GENERATED
MAN

Diego Trujillo

Diego Trujillo was born in **Mexico City** in 1986 where he studied both photography and biology. In 2010 he moved to London, U.K. to take a Master in Arts in the **Design Interactions** department at the **Royal College of Art**.

Trujillo's work deals with science and technology as a conceptual space. Rather then trying to use science in design, his work uses design to talk about issues that emerge from our contact with constantly developing technology. The time Diego spent in science can be seen through the thorough research that accompanies his projects. His recent work relates to how technology can have an effect on our identity, our freedom and our perception of the world.

A multidisciplinary background has provided him with a wide range of skills including photography, 3D graphics, electronics and programming. This has allowed him to present work in a wide variety of mediums.

Generated Man

Generated Man is a software project about identity on the internet. Increasingly, more internet services track you while you are online in order to create a 'personalized' web experience. Tracking means that you end up being **portrayed** by statistics.

The way **Google** chooses which search results and advertisements to show is probably the most famous example of this. To decide what you are interested in, Google creates a digital persona based on your previous web behaviour. This data is analogous to a character sheet used in **role playing games**.

Generated Man flips the Google system around, starting with a **character** sheet and ending with a personality. The values on the sheet are fed into Google's **search engine**, which suggests websites based on what it computes as the **personality** of the artificial web user. The generated **personality** is presented as a collection of models that the software chooses from Google's **3D Warehouse**.

Introduction

I grew up in the 1990s, part of the last generation to have an internet free childhood, probably the last generation that will be able to imagine life without this technology and that will remember how quickly it got into our lives. I get a strong nostalgia whenever I hear the sound of a dialing modem or encounter an old, unmaintained website crowded by animated gifs, bullet point lists and midi background music.

The technology has changed a lot since it became commercially available. We can now experience connection speeds that were previously unthinkable, along with very high quality sound and video streaming live from a remote place where they are stored. We have also been given tools that make it easier to contribute to the digital universe, things like blogs and social media allow nearly everybody to broadcast a message to whoever is interested. This has created an internet that is rich and diverse but also hard to filter, kind of like finding a peice of straw in a needle stack, in which every needle seems relevant and informational.

The web experience 15 years ago was relatively straight forward. You sat down on your computer, opened a web browser (probably Netscape Navigator), went to a website and saw whatever you were looking for and nothing else. No matter how many times you used a service or what you typed into it you would always get the same thing. If a friend of yours used the service in the exact same way, he too would see the same content that you did. We were all used to saying: "search for 'whatever' on Yahoo! And then click on the third link, that's the site I'm talking about."

That experience has changed recently. Internet based services have developed efficient information gathering systems, giving them detailed information about their costumers and allowing their products to react to individual users. This can be seen all over the web. From the advertisements that follow you everywhere to the results people receive based on their previous on-line behavior. In short, internet services are becoming personalized, a profile is built for you in order to decide what the internet displays and what you see first.

Personalization sits at the core of this work. In order to tailor a service we need an idea of who we're cutting the fabric for. On-line services gather data about what individual users do; the data gathered then becomes a representation of the user. This representation determines what that person is interested in and is used to tune the way a service behaves. The mere act of creating a representation of someone takes us into a never ending debate around identity. A lot can and has been said about identity in the digital age, ranging from avatars in on-line games to identity theft crimes worthy of a tabloid's front page. These forms of digital identity differ a lot from the ones that a service creates for us, as in a game or in a crime we can choose who we become. However, we cannot choose the identity that a service gives us, we can only hope that it is close enough to who we really are and that we have been behaving accordingly.

The last generation with an internetless childhood was also a generation that played table top role playing games (such as Dungeons and Dragons). An extremely nerdy form of social interaction, in which players take the role of characters fitted to a specific universe -medieval, magical, sci-fi, etc. Players then try to overcome problems presented by a story-teller (usually the geekiest person present), they must do so based on their character's talents and by rolling many sided dice.

www.google.co.uk

Google Sorry...

We're sorry...

... but your computer or network may be sending automated queries. To protect our users, we can't process your request right now.

See Google Help for more information.

© 2009 Google - Google Home

While thinking about recommendation algorithms, I realized that they operate in a similar way to role playing games (RPGs). An algorithm creates a virtual version of its users, it then filters the world for them in order to create an engaging universe. In an RPG the player creates an imaginary version of himself and a story teller generates a universe that is fun and interesting for that character. This analogy between algorithms and games was the paradigm that drove the design of Generated Man's algorithms which includes creating characters, rolling dice and navigating through a service's universe.

Generated Man is a computer program that explores the artificial personalities that computers create for us. The software flips the standard system around. Instead of creating an artificial representation through web behavior, it automatically browses the internet based on a character that is given as an input. My approach to the question concerning the fidelity of an on-line persona moves away from the theoretical. Instead of presenting all existing literature on identity issues and representational algorithms, I have taken experimental means to research the question. This book is a tour of my thought process and my experimental results.

The word experimental in this context has a scientific tone to it. It means that Generated Man provided a controlled way for me to interact with a self regulated system. Much like a scientific experiment allows researchers to understand nature's underlying mechanisms, Generated Man allowed me to take a look at how an internet created personality evolves. The results of these trials are alternated between each section of writing. This is presented by a series of internet found 3D model which the program has chosen to represent a character, creating an object based portrait of a person.

The portraits produced by Generated Man's system have provided more than a a humorous understanding of how we are represented on the internet (or misrepresented in some cases). They have also made me ask myself how we construct an idea of a real person based on what we know about them. I feel the project has successfully transcended its original niche of on-line personalities and moved to a space in which broader questions.

I have tried to keep the text as far away from the software technicalities in order to concentrate on the larger philosophical questions I have to answer, never the less they do come up at some points. I hope this book will provide an interesting and enjoyable look at the conceptual, technical and artistic journey I have taken.

Previous Page
Google Blocks
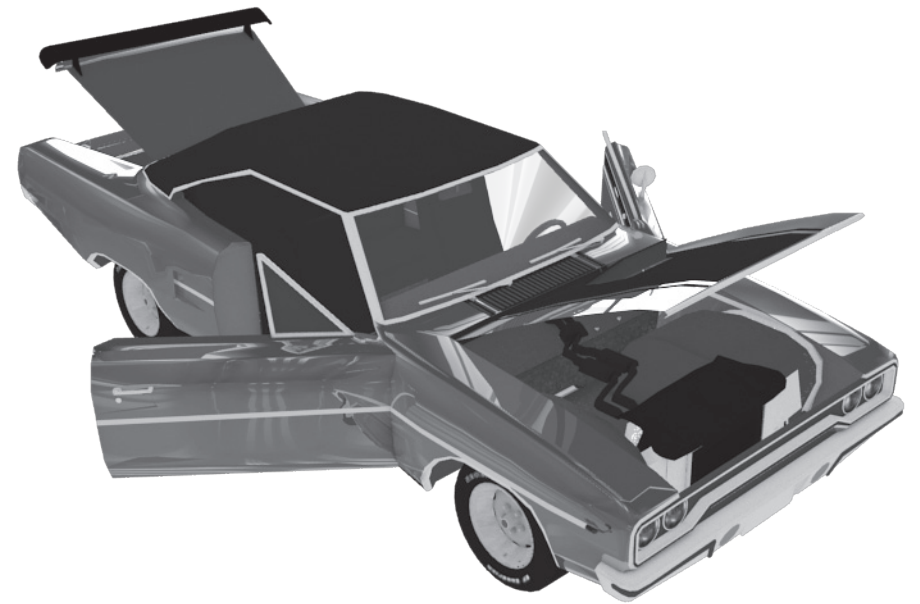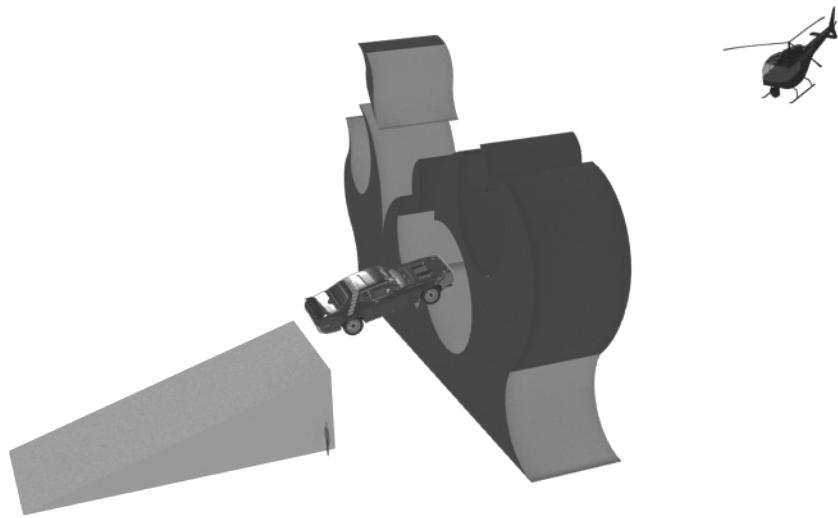Generated Man

