Inspiring healthy habits: data science at WW

Carl Anderson

carl.anderson@ww.com
@leapingllamas

Outline

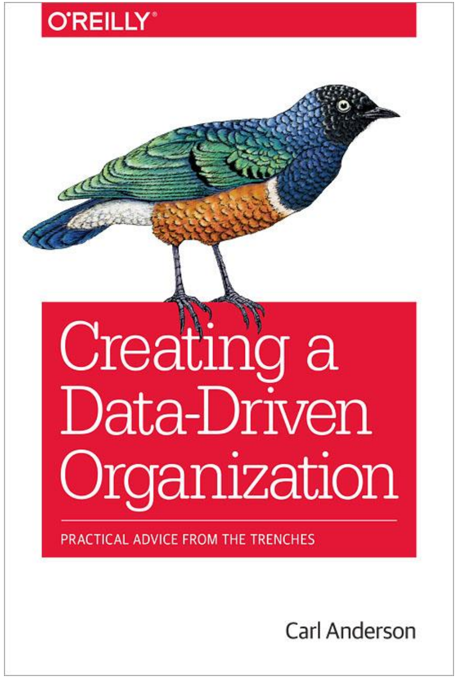
- Intro to WW: purpose, program, type and scale and data
- Behavioral Nudges
- WW Data Products
- Primrose: how we develop and deploy ML models
- Q&A

About Me

Data Strategy

Data Science

Business
Intelligence
Engineering

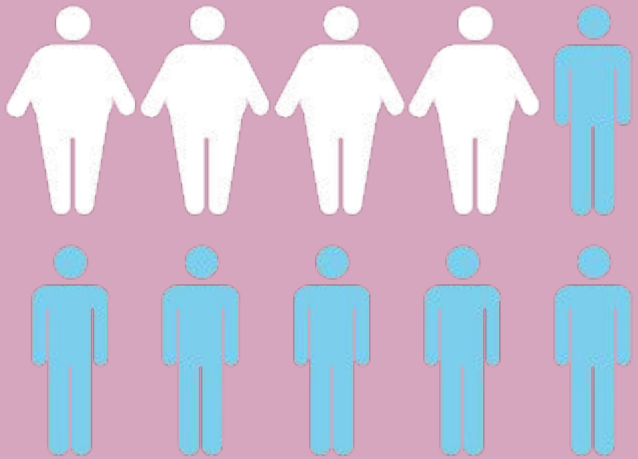


KAISER PERMANENTE®

Durham University · University of York · The University Of Sheffield. · AARHUS UNIVERSITET · Duke University · Universität Regensburg · Georgia Tech

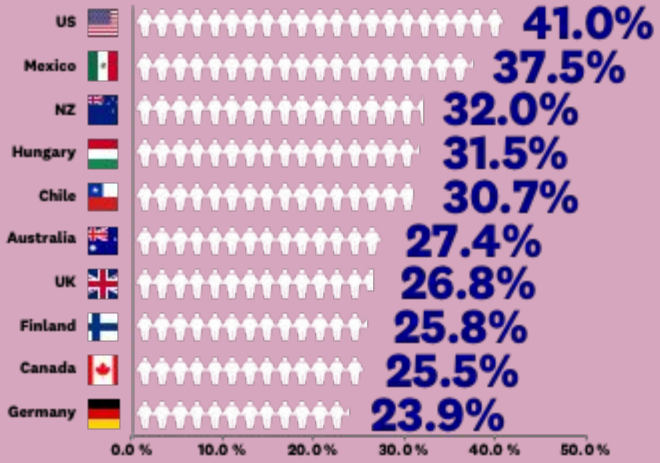

TIME
The Obesity Epidemic in America

40% of adults on the planet are already overweight or obese and more are joining their ranks every day



While the world's bulging waistlines are driven by economic success—wealthy populations eat more—obesity's estimated cost of $2 trillion a year worldwide is now almost as much of a financial burden as smoking.

Top 10 countries ranked by percentage of adult population that is obese in 2015 or nearest year



Source: TIME Health, New England Journal of, Medicine Institute of Health Metrics and Evaluation, OECD (2017)

HEALTH PARADOX #1

We spend more time and
money than ever before on wellness,
but we've never been
more unhealthy.

HEALTH PARADOX #2

Despite all the advances
in science and food production,
eating healthy
has never felt more complicated.

Confusing headlines have left many in a fog and unsure what to do

5 Reasons You Need to Count Calories	vs	5 Reasons To Never Count Another Calorie
Why Sugar isn't the Bad Guy	vs	10 Disturbing Reasons Why Sugar is Bad for You
Why Gluten is Evil *Grains and Autoimmune Disease*	vs	JUNK SCIENCE: Gluten is Not Bad For You
THE SECRET IS OUT: THE REASON RED MEAT IS BAD FOR OUR HEALTH	vs	RED MEAT: It Does a Body Good!
THE ORGANIC FOOD LIE	vs	New study finds organic foods are healthier than conventionally grown foods

People want inspiration, not just information.

People want healthy habits that fit their lives.

Today, healthy is the new skinny.

It’s not about lifestyle, it’s about livability.

Community is essential.

People crave purpose.

OUR PURPOSE

We inspire healthy habits for real life.*

* For people, families, communities, the world—for everyone.

WW: Wellness Ecosystem

TENET #1

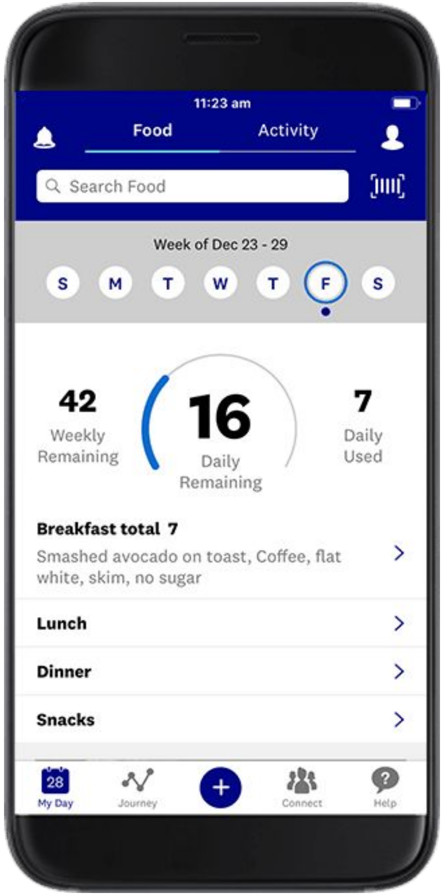
We help you build powerful habits, rooted in science.

TENET #3

We enable you to find and form inspired communities.

TENET #2

We know you and meet you where you are.



Wellness that Works.™

SmartPoints is about health, not just calories

All calories are NOT created equal.

300 calories

Greek yogurt & fruit and peanut parfait

Scrambled eggs, Canadian bacon, avocado, tomato and English Muffin

Pancakes with chocolate chips and maple syrup

Nutritional science to make healthy eating simpler

Nutrition Facts

Instant Oatmeal, Flavor Variety

Serving Size: 1 packet (43g)

Amount Per Serving	
Calories 160	Calories from Fat 25
	% Daily Value*
Total Fat 2.5g	**4%**
Saturated Fat 0.5g	**3%**
Trans Fat 0g	
Cholesterol 0mg	**0%**
Sodium 260mg	**11%**
Total Carbohydrates 32g	**11%**
Dietary Fiber 3g	**12%**
Sugars 9g	
Protein 4g	
Vitamin A	20%
Vitamin C	0%
Calcium	8%
Iron	10%

* Percent Daily Values are based on a 2000 calorie diet.

SmartPoints nudges you towards a healthy eating pattern with more fruits, vegetables and lean protein, and less sugar and saturated fat.

- **Calories** establish the baseline.
- **Sugar** and **Saturated Fat** increase the SmartPoints value.
- **Protein** lowers the SmartPoints value.
- Foods that form the foundation of a **healthy eating pattern** have SmartPoints value of **zero**.

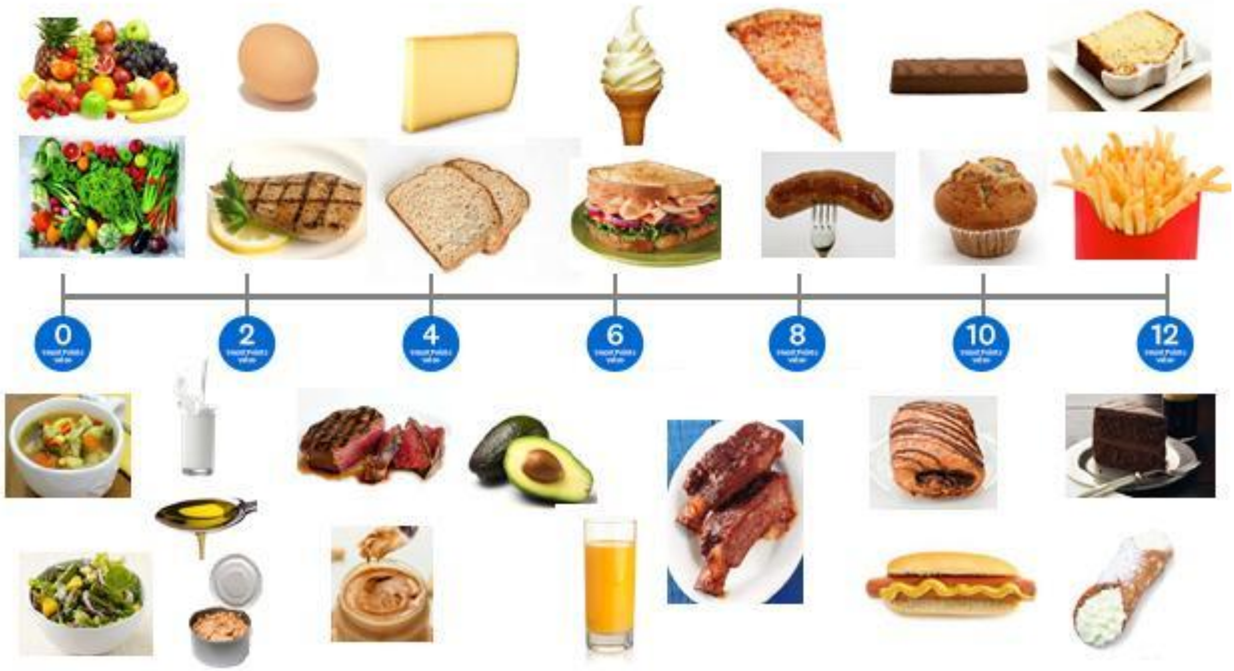
ZeroPoint™ Foods

ZeroPoint foods form the foundation of a healthy eating pattern and have a low risk of overeating.

They don't have to be weighed, measured, or tracked.

- Vegetables
- Fruits
- Skinless chicken breast
- Skinless turkey breast
- Non-fat plain yogurt
- Eggs
- Beans, peas, lentils, tofu and corn
- Fish & seafood

Everything is on the menu

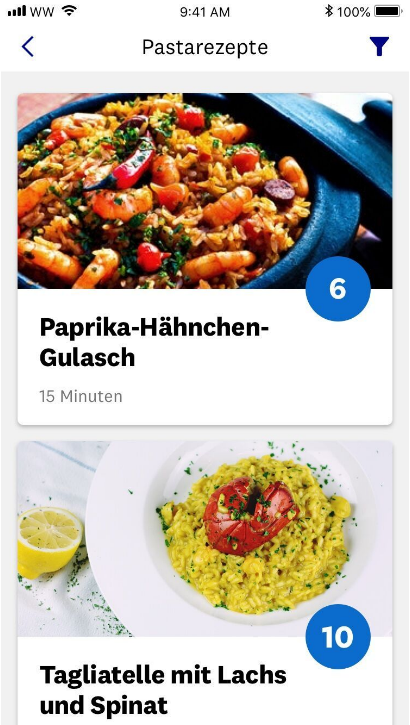


App


WW
9:41 AM
100%
Community
Was machst du gerade ...?
Beliebt
Neu
Abonniert
ga-diva16
Healthy in Heels
TGIF! Ich bin gerade dabei ...
5
Mein Tag
Mein Weg
Community
Hilfe


WW
9:41 AM
100%
Lebensmittel
Aktivität
Suchen
12 von 28
FitPoints
15.237
Schritte
126
Minute
5
S
M
D
M
D
F
S
Woche von Di 09.03. – Sa 15.03.
Du möchtest dein Ziel anpassen?
5
Mein Tag
Mein Weg
Community
Hilfe

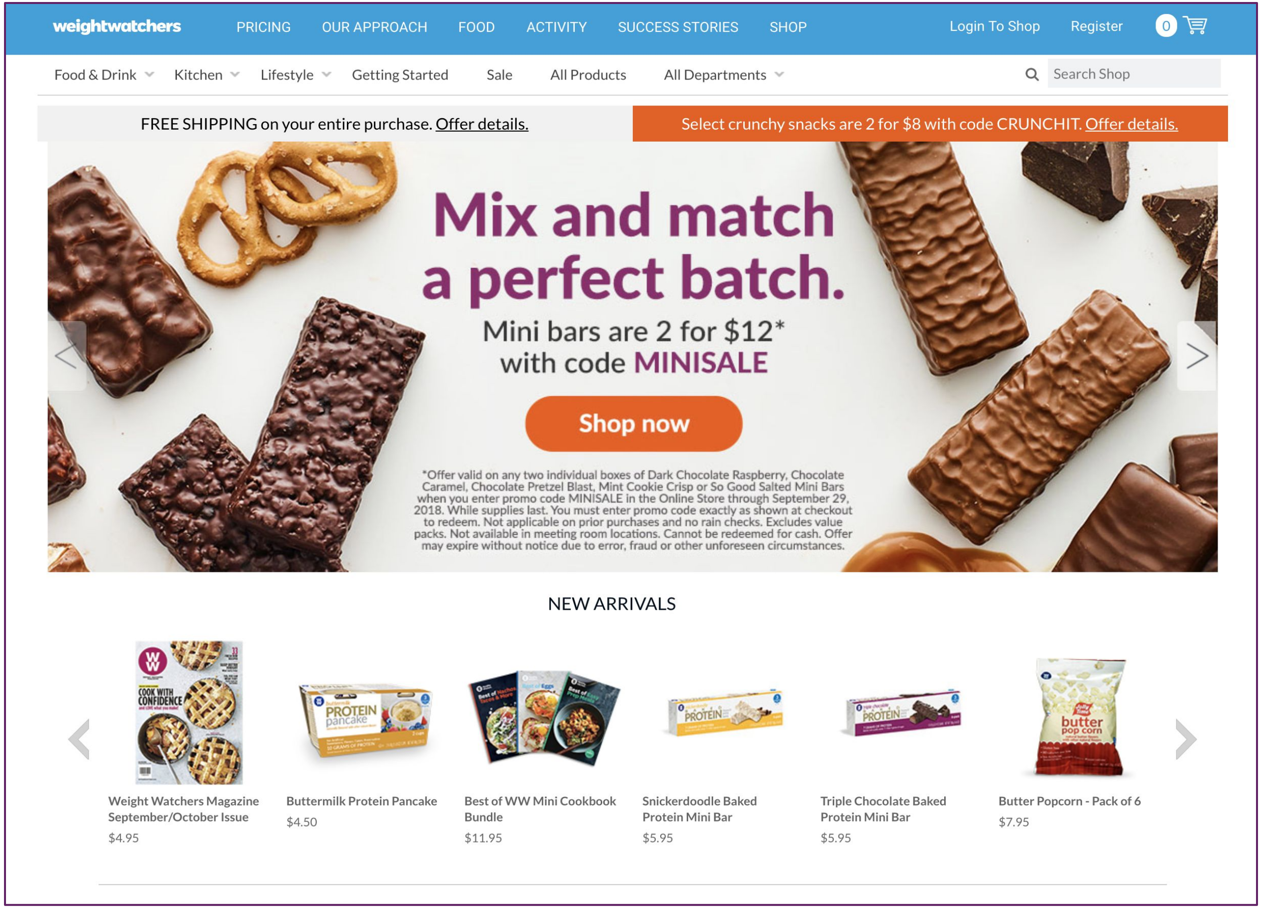

WW
9:41 AM
100%
Mein Weg
Mein Gewichtsverlauf
101
98,5
25.01. 01.02. 08.02. 15.02. 22.02. 29.02. 07.03. Heute
Tage im grünen Bereich
Mai
S M D M D F S
8 9 10 11 12 13 14
15 16 17 18 19 20 21
Mitteilungen
5
Mein Tag
Mein Weg
Community
Hilfe


WW
9:41 AM
100%
Pastarezepte
6
Paprika-Hähnchen-Gulasch
15 Minuten
10
Tagliatelle mit Lachs und Spinat

WW Studio

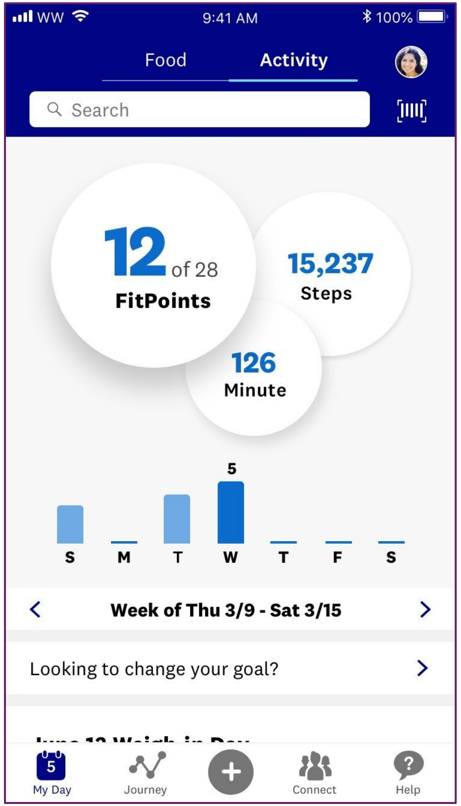
30,000 meetings per week globally

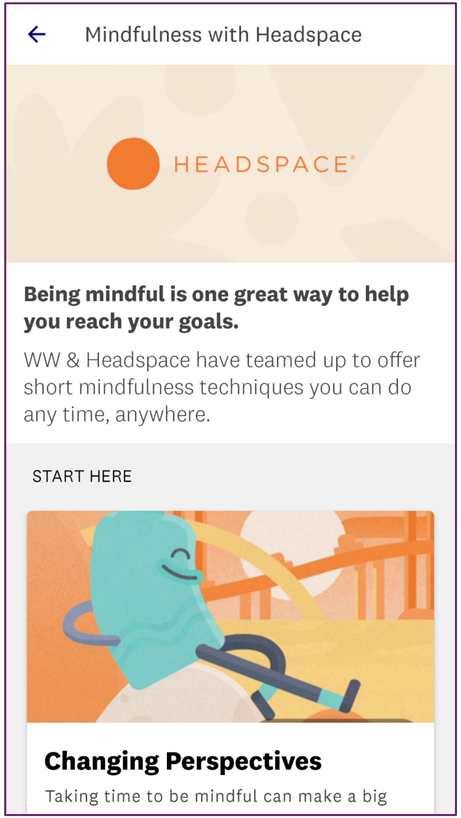
eCommerce



weightwatchers
PRICING
OUR APPROACH
FOOD
ACTIVITY
SUCCESS STORIES
SHOP
Login To Shop
Register
0
Food & Drink
Kitchen
Lifestyle
Getting Started
Sale
All Products
All Departments
Search Shop
FREE SHIPPING on your entire purchase. Offer details.
Select crunchy snacks are 2 for $8 with code CRUNCHIT. Offer details.
Mix and match a perfect batch.
Mini bars are 2 for $12* with code MINISALE
Shop now
*Offer valid on any two individual boxes of Dark Chocolate Raspberry, Chocolate Caramel, Chocolate Pretzel Blast, Mint Cookie Crisp or So Good Salted Mini Bars when you enter promo code MINISALE in the Online Store through September 29, 2018. While supplies last. You must enter promo code exactly as shown at checkout to redeem. Not applicable on prior purchases and no rain checks. Excludes value packs. Not available in meeting room locations. Cannot be redeemed for cash. Offer may expire without notice due to error, fraud or other unforeseen circumstances.
NEW ARRIVALS
Weight Watchers Magazine September/October Issue
$4.95
Buttermilk Protein Pancake
$4.50
Best of WW Mini Cookbook Bundle
$11.95
Snickerdoodle Baked Protein Mini Bar
$5.95
Triple Chocolate Baked Protein Mini Bar
$5.95
Butter Popcorn - Pack of 6
$7.95

