

# Voiceflow Keynote: Demystifying Multimodal Design

**CHERYL PLATZ**

Principal Designer & Owner, Ideaplatz

[cheryl@ideaplatz.com](mailto:cheryl@ideaplatz.com)

[@ideaplatz](#)

October 2021

Ideaplatz 



# Before we begin...

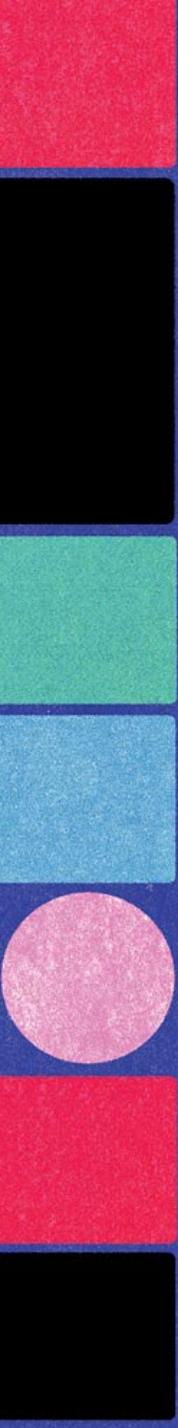
- This talk is being delivered within the traditional territory of Coast Salish peoples, specifically the Duwamish Tribe. Despite hundreds of years of broken treaties and setbacks, the Duwamish people continue to shine a light here in their ancestral lands as the host tribe of Seattle, all while continuing to petition for formal government recognition. I am grateful to the Duwamish tribe, past and present, and honor their land itself.
- Thank you for my 17 years in Seattle!



# Computer, who is Cheryl?

- Author of Design Beyond Devices
- Original UX Designer on Echo Look team
- Original designer for Alexa Notifications
- Voice, systems, and multimodal design: Alexa, Cortana, Windows Automotive, Dynamics Power Virtual Agents
- Shipped one of the first speech-enabled Nintendo DS games (Disney Friends)
- Creator of several Alexa skills





# The future is multimodal, because humans are multimodal.

A **mode** (in this context) is a type of communication, and humans communicate using their senses.

A **multimodal interaction** is an exchange between a device and a human being where multiple input or output modalities may be used simultaneously or sequentially depending upon context and preference.



# What is multimodal design?

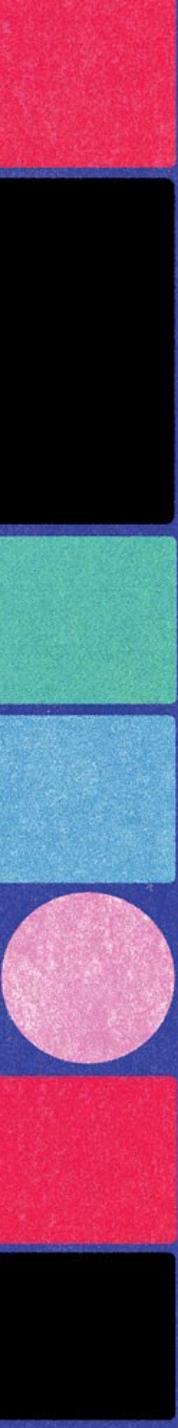
## DEFINITION

Multimodal design seeks to coordinate the delivery of multiple input and output stimuli to create a flexible, coherent experience for our customers.

## PRACTICE

Multimodal design is an additional layer of design rigor added on top of our existing modality-specific designs, like voice UI (VUI) designs.





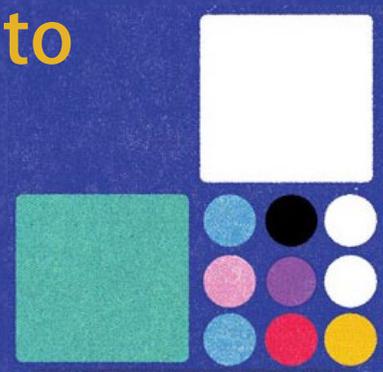
## WHY INCLUDE BOTH “MULTIMODAL” AND “CROSS-DEVICE” EXPERIENCES?

Wasn't just multimodality  
complicated enough for one  
book?

It's short-sighted to assume ANY  
experience exists in a vacuum. **Our  
customers are swimming in devices.**

**Even websites are cross-device now:**  
most websites must function on  
desktop and mobile, which means  
interruption, context, and  
notifications become relevant.

The limits of multimodality on one  
device **may cause a customer to  
turn to another device.**

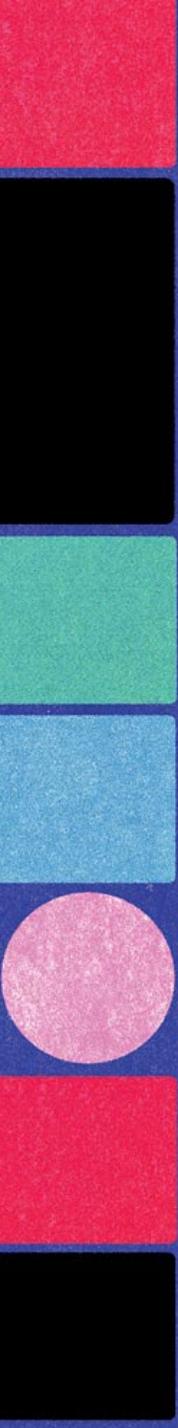


# What will we cover today?

- Multimodal Design Fundamentals
- Capturing Customer Context
- Transition Strategy
- Adding Visuals to Voice

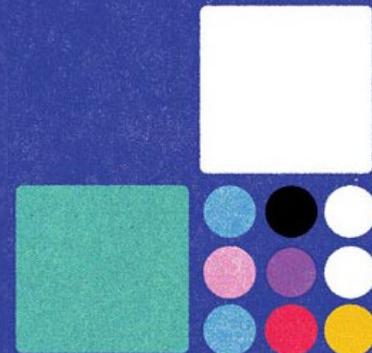
It's OK if you're not a voice design expert! We'll start with a light refresher, and much of this is systems thinking work that will apply regardless of experience level.





FUNDAMENTALS

# Multimodal Design



## Modality Description



### Visual

Projection or rendering of a stimulus that will be interpreted over optical channels— from books and e-readers to GIFs and videos.



### Auditory

The use of acoustic waves to communicate meaning: music, sound effects, or language.



### Haptic

Communicating meaning with changes to the physical environment: pressure, vibration, force feedback, or direct manipulation like taps or clicks.



### Kinetic

Communication based on movement or orientation in space.



### Ambient

Inferred meaning driven by environmental or biometric conditions: temperature, heart rate, lighting, etc.

## COMMUNICATION MODALITIES

As defined in the book; there is some debate about how to split the Kinetic and Ambient categories.

Note that these don't align 1:1 to the human senses.



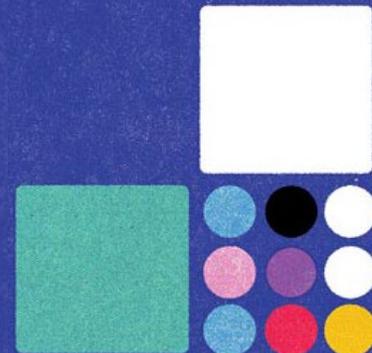
# Dimensions driving multimodality

## How rich is your information?

- Low information density  
Temperature, sports score
- High density  
Book, movie times, 10-day forecast

## How close is the device to the customer?

- Close proximity  
Wearable to arms-reach
- Long range  
3-10 feet



# THE SPECTRUM OF MULTIMODALITY

By plotting information density and proximity on a grid, you can place all current and future experiences in one of four categories.