OLIVER
Brown

Archetypal Londoner
(fictional person)

Allsop Arms
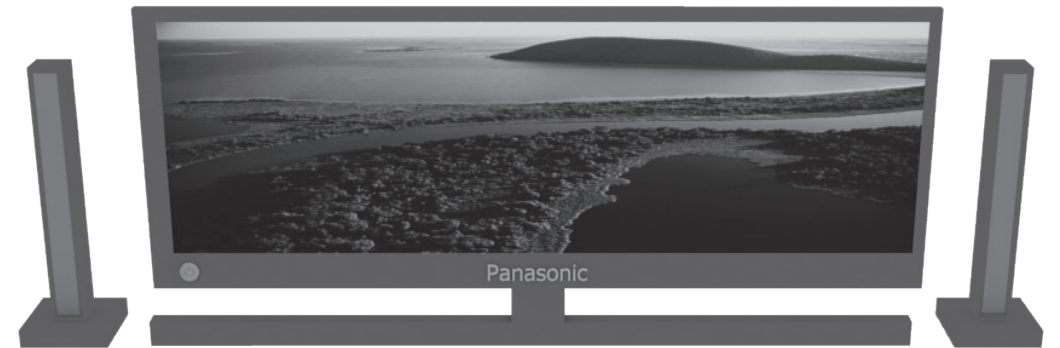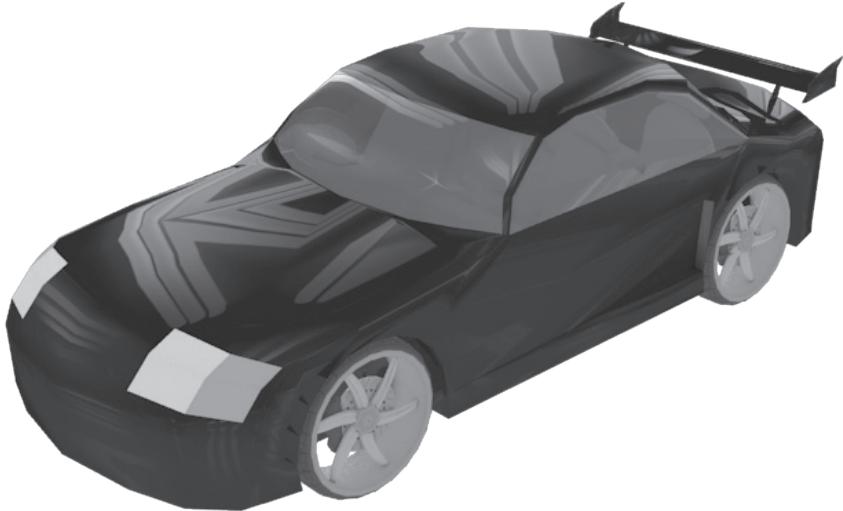
Car

Car Jumping Through
Fire!

TV

Car Battery

Roast Turkey

CSI Pinpall Machine

Car

TV

TV

Computers have always been good at keeping ridiculously long datasets, each bit added to hard drive memory is rapidly seen as potential space to store more information into our ever growing sea of databases. Humans have learnt to value records throughout the years, particularly when studying complex phenomena which are hard to predict; such as changes in the weather , economics and statistics-driven branches of the social sciences.

The internet has always heavily relied on databases of these sorts. Search engines have indexes of websites containing certain words and topics, email services must have a list of all their users associated with their passwords and stored emails, most periodicals have an archive of past on-line publications, and so on. In the last few years individual user behavior have also been turned into databases. Services like Google, Facebook, Amazon, Netflix and StumbleUpon largely depend on knowing previous web usage tendencies of individual users. What this means is that on the internet, you are a folder containing information on what you have been doing on-line.

The way services do this is not straight forward. A service must first be able to recognize its users in order to know who's habits it is storing. There are several official ways in which this is done, including looking at an IP addresses, cookies and signing in with an account. There is also a great deal of mythology on how any given service keeps track of its users. According to rumours, a system can determine who you are by monitoring things as sophisticated as how fast you're typing, how many spelling mistakes you make or who have you been e-mailing. There are several reasons why services would want to keep track of their users. The first one is to improve their products based on what users have been doing on them. Another motive is to tailor the information sent to specific user interests, or to sell targeted advertising spaces which provide an income stream for free of charge services such as Google and Facebook.

Google has justified their reasons to track you in their Good To Know campaign, a series of large format signs posted on the walls of the London Underground and elsewhere. The posters explained, in a soothing tone, why it was good for you to hand in information and how it makes all off their services better.

Say you go to the same coffee shop every morning for a latte and the same barista makes it for you every day. Chances are they'll know your order before you even walk through the door. Websites, including Google, have learned a lot form this relationship. We've learned that we can serve you better.

When the web first started, it was a set of static pages that looked the same to everybody. Nowadays, the web has become even more useful because websites can know something about you that helps them guess what you would like to view. For example, they can remember whether you want them in English or French, can suggest books or films that you might enjoy based on what you've viewed in the past, and can store your delivery address ready for your next purchase.

There are many clues that website can guess your preferences, including cookies, IP addresses and signing in with an account.

Google's Good to Know Campaign

There are several things to comment about "Your data on-line". The first one is that it is used to know you better. Essentially this means that a website is trying to make assumptions about your personality based on what words you've written and where you have clicked. This probably is a representation of someone to a certain extent, but it can hardly be said to be a reflection of your complete personality.

The analogy of the cafe barista provides a useful space to further discuss this. If I walk into the same coffee shop everyday my conversation with the barista will not go beyond "give me a latte" and possibly the weather. His way of making my coffee will not depend on what newspaper I bring to the shop, what other coffee shops I've visited and what I've bought there. The worker in this coffee shop is not affected by what I do in other places, is response to me is based only on what I've done in this shop. Therefore he does not try to guess my personality, he merely needs to remember one thing about me; how I like my coffee.

Google's services are nothing like that, especially since their new privacy policy kicked in on March 2012. What Google does now, is that whatever you tell one of its services will affect what the others display. If I do a Google search for "plumbers in London" it is possible that when I visit YouTube I'll see videos related to plumbing first. Rather than a coffee shop worker, Google is more like a shopping mall in which all the shop owners call each other and spread gossip about what you did in their premises. As I said before, this includes all of Google's services, so what I write on my emails affects what search results I see, what YouTube videos I get, what advertisement gets sent and which points of interest are shown when I look at a map; all based on what the shop owners at the mall told each other I like.