Activity & Mindset



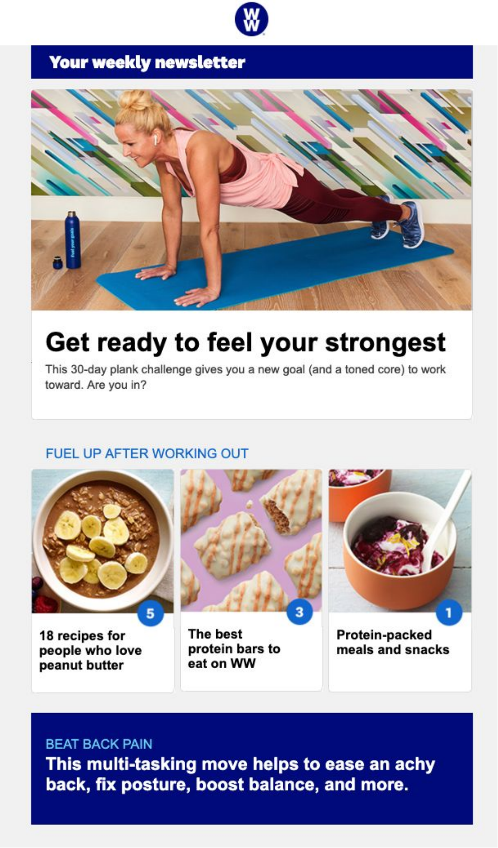


Voice & AI

Google Assistant

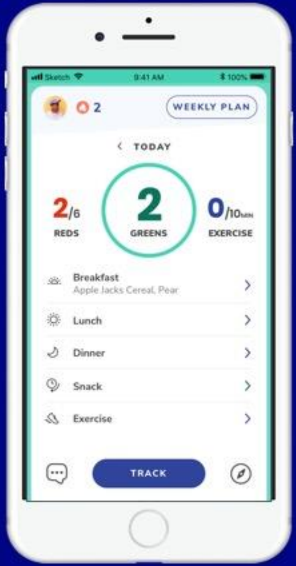
Alexa

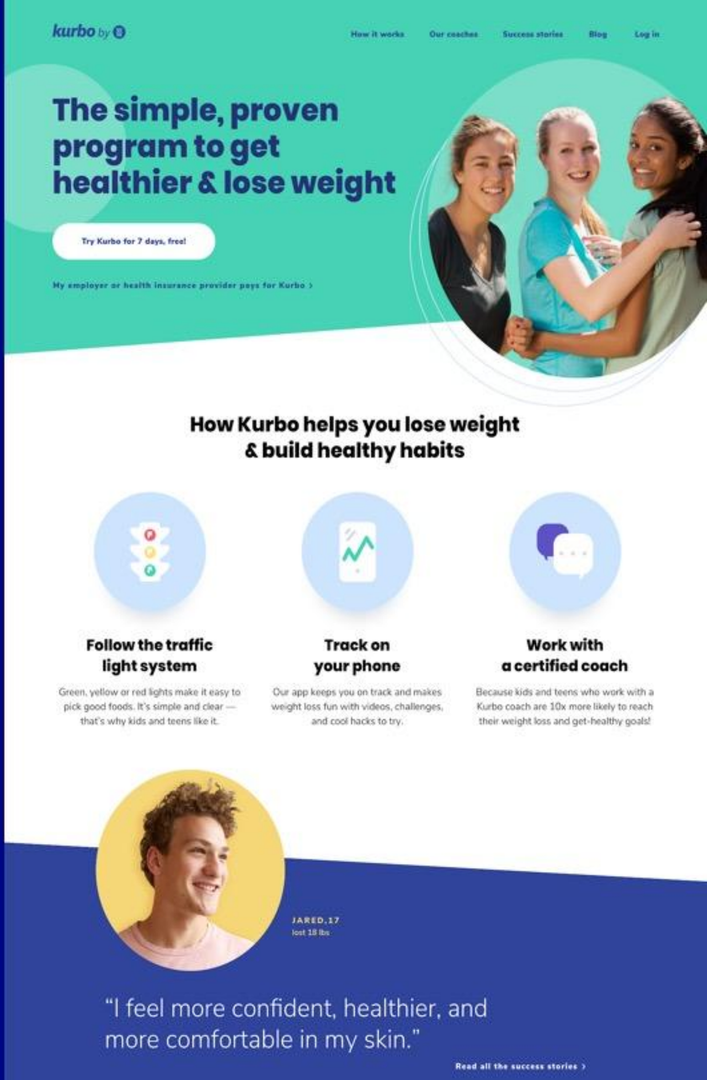
Newsletter


Your weekly newsletter
Get ready to feel your strongest
This 30-day plank challenge gives you a new goal (and a toned core) to work toward. Are you in?
FUEL UP AFTER WORKING OUT
5
18 recipes for people who love peanut butter
3
The best protein bars to eat on WW
1
Protein-packed meals and snacks
BEAT BACK PAIN
This multi-tasking move helps to ease an achy back, fix posture, boost balance, and more.

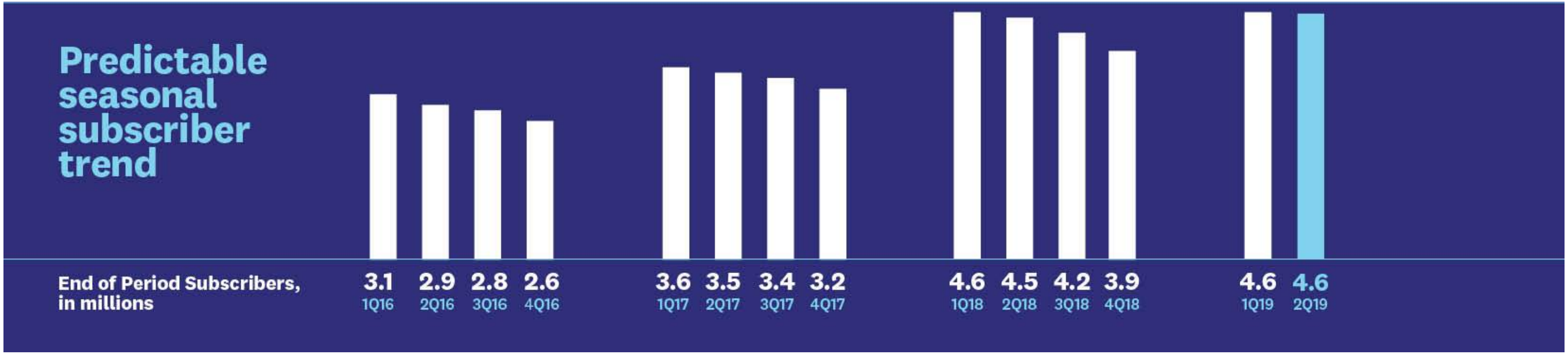
kurbo by WW

Proven digital health solution for teens and children





Dynamics and Scale



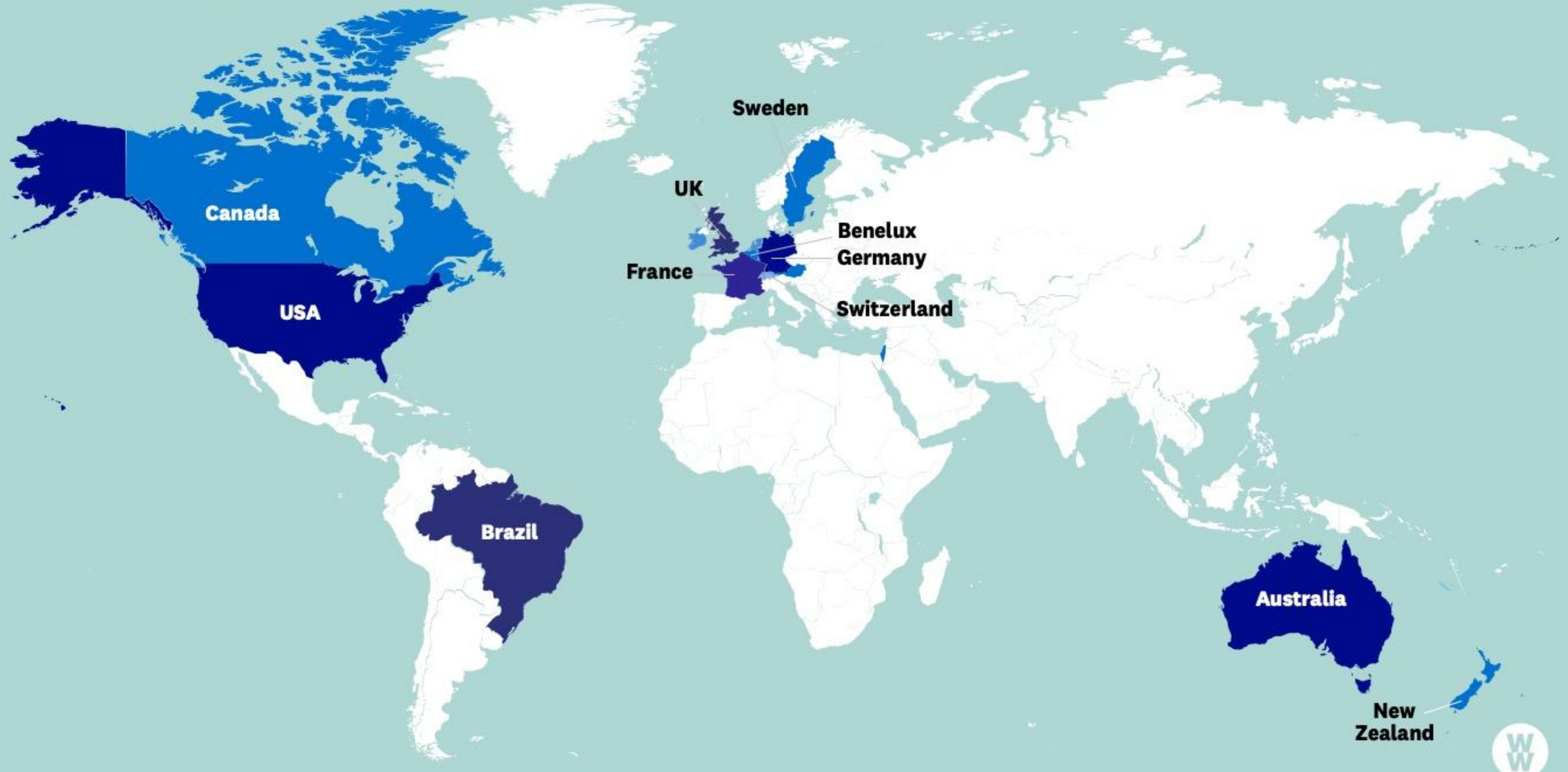
In Q2,

Social network:

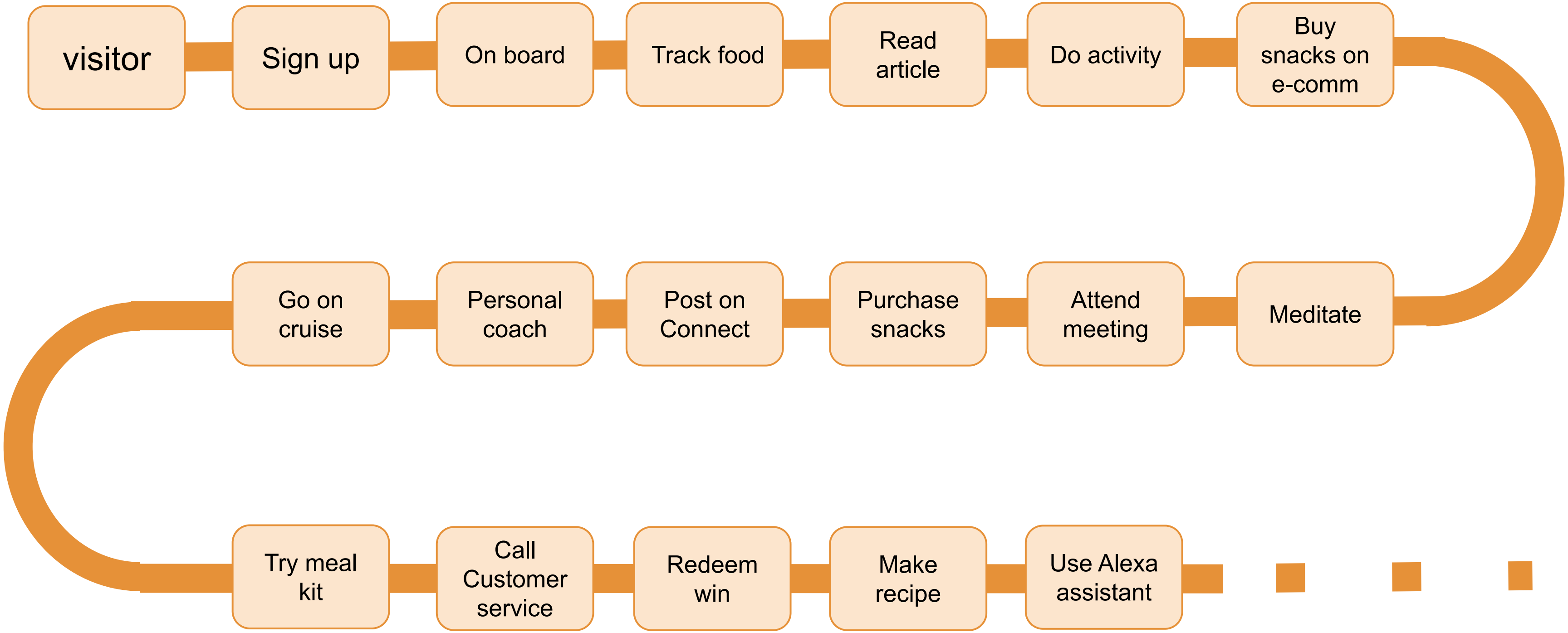
- 2 million posts
- 14 million comments
- 70 million likes

1 million members tracked a physical activity

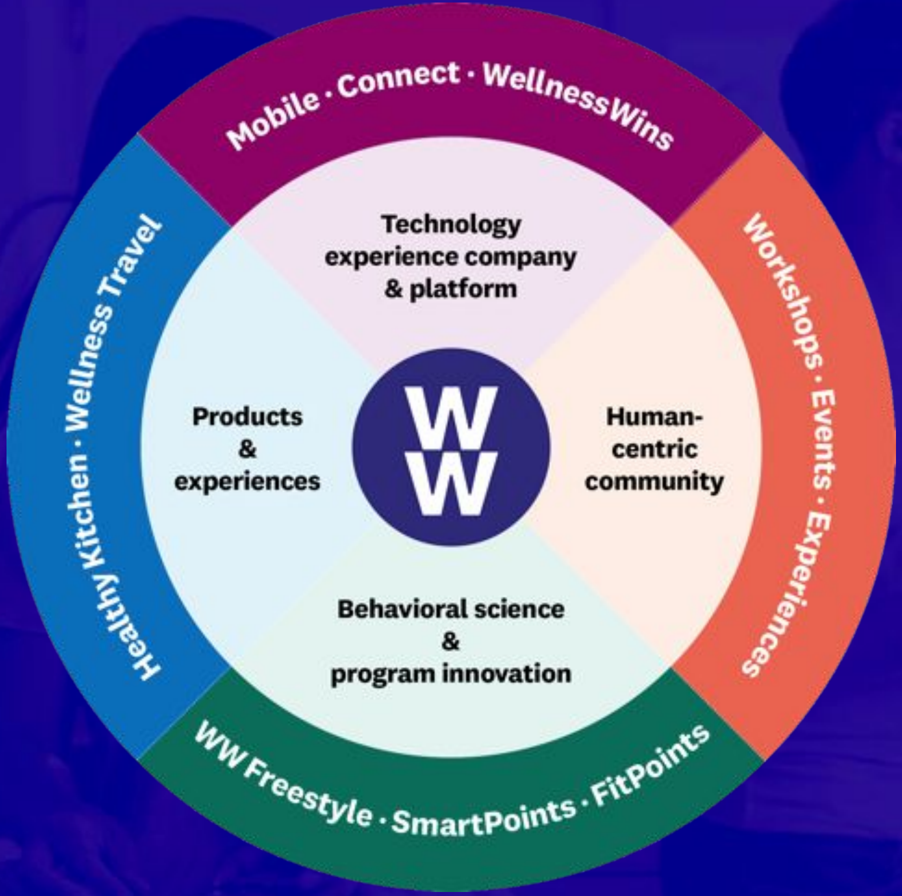
Our geographic footprint



Member Journey (Illustrative)



Almost none of this is personalized!



Big data:

- Food
- Activity
- Exercises
- Challenges
- Social network
- Workshops
- Personal Coaches
- CRM
- Fulfillment
- Meal kits
- Supermarket foods
- E-commerce
- Cruises

...for 56 years

Scale of Data

- Nackers et al (2013) showed that fast (≥0.68 kg/week) weight loss in the first month predicts higher weight loss success at 6 months than slow (<0.23 kg/week) or moderate initial weight loss
- Sample size: 298
- We checked our weigh in data to compare these results to what we observe in our member base

Nackers, Ross & Perry (2013). The Association Between Rate of Initial Weight Loss and Long-Term Success in Obesity Treatment: Does Slow and Steady Win the Race? *Int J Behav Med*. 17(3):161-167.

Scale of Data

- Members considered:

  1) started and ended their membership between Apr 2017 and Apr 2019

  2) were members for at least 6 months

  3) weighed in in their first week, fourth week and sixth month

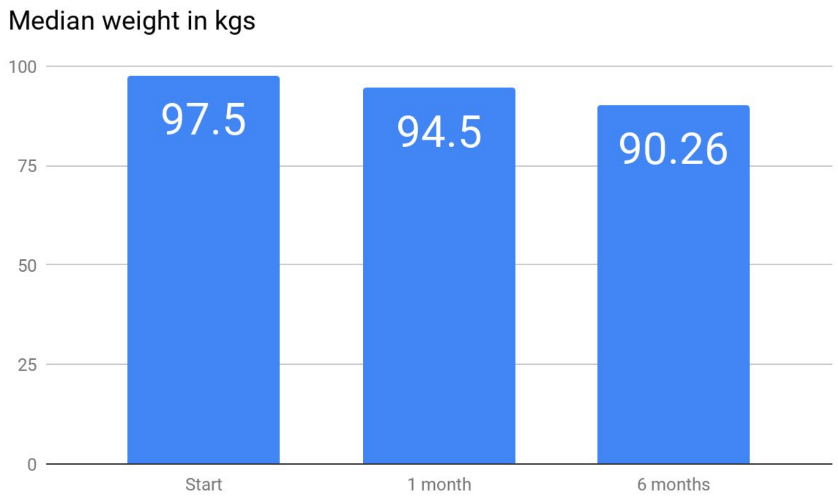
  4) were obese at the beginning of their membership (BMI > 30)

- For all analyses (mostly) unfiltered self report data was used.

Sample size: 211,000 members!