RICH INFORMATION

QUADRANT 2  
**Anchored**

Experiences with rich physical presence where a customer is usually nearby.

*Fire TV, Xbox One, Cortana on PC*

QUADRANT 1  
**Adaptive**

Experiences that support both close proximity and long-range interactions.

*Echo Show, Facebook Portal, Google Nest Hub*

CLOSE PROXIMITY

LONG RANGE / FAR FIELD

QUADRANT 3  
**Direct**

Customer and device must be in direct contact or extreme proximity for use.

*Fitbit, Google Glass, Hololens, Apple Watch*

QUADRANT 4  
**Intangible**

Hands-free experience where close proximity to the device is not required.

*Echo (original) Google Home*

SCOPED INFORMATION

@IDEAPLATZ

You must understand your customer's context to know what interaction model makes sense in the moment.



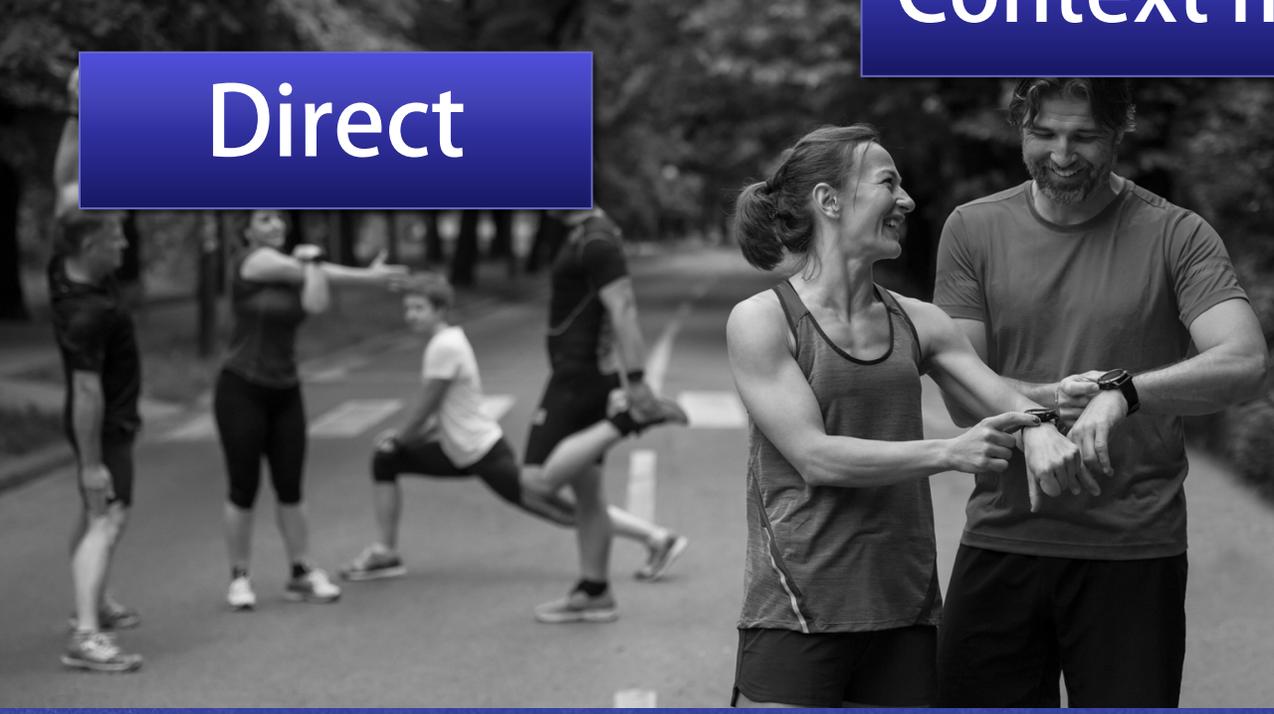


Anchored



Adaptive

Context matters.



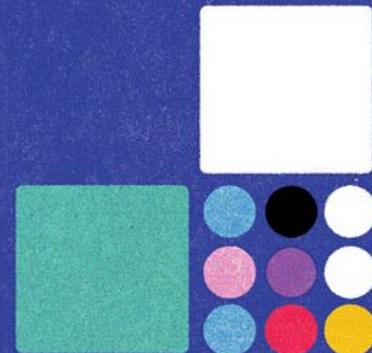
Direct



Intangible

EXPANDING YOUR UNDERSTANDING

# Capturing Customer Context



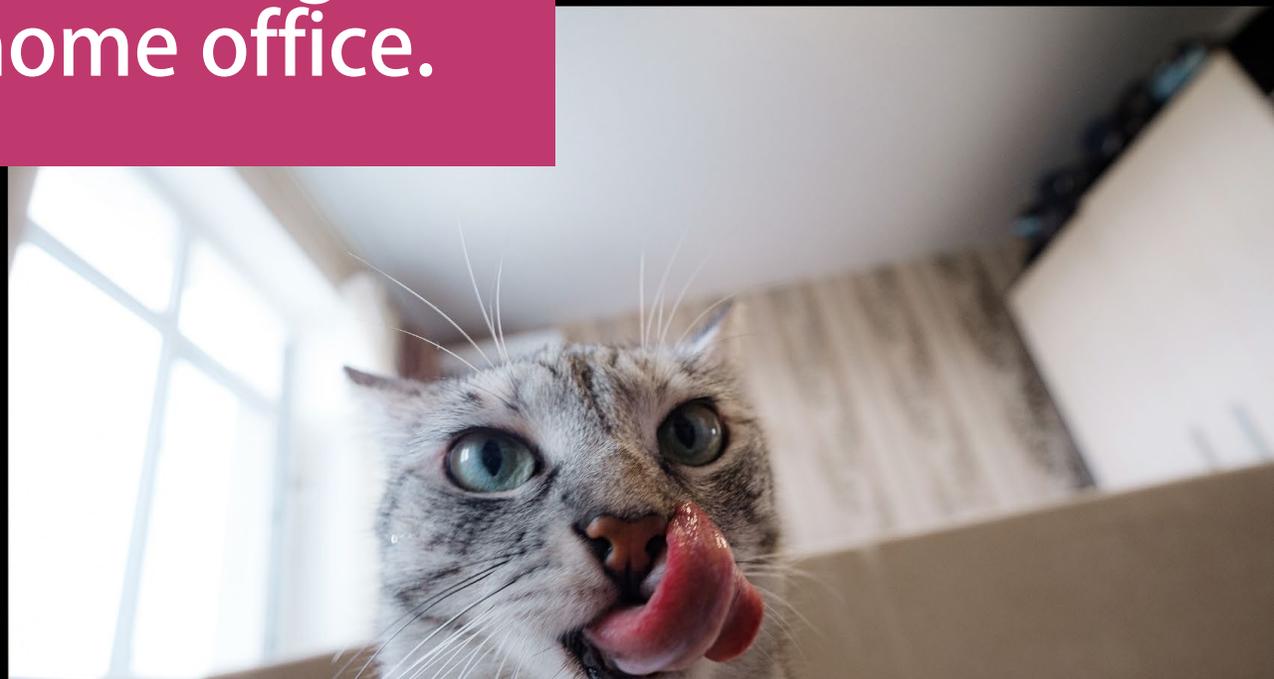


**What we thought we knew about  
the world as researchers and  
designers has changed.**





We've all learned  
there's no such thing as  
a "typical" home office.



# Use CROW to help you define and capture customer context.

CROW is a **storytelling shorthand** we use at my improv theater (Unexpected Productions) to quickly define the essential elements of a scene.