Below
Google's new
private policy
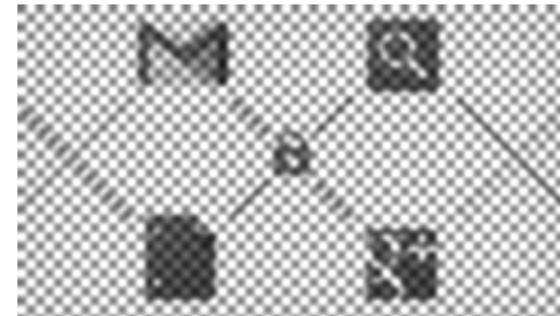
Easy to work across Google

Our new policy reflects our desire to create a simple product experience that does what you need, when you want it to. Whether you're reading an email that reminds you to schedule a family get-together or finding a favorite video that you want to share, we want to ensure you can move across Gmail, Calendar, Search, YouTube, or whatever your life calls for with ease.

Tailored for you

If you're signed into Google, we can do things li suggest search queries – or tailor your search results – based on the interests you've express Google+, Gmail, and YouTube. We'll better understand which version of Pink or Jaguar you searching for and get you those results faster.

There are many existing discussions on the ethical and practical implications of data gathering and filtering. There is a justified concern in terms of privacy, if a company's algorithms are reading our emails and storing everything we search for, what guarantee is there that a privacy policy will not change in the future and our information will be shared with the world?

The other common question is the desirability of filtering. It is true that the internet has grown massively and that finding what your looking for is often next to impossible. However, having software that automatically decides what you are not interested in could mean that your access to information is limited and that a comfortable little bubble is created for you, what Eli Pariser calls The Filter Bubble. Filter bubbles raise broader debates related to freedom of speech and corporate control over the media which I will not go into, albeit interesting as they are.

There is however a less political and technical issue to discuss. This is the relationship between profile making and identity. The moment a service starts tracking users it must give them an identity in order to recognize who they are, the most basic (form of identity on-line is an IP address, data on cookies and an account name). None of these things attempt to communicate anything about the user, they are just a name tag to know what a single computer is doing. The question concerning identity emerges when search results, website visited and words sent in emails are attached to a computer's name. That computer plus the added information attempt to become a representation of the user controlling it. The user thus becomes a series of numbers and words in a digital, computer readable file.

This would still not make a difference to the users working at their desks. When the user starts to notice, when the on-line universe starts to distorted and fit an automatically generated binary version of himself. The interaction between the user and the internet is now mediated by an artificial version of the user. A being similar to an avatar or a character from a role-playing game, with the major difference that users do not choose this representation but embody it automatically. Having no control of the digital being users become, it is not hard to think that their replicas are not necessarily precise copies of who they are.

The interaction between a user and a software mediated universe holds a resemblance to role-playing games (RPGs), in which players exist in an imaginary universe. Their only access to that world is mediated by another player and by a book of rules. To illustrate the similarities between the internet and RPGs we must first look at how RPGs work.
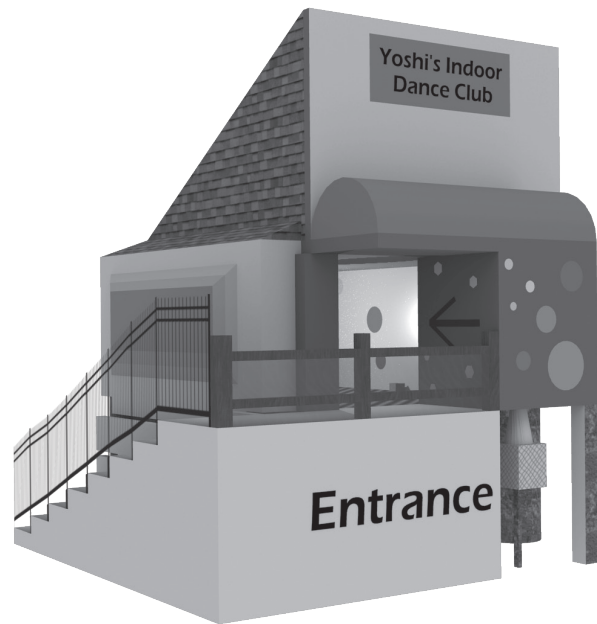
Above
Buffalo 66,
Generated Man search
for Tim Sargent

TIM
Sargent

Artist
(real person)
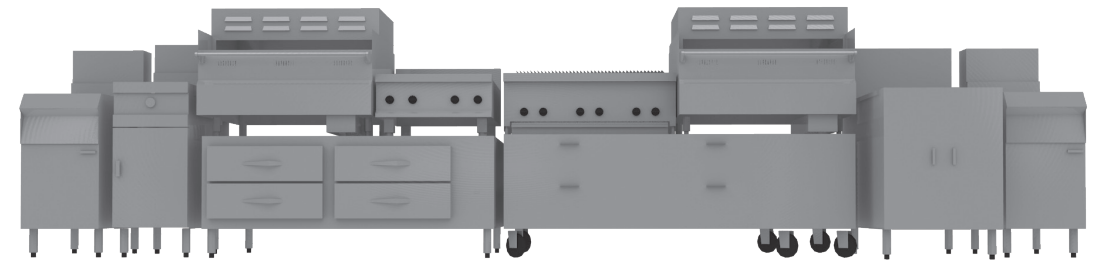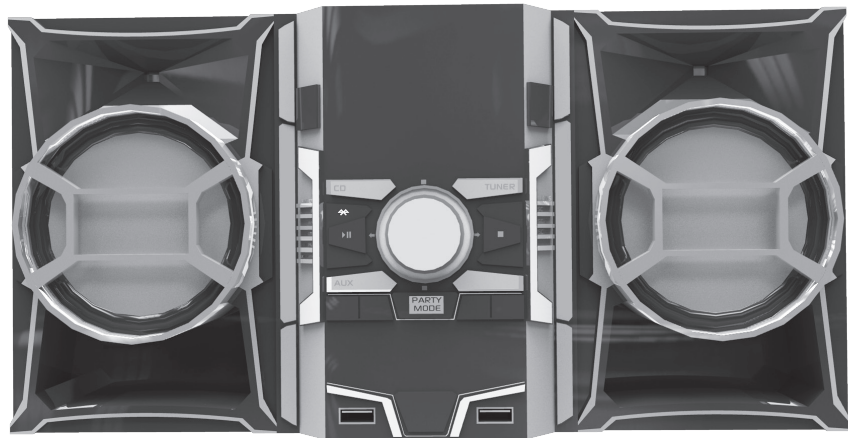
Yoshi's Indoor
Dance Club
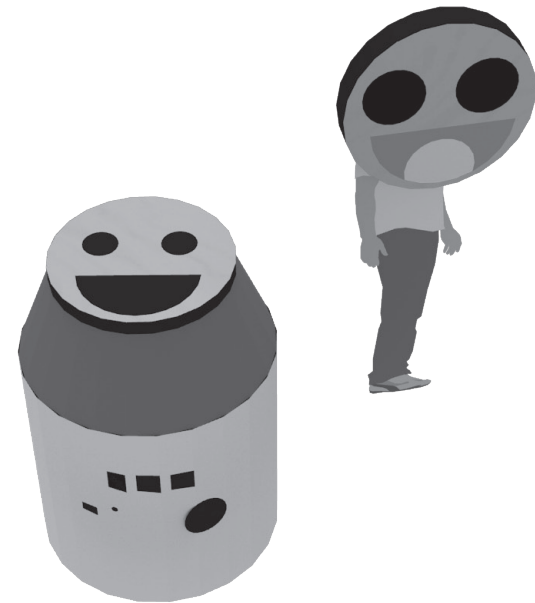
Bike