7kg median weight loss after 6 months

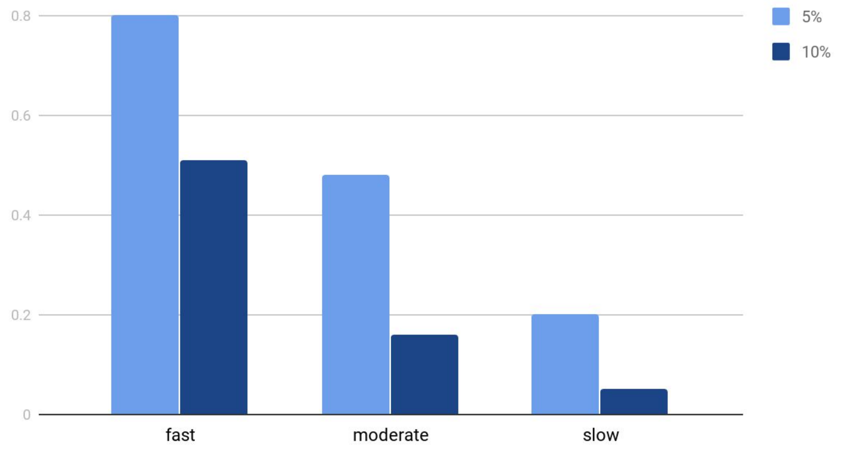
- 211k members
- 54% digital (sampling bias)
- Mostly female
- Median birth year 1971
- Median start BMI: 35
- Median BMI @ 6 months: 32.5



Significant effect of initial weight loss rate on weight loss success

- Weight loss defined as 5% or 10% of initial start weight lost
- Initial weight loss speed:
  - Fast (≥0.68 kg/week) : 116,107
  - Moderate (<0.68 & >0.23 kg/week): 61,204
  - Slow (<0.23 kg/week): 34,107

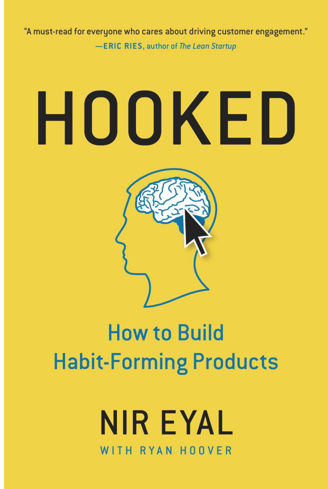
Proportion of members who lost 5% or 10% initial body weight at 6 mo



*all differences statistically significant

Nudges & behavioral change

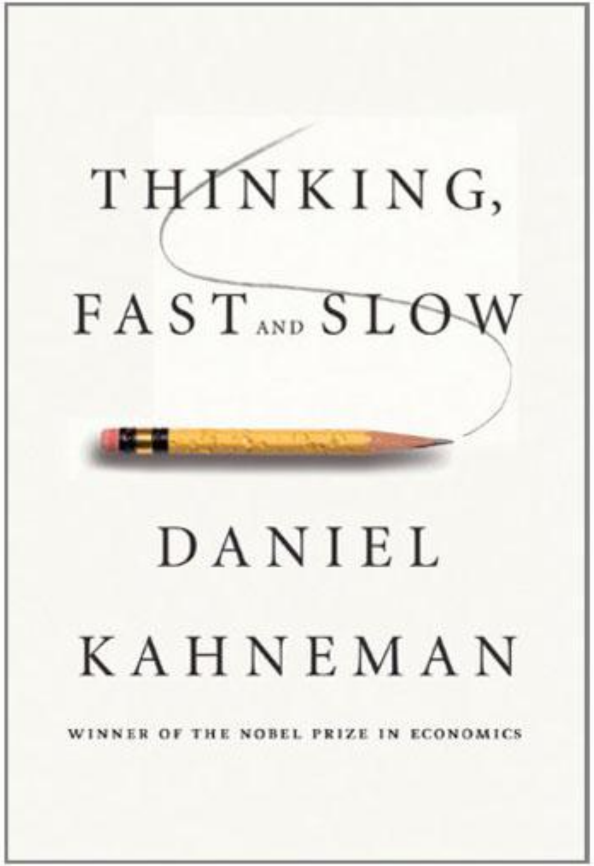
Healthy Habits

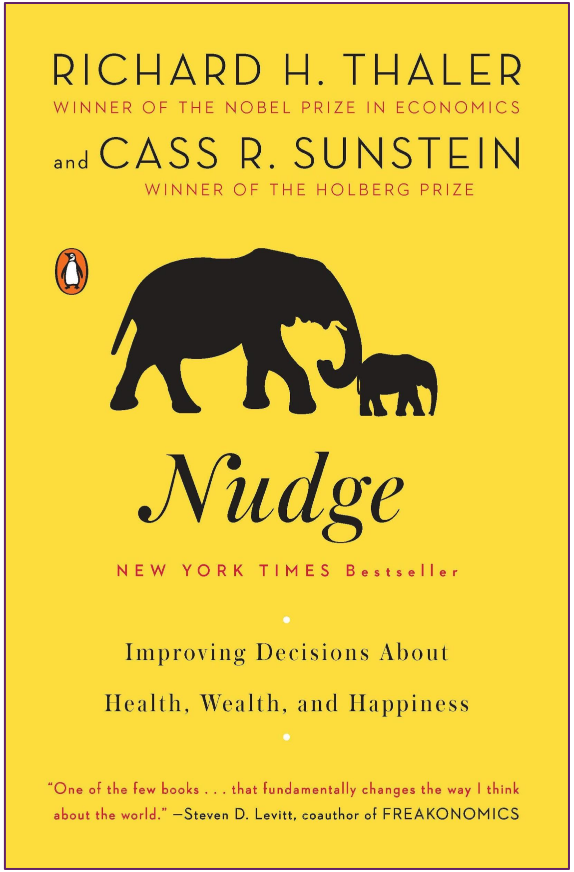


Habit:

"Automatic behaviors triggered by situational cues"

"Habit-forming companies link their service to the users' daily routines"


THINKING,
FAST AND SLOW
DANIEL
KAHNEMAN
WINNER OF THE NOBEL PRIZE IN ECONOMICS


RICHARD H. THALER
WINNER OF THE NOBEL PRIZE IN ECONOMICS
and CASS R. SUNSTEIN
WINNER OF THE HOLBERG PRIZE
Nudge
NEW YORK TIMES Bestseller
Improving Decisions About
Health, Wealth, and Happiness
"One of the few books . . . that fundamentally changes the way I think
about the world." —Steven D. Levitt, coauthor of FREAKONOMICS

"Any addition to or modification of the environment that influenced consumers in a predictable way, without changing economic incentive"

Altering environment by changing presentation of options — called **choice architecture**

"Any aspect of the choice architecture that alters people's behavior in a predictable way (1) without forbidding any options or (2) significantly changing their economic incentives. Putting fruit at eye level counts as a nudge; banning junk food does not."

Nudging & Choice Architecture

i**N**centives
Understand mappings
Defaults
Give feedback
Expect error
Structure complex choices

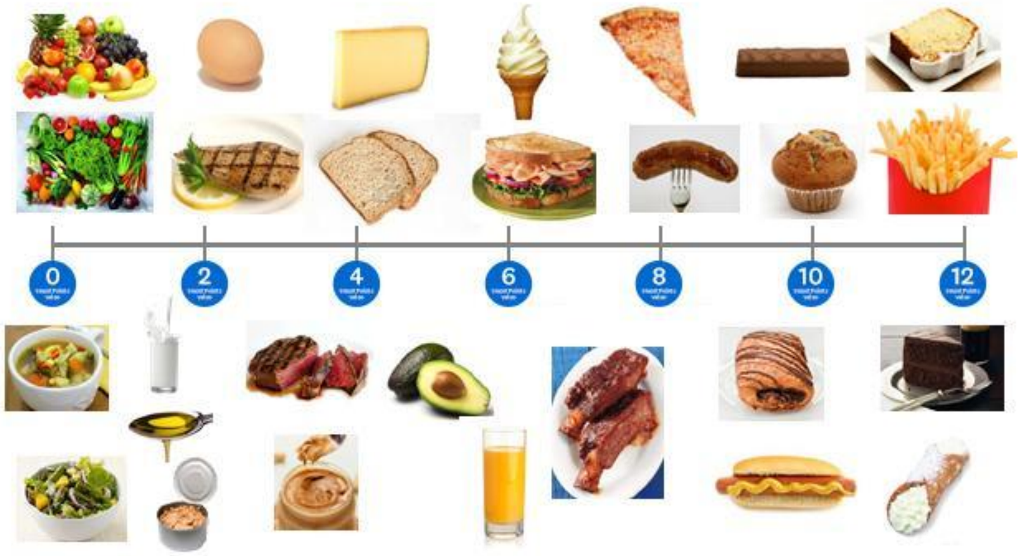
This is Thaler & Sunstein's framework. Instead, I will use Blumenthal-Barby & Burroughs

Nudging & Choice Architecture

Category	Explanation	Examples
Priming	Subconscious; physical, verbal, sensational	Place unhealthy options out of sight or farther away in cafe
Salience	Informational; attention grabbing; emotional	Calorie label; graphic image on cigarette cartons; **recommenders**
Default	Pre-set default choice; good option for do-nothing behavior	Automatic opt-in, have to explicitly change or opt-out: benefits, organ donation
Incentive / Gamify	Reward or punish for behavior; recognition	Badge; coupon; $$?
Commitment / Ego	Get someone to make a commitment; leverage ego, pride	Sign up for 5k; invest (pay for membership); share with friends
Norms / Messenger	Use other to establish a norm and for consumer to compare themselves	80% of (other) people are organ donors; most people wear seatbelts

Blumenthal-Barby & Burroughs. (2012). Seeking better health care outcomes. The ethics of using the nudge. Am. J. Bioethics 12(2):1-10.

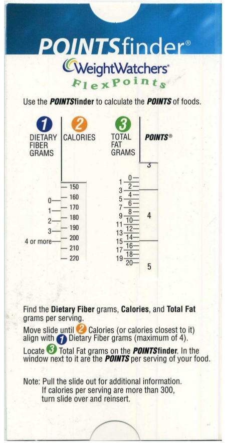
Everything is on the menu

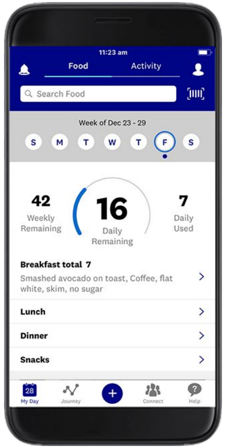


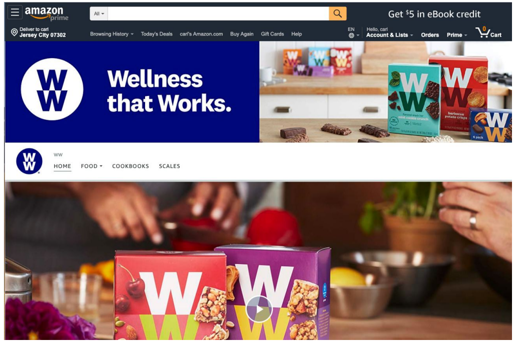
Members are empowered with lots of choices.
We are not dictating diet, exercise regime etc. Hence, these are nudges

Priming

Visibility, accessibility, available







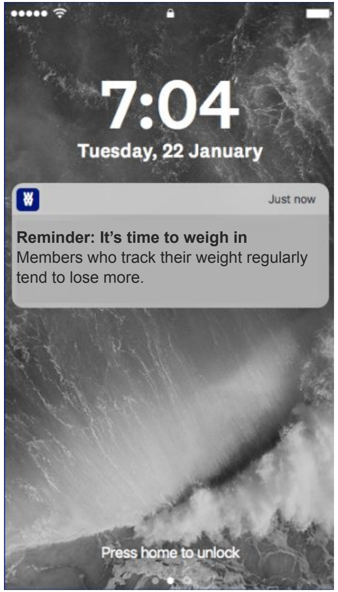
Further prime by reducing friction for key actions such as tracking

Priming

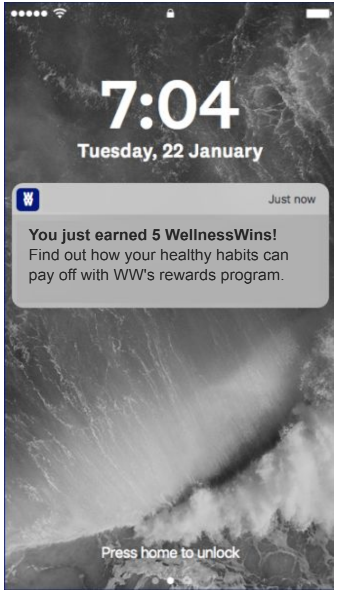
Visibility, accessibility, available

We plan to leverage notification nudges, similar to these, to help keep people on track

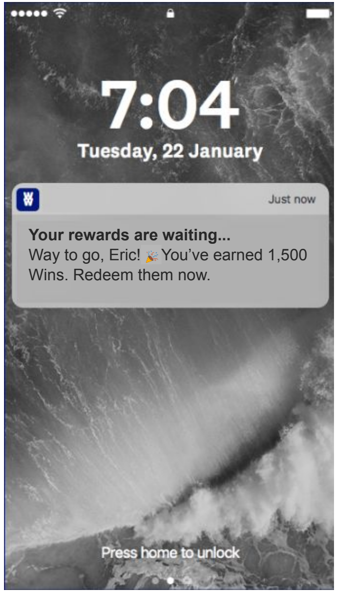
Weight Tracking Reminder



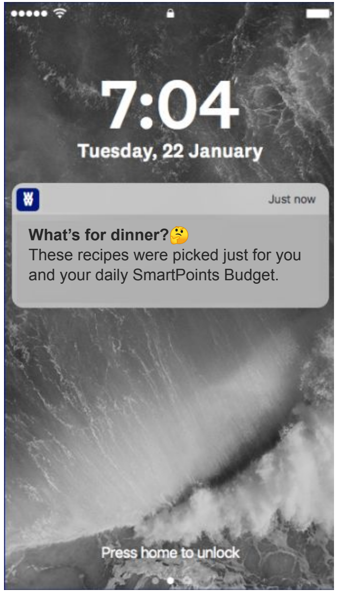
Wellness Wins Awareness



Wellness Wins Redemption

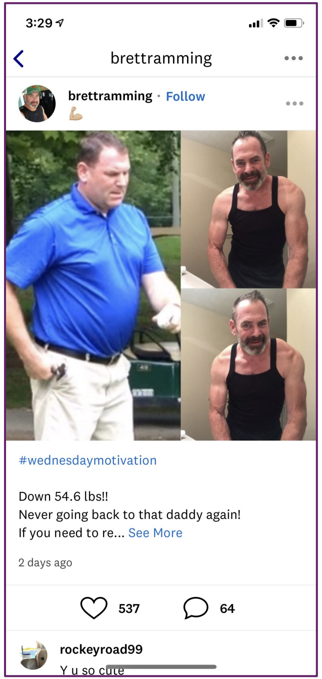


Dinner Recommendation



Priming

Visibility, accessibility, available

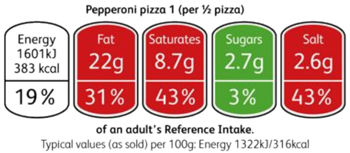


Connect social network: very highly supportive, lots and information from staff and fellow members. Priming and saliency.

Saliency

meaningful, relevant info

- Points on very large number of foods
- Clear “mappings” make it easier to make good choices:
    - instead of calorie counting, 300 cal (or is it kJ, kcal) → 5 points