CROW stands for:

- Character
- Relationship
- Objective
- Where



# C: Character

## What defines your customer?

Dimension	Definition	Example questions
Attributes	Fundamental traits, mannerisms, and habits. Skills, quirks, gender identity, preferred pronouns, disabilities, communities.	<ul style="list-style-type: none"><li>• How does your customer define their own identity to themselves and others?</li><li>• Which of your customer's attributes are underrepresented, and how might that affect them?</li></ul>
Attitudes	Emotions and reactions to outside stimuli—other people, objects, or situations.	<ul style="list-style-type: none"><li>• Would this customer have any preconceived opinions or learned behaviors that they would bring to bear on this experience?</li></ul>
Choices	The actions you take, based on your beliefs and attitudes.	<ul style="list-style-type: none"><li>• Why would a customer choose to seek out your experience? Did they have a choice at all?</li></ul>



# R: Relationships

## What connects your customer?

The closer you are to someone (or something), the more likely you are to get emotional about it. Relationships drive satisfaction - and frustration.

### Human to device

- Device ownership
- Anthropomorphization
- Emotional attachment
- Financial investment
- Self-expression

### Human to business

- Communication channels
- Perception
- Choice
- Market conditions
- Expectations

### Human to human

- Cooperative use
- Sequential use
- Trust
- Identity
- Competition



# O: Objective

## What drives your customer?

What have you defined as your customer's objective?

Is that **truly** their end goal, or simply a sentence written to get the customer to your feature?

Has your product team assumed that their solution stands alone? Is that **true**, or are you part of a **larger, device-agnostic human objective** that might span multiple experiences?



# DON'T OBSTRUCT THE OBJECTIVE

Timing and context matter when considering customer objective. Even if you're offering something of value *later*, you might be obstructing objective in the moment.

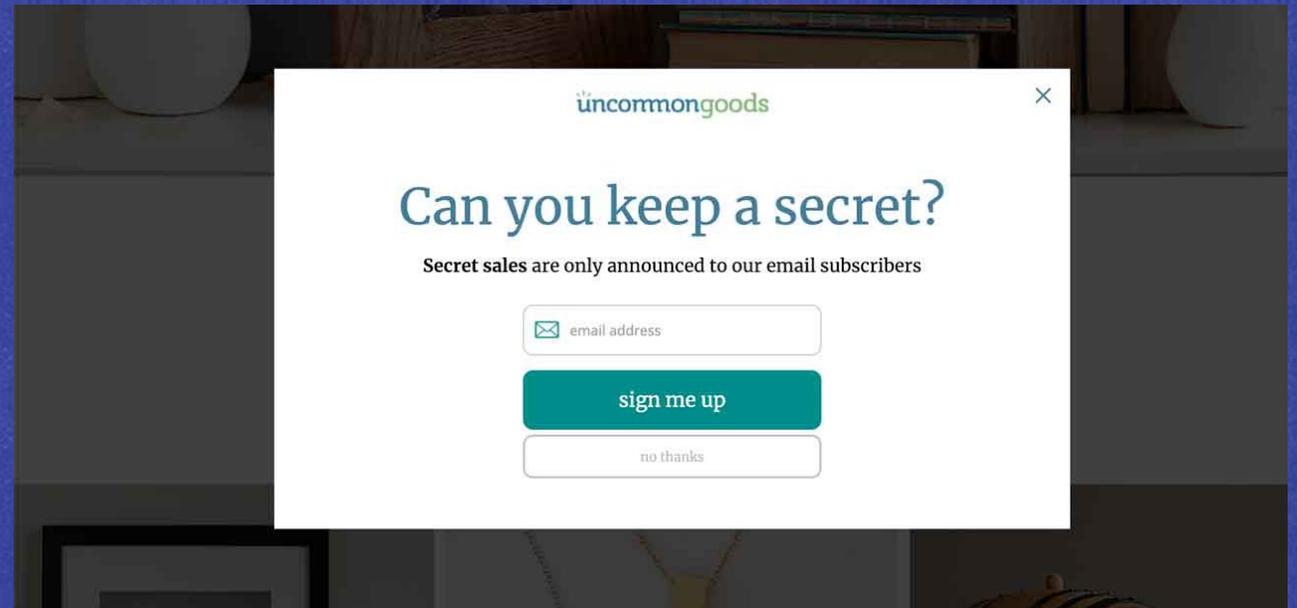
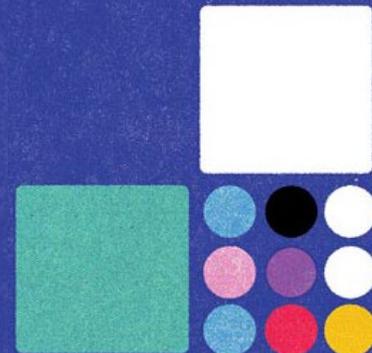


Image from <https://www.nngroup.com/articles/popups/>



# W: Where

## What surrounds your customer?

- Where will your customer be when they want to interact with you?
- Will they be seated? Standing? Moving?
- What is in arm's reach?
- What devices will be available?
- Who else will be in those environments?
- Are there distractions in the environment?
- Will customers expect to continue this experience between locations or devices?

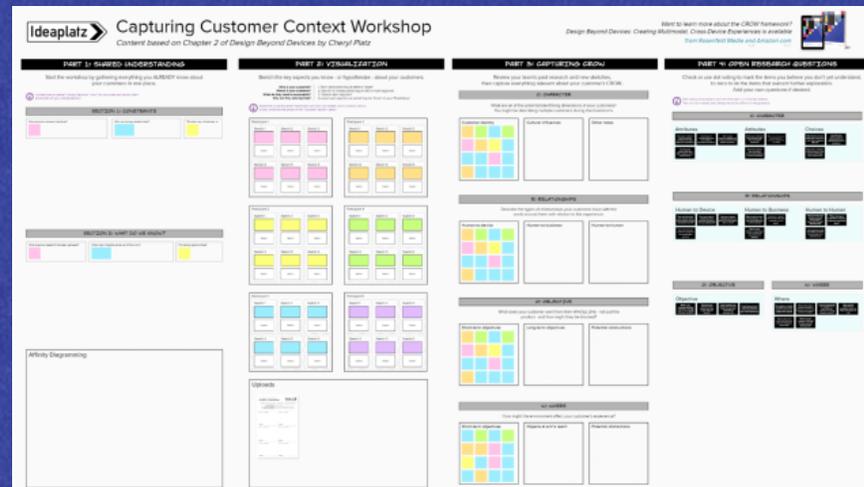




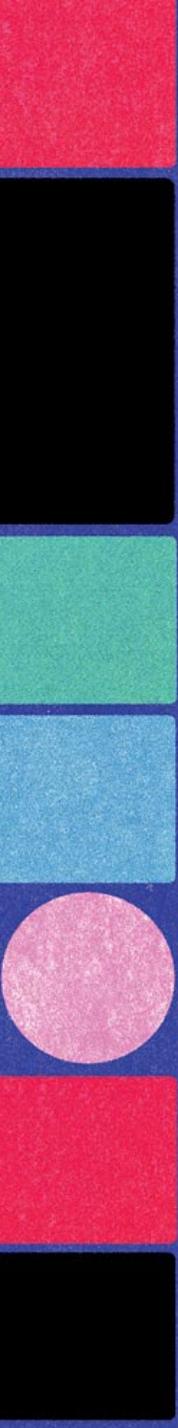
# Also available at <http://bit.ly/DBD-Ideaplatz>:

## MURAL template for Capturing Customer Context workshops

## Interview guide with CROW-inspired questions

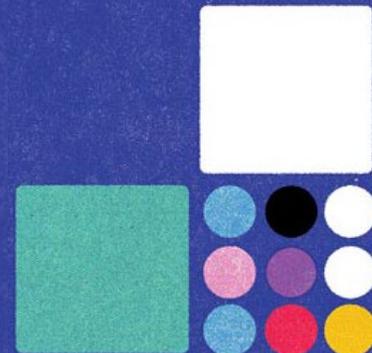


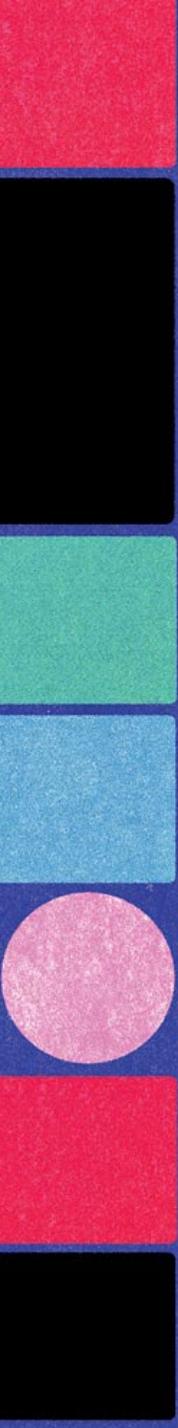
So you understand your customer context. What changes about the design process in multimodal and cross-device design?



COPING WITH TIME

# Transition Strategy

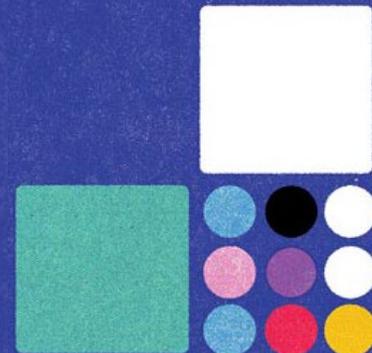




**TRANSITIONS WILL  
MAKE OR BREAK  
YOUR MULTIMODAL  
EXPERIENCE.**

We're so used to designing for the moments when things are fixed – but what about the shifts?

- Between modalities
- Between network connections
- Between devices



# Transitioning between modalities

## VOLUNTARY

The customer initiated the transfer, motivated by the belief the new modality will be easier or more appropriate.

## INVOLUNTARY

The current modality is deemed insufficient by the system, and the customer will be forced to switch to continue their activity.



## PAY SPECIAL ATTENTION TO YOUR INPUT AND OUTPUT CLIFFS

Cliffs are the rough “edges” between interactions. (Table 9.1 in Design Beyond Devices.)

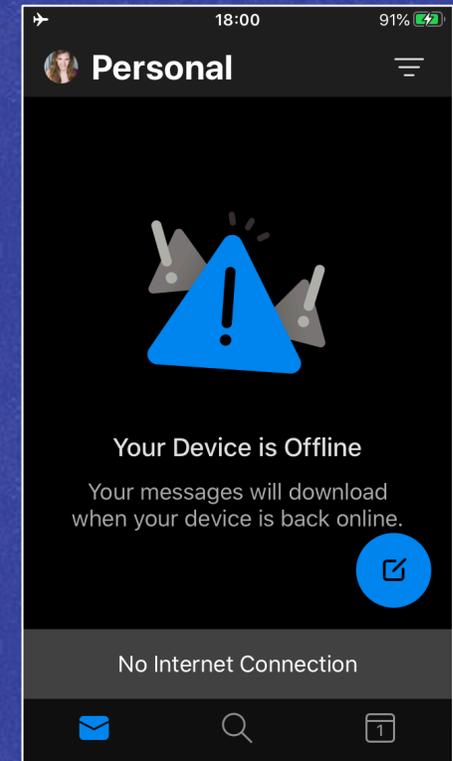
Cliff Type	Description
Input Transition	A customer transitions between two or more input modalities during a single activity.
Output Transition	The system changes the way it communicates with a customer in the middle of an activity.
Input/Output Mismatch	A system responds to a customer request using an output modality that does not match the input modality the customer used to make the request.



# Network connection transitions

Many apps fail to fully and properly consider the experience their customers will have when moving between network connections.

- Intentional connection loss
- Unstable connection
- Insufficient connection
- Proximity loss
- Connection failure



# Why enable transitions between **devices**?

## ENVIRONMENTAL CHANGES

- Commuting
- Traveling
- Mealtime
- Exercise
- Daytime/nighttime shift

## DEVICE SUITABILITY

- Data entry vs. consumption
- Physical comfort
- Temporary interference
- Multitasking



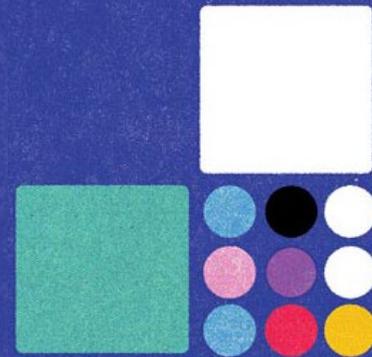
# COMMON DIRECTED TRANSITION ARCHETYPES

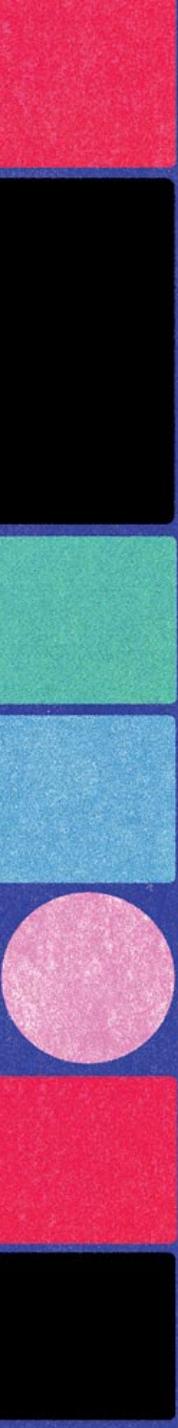
Excerpt from Table 9.5 in Design  
Beyond Devices

Archetype	Motivation
Environmental	Your customer is moving between physical spaces or adapting to changing conditions.
Device Preference	Your customer has multiple ways to complete a task and prefers different devices for different tasks.
Device Suitability	Your customer's goals have changed, and they have hit the limits of capabilities on a particular device.

DESIGN & DELIVERY

# Adding Visuals to Voice

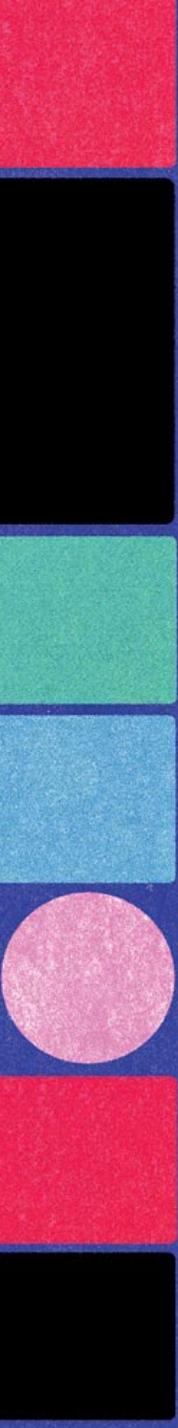




What does your customer see when they're interacting with your device?