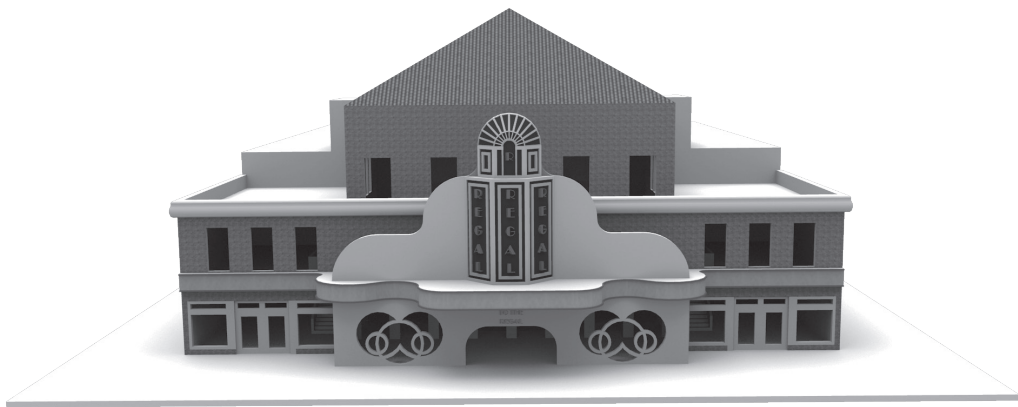Dancing Tango

Kitchen

Boom Box

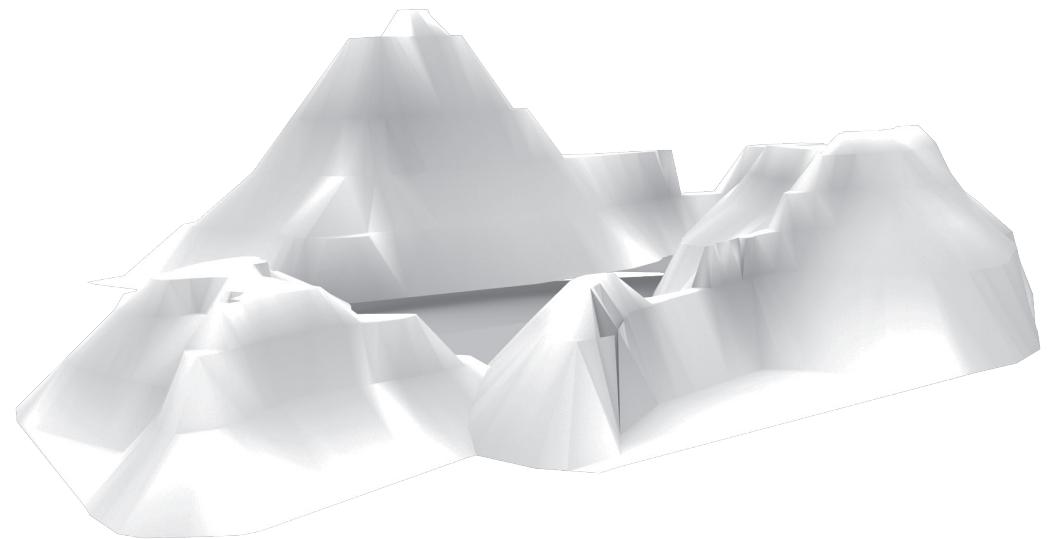Happy Pills

Desert Plants 2D

RJ Music Technology
Master Mind

Regal Cinema

Snow Landscape

RPG

The instant a service creates a user profile it is generating a doppelgänger of the person using it. This virtual persona is essentially a collection of data built from your web usage habits. We can imagine the data includes your name, email address, location, gender, age, interests, music taste, marital status, favorite pastime, hobbies, taste in food, things bought, etc. Basically anything we can learn of you from what you've searched for and visited on-line. We can imagine all this information ends up in a never ending file categorizing every aspect of your personality. In summary, you become a data set.

People have been represented by data sets long before the internet was even thought to be possible. Telephone books, medical records, census, national security numbers and vote registries are all examples of data used to represent people. However, existing in these data forms did not change the way in which you saw it.

Role playing games use a similar data based system for players to determine their character's traits. Character sheets, as they are called, contain everything there is to know about a character. Their name, age, gender, race, clan, physical appearance, magic powers, skills and abilities, all their possessions and their wealth. In order to better understand the link between RPGs and tailored services the following pages will explain how table top role playing works.

Character Sheet

SUITABLE FOR CHARACTERS OF ANY RACE OR CLASS (INCLUDING MULTICLASS CHARACTERS)

DUNGEONS & DRAGONS

CHARACTER NAME _____

PLAYER NAME _____

CLASS AND LEVEL _____

ECL _____ RACE/TEMPLATE _____ SIZE _____ GENDER _____

ALIGNMENT _____ RELIGION/PATRON DEITY _____

HEIGHT _____ WEIGHT _____ LOOKS _____

ABILITY SCORES

	TOTAL		BASE SCORE + RACE/LEVEL	ENHANCEMENT BONUS	MISC. FIELD/BLESS	MISC. MODIFIER
STR STRENGTH		=	+	+	–	STRENGTH MODIFIER
DEX DEXTERITY		=	+	+	–	DEXTERITY MODIFIER
CON CONSTITUTION		=	+	+	–	CONSTITUTION MODIFIER
INT INTELLIGENCE		=	+	+	–	INTELLIGENCE MODIFIER
WIS WISDOM		=	+	+	–	WISDOM MODIFIER
CHA CHARISMA		=	+	+	–	CHARISMA MODIFIER

SPEED _____

GRAPPLE MODIFIER _____ = _____ + _____ + _____
TOTAL   BASE ATTACK BONUS   STRENGTH MODIFIER   SIZE MODIFIER

SAVING THROWS

	TOTAL		BASE SAVE	ABILITY MODIFIER	MAGIC MODIFIER	MISC. MODIFIER	TEMPORARY MODIFIER
FORTITUDE (CONSTITUTION)		=		+	+	+	+
REFLEX (DEXTERITY)		=		+	+	+	+
WILL (WISDOM)		=		+	+	+	+

ARMOR CLASS

AC _____ = 10 + _____ + _____ + _____ + _____ + _____ + _____
TOTAL   ARMOR BONUS   SHIELD BONUS   DEX MODIFIER   SIZE MODIFIER   NATURAL ARMOR   DEFLECTION MODIFIER   MISCELLANEOUS MODIFIERS

TOUCH AC _____   FLAT-FOOTED AC _____

ARMOR WORN		MAX DEX	ARMOR CHECK PENALTY	WEIGHT
SHIELD CARRIED		MAX DEX	ARMOR CHECK PENALTY	WEIGHT

COMBAT OPTIONS

HIT POINTS

BASE ATTACK BONUS

WEAPON	ATTACK BONUS	DAMAGE	CRITICAL
RANGE INCREMENT	TYPE	NOTES/AMMUNITION	
WEAPON	ATTACK BONUS	DAMAGE	CRITICAL
RANGE INCREMENT	TYPE	NOTES/AMMUNITION	
WEAPON	ATTACK BONUS	DAMAGE	CRITICAL
RANGE INCREMENT	TYPE	NOTES/AMMUNITION	
WEAPON	ATTACK BONUS	DAMAGE	CRITICAL
RANGE INCREMENT	TYPE	NOTES/AMMUNITION	

INITIATIVE MODIFIER _____ = _____ + _____
DEX MODIFIER   MISC. MODIFIER

CONDITIONAL MODIFIERS

SPECIAL DEFENSES
_____
_____
_____
_____

There is a great variety of role playing games, however they all work in pretty much the same way. Most games have a core book, this book can be seen as a very long instruction booklet. It contains all practical aspects of game-play as well as providing a description of the universe in which the game takes place. If the game is Sci-Fi themed it will contain all the planets that exist and their major cities. If it is a medieval magic game, it will contain a description of the mystical land with all its buried treasures and shrines. The book also explains what creatures or races the player can choose from. Each race has certain powers and advantages along with specific weaknesses. For example, dwarfs could be very good fighters but poor wizards. Likewise a halfling may be terrible at melee combat but could be a powerful magician.