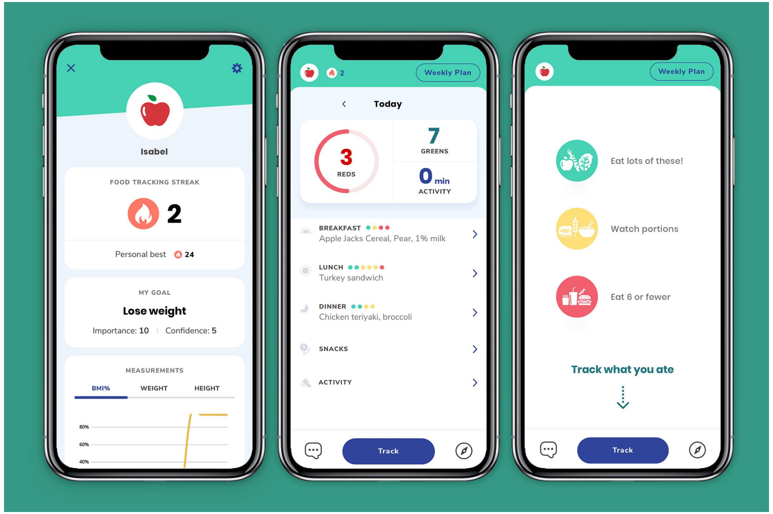
Saliency

meaningful, relevant info

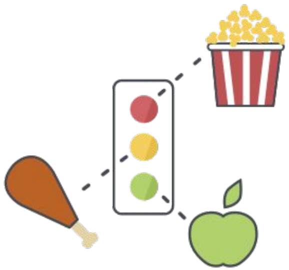
- Clear budget
- Clear progress

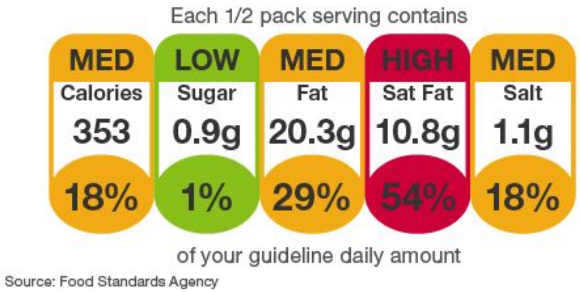
Saliency

meaningful, relevant info



Kurbo traffic light system



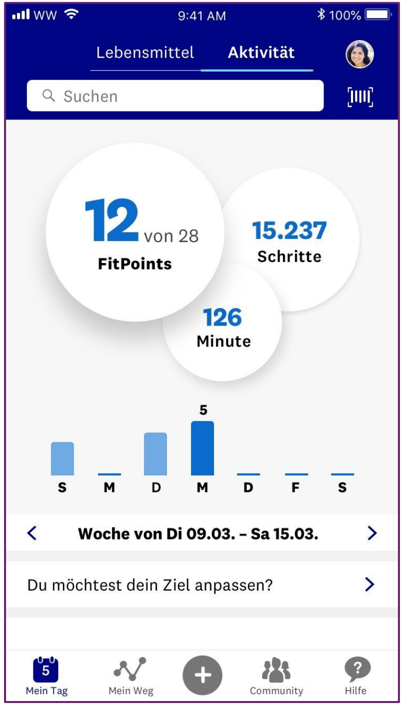


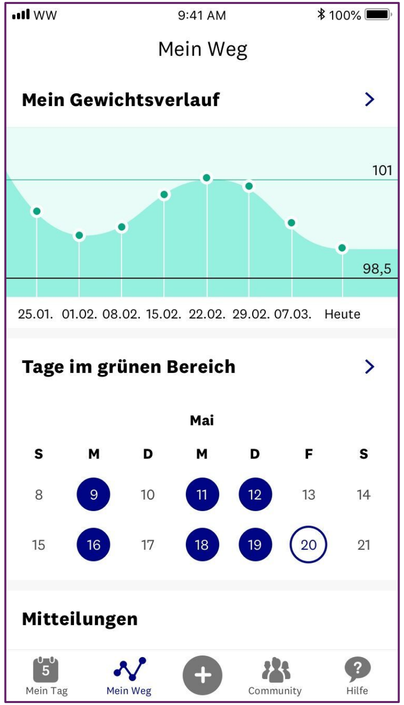
UK nutrition labels

Saliency

meaningful, relevant info

Activities are also pointed: FitPoints





Saliency

meaningful, relevant info

Tracking and wearables

Quantified Self

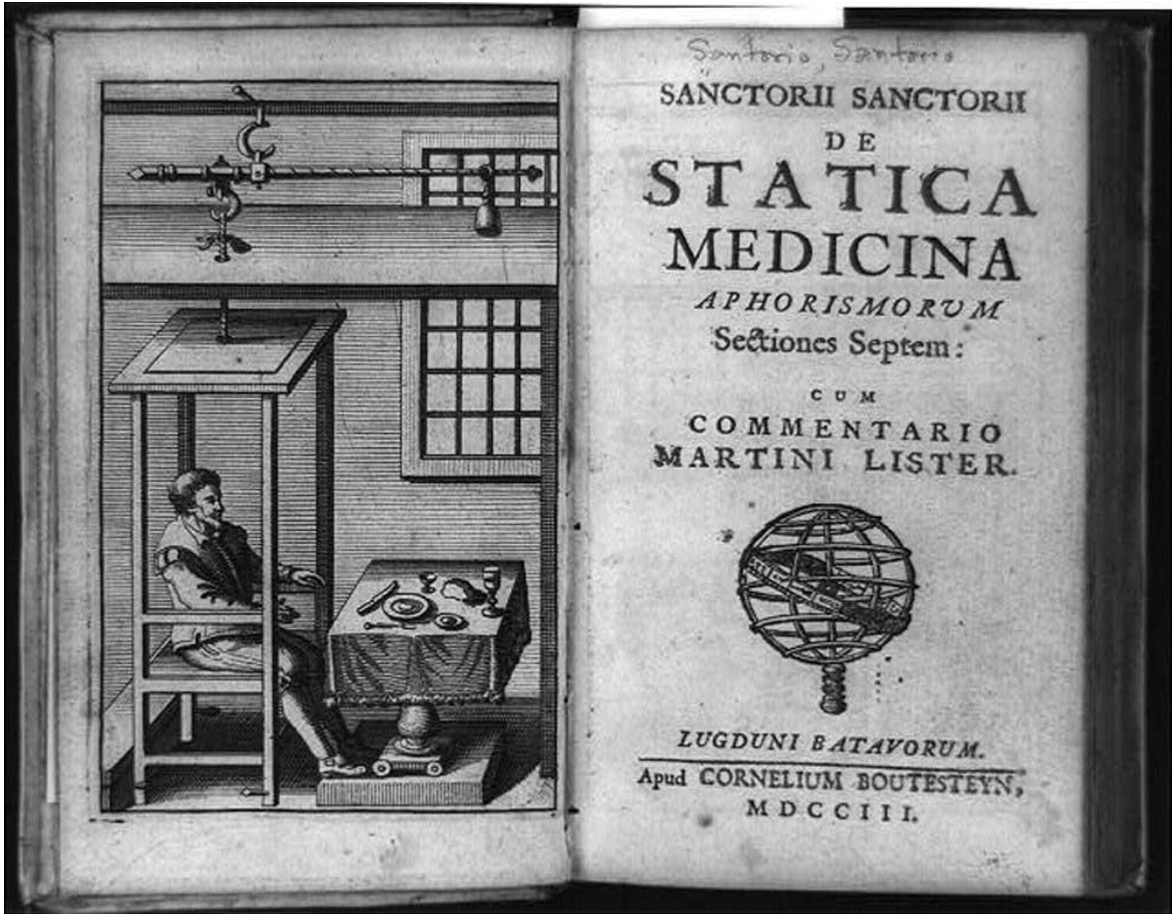
self knowledge through numbers

Kevin Kelly

Gary Wolf

Girolamo Cardano
1501-1576

Santorio, Santorio
1561-1636

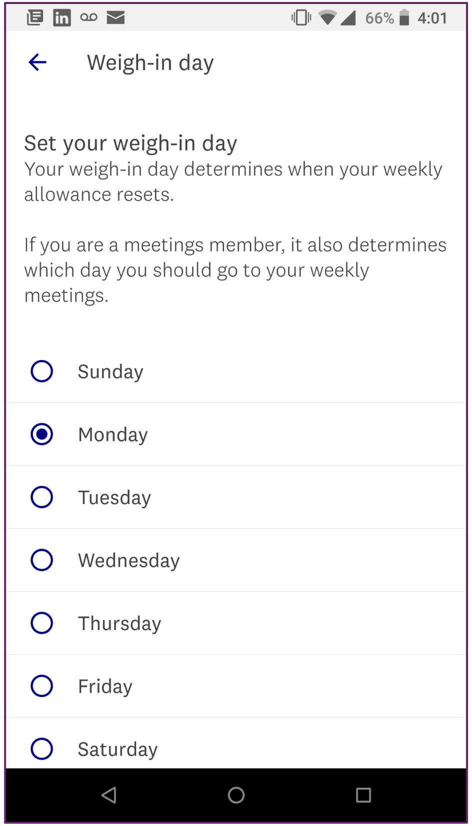


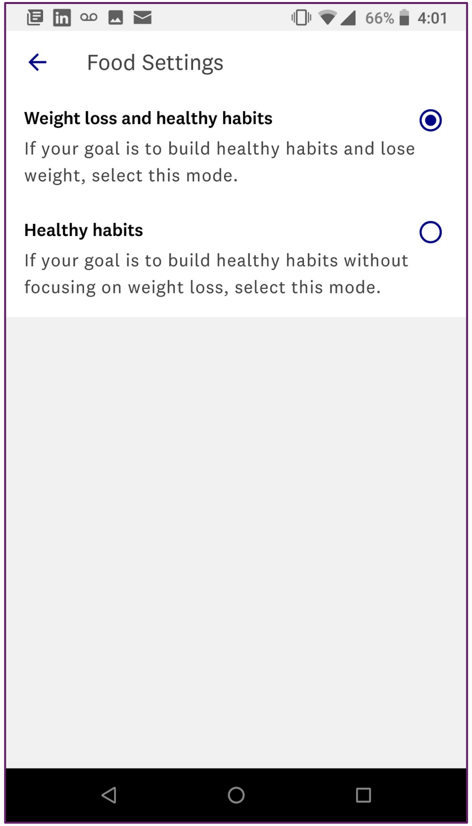

tracker
WAKE BY 6AM
MIRACLE MORNING
WORKOUT
GOAL CHECK
HYDRATE
BUDGET REVIEW
BLOG POST
YT VIDEO
INSTAGRAM
"ME DAY"
BED BY 11PM
MOOD
AM
PM
HAPPY
EXCITED
STRESSED
RELAXED
PRODUCTIVE
GRATEFUL
UGH!

Wellness that Works.™

Defaults

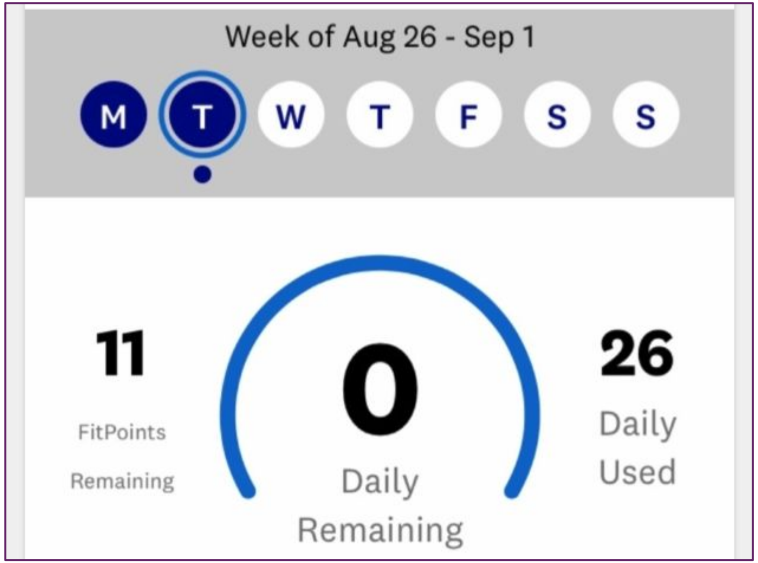
- Defaults should be good, fair, equitable...
- Easily changed



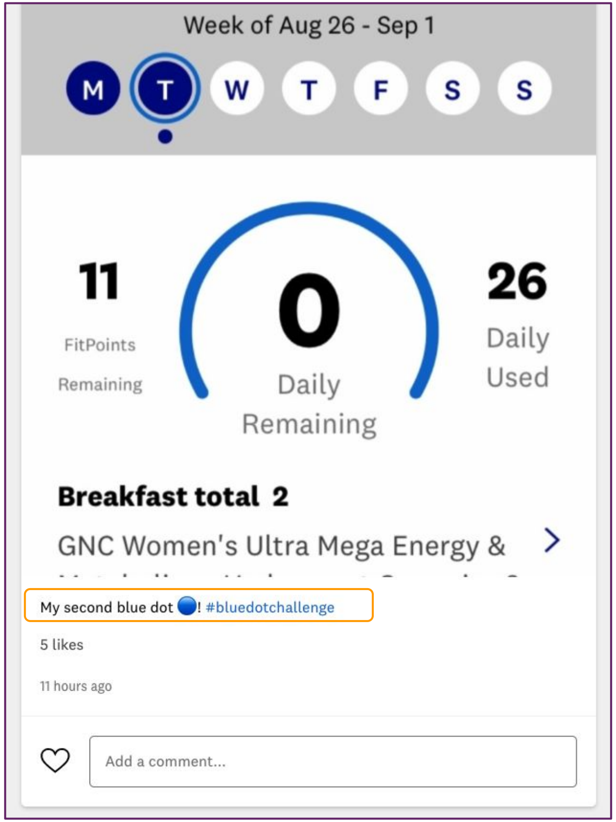


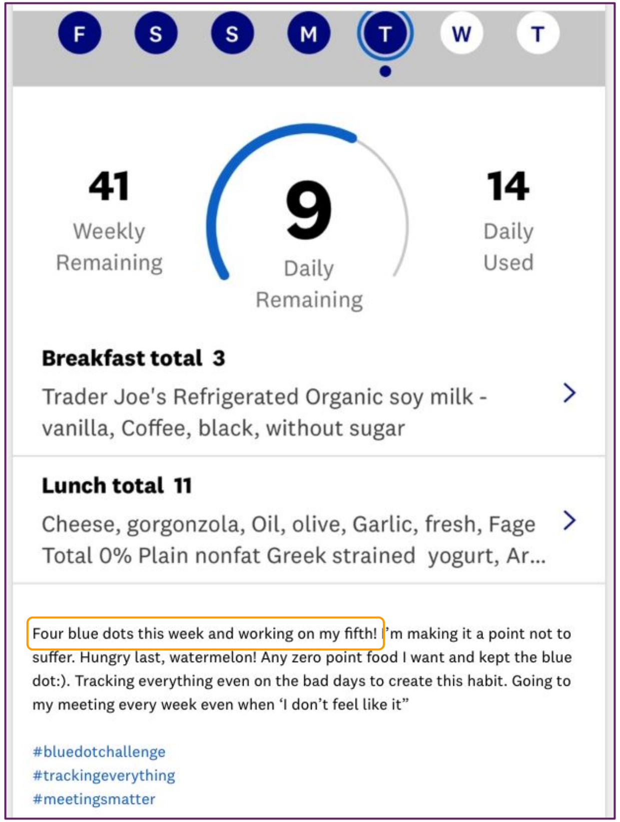
Incentives / Gamify

Incentives / Gamify



Blue dot: daily, weekly, monthly





Wellness that Works.™

Incentives / Gamify

WellnessWins™

A first-of-its-kind program that rewards members for building healthy habits.

You earn "Wins" for:

- Tracking meals (breakfast, lunch, dinner)
- Tracking activity
- Tracking weight
- Attending workshops

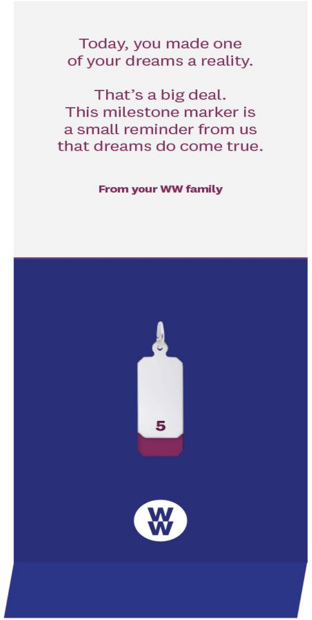
Incentives / Gamify

"I've got my keyring with all my little bangles hanging off of it. I love that thing. It might seem stupid but it was just fun to get those rewards along the way, a physical manifestation of your success"

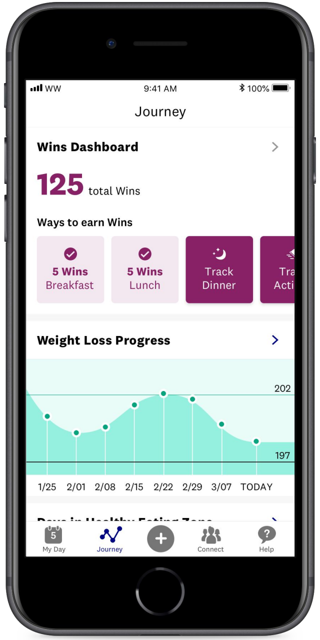
WellnessWins celebrates **outcomes**

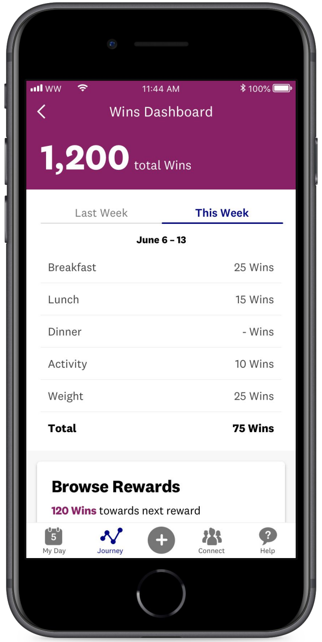
Milestones are rewarded for:

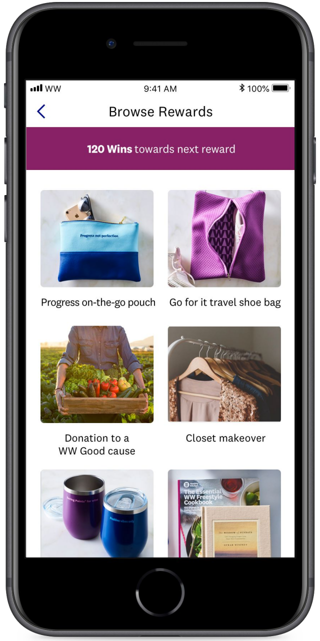
- Weight loss: 5, 10, 25, 50, 75, 100, 125, 150, 175 & 200 lbs
- Goal weight



Incentives / Gamify







Incentives / Gamify

Tier 1 (1,500 Wins)