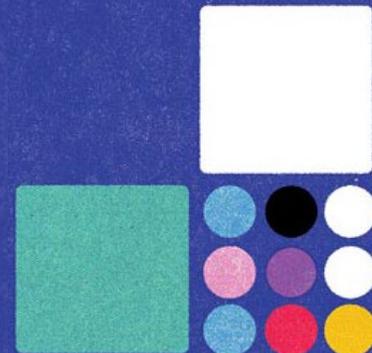
And how will your chosen interaction model impact your design?

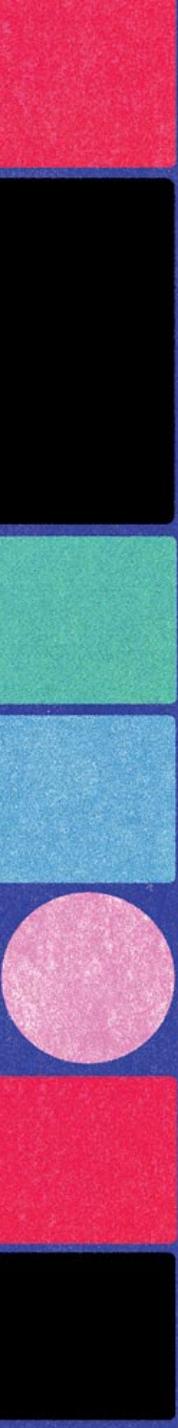




You're still going to need to  
do full visual designs for  
multimodal experiences.

MULTIMODALITY JUST ADDS ONE  
MORE LAYER OF COMPLEXITY.





# Surprisingly, your choice of visuals should be informed by your microphones.

## FAR-FIELD MICROPHONES

You can't assume the customer is looking at the device, so visuals shouldn't be **required** to complete key tasks if you're truly voice-forward.

## NEAR-FIELD MICROPHONES

Since your customer is so close to a microphone, they can see whatever the device has built in: LEDs, displays, etc.



# THE SPECTRUM OF MULTIMODALITY

By plotting information density and proximity on a grid, you can place all current and future experiences in one of four categories.



RICH INFORMATION

QUADRANT 2  
**Anchored**

Experiences with rich physical presence where a customer is usually nearby.

*Fire TV, Xbox One, Cortana on PC*

QUADRANT 1  
**Adaptive**

Experiences that support both close proximity and long-range interactions.

*Echo Show, Facebook Portal, Google Nest Hub*

CLOSE PROXIMITY

LONG RANGE / FAR FIELD

QUADRANT 3  
**Direct**

Customer and device must be in direct contact or extreme proximity for use.

*Fitbit, Google Glass, Hololens, Apple Watch*

QUADRANT 4  
**Intangible**

Hands-free experience where close proximity to the device is not required.

*Echo (original) Google Home*

SCOPED INFORMATION

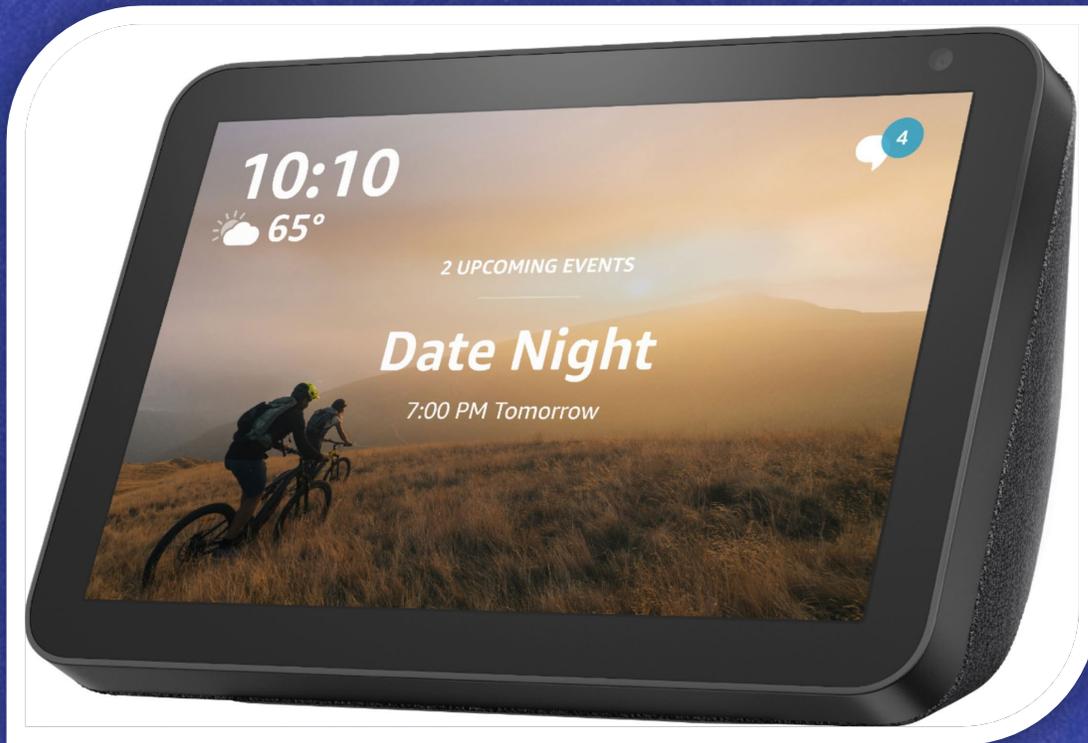
@IDEAPLATZ

QUADRANT 1:

# ADAPTIVE experiences

(Echo Show, Google Nest Hub etc.)

PROXIMITY: Far (or close)  
INFO DENSITY: Medium/High



- Generally support both close and far scenarios
- Scenarios are completable remotely but proximity unlocks new options
- Customers can choose how to interact in some or all situations



QUADRANT 2:  
**ANCHORED experiences**  
(Smart TV, home computer, VR, etc.)

PROXIMITY: Close  
INFO DENSITY: High



- Customer is likely to be in arms' reach of a controller or screen
- Close proximity means high density displays are supported
- Voice is usually supplemental due to high visual load



QUADRANT 3:

## DIRECT experiences

(Smart watches, Google Glass, AR, etc.)

PROXIMITY: Close  
INFO DENSITY: Low



- Device is usually attached to or in immediate proximity to customer
- Proximity allows use of additional ambient input sensors
- Small, constrained screens force low information density



QUADRANT 4:

## INTANGIBLE experiences

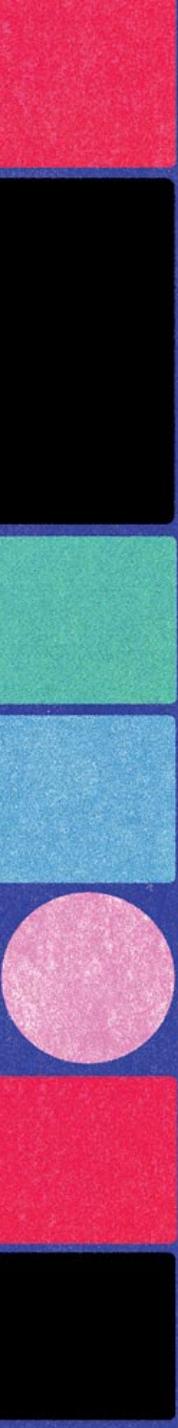
(Smart speakers, etc.)

PROXIMITY: Far  
INFO DENSITY: Low/Medium



- Customer is rarely near the device
- Customer does not have to look at the device to interact
- All interactions should be completable without physical interaction due to lack of proximity
- Lower information density





Multimodal flows are used to show how customers transition between multiple modes of interaction.

WHETHER THOSE TRANSITIONS ARE SYSTEM-DRIVEN OR CUSTOMER-DRIVEN.