After choosing a race players decide what class their character's belong to. A class is similar to a profession in an adventure world includes things like: wizard, thief, bard, paladin, dragon tamer, etc. Whilst in a Sci-Fi universe it would be composed of categories such as pilot, soldier, Jedi, bounty hunter, spy. The race and the class together provide the character with a unique set of abilities that only players of that race and class share.

Once the type of creature and its class have been defined the player fills in the rest of the character sheet. Most of the remaining fields are shared between players and include attributes and abilities of the character. For example, how good looking he is, how charismatic, his ability to speak French, how empathetic and manipulative he can be and so on. At the end of the character creation procedure players usually add extra things to support their character's back story. A character may be a veteran from an old and romanticized war and to emphasize this, I might say that he limps on his right leg as a result of an ancient injury.

Creating characters is a long process and is considered part of the fun in RPGs As a player you are making a projection of yourself to exist in an imaginary universe which provides you with freedoms inexistent in the real world.

The players interact with this world through a story teller that tells them where their characters stand and what they see. The storyteller has the most important role in the game; he must be familiar with all the game rules and have an overall idea of the in-game universe. He must also be creative and literate enough to make up a story which the players will enjoy and he must know how to react to what the players choose to do with their characters.

Game-play consists of all the participants sitting around a story teller. He then starts to narrate what the characters are seeing and may even tell a player what he's feeling. Something along the lines of "You are awoke by the cold of the stone floor on which you are sleeping. You are in a dark room with no windows, it is moist and there is a draft of wind blowing from under the door, you hear talking outside but cannot make out the words." The players must then tell the story-teller what they choose to do. The story-teller's role at this point is to look at their character sheets and see if they have the necessary skills to do things.

CO-EVOLVING NPC

Name: TIM
Age: 24
Gender: M

Ocupation: Student
City: London
Country: UK

Languages Spoken:
English

TASTE

FILM 3x

Buffalo 66
Robert legacy
predator.

MUSIC 3x

OMAR Souleyman
gamelan
Leghor acid crew

FOOD 3x

lentals
buckwheat
haricot beans

INTERESTS

Dance	●○○○○	1
Music	○○○○○	10
Bike	○○●○○	13
colour	○○●○○	17
cooking	○○●○○	21
≠ Stories	○○○○○	26

folk	○○○○○	31
E Smell	○○○○○	33
taste	○○●○○	35
landscape	○○○●○	39
trees	●○○○○	40
	○○○○○	

40+

KEY WORDS

buddist
happy   joy
proposal
Mutualism
Cybernetics

15x
Speed
aethics
Spirits.

Let's suppose a character has the power to read minds and attempts to read the mind of the mysterious people outside the stone floored hut. Even if he has the skill there is no guarantee that he will perform well, so the player must roll a number of dice proportional to how good he is at reading minds (e.g. if he has 3 mind-reading points he's allowed to roll 3 dice). The outcome of the dice roll defines how well the skill worked and the story-teller must continue his narrative taking this into account.

The role of the story-teller is that of a mediator between the world and the character. The way he describes particular parts of the universe and how they are perceived is based on his understanding of the player's characters. The whole way in which he constructs the world must be appealing and challenging to the players. However, the only ways a story-teller has of knowing a player, is through their character sheet and based on previous reactions to other parts of the story.

At this point we can start to draw similarities between internet filtering and role playing games. On the internet we have a personal profile that a service fills in for us based on our previous interaction with it. It then uses the information on this profile to deliver the parts of the world we are interested in. The filtering service acts as a digital story-teller, looking at our profile and then rolling dice to decide what to show us next. The big difference between this world construction and an RPG's universe is that in the game we decide what our character is whereas on the internet it is being decided for us.

Generated Man is designed to tell us how our character is perceived on the internet. It uses an interaction between both character creating procedures. A user creates a character which starts to browse the internet, then as a response to this, a service's algorithms create a user profile for that character and starts to craft a universe for it. As time passes both the initial character created by the user and the profile created by the service learn from each other, causing a feedback loop that generates narrowed down identities for the characters.

Generated Man

After spending sometime submerged in an artificial world followed by the same advertisement, and starting to accept the comfort created by familiar search results I began to question how the system perceives humans. What kind of image does it construct of us with its constant gathering of data? I imagined the data to be a cabinet of curiosities, such as the ones created as displays of wealth and power during the XVIII and XIX centuries. By looking into these private collections one could get valuable information about the collector. The digital successor to this, constructs a little cabinet with all the web pages someone has visited, the things he has bought, trips, dates, what news papers he's read. This is the character someone ends up being more then a humanoid with complex psychologies he ends up being a collection of thing he has gathered, a very long log of websites he's seen and places he has visited.

To better understand what image of us is being created I decided to construct a computer program. The design of the program deviates from standard software design criteria. The algorithm this program follows was not made thinking in terms of optimization and hardware performance. Instead, it was manufactured to reflect my understanding of the process by which we are assigned a personality. Being almost impossible to get the directory that represents you on an internet service, I decided to reverse the process. Rather then having a character as an output it would be the initial data used by the program to interact with the internet.

It was at this point that the link to role playing games became obvious; I would feed my program a character, with certain interests and keywords. The software would then creates search terms out of these words and feed them into a recommendation algorithm. The original data I fed in, would serve as a comparison point to whatever the internet spat out, allowing me to gain insight into what the user-profile making computation was doing. The first step towards writing this program was to define the shape of the input data.

Given my inspiration from role playing games, I decided that the format of the input data should be heavily based on a character sheet. The user is asked to fill in a form-like sheet divided into four sections: Personal Details, Taste, Interests and Key Words. Like in a table top RPG the user can decide to provide a realistic representation of him or herself or make up a more interesting alter ego.

The personal details section contains information used mainly to identify the character as well as to generate search terms based on the city and country of the user. The taste category is divided into 3 sections: Film, Music and Food. On each of these sections the user must enter three things that he likes related to an specific header. For example, under film he may enter "Star Wars", "George Lucas" and "X-Wing". For music "Imperial March", "John Williams" and "Soundtrack" and finally for food "Pizza", "Wookie Cookies" and "Fried Chicken".

After the the user's tastes have been captured it is time to fill in his interests. The interests section of the character sheet is the most important and complicated one to fill. The user is asked to write something that he's interested in, which can either be a single word or a phrase. Afterwards, he must rate from one to five how much he is interested in that topic. There are 40 points to be spent in this section Each new interest starts with a rating of one and costs one point and additional points the user adds are subtracted from the initial 40. In the end, this section may contain a minimum of eight interests, all rated five; or a maximum of 40 interests, all rated one. Interests come up more often when making search terms and are weighed higher than taste and keywords when rating a website.

The last section to fill in is the Keywords one. Keywords look very similar to interests they should be a word or a phrase representing something liked. The big difference is that keywords are not rated. They are just there to become search terms and have a weaker influence than interests when rating a website.

Once the input data is ready it is fed into the program. The Generated Man algorithm has four general steps. First it loads a character sheet, it then uses Google to search the internet based on the words the sheet has. It rates the websites in Google's search results and clicks on the highest rated one. Finally it chooses a 3D model from Googles 3D warehouse to represent the constructed search.

Loading the character sheet only happens when the program is started. It is a very simple process in which the words in the sheet are turned into categorized computer usable variables. The key point is that they can be recognized as personal details, taste, interests and keywords as, these categories will be used to navigate on-line.

The most complicated and important part in the program is searching Google. It is through these searches that Google creates a vision of who the input character is. It is also at this stage that the character adds new words based on what Google recommends too.