6 months premium subscription to Lifefyle

Headspace 90-day starter membership

Five free day passes at The Gym

Donate your Wins to someone in need

"Positive vibes only" hat

"Go for it" travel shoe bag - blue

"Unstoppable" hat

"Progress" zip pouch

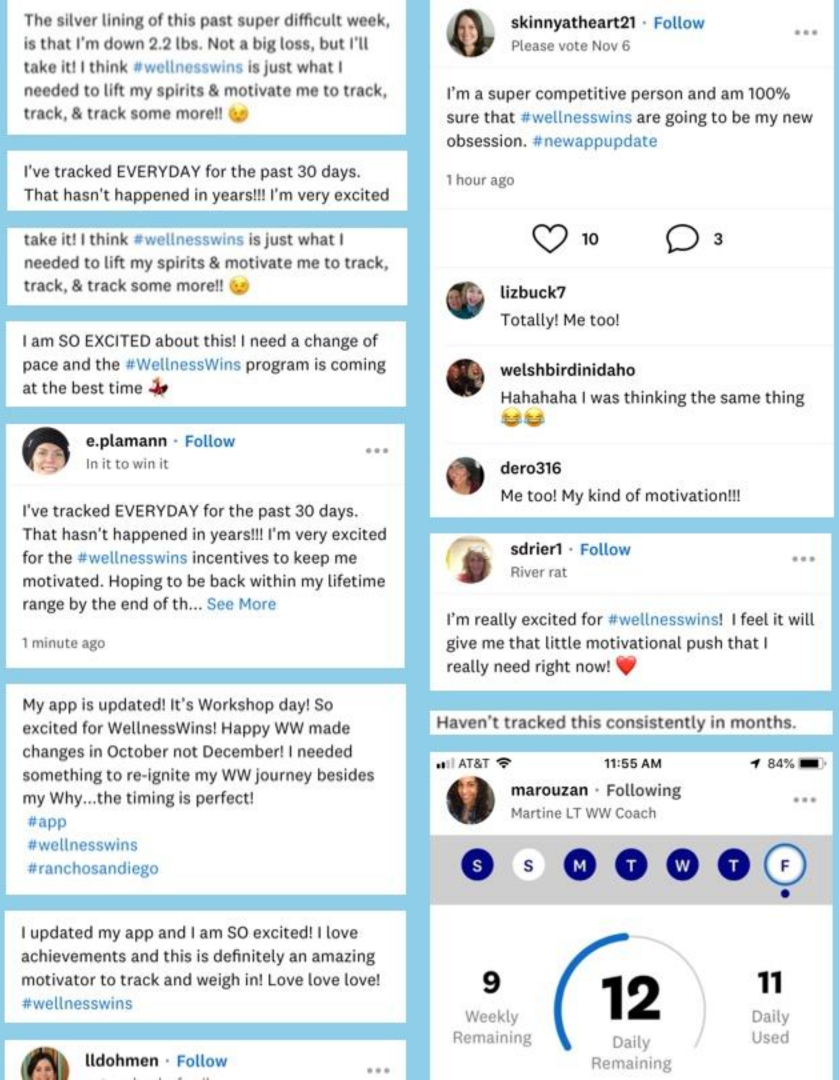
"My Story" journal

"WellnessWins" tumbler

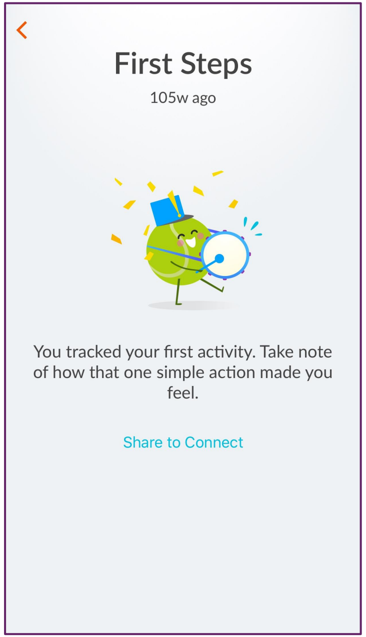
Sunglasses case

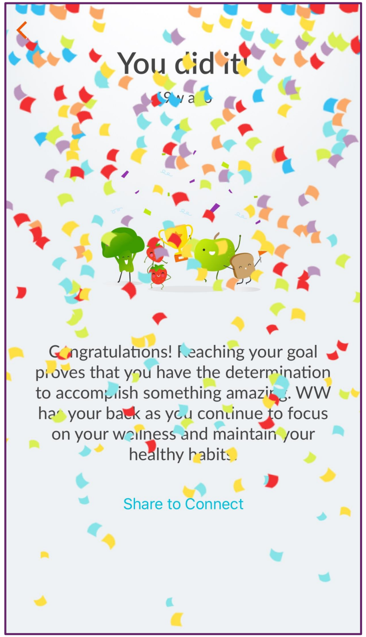
WellnessWins is motivating

We are holding members accountable while making it fun to earn



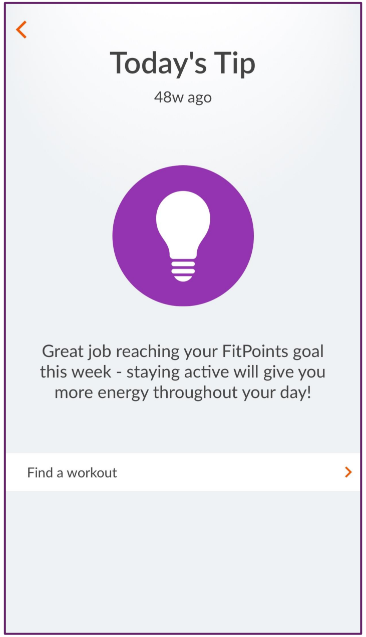
Incentives / Gamify

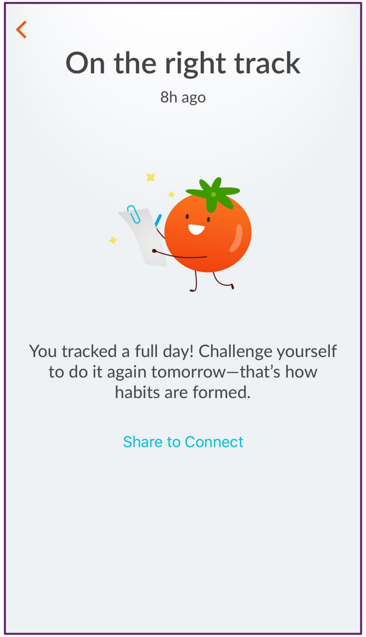


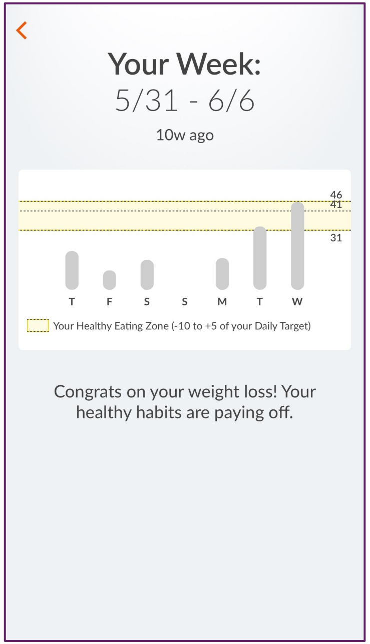


Recognition helps motivate

Incentives / Gamify



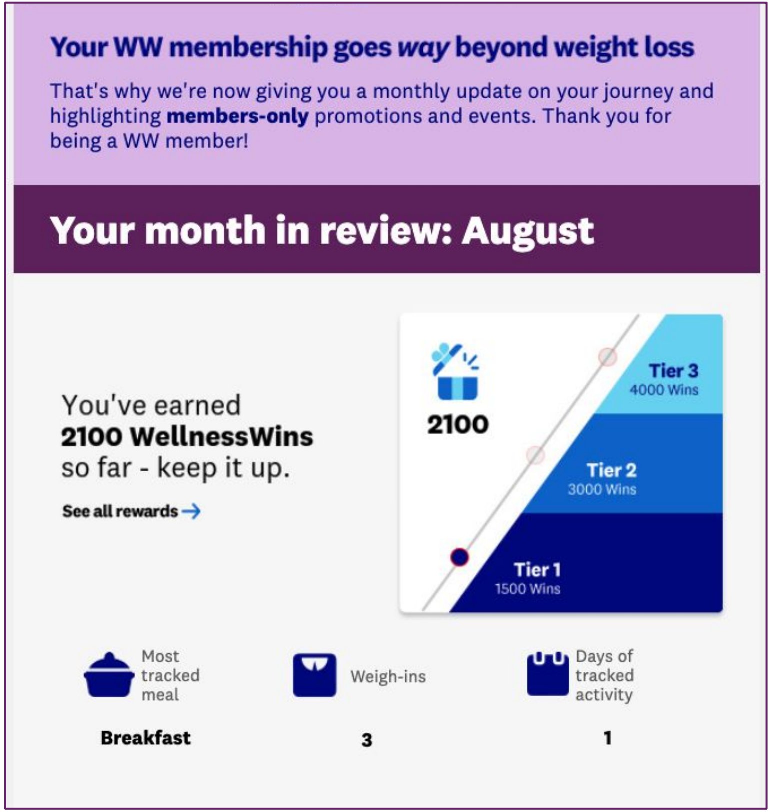




Recognition helps motivate

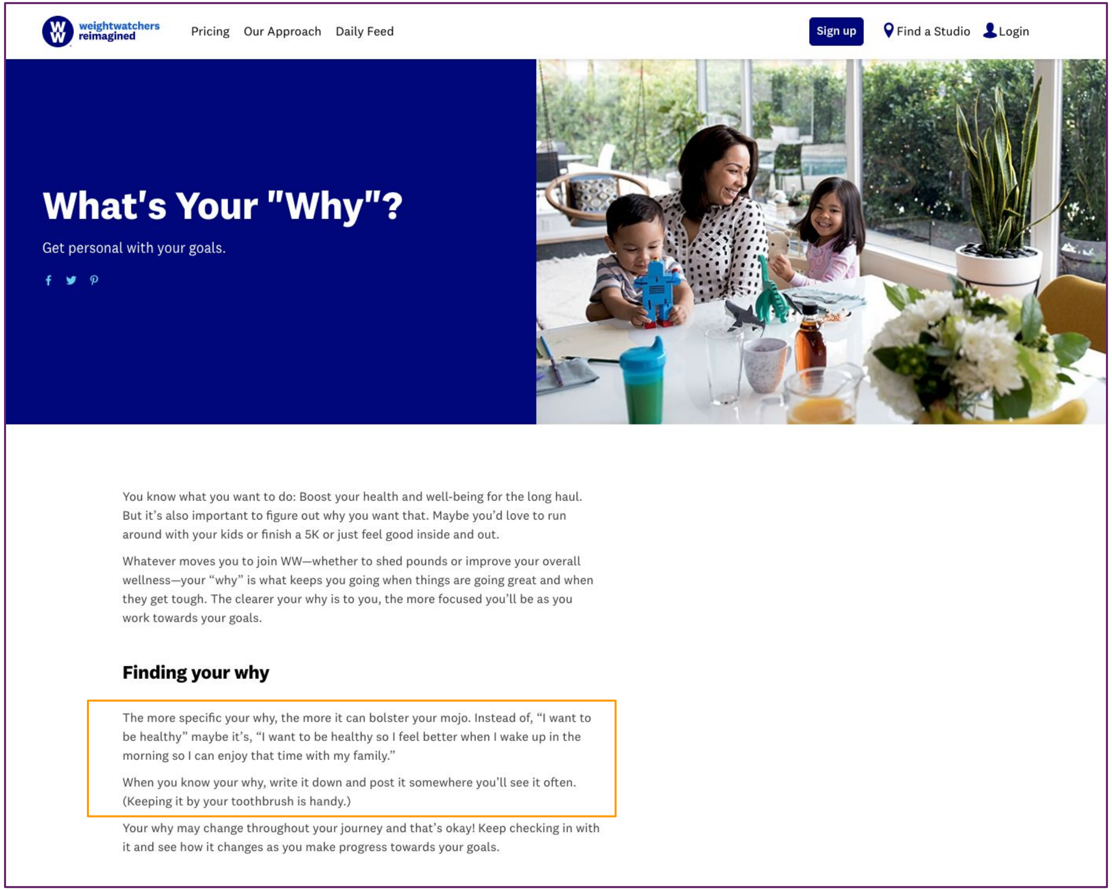
Incentives / Gamify

Incentives / Gamify



Motivation

- Member is doing the work.
- Important for them to remind *themselves* why they are doing this



weightwatchers reimagined

Pricing Our Approach Daily Feed

Sign up Find a Studio Login

What's Your "Why"?

Get personal with your goals.

You know what you want to do: Boost your health and well-being for the long haul. But it's also important to figure out why you want that. Maybe you'd love to run around with your kids or finish a 5K or just feel good inside and out.

Whatever moves you to join WW—whether to shed pounds or improve your overall wellness—your "why" is what keeps you going when things are going great and when they get tough. The clearer your why is to you, the more focused you'll be as you work towards your goals.

Finding your why

The more specific your why, the more it can bolster your mojo. Instead of, "I want to be healthy" maybe it's, "I want to be healthy so I feel better when I wake up in the morning so I can enjoy that time with my family."

When you know your why, write it down and post it somewhere you'll see it often. (Keeping it by your toothbrush is handy.)

Your why may change throughout your journey and that's okay! Keep checking in with it and see how it changes as you make progress towards your goals.

Willpower

"Weight Watchers, for example, coaches dieters to use an array of self-deployed situational and cognitive strategies and, in addition, sponsors in-person meetings, communicates social norms, and provides a phone app to track eating and exercise"

Fig. 2. Illustrative examples of approaches aimed at reducing self-control failures. Approaches are classified as situational versus cognitive and self-deployed versus other-deployed. MCII = mental contrasting/implementation intentions.

"

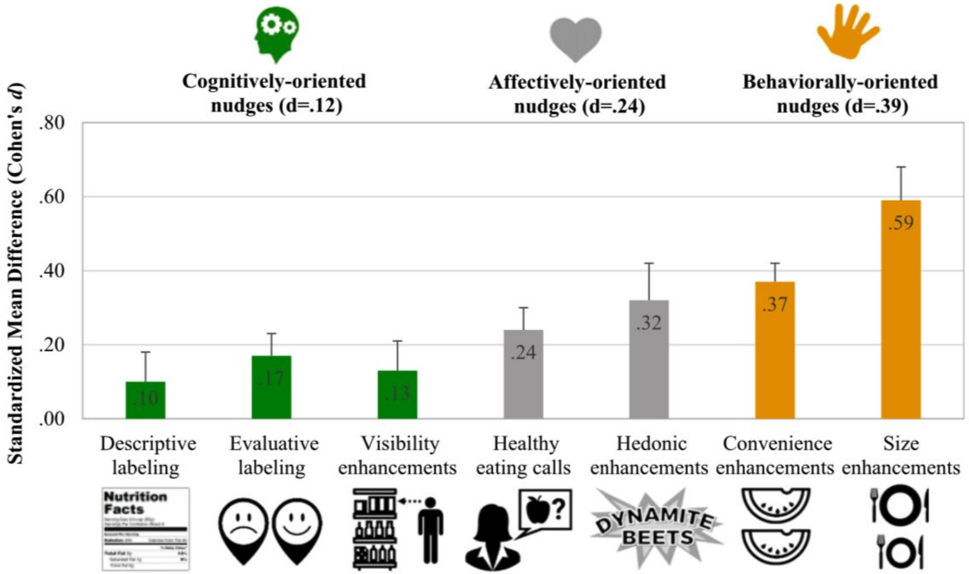
Table 1. How Strategies for Reducing Self-Control Failure Might Be Applied to Increasing Healthy Eating

Strategy	Example
Self-deployed situational strategies	
Commitment devices	Decision maker commits to eat a serving of fruit and vegetables at dinner every night, asks spouse to serve as a referee, and puts money on the line that will be forfeited to spouse in case he or she fails to meet this commitment.
Temptation bundling	Decision maker listens to a favorite music album only when cooking dinner from scratch (rather than eating fast food).
Situation modification	Decision maker stocks up on bags of Halloween candy for trick-or-treaters—but only candy that she does not like.
Behavior therapy	Decision maker works with a therapist, learning to identify triggers that result in junk food binges (e.g., deadlines at work) and also alternatives (e.g., taking a walk) that can meet the same needs (e.g., stress relief).
Self-deployed cognitive strategies	
Goal setting	Decision maker decides: "I will eat a fruit or vegetable with every meal!"
Planning	Decision maker plans: "If it is 8 a.m., then I will look in the refrigerator for some fruit to have with my breakfast."
Mental contrasting with implementation intentions	Decision maker thinks: "The best outcome of eating healthy is that I will have more energy. The obstacle that stands in the way is that I don't have time to go shopping. My plan is: 'If it is Saturday morning, then I will take a nice walk to the grocery store to buy fresh fruit that I'll then eat.'"
Psychological distancing	Decision maker reframes situation using third-person perspective: "Angela is hungry and has a choice between a bag of potato chips and an apple. Which should she choose?"
Mindfulness	Decision maker introspects: "I notice that I'm craving potato chips. I accept that I have this urge. I may or may not act on it."
Cognitive therapy	Decision maker works with a therapist, learning to ask, "What thoughts lead me to snack on potato chips in the afternoon? Do I think, 'I can't resist junk food. I have no self-control at all!' And is that a reasonable thought? Or am I exaggerating?"

Duckworth, Milkman, & Laibson. (2018). Beyond willpower: strategies for reducing failures of self-control. Psychological Science in the Public Interest. 19(3) 102–129

Effect Sizes

Figure 3. (Color online) Effect Sizes by Nudge Type



Note. Error bar represents standard error.

Meta-analysis of 90 articles + 96 field experiments (299 effect sizes), average effect of healthy eating nudges of Cohen's d=0.23.

= 124 kcal change in a daily intake
Or -7.2%

8 tablespoons sugar / day

Cadario & Chandon. (2019). Which healthy nudges work best? A meta-analysis of field experiments. *Marketing Science*. DOI: 10.1287/mksc.2018.1128

Summary

Highly primed experience

easy	**SmartPoints, FitPoints**
available	**In your pocket, AMZN, neighborhood**
supportive	**community**

Highly saliency

progress
food, fitness
tips

Incentives / Recognition at multiple temporal scales

Initial	First track, first barcode scan
Daily	Blue dot
Weekly	Weight check in
Continuous	streaks
Monthly	review
Event	Wins, milestone
Annual	Annual review???

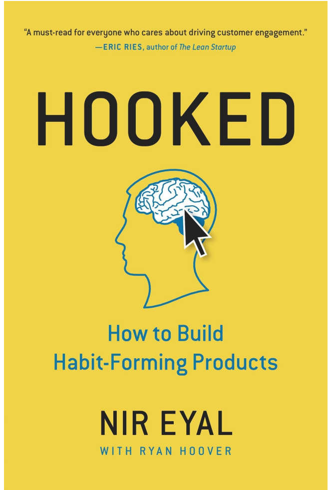
Multiple types of recognition

Non tangible	Badges, kudos
Peers	Connect
Goods in kind	wins

		Initial	Daily	Weekly	Monthly	Milestone	Year
Food	Saliency		MyDay Blue dot	MyDay Blue dot	MyDay Blue dot Newsletter Month in review	badges/tips?	
	Reward & Recognition (R&R)	badges	#bluedotchallenge Badges Streaks	#bluedotchallenge Badges recommenders	#bluedotchallenge Badges Month in review	Wellness Wins badges	
	Defaults			SP budget?			
Fitness	Saliency		MyDay	Myday	Newsletter Month in review		
	R & R	badges	badges	badges	Month in review		
Weight	Saliency			Coaches / Meetings	Month in review		
	R & R	badges	badges	Badges Coaches / Meetings	Month in review	Wellness Wins badges	
	Defaults			Check in day			

The Hook


Trigger
Action
External
Internal
Investment
Variable
Reward


"A must-read for everyone who cares about driving customer engagement."
—ERIC RIES, author of The Lean Startup
HOOKED
How to Build
Habit-Forming Products
NIR EYAL
WITH RYAN HOOVER

Data products at WW

Data products at WW

Growth

Churn model
Return model
LTV models
Single Member View

Program

Recipe recommender
Similar recipes
Auto-tags
Clustering member foods
Composite foods

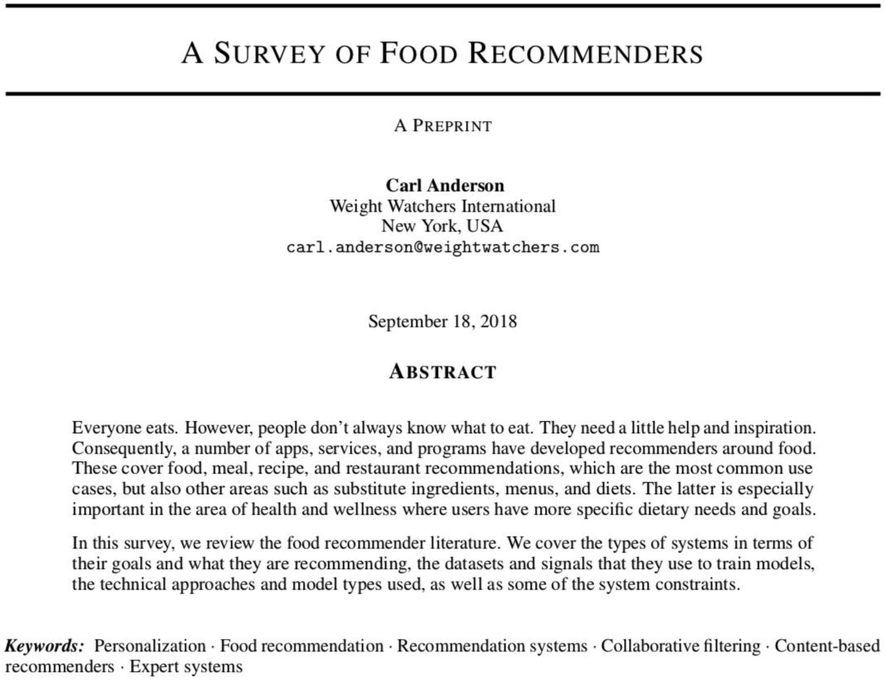
Social Network: Connect

Personalized feed
Groups search
Who to follow

Infra-structure

APIs
Primrose

https://arxiv.org/pdf/1809.02862.pdf


A SURVEY OF FOOD RECOMMENDERS

A PREPRINT

Carl Anderson
Weight Watchers International
New York, USA
carl.anderson@weightwatchers.com

September 18, 2018

ABSTRACT

Everyone eats. However, people don't always know what to eat. They need a little help and inspiration. Consequently, a number of apps, services, and programs have developed recommenders around food. These cover food, meal, recipe, and restaurant recommendations, which are the most common use cases, but also other areas such as substitute ingredients, menus, and diets. The latter is especially important in the area of health and wellness where users have more specific dietary needs and goals.

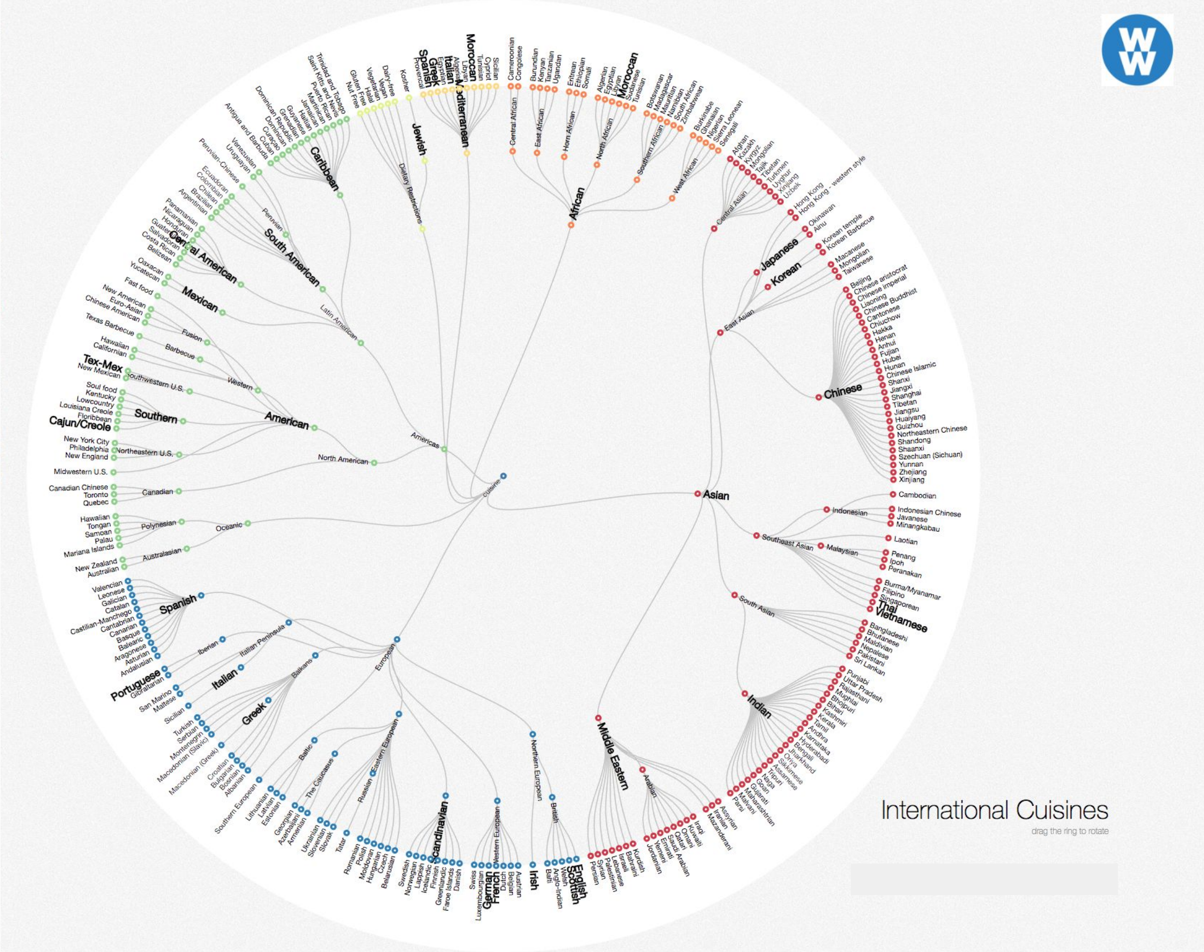
In this survey, we review the food recommender literature. We cover the types of systems in terms of their goals and what they are recommending, the datasets and signals that they use to train models, the technical approaches and model types used, as well as some of the system constraints.