# Choose your flow technique:

## SIMPLE FLOW

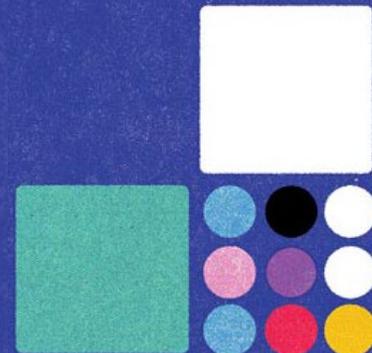
Useful to depict a single path through an experience, with minimal branching possibilities.

## TRIPLE FLOW

Depicts a few branching possibilities and makes it clearer when each modality updates.

## SWIM LANES

Best for situations where input and output may change separately, or fluidity and complexity are high.



# Example simple flow for a single intent

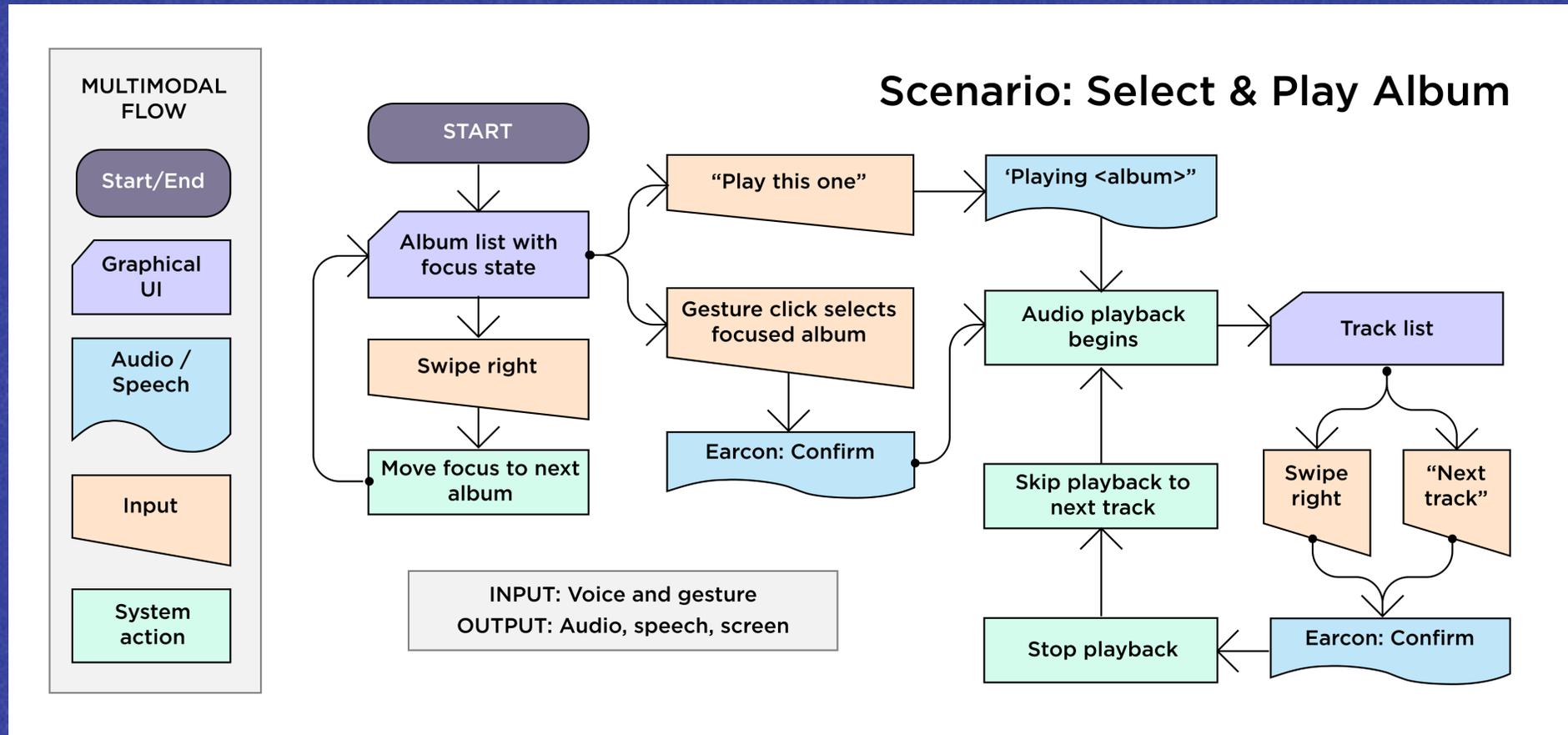
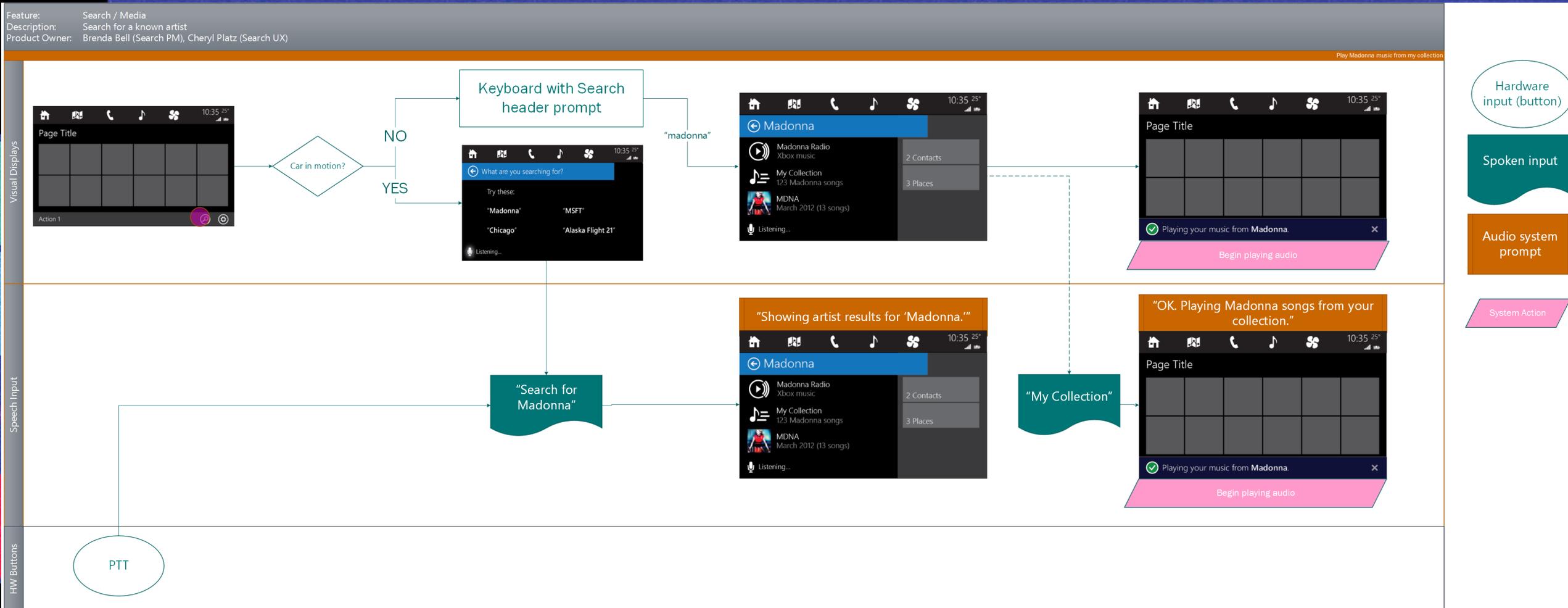