At the beginning of the Google search phase a query is generated. To do so, the program uses the categorized words that it got from loading the character. It needs to decide if it will search for interests, keywords or tastes. This decision is made through a psudo-random number generator. The program basically throws dice, as it would do if it were playing Dungeons and Dragons. It rolls one four sided dice to decide what set of words to choose from. If it gets a 1 it looks for a taste, if it gets a 2 it looks for a keyword. If it gets a 3 or a 4 it looks for an interest, this makes interests appear with a probability of 50%.

Likewise, the program needs to decide if it will look for a phrase by its self or if it will add other words. It can add words it has learnt on previous searches or it can add a prefix from a predefined list. The available prefixes are 'good', 'bad', 'review', 'free', 'best' and location, constructed by combining the city and country given in the character sheet.

The generated query, which is already a reflection of the character, is fed into Google's search engine, one letter at a time to simulate typing. The search results that come up are then analyzed by the program. The first thing that happens is that all the words in the search results are sorted according to how many times they appear. The words in the query and some very common English words are ignored from this list. Once the program has an ordered record of all the words, it associates the most common word to the words in the search terms. For example, if the search terms were "Food red meat" and the most common word in the record was "Nutrition" then the word nutrition would be associated with the words food, red and meat.

What this association means is that once the program starts to repeat search terms, it can make them more specific by adding words associated to the ones on the characters sheet. As the program adds more words to its association dictionary some of this words can become part of the character sheet, this means that they start to be used to rate websites and not only to create queries.
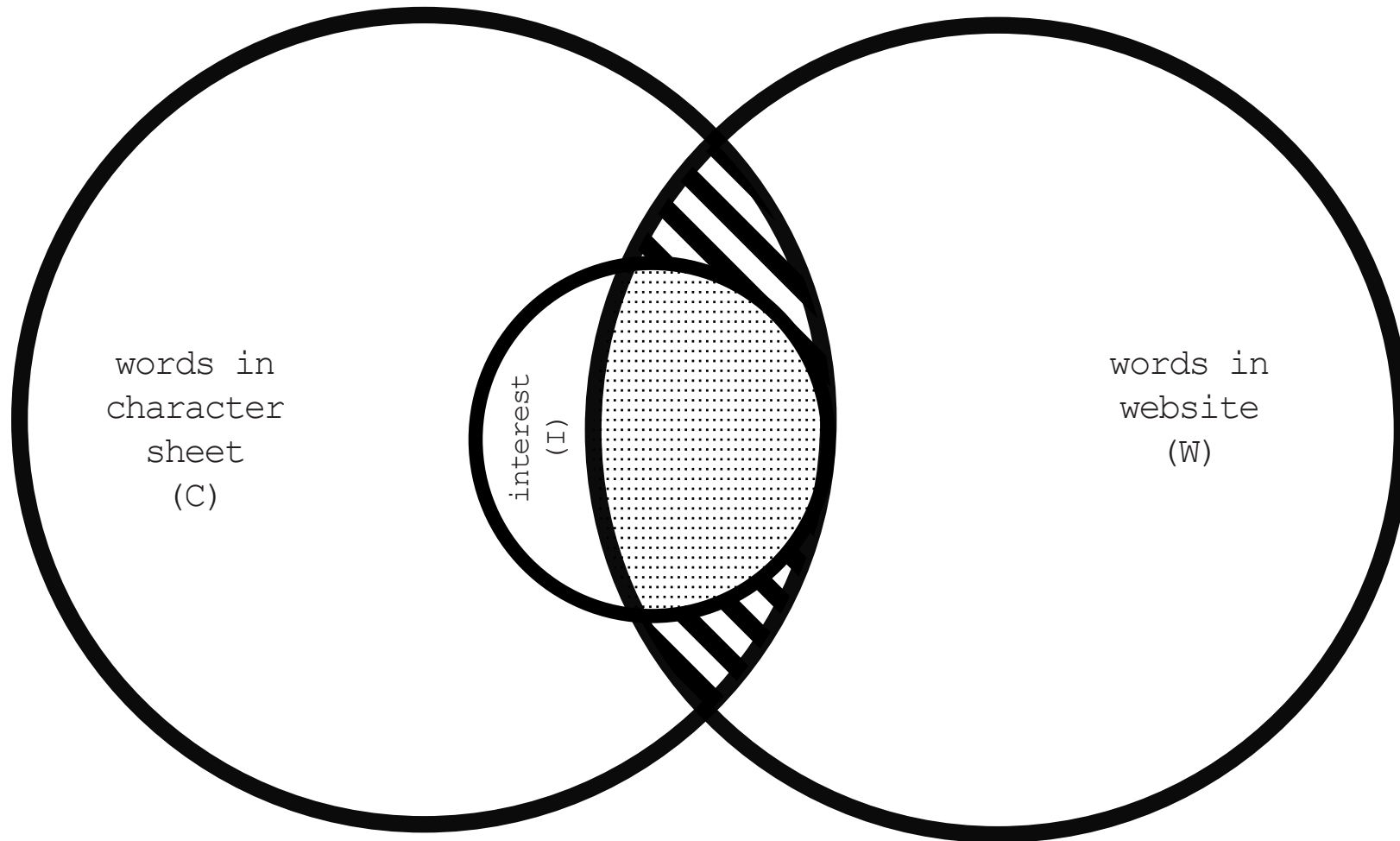
New words are integrated into the character sheet every time the number of words in the association dictionary duplicates. This means that at the beginning, when we first run the program it will be easy to get new keywords and interests. However, as time passes it will become harder for the character to like new things, it will become more secure of what it likes and will repeat it searches more often. This will cause his Google counterpart to do the same, to create an ever more specific vision of who this character is.

From a role playing game perspective, the association dictionary acts as experience points. Every time you perform a search you get more words, every time the number of words is duplicated you tweak your character sheet a bit, as it happens when you level-up in an RPG. In Generated Man, levelling up means integrating the two most common associated words into the keywords section of a character. It also means that the two keywords with the highest number of associated words turn into interests. Finally, any interest that has been associated while looking for other things, will gain an extra point, making the character more interested in that particular topic.

Previous Page
Top Gear Car Race,
Generated Man search
for Oliver Brown

Rating (R) System

$$R = \frac{|C \cap W|}{|W|} \; \left(\sum_{i \in W} n_i\right)$$
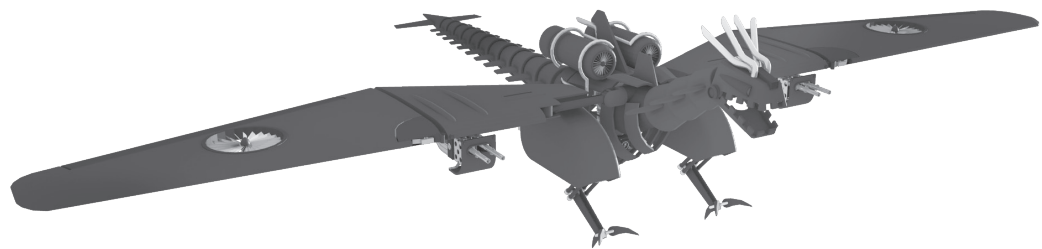
n: I→{1,2,3,4,5}

words in
character
sheet
(C)

interest
(I)

words in
website
(W)