Keywords: Personalization · Food recommendation · Recommendation systems · Collaborative filtering · Content-based recommenders · Expert systems

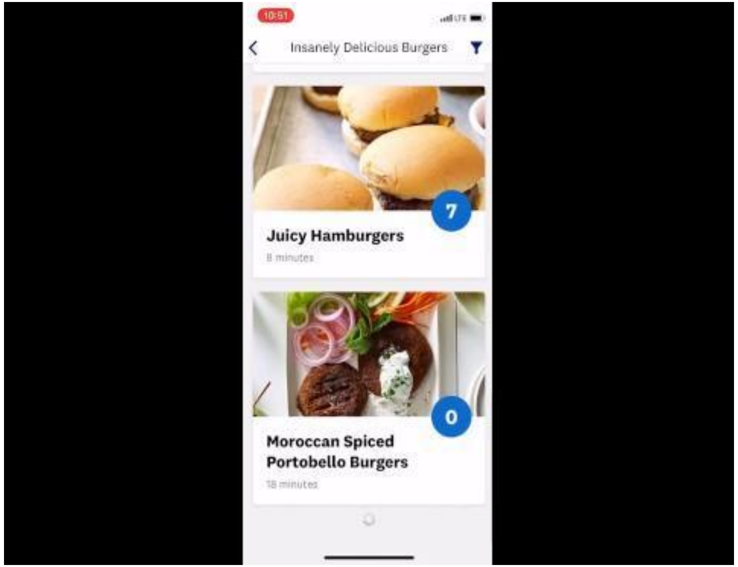
See also Trattner, C., and Elsweiler, D. (2017.) Food Recommender Systems: Important Contributions, Challenges and Future Research Directions. https://arxiv.org/abs/1711.02760

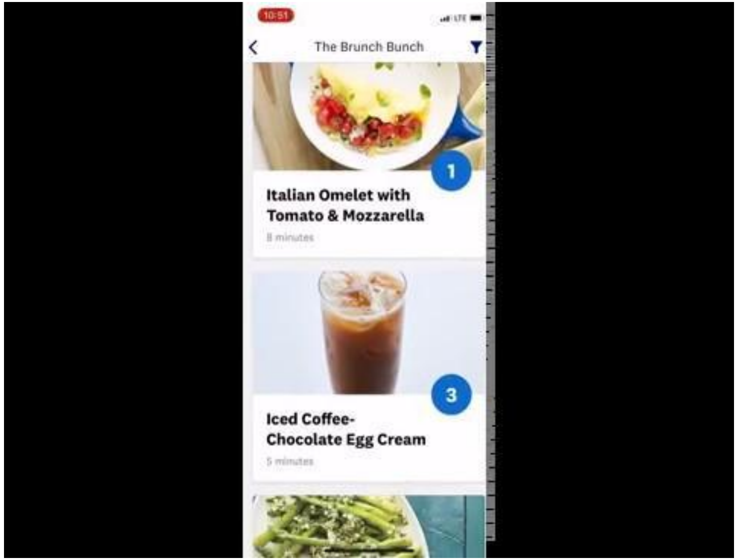
Table 1: High-level summary table that highlights the breadth of food recommender space, covering what is being recommended to whom, how, and why. * represents more speculative examples.

Dimension	Examples
Who are the users?	**Hungry people**: you might like to order this meal **Cooking enthusiasts**: you might like to make this recipe **Health-conscious**: you'll love this healthy, nutritious lunch **Dieters**: this is a low-calorie but filling and healthy meal **Patients**: doctors suggest that you follow this diet
What is being recommended?	**Ingredient**: you can substitute butter with sour cream for reduced fat and calories **Food**: we think you'll like these summer rolls **Meal**: we think you'll like this chicken breast plate with rice and broccoli **Recipe**: try this pecan pie recipe **Recipe collection**: here is a set of salad recipes you'll love **Restaurant**: you have to try Danny's Pizza **Cuisine***: as you like Thai, you might like Indonesian food too **Diet / menu / meal plan**: this is a low-sodium diet that ought to work for you
When is it being recommended?	**Realtime**: where should I eat now; what's near me? **Batch**: here is your weekly email of recipes, just for you
Why is it being recommended?	**Taste**: here is something you might like to eat / make / order **Health & wellness**: to help people become or remain healthy, to help people lose weight, and to help patients recover


International Cuisines
drag the ring to rotate
Cuisine
Americas
Asian
European
African
Middle Eastern
Jewish
Dietary Restrictions
Mediterranean
Caribbean
South American
Latin American
Central American
Mexican
American
North American
Southern
Tex-Mex
Cajun/Creole
Canadian
Oceanic
Polynesian
Australasian
Central Asian
East Asian
Japanese
Korean
Chinese
Southeast Asian
Indonesian
Malaysian
Thai
Vietnamese
South Asian
Indian
Arabian
Spanish
Iberian
Italian Peninsula
Portuguese
Italian
Greek
Balkans
Baltic
The Caucasus
Eastern European
Scandinavian
Western European
German
French
Northern European
British
Irish
English
Scottish
Moroccan
Central African
East African
Horn African
North African
Southern African
West African

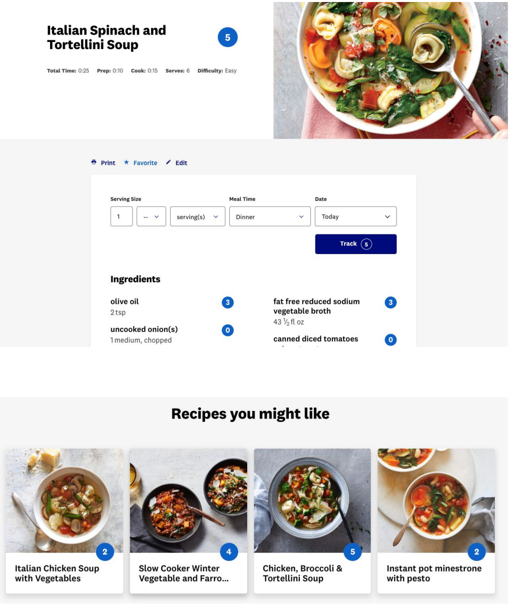
Food is at the core of our product



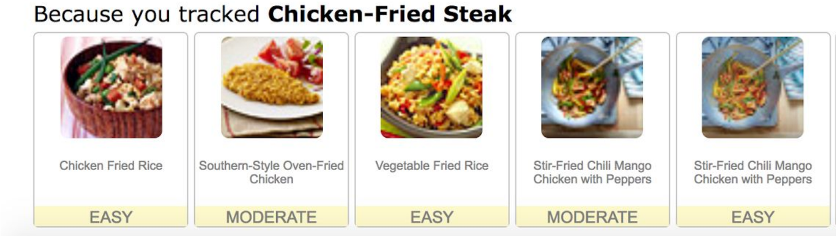


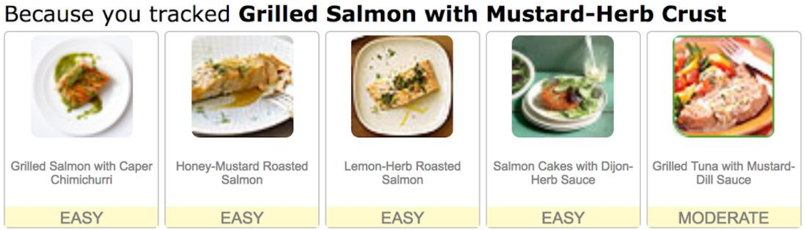
Recipe Recommendations

Similar Recipes



Dinner Recommendations






```
# note: push tokenization and and handling of ngrams down to tokenize in concrete classes
self.tfidf = TfidfVectorizer( tokenizer=self.tokenize)
self.term_document_matrix = self.tfidf.fit_transform( self.docs)
def cosine_similarity_matrix(self):
    return cosine_similarity(self.term_document_matrix)
```

Similar Recipes Flow

US WW Recipes → Similar Ingredients / Similar Names → Filters

Only recipes with images

document = ingredient list or name string

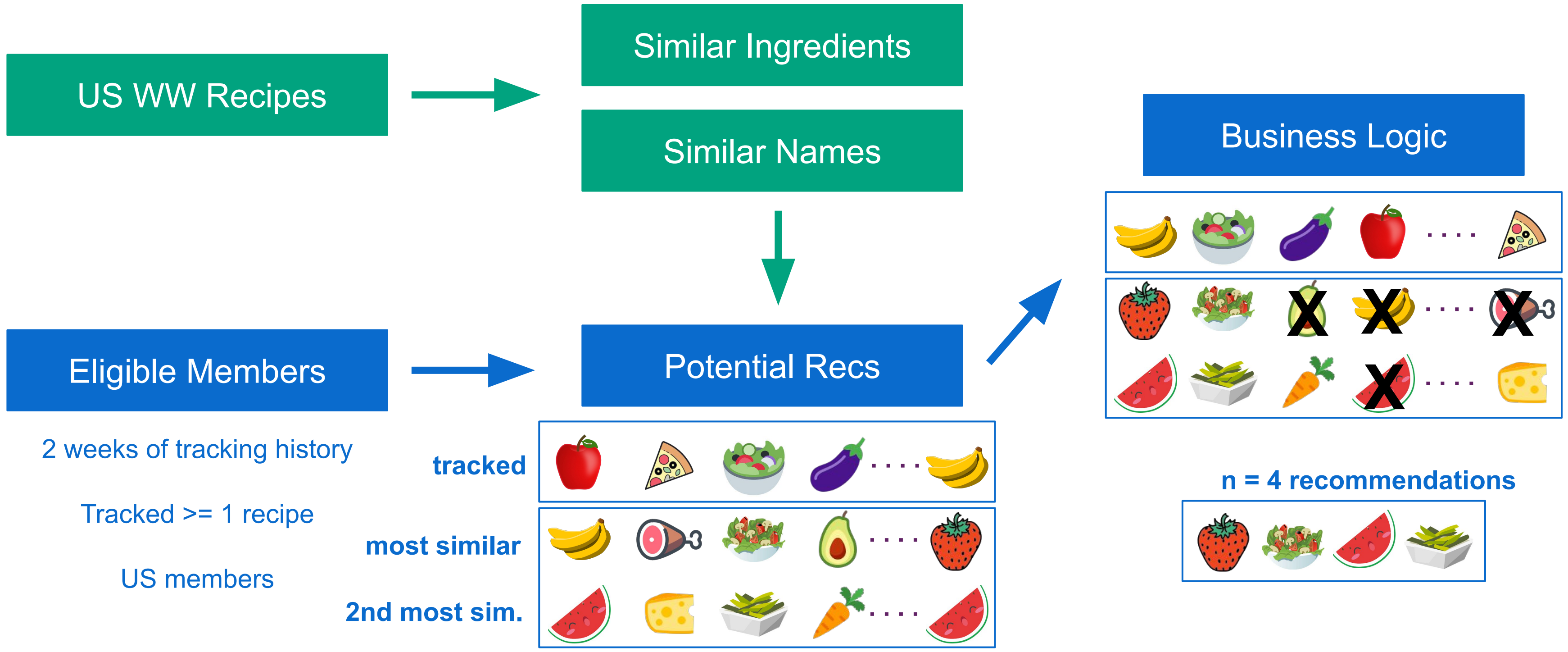
lemmatize, tokenize, TF-IDF

Cosine similarity

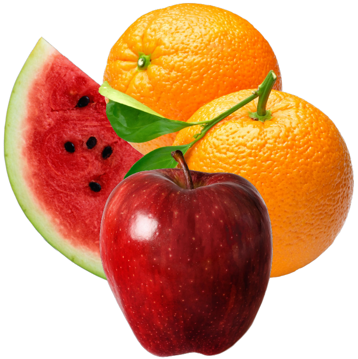
Rank

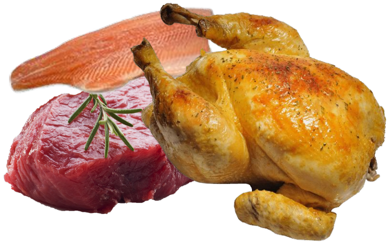
dietary
course
cuisine
main ingredient

Dinner Recommendations Flow



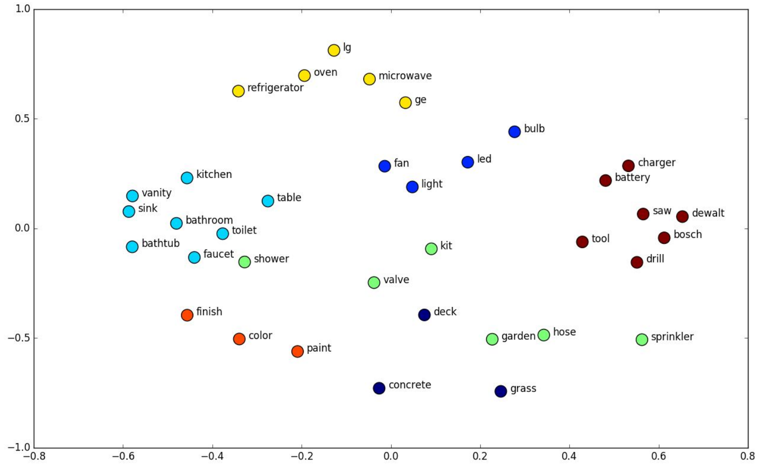
Food Embeddings





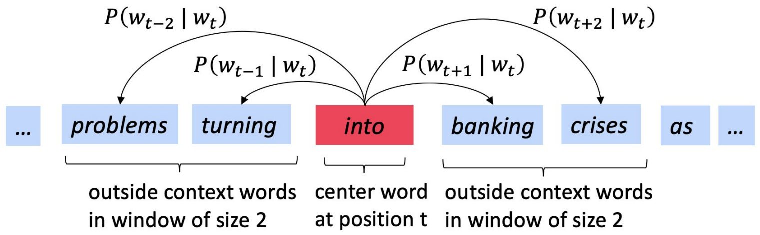
- **Motivation:** want to learn a space of foods where similar foods are located near each other
- Applications
  - Recommend low point substitute foods
  - Input into recipe recommender
  - Classify new foods and users
- How to do this? **Word embeddings**!

Word Embedding Overview



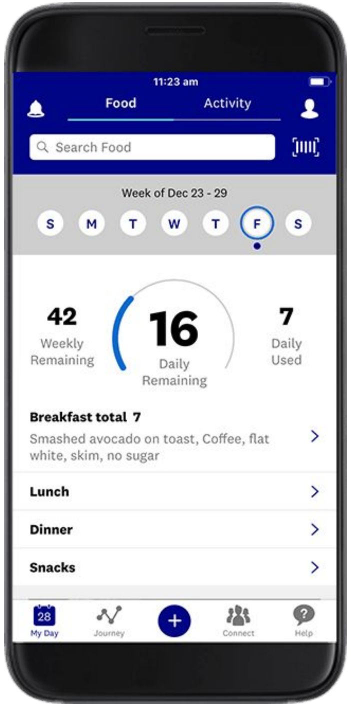
- Dense real-valued vectors representing word meaning
- **Idea:** words with similar meanings are grouped together in the embedding space
- Many forms of meaning are conflated since there is only one representation per word

FastText Behind the Scenes



- Learns embeddings using either Skip-gram or CBOW algorithms
- But learns representations for sub-word units rather than entire words
- Representations for whole words are composed from subword representations

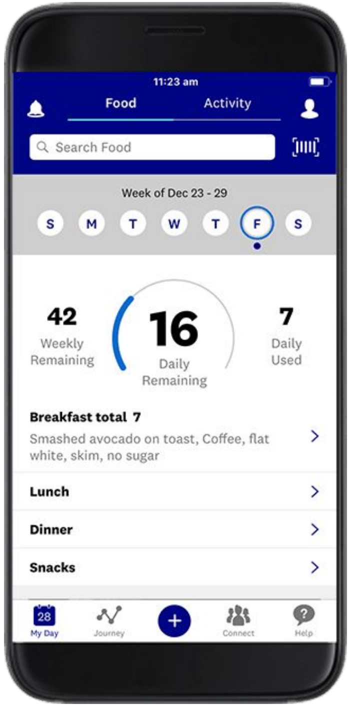
Preliminary Attempts



- **Attempt 1 (using food log data) [1]:**
  - Split food names into tokens
  - Each food name = 1 document
  - Average token embeddings for food name
  - Append calorie-normalized nutritional info
  - Did not work well, but might work with better preprocessing
- **Attempt 2 (using recipe data) [2]:**
  - Context = recipe ingredients
  - Each recipe = 1 document
  - Did not work well, recipe data too small

[1] Diet2Vec: Multi-scale analysis of massive dietary data; Tansey et. al 2016
[2] Cooking up Food Embeddings; Sauer et. al 2017

Final Attempt



- Context = ordered food entries, grouped by user id and time of day (meal type) over one week
- Preprocessed data same way as in [2]
- Each "word" in a document is a whole food name
- Best result from using subword unit modeling
- Other ideas: filtering for power users

(UUID=1234, breakfast, week1) = [Monday breakfast, Tuesday breakfast, Wednesday breakfast,...]

= [coffee, toast, jam, apple, coffee, orange_juice, tea, cereal, 2%_milk, banana,...]

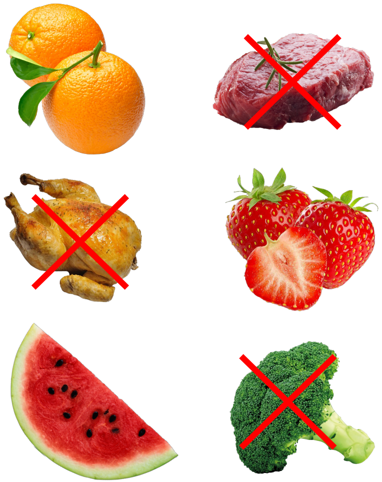
Will learn associations among items:

within meal: cereal↔2% milk, cereal↔whole milk

among meals: apple↔banana, coffee↔tea

Post-processing for Substitute Extraction

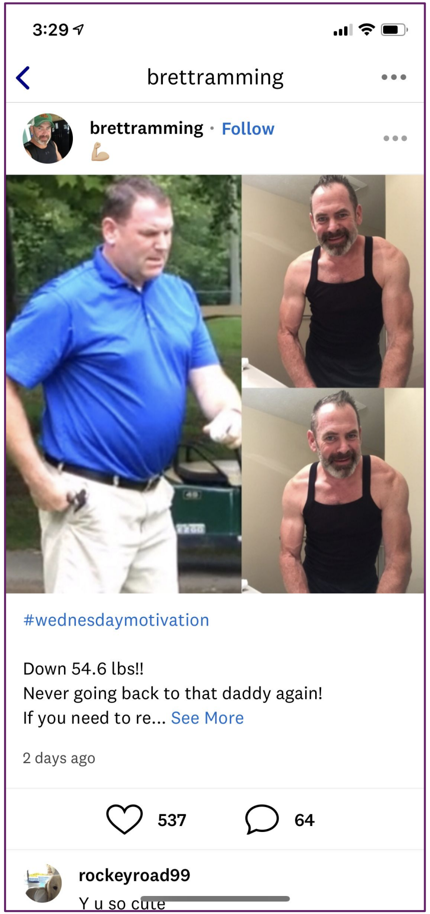
Query: 

Results: 

- One of the main goals of the project was to extract substitute food items
- Food data contains category information
- Simply eliminate results from NN list that are not in the same category

Personalizing Social Network

Connect


3:29
brettramming
brettramming · Follow
#wednesdaymotivation
Down 54.6 lbs!!
Never going back to that daddy again!
If you need to re... See More
2 days ago
537
64
rockeyroad99
Y u so cute


3:28
bethawny
bethawny · Follow
Beth #wwconnectcrew: groups;in my...
#throwbackthursday Italy edition
I was 25 years younger and over 100 pounds heavier back in 1994 when I visited Italy. I remember how hard it was to get... See More
1 day ago
926
70
auntmarci
Thanks so much for always sharing your


3:28
◂ Search
bethawny
bethawny · Follow
Beth #wwconnectcrew: groups;in my...
#facetofacefriday Italy edition
And look I have a neck, too!!
25 years ago i was 100 pounds hea... See More
8 hours ago
373
22
sabinec58

Seeking positivity

Getting help

Sharing goals

Encouraging others

Making friends

Building a brand

Qualitative research from our experience research / analytics teams

Seeking positivity

I want to feel good and see inspirational and useful posts.

Getting help

I post when I have questions or need encouragement.

Sharing goals

I post to show what I did or am going to do.

Encouraging others

I give back the support I received to those who need it.

Making friends

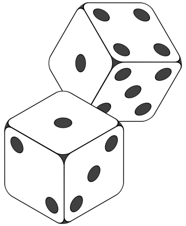
I build and invest in meaningful relationships.

Building a brand

I create content to grow a large following.

“Hidden” agendas may change throughout the day or over the course of a member’s journey

Personalized Feed



Your personalized feed of recommended posts

Video
Video
Content-based
Before / After

...

Collaborative filter | Content-based | Videos | Before / After | Popular

Personalized Feed

D1 D2 D3 D4 D5

The Multi-Armed Bandit Problem

- We have d arms. For example, arms are ads that we display to users each time they connect to a web page.
- Each time a user connects to this web page, that makes a round.
- At each round n, we choose one ad to display to the user.
- At each round n, ad i gives reward $r_i(n) \in \{0, 1\}$: $r_i(n) = 1$ if the user clicked on the ad i, 0 if the user didn't.
- Our goal is to maximize the total reward we get over many rounds.