Figure 12.1 from *Design Beyond Devices: Creating Multimodal, Cross-Device Experiences*

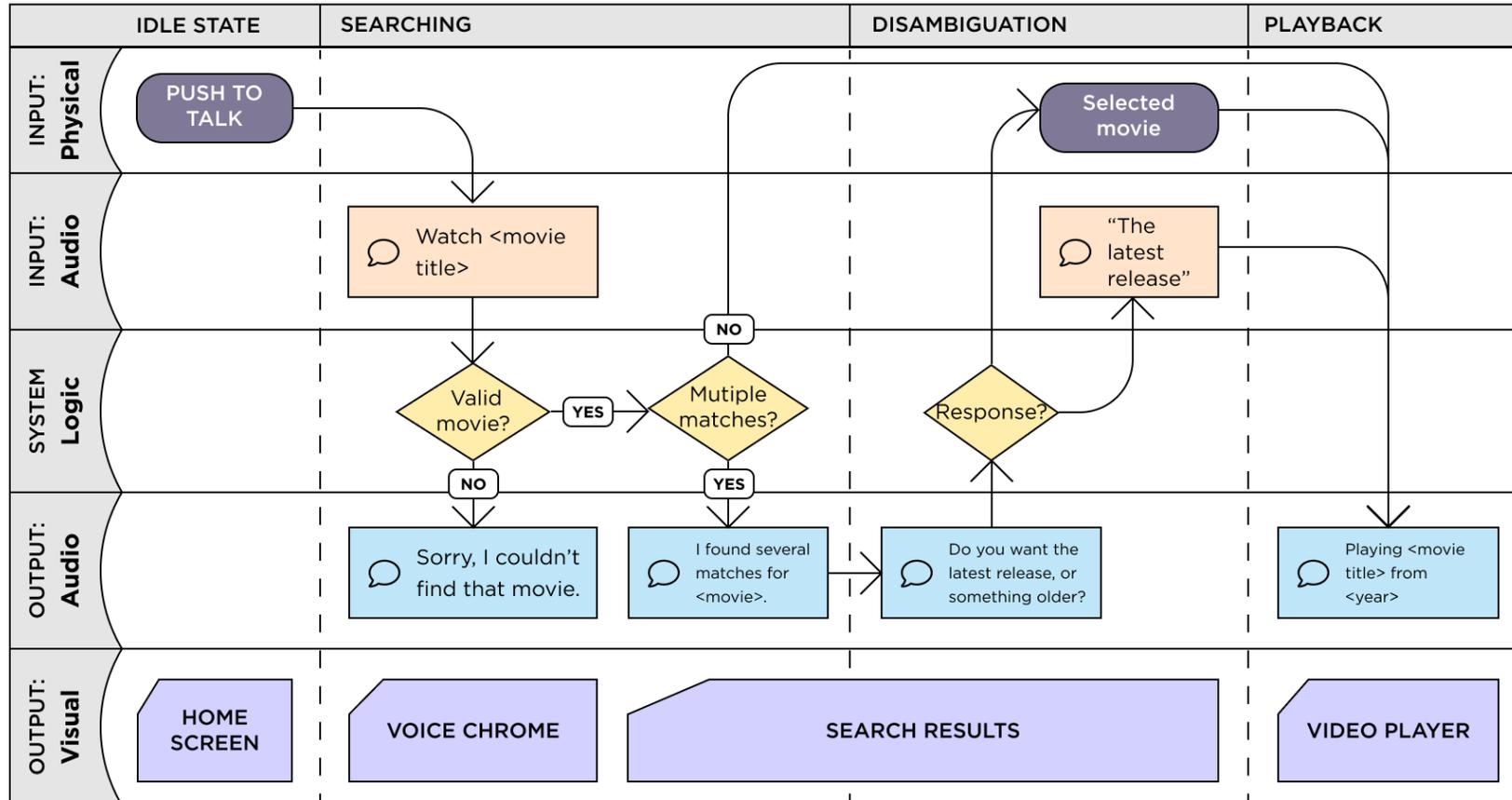
# Windows Automotive “triple flow”



# Multimodal swim lanes: Movie search

## STREAMING MEDIA DEVICE WITH VOICE-ENABLED REMOTE

INPUT: Remote, Voice  
OUTPUT: Audio, Screen



Fully multimodal swim lanes offer dedicated input and output channels per modality.

Figure 12.4 from *Design Beyond Devices: Creating Multimodal, Cross-Device Experiences*



# Multimodal flows: Example patterns

There are no industry standards, but this shorthand has served me well at multiple companies. Adapt and change as needed. Colors are optional.

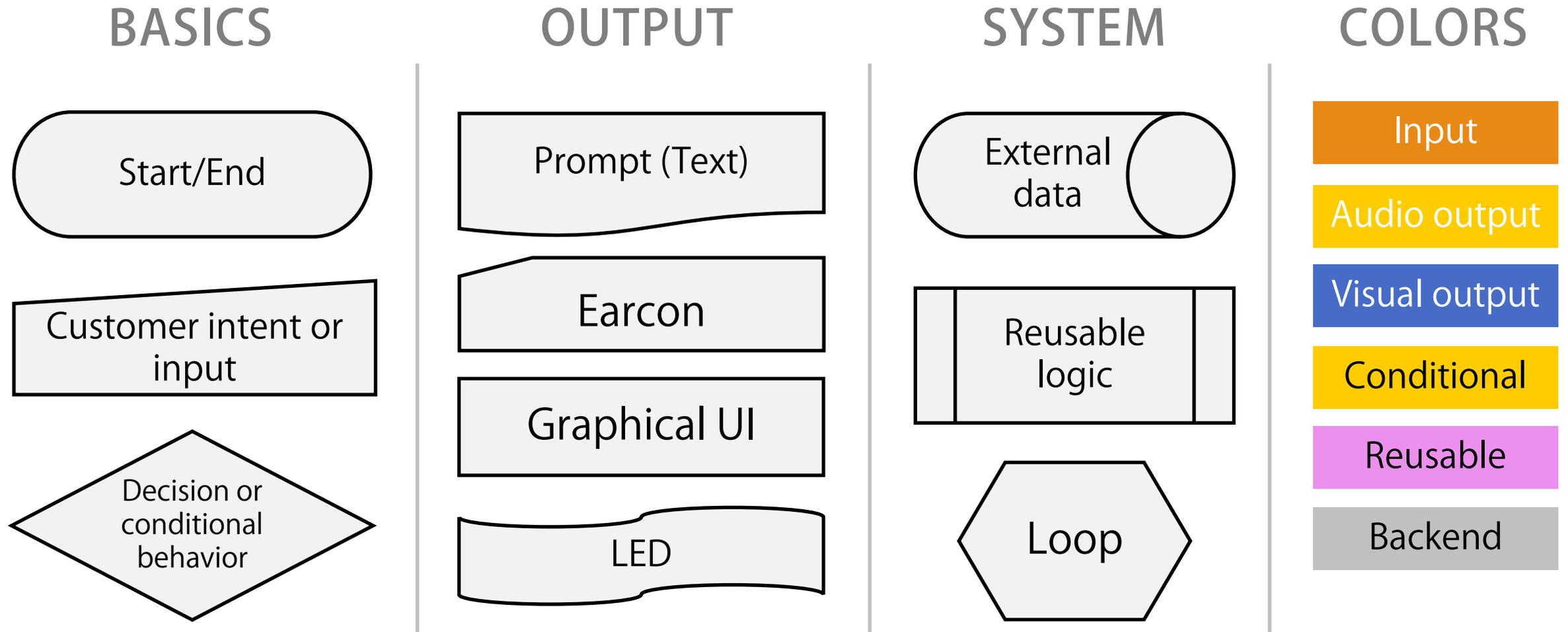


Figure 12.2 from [Design Beyond Devices: Creating Multimodal, Cross-Device Experiences](#)  
@IDEAPLATZ

# Coping with the complexity

You're still going to need your full voice UI designs and your full visual UI designs.

