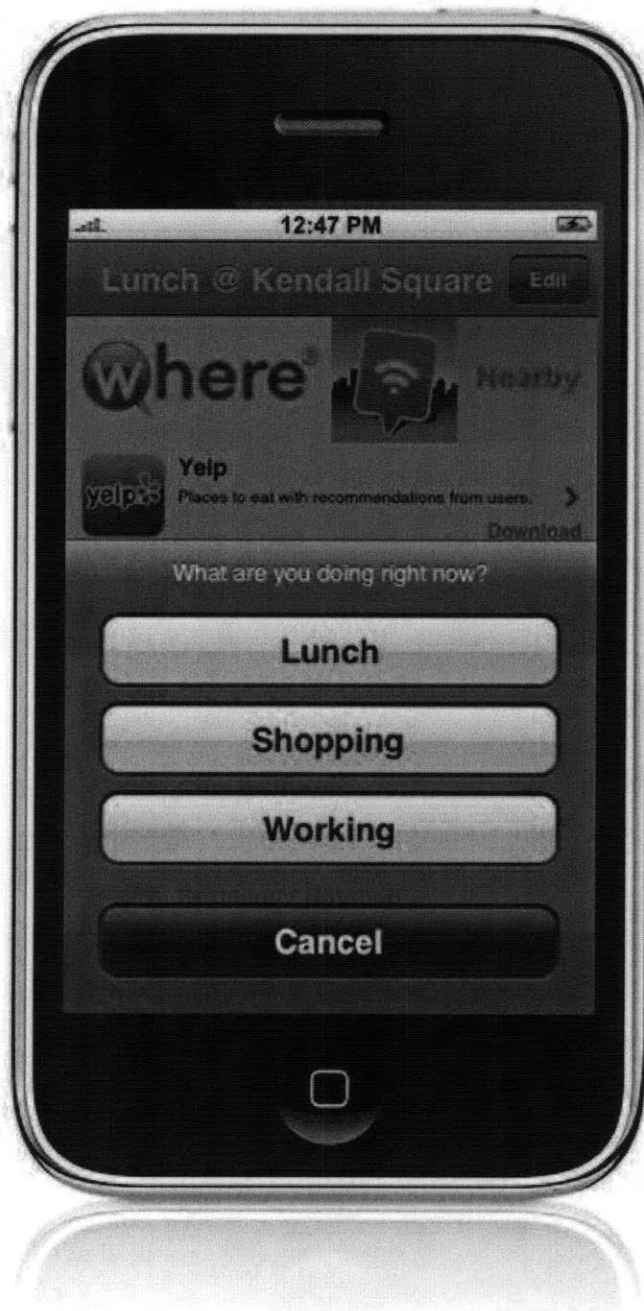


Figure 4-10: Purpose Correction Screen

# Chapter 5

## Contributions and Future Research

In this thesis, I have described the concept of Localized Keyword Distributions and the implementation of Concierge, a iPhone demonstration application. I elaborated on how a prototype version of Concierge might be implemented using an inference technique to extract highest incidence keywords from a hyperspace location.

Localized Keyword Distributions provide a framework with which we can represent location-tied keywords. We can source these location-tied keywords through a multitude of sources, including mobile search engine queries, twitter messages, and social media status text. We can then perform inferences that detect the most likely keyword given the current location and time. Using these keywords, we can utilize an SVM classifier to get the most likely user purpose. Here we incorporate the use of a Bayesian Inference Engine using the Gittins Index to determine the most relevant content variants suited for that purpose. I discussed a variety of content categories that can be provisioned using this approach, including the discovery of mobile applications, increasing contact with local businesses, and providing for more accurate social statuses.

Moving forward, the next step would be implementation of the live inference engine into mobile devices and incorporation of localized keywords sources. From here we have several avenues of investigation:

### **5.0.1 Localized Applications**

The concept of Localized Applications takes the Apps page of Concierge a step further. Instead of merely displaying a listing of applications that could be downloaded, the most relevant applications are automatically preloaded on the device. What is important to note that these new "applications" we are discussing are not the traditional applications we would find on the App Store. These applications are highly customized tools that the proprietor of that location has specified to be loaded up.

One example would be walking into a bookstore and the device automatically loading up a query engine designed to help customers search for books. You might walk into a cafe and be able to order directly using your phone without having to wait in line and simply pick up your coffee when the order is fulfilled.

The key point lies with seamless engineered experiences that happen with minimal friction. Each of these Localized Applications are designed professionally by the business or establishment you are currently at, extending the customer experience to include localized mobile exchanges. This platform has the capability to greatly magnify the self-serving capacity for a business as customers would have instantaneous access to resources.

### **5.0.2 Receptivity and Privacy Concerns**

One important aspect that is currently unknown is human response to a digital assistant fully cognizant of their purpose. At best, the system may be heralded as an incredibly useful tool indispensable for effective mobile usage, or at worst, a serious breach in privacy. The degree Concierge will utilize all data available on the phone is unprecedented, warranting a serious look at how to ensure good receptivity. It will be important to assure users that all data collected is immediately depersonalized when sent to the aggregate central database. The user would be able to then selectively choose to whom to reveal their current purpose and other identifying information.

# Bibliography

[1] Lee, C., User Adaptive Web Morphing: An Implementation of a Web-based Bayesian Inference Engine with Gittins' Index. Massachusetts Institute of Technology, Sloan School of Management, June 2008.

[2] Wang, K., Morphing Content in Mobile Applications. Massachusetts Institute of Technology, Sloan School of Management, June 2009.

[3] Hauser, J., Urban, G., Liberali, G., and Braun, M. Website Morphing. Marketing Science, 2009.

[4] Tokoro, Y. Improving customer experience through advocacy and morphing: A web application for suruga bank. Sloan fellows master thesis, Massachusetts Institute of Technology, Sloan School of Management, June 2008.

[5] Urban, G. Don't Just Relate! Advocate. Wharton School Publishing, 2005.

[6] Farrahi, K., Gatica-Perez, D. What did you do today?: discovering daily routines from large-scale mobile data. Proceeding of the 16th ACM international conference on Multitmedia, (pp. 849-852), 2008.