Thompson Sampling Algorithm

Step 1. At each round n, we consider two numbers for each ad i:

- $N_i^1(n)$ - the number of times the ad i got reward 1 up to round n,
- $N_i^0(n)$ - the number of times the ad i got reward 0 up to round n.

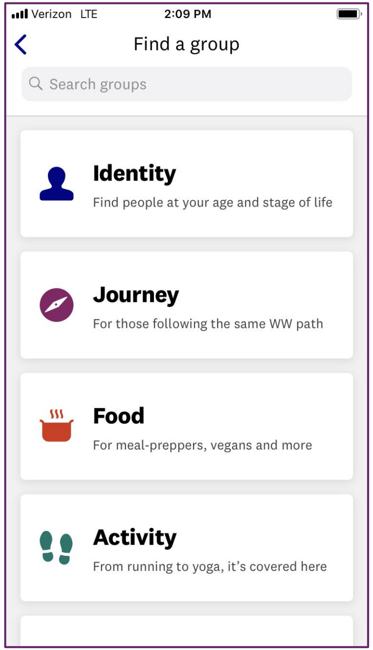
Step 2. For each ad i, we take a random draw from the distribution below:

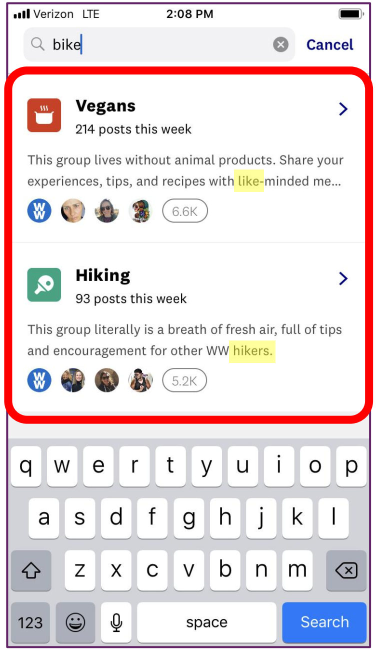
$$\theta_i(n) = \beta(N_i^1(n) + 1, N_i^0(n) + 1)$$

Step 3. We select the ad that has the highest $\theta_i(n)$.

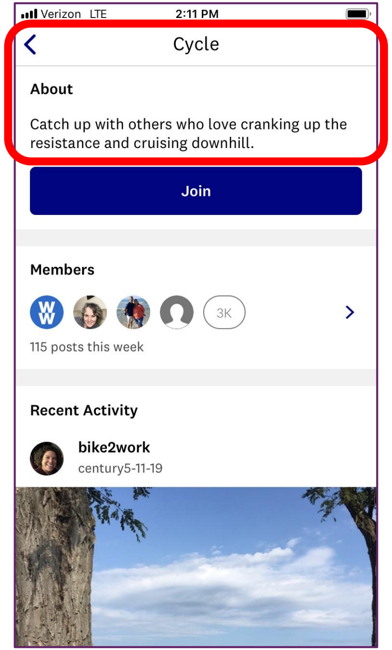
Group Search

Example: I love biking. Is there a group for this?





Issue: search only uses title and description, not content



Solution: we've provided top 100 terms + top 100 hashtags per group

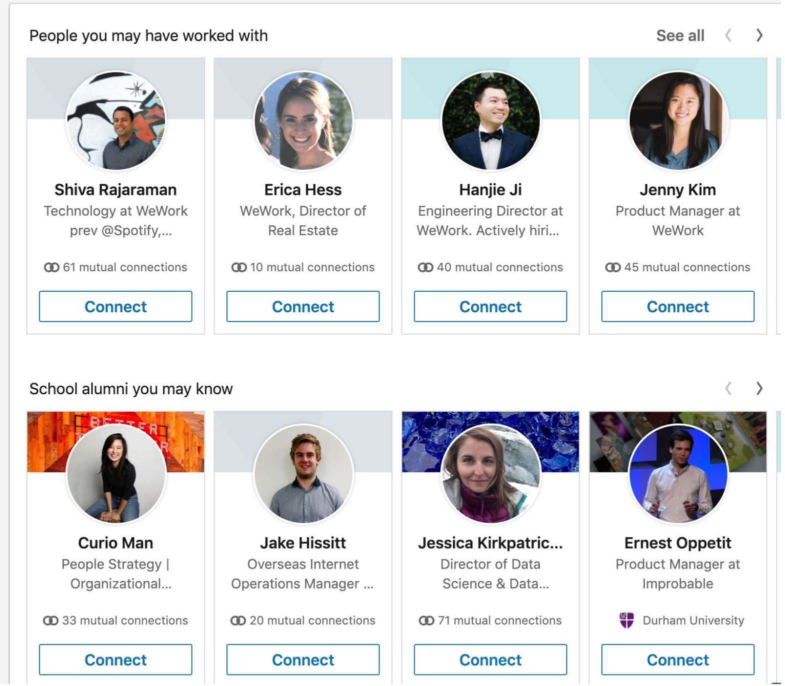
Cycle	Camping	Brides
rider	campground	bride
cyclist	tent	bridal
cycling	camping	alteration
ally	camper	bridesmaid
century	trailer	venue
biking	teardrop	seamstress
output	campfire	engagement
cadence	camped	ceremony

...

...

Who To Follow

Who To Follow



Who To Follow

light fm

arXiv.org > cs > arXiv:1507.08439

Search...

Help | A

Computer Science > Information Retrieval

Metadata Embeddings for User and Item Cold-start Recommendations

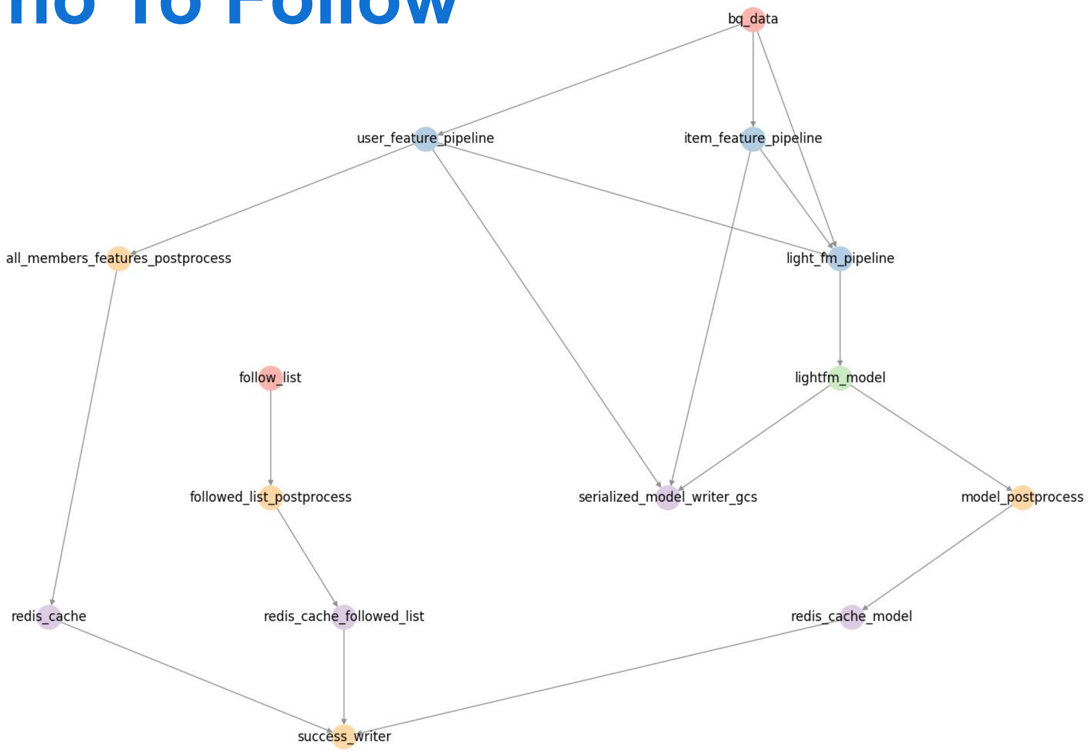
Maciej Kula (this http URL)

(Submitted on 30 Jul 2015)

I present a hybrid matrix factorisation model representing users and items as linear combinations of their content features' latent factors. The model outperforms both collaborative and content-based models in cold-start or sparse interaction data scenarios (using both user and item metadata), and performs at least as well as a pure collaborative matrix factorisation model where interaction data is abundant. Additionally, feature embeddings produced by the model encode semantic information in a way reminiscent of word embedding approaches, making them useful for a range of related tasks such as tag recommendations.

Subjects:	**Information Retrieval (cs.IR)**
ACM classes:	H.3.3
Cite as:	arXiv:1507.08439 [cs.IR]
	(or arXiv:1507.08439v1 [cs.IR] for this version)

Who To Follow

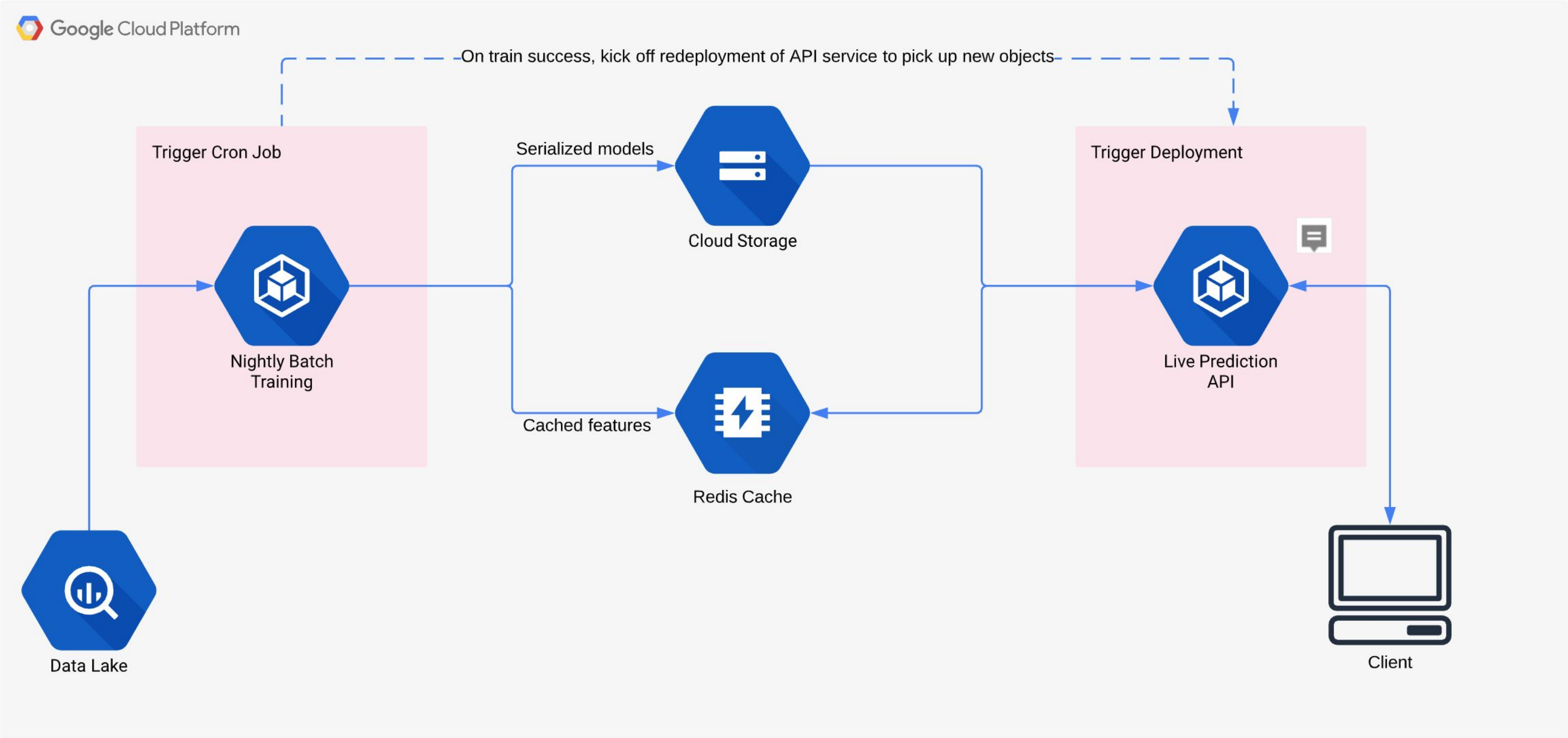


Features:

- Demographics: age, gender…
- Location
- Membership: type…
- Goal / Weight
- Tags, interactions
- Groups

…

Who To Follow



Primrose

Taking Stock of our own challenges

What would make a good recommender system **at WW?**

Slow serialization

but our medium data
can be kept in RAM…

No live features

but we know Docker, k8s…

Easy onboarding

mono repo with config as code…

Primrose has features to address each design consideration

Primrose: (*Production In-Memory Solution)* framework for solving WW's most common use cases, caching batched predictions with machine-learning engineering baked-in.

Data science	Infrastructure	People
Python **in-memory DAG** runner, with **no serialization** between nodes of the DAG.	DAG is defined as **configuration-as-code** approach -- one container for all models	Abstract ML and data manipulation operations, data scientists can easily **extend the framework**

Primrose: a framework for simple, quick modeling deployments

and we open sourced it....

Python Software Foundation [US] | pypi.org/project/primrose/

Apps | Calendar | Email | The Wave > Home | Oracle Application... | FB Workplace | Springboard | Das... | Identity now (Sail... | JIRA | Confluence | Other Bookmarks

Search projects | Help | Donate | Log in | Register

primrose 1.0.6

`pip install primrose`

Latest version

Last released: Sep 11, 2019

Primrose: a framework for simple, quick modeling deployments

Navigation

- Project description
- Release history
- Download files

Project links

- Homepage
- Source
- Documentation

Statistics

GitHub statistics:

- Stars: 9
- Forks: 0
- Open issues/PRs: 3

View statistics for this project via Libraries.io, or by using Google BigQuery

Meta

Project description

Overview

build passing | python 3.6 | 3.7 | pypi package 1.0.6 | license Apache 2.0 | docs up

Primrose at a glance

`Primrose` is a simple **Python** framework for executing **in-memory** workflows defined by directed acyclic graphs (**DAGs**) via configuration files. Data in `primrose` flows from one node to another while **avoiding serialization**, except for when explicitly specified by the user. `Primrose` nodes are designed for **simple batch-based machine learning workflows**, which have datasets small enough to fit into a single machine's memory.

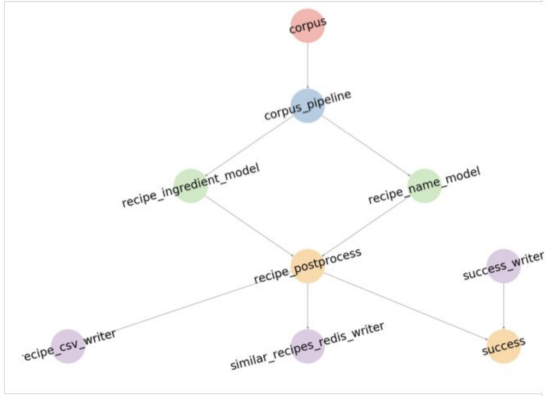
Table of Contents

We suggest reading the documentation in the following order:

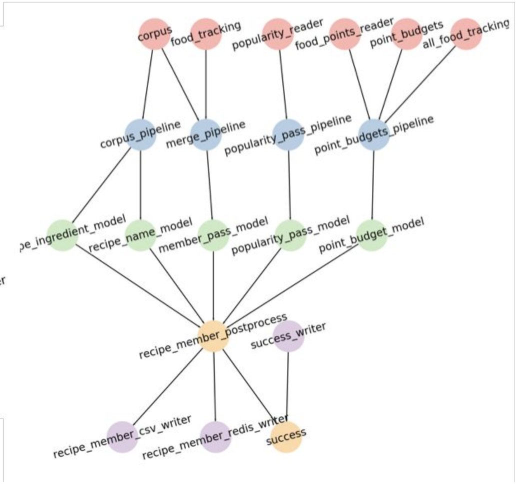
- Overview and motivation for `primrose`—this file.
- Getting Started: run your first `primrose` jobs.
- DAG Configurations: `primrose` adopts a configuration-as-code paradigm. This section introduces `primrose` configuration files.
- Metadata: this covers more advanced options of the configuration files.
- Command Line Interface (CLI): run commands using the CLI.
- Developer Notes: how to create your own new Node classes.
- DataObject: a deep dive into `DataObject`, the core data handling and book-keeping object.

Primrose jobs are executed as Directed Acyclic Graphs (DAG)s in python

Similar recipe DAG example



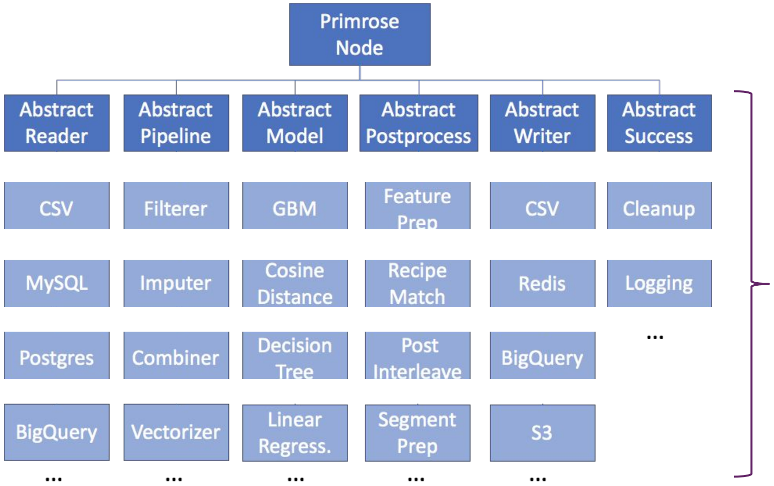
Recipe Recommendations



Flexibility: any number of operations allowed in a single DAG, across any python library

Data and functions are passed between nodes in an object that understands how to extract the correct data for each node

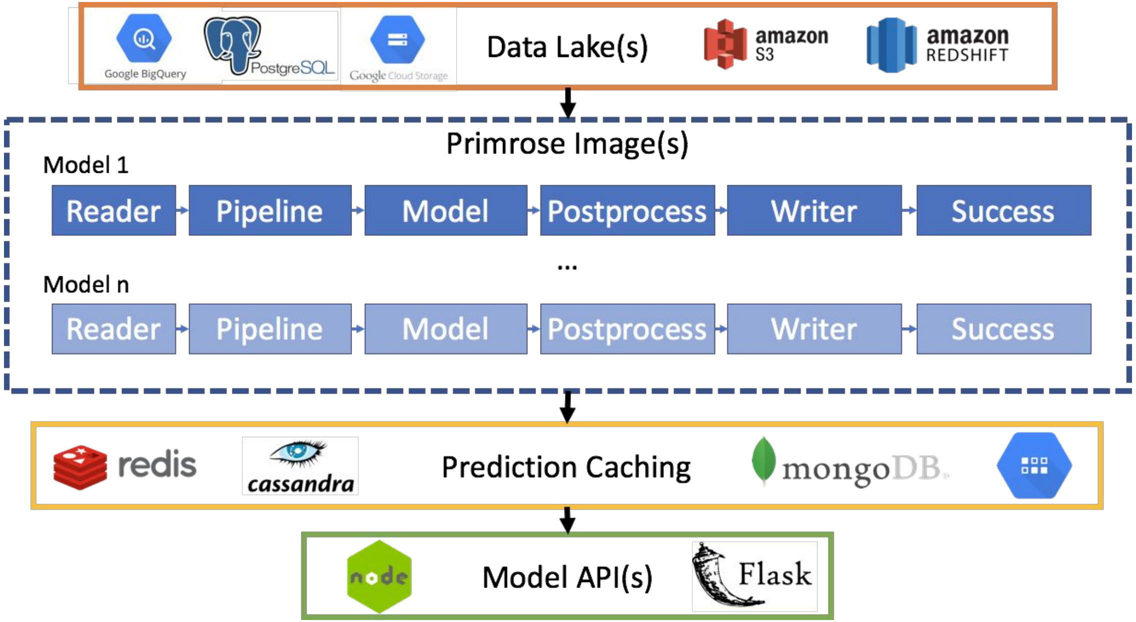
DAGs are composed of implementation agnostic, extensible nodes for data science