ALEX
Woods

10 Year Old Boy
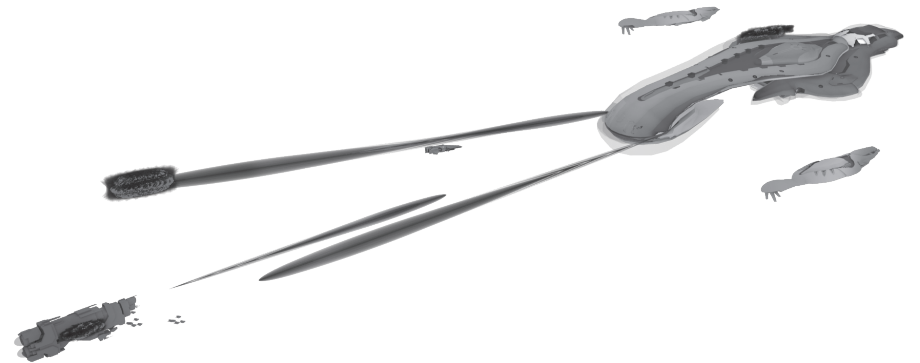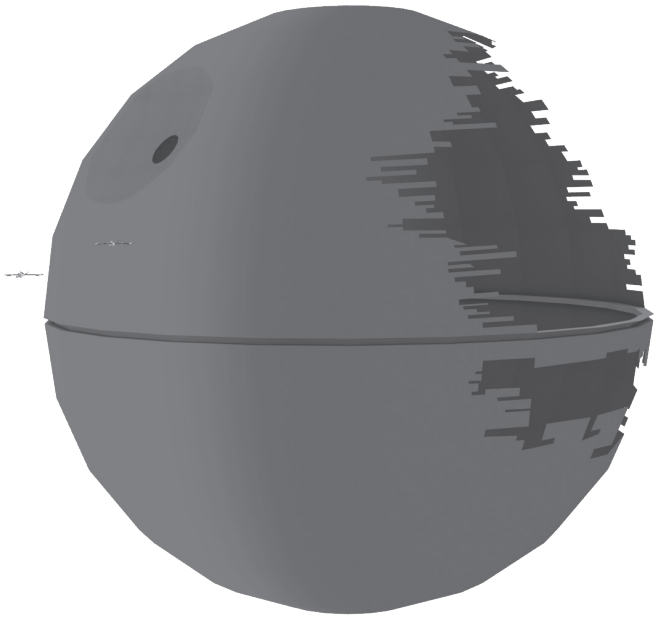(fictional person)

Mech War Dragon

Fire

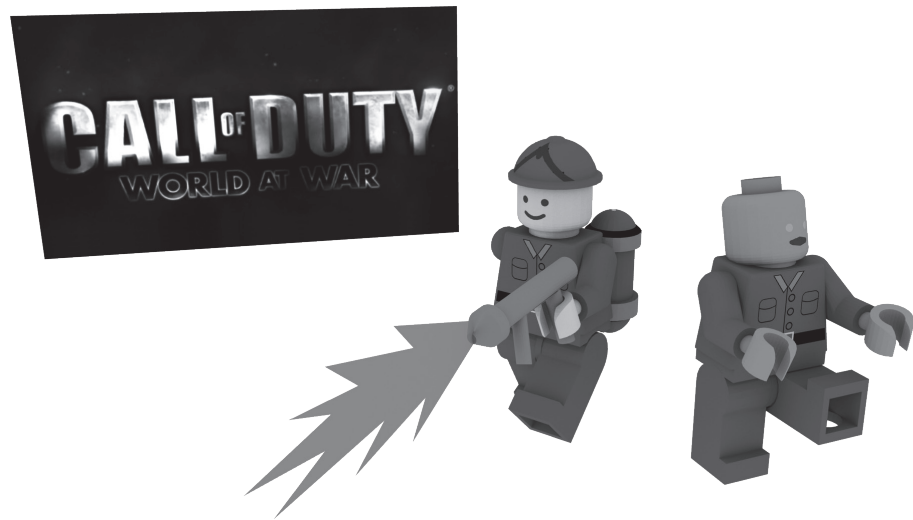Lego Han Solo

Halo Spacebattle

Death Star Return
of the Jedi

Discovery Taking
Off
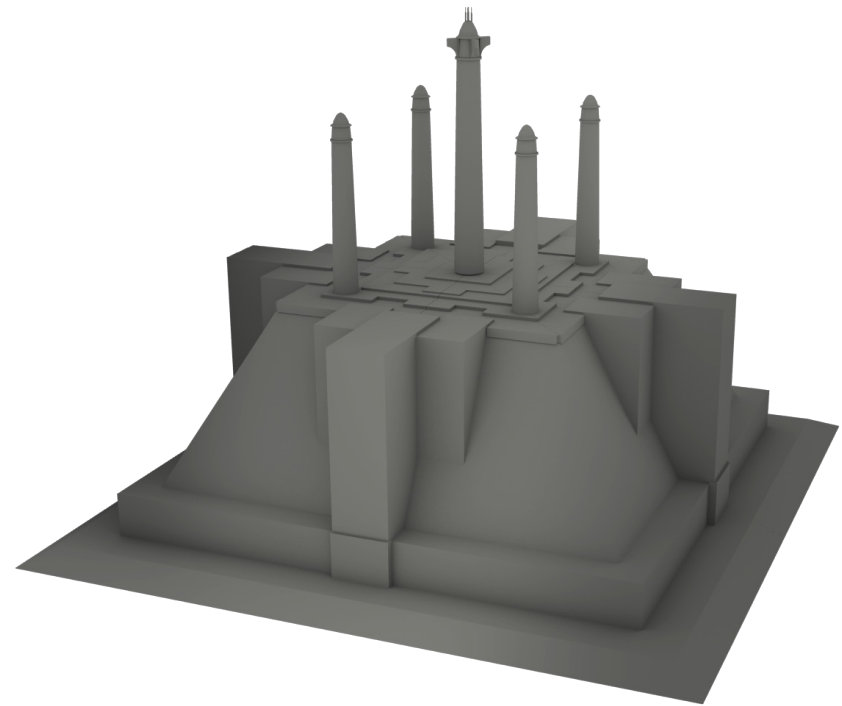
Fire Truck
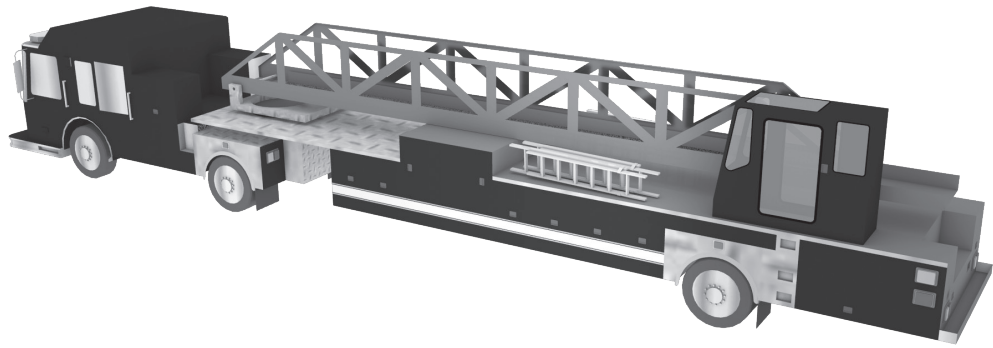
Take That Bots

Coevolution

From the previous section we know that Generated Man works through a dialogue with Google. The interesting dynamics arise when Google responds to the information provided by Generated Man and when a character learns new words from Google.

When a Generated Man character is first started, Google knows nothing about it until it begins to search for words from its character sheet and at this point, Google creates a profile for it. The algorithms on Google's side start to collect and process all the web searches and sites visited by the character. After a short while, Google starts to tailor its search results to fit an approximation of what the character wants to see.

On the character's side of the system, it is searching for certain words and is learning related words from Google's search results. Before Google has a clear profile for a character, the words displayed in the search result reflect broader internet culture rather then the character's Google profile. As Google starts to tailor results, the new words the character can learn are all words that have been somehow filtered to match its 'personality'.

The next time it searches, the new words in its vocabulary are fed back into Google, providing new search results from which the character must choose a website to visit. This searching, learning and clicking cycle creates a feedback loop in which both Google and Generated Man react to, a modified version, of what they send out.