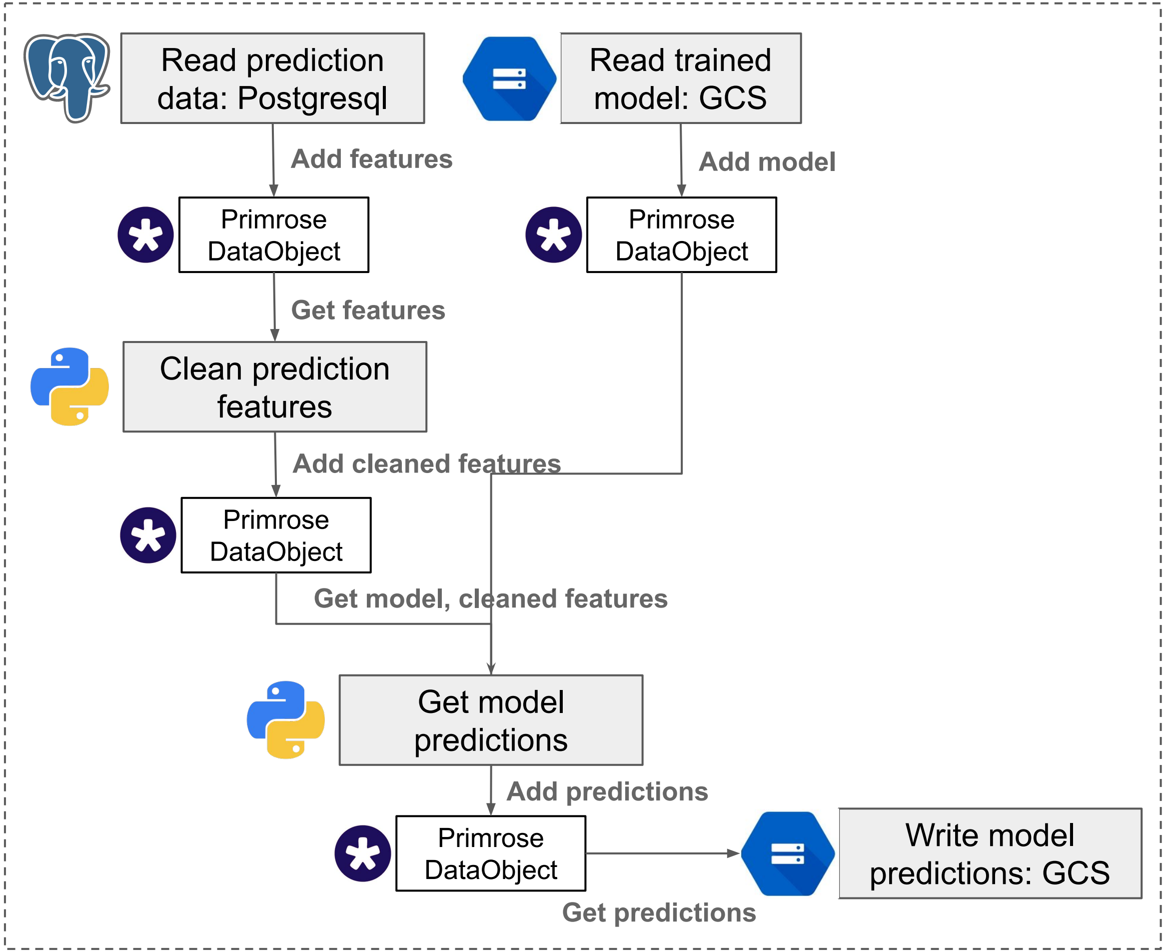
Data scientists can write any class that matches the abstract interface & incorporate in their DAGs

Data scientists can write individual nodes using *any Python framework* or library they choose

Primrose is run like an ETL pipeline in a single docker container for each configuration



Single Primrose image

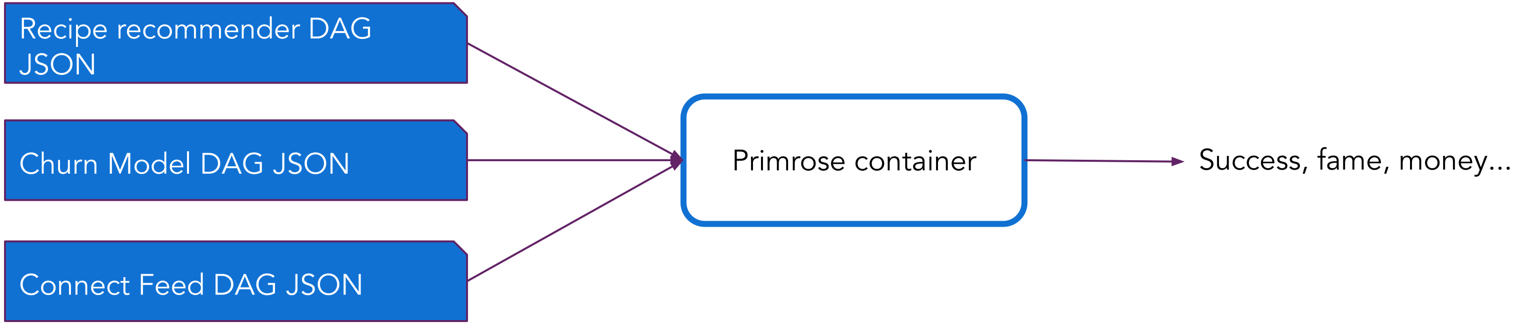


For simpler deployments: Primrose uses a “configuration as code” approach

Object configuration and DAG structure are build in a configuration JSON

Primrose validates the configuration and instantiates the correct classes at runtime

Different outputs and results for each DAG



Primrose config snippet: cluster with KMeans

```
"kmeans_cluster_model":{
  "class": "SklearnClusterModel",
  "mode": "train",
  "features": ["x1","x2"],
  "model": {
    "class": "cluster.KMeans",
    "args": {"n_clusters": 6, "random_state": 42}
  },
  "destinations": ["write_data", "write_model"]
}
```

Primrose config snippet: use DBSCAN instead

```
"dbscan_cluster_model":{
  "class": "SklearnClusterModel",
  "mode": "train",
  "features": ["x1","x2"],
  "model": {
    "class": "cluster.DBSCAN",
    "args": {"min_samples": 3}
  },
  "destinations": ["write_data", "write_model"]
}
```

Primrose job in cloud

Primrose container

Same container & build!

Similar Recipes Flow

US WW Recipes → Similar Ingredients / Similar Names → Filters

Only recipes with images

document = ingredient list or name string

lemmatize, tokenize, TF-IDF

Cosine similarity

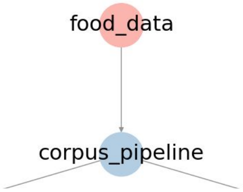
Rank

dietary
course
cuisine
main ingredient

Productionalize in Primrose DAG

Google BigQuery Data lake Reader

NLTK + Custom Lemmatization




```
"food_data": {
  "class": "BQSqlReader",
  "query_json": [
    {
      "query": "sql/recipe_recommender/recipe_data_us.sql",
      "parameters": {"market_id": "US"...}
    }
  ],
  "bucket_name": "user-model-storage",
  "ds_dataset": "ds_wwi_views",
  "project": "wwi-data-playground-3",
  "gcs_project": "wwi-data-playground-3",
  "debug": true,
  "read_cache": true,
  "write_cache": true,
  "destinations": [
    "corpus_pipeline"
  ]
```

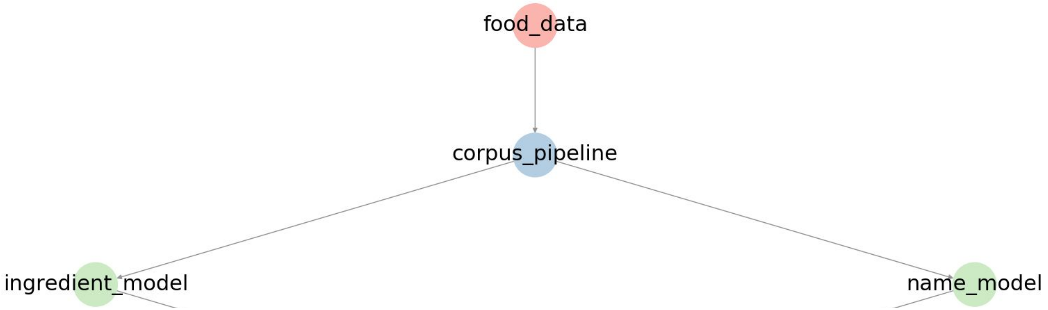
```
"corpus_pipeline": {
  "class": "SimilarRecipesPipeline",
  "nonfood_image_exclude_list": "data/non_food_image_exclude_list.txt",
  "filter_out_non_image": true,
  "remove_dupe_ingredients": true,
  "filter_in_dinner": false,
  "is_training": true,
  "popularity_scaling_min": 0.4,
  "popularity_scaling_max": 0.6,
  "destinations": [
    "ingredient_model",
    "name_model"
  ]
}
```

Productionalize in Primrose DAG

Google BigQuery Data lake Reader

NLTK + Custom Lemmatization

Sklearn TF-IDF + cosine similarity




```
"ingredient_model": {
  "class": "RecipeIngredientSearchEngine",
  "mode": "predict",
  "id_key": "recipeID",
  "doc_key": "ingredient_string",
  "destinations": [
    "recipe_postprocess"
  ]
},
```

```
"name_model": {
  "class": "RecipeNameSearchEngine",
  "mode": "predict",
  "id_key": "recipeID",
  "doc_key": "displayName",
  "destinations": [
    "recipe_postprocess"
  ]
}
```

```
"corpus_pipeline": {
  "class": "SimilarRecipesPipeline",
  "nonfood_image_exclude_list": "data/non_food_image_exclude_list.txt",
  "filter_out_non_image": true,
  "remove_dupe_ingredients": true,
  "filter_in_dinner": false,
  "is_training": true,
  "popularity_scaling_min": 0.4,
  "popularity_scaling_max": 0.6,
  "destinations": [
    "ingredient_model",
    "name_model"
  ]
}
```

Productionalize in Primrose DAG

Google BigQuery Data lake Reader

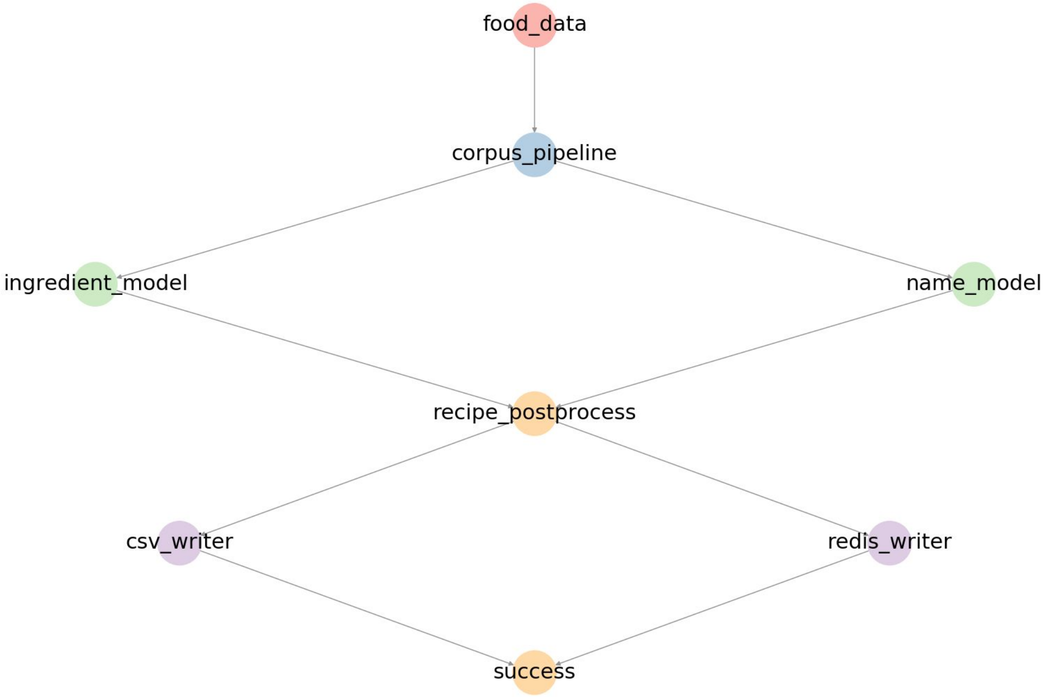
NLTK + Custom Lemmatization

Sklearn TF-IDF + cosine similarity

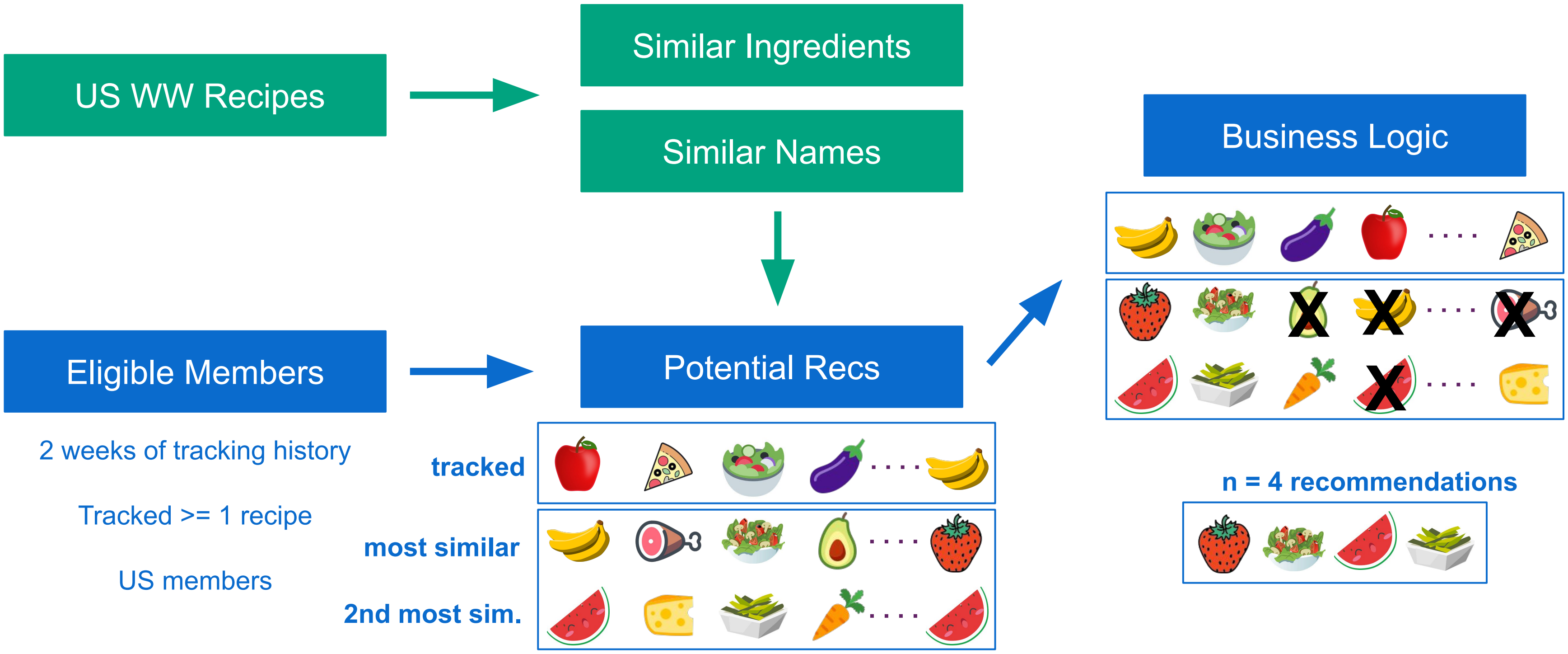
Business Logic (filters)

Write to GCS Bucket and Google MemoryStore

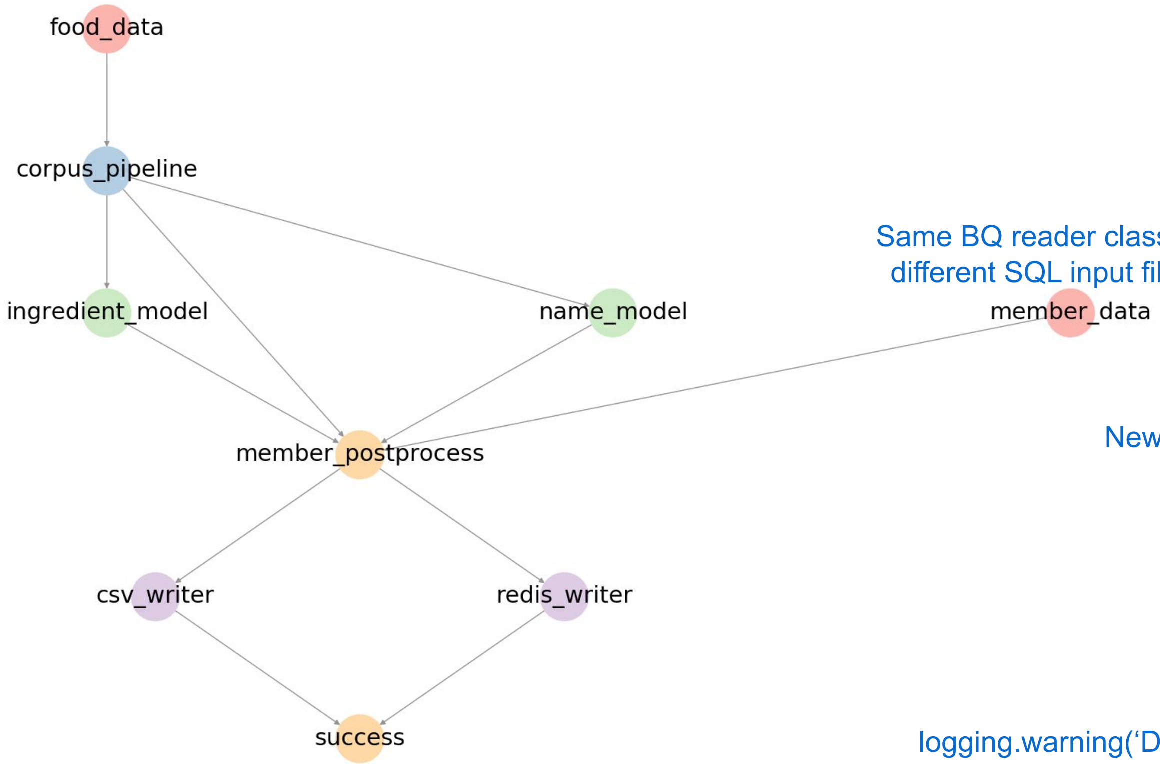
Success!
logging.info('Your newbie DS has written production quality code.')



Dinner Recommendations Flow



Productionalizing is easier the second time



Same BQ reader class,
different SQL input file

```
"class": "BQSqlReader",
"query_json": [
  {
    "query": "sql/recipe_recommender/member_food_tracking_us.sql",
    "parameters": {"market_id": "US"...}
  }
],
"bucket_name": "user-model-storage",
"ds_dataset": "ds_wwi_views",
"project": "wwi-data-playground-3",
"gcs_project": "wwi-data-playground-3",
"debug": true,
"read_cache": true,
"write_cache": true,
"destinations": [
  "member_postprocess"
]
```

New postprocess class to sort, filter and interleave potential recommendations

```
"member_postprocess": {
  "class": "RecipeMemberPostprocess",
  "destinations": [
    "csv_writer",
    "redis_writer"
  ]
}
```

Success!
logging.warning('Data Scientist is developing software engineering skills.')

Primrose has features to address each design consideration

Primrose: (*Production In-Memory Solution)* framework for solving WW's most common use cases, caching batched predictions with machine-learning engineering baked-in.

Data science

Python **in-memory DAG** runner, with **no serialization** between nodes of the DAG.

Infrastructure

DAG is defined as **configuration-as-code** approach -- one container for all models

People

Abstract ML and data manipulation operations, data scientists can easily **extend the framework**

Wrap Up

Nudges:

- Once is not enough: nudge different times, channels, timescales
- Recognition really important: Nudge before, recognize after
- Holistic view: challenges, community, personality

Primrose:

- In-memory, config-as-code, extensible
- Helped our new team be productive and get models into prod
- Available today

Questions

- carl.anderson@ww.com
- @leapingllamas

- **Food RecSys**: https://arxiv.org/abs/1809.02862
- **Primrose**: https://github.com/ww-tech/primrose
- **Tech blog**: https://medium.com/ww-tech-blog

Hiring: especially data scientists in Toronto