In more practical terms this means that the character's personality is being reassured as it interacts with Google. It searches for an interest on its character sheet, Google recommends some results and it learns from those recommendations. The recommendations in the search results are built on a distorted mirror image of the character, therefore the words it learns come from a narcissistic chat with him self. This causes his existing interests to become more specific and provides closely related new interests.

As the new words are used to rate websites, the character will highly rate websites containing Google taught words. Google's reaction to this, is to tell its own version of the character that it likes websites similar to those, hence, Google will displays related sites next time there is a search for similar topics. By looking at the words the character has associated with previously known words, it is possible to start constructing an idea of what the character's taste is.

Lets look at a real example of this happening. The character Tim Sargent, a real person, has the word "dance" in his character sheet. After a while the associated words were the following:

dance:net, belly, people, technique, exotic, music, pole, ever, play, fitness, workshop, someone, ballet, schools, classes, html, perform, job, world, rude, forms, list, traditional, presented, shoes, several, video, performed, without, surprise, movement, delivery, type, girl, leading, watch

Some of these words are pretty meaningless, things such as "ever", "someone" and "without" apply to dance as much as they apply to anything else. However, the rest of the words do describe what Google (and internet culture to an extent) think of the character performing this searches. The words "exotic", "world", "traditional", "performance" and "music" are associated to some more of this character's interest. For example, the world "folk" from the sheet has associated some matching words:

folk:music, traditional, design, little, acoustic, popular, song, well, world, culture, find, home, festival,year ,subject, history, live, creative, official, rustic, scene, album, known

The character's image drawn from this words is that of a man interested in foreign cultures and their traditions. The meaning of the words folk and dance have been altered to fit Tim Sargent's interests. By looking at their adjusted meanings, a lot of information on how he is perceived on the internet comes across.

Under Google's new privacy policy the meaning it gives one word is shared across all of their services, meaning that the words typed into the search engine will be interpreted in the same way when typed into any other of Google's websites. By repeating the same search on Google's 3D Warehouse it is possible to see Google's interpretation of a word represented by a 3D model.

Creating a collection of 3D models does not lay entirely on Google's side, once again it is a cooperative process between Google and Generated Man. The 3D Warehouse offers some models matching search criteria, however it is down to the character to choose which one represents him, this choice is down to the words in the character sheet.

Time plays an important role in this co-evolutionary process. For Google to create an image of the character it must first have gathered enough information about the character. In other words, the character must have performed enough searches and clicks. The same applies to Google's influence on the character sheet. At first there are only user written words on the sheet but after enough searches new words are added from the association dictionary. As Google evolves to match the character, the character also evolves to please Google. The comparison with biological co-evolution is unavoidable.

The character and Google are symbiotic life forms, as such, the changes in one demand that the other adapts. As Google struggles to create a comfortable environment for the character, the character works towards becoming part of this environment. Google becomes a zoo keeper and the character becomes a tame animal, pacing around his cage, knowing every detail of it, until it is convinced that this artificial habitat contains all of the universe. When hungry, it only needs to demand food, which the zoo keeper will provide, preprocessed, medicated and cut down to bite sizes which the caged animal will have no problem understanding and digesting. If the character starts to do unexpected things, then Google must modify the created bubble to new habits.

The real life system is much more complicated than the metaphorical virtual zoo. Google has limited resources to construct a universe for constantly changing user behaviours. Google cannot provide any website it does not have in its indexes. The same applies to the 3D models used to represent a person, if there is no model for the search "good concert festival" then that phrase has no graphical representation. It could be argued that the lack of an object is also part of how the internet sees people. If there is no results for what a user is interested in, then that interest will simply be ignored and there will be holes in his virtual personality.
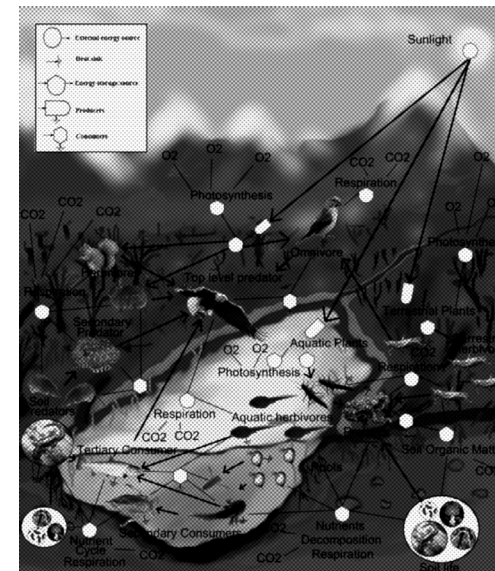
Closing Remarks

Internet services are likely to track users more and more as the technology matures. Eventually we'll be as used to tracking as we are to every other technology that we've taken into our lives. The popularity of smart phones and omnipresent mobile networks already allow tracking to happen outside of our homes. Your assigned online character no longer sits browsing on his desk but actually goes out to places, which end up being part of his constructed personality.

As the virtual world becomes more real, it starts to hold representations not only of people, but of everything else too. Our representations as semi-fictional characters will start to inhabit semi-fictional worlds. This merging of physical and digital reality makes me wonder what forms this half fiction will take in the future. When the real and the virtual coexist spatially (as they already do in sat-nav systems) will our mismeasured user profiles follow us into an off-line world? Or will we be able to turn our characters off and stand away from the dice and the story tellers that mediate our interaction with a fitted universe?

This project has shown our representations are not necessarily accurate. Being unable to manually tweak them, as one does in an RPG, means that we are imposed some categories to which we don't really belong. Profiling itself is not bad, as I've repeatedly stated in this book; it allows services to improve and stay free of charge. When profiles become dangerous is when they limit user freedom by forcibly tailoring content, or when privacy policies change and your information is shared in ways that affect your life.

All these problems are already being addressed and discussed by academic groups, policy makers and everything in between. While I was writing this paragraph, on May 28th 2012, the technology blog Ars Technica published a note on a law to be passed in Britain. The "cookie law" as it has been called, will prohibit tracking without consent. This means that any website wanting to create or add to an on-line profile must first notify its users. This is a positive first step into creating awareness about tracking and informing people how the information sent from their computers alters what they are viewing.

Acknowledging and questioning the fiction that intrinsically populates cyberspace seems the next essential step in the discussions around profiling. For the technology to grow healthy and useful developers, users and governments have to understand and accept. That information gathered on browsing habits is not the person browsing but a skewed representation of their interests. If, on the other hand, we deny this fiction and allow our characters to be used to tailor our world, or if we give them legal value, then the parts of our world that must remain real will conflict with the technology, which will grow dangerous and manipulative.

Concept
Diego Trujillo

Book Design
Ilona Gaynor

Thankyou
Nina Pope
James Auger

Anthony Dunne
Fiona Raby
Noam Toran

Tim Olden
Brigitte Lelievre

Tim Sargent
Jaemin Paik
Ilona Gaynor
Naomi Elliot

3D Warehouse
Python
Mechanize
BeautifulSoup
Selenium

I would like to thank the tutors in the **Design Interactions** department at the **Royal College of Art** for their guidance throughout this project. Many thanks to **Tim Olden** and **Brigitte Lelievre** too for all their technical and spiritual support. Thanks to **Tim Sargent**, **Jaemin Paik** and **Ilona Gaynor** for taking the time to fill in some test character sheets. An extra thankyou to you Ilona for designing this book. Thankyou to **Naomi Elliot** for taking the time to proof read this book.

Finally I would like to thank all the modellers from the **3D Warehouse** for sharing their models with the world, as well as the developers of **Python**, **Mechanize**, **BeautifulSoup** and **Selenium** for releasing their work under free software licences.

Previous Page
Roast Dinner,
Generated Man search
for Oliver Brown

© 2012 Diego Trujillo