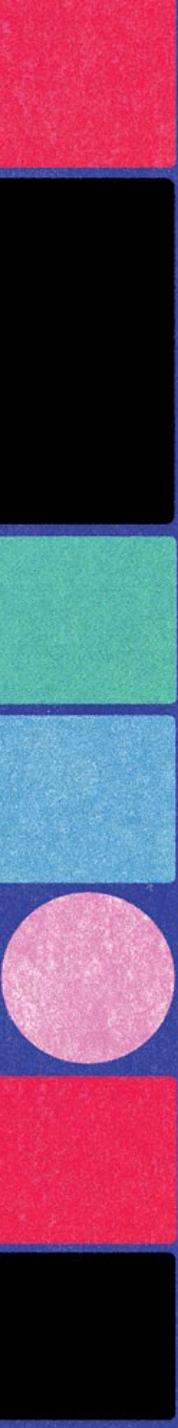
These swim lanes depict the ways customers may move between modalities in systems that support those transitions.

Add these flows as an additional deliverable.

## FOCUS ON FLOWS WITH THESE ELEMENTS:

- Customer-driven transitions between modalities
- System-driven transitions between modalities
- Handoffs to other devices
- System state changes
- Events





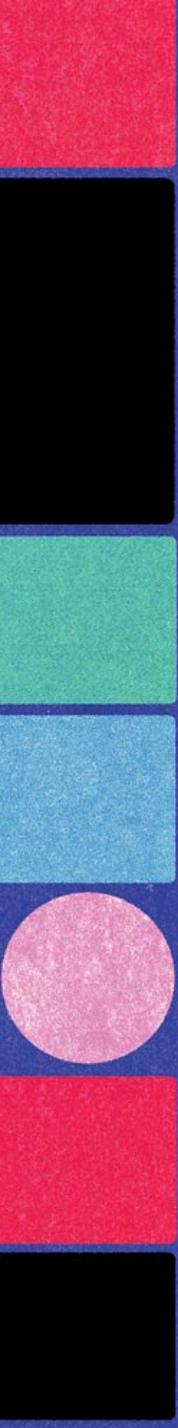
# Keep in mind: intangible and adaptive experiences should always allow voice completion of tasks.

Visuals can be additive, but never REQUIRE visual interaction in these two quadrants.



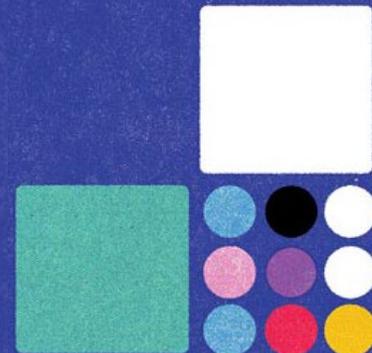
# CLOSING





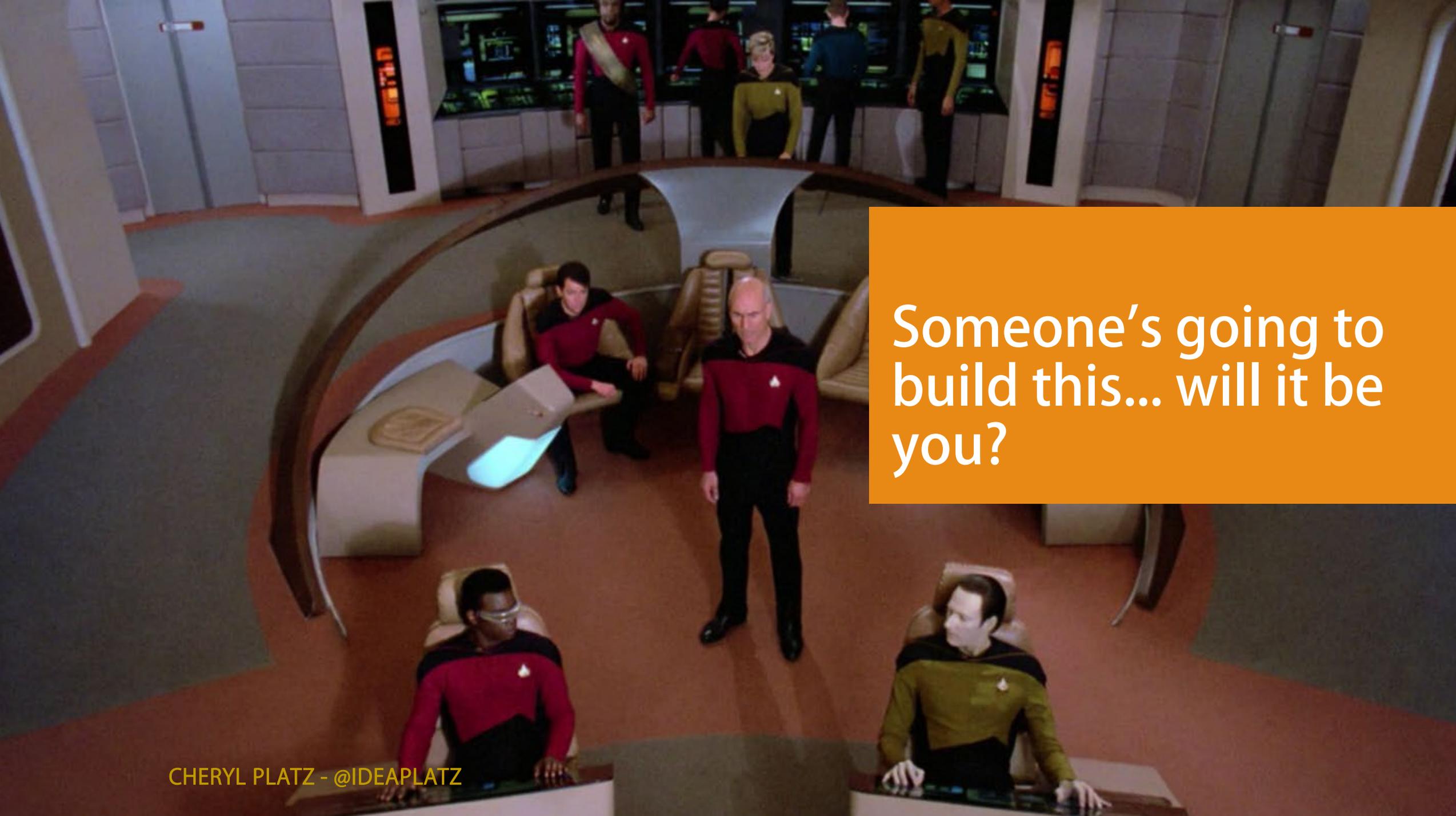
TO EXTEND YOUR  
VOICE DESIGNS FOR  
MULTIMODALITY,  
CONSIDER:

- What input and output modalities will you support?
- Where does your experience fall on the spectrum of multimodality?
- What transitions will be key in your experience?
- How will your chosen interaction model impact your design?
- Which flows are most critical to document from a multimodal perspective?



As I tell folks outside the software industry, I hope this book will be the design manual for folks who want to design the bridge of the Starship Enterprise.





Someone's going to build this... will it be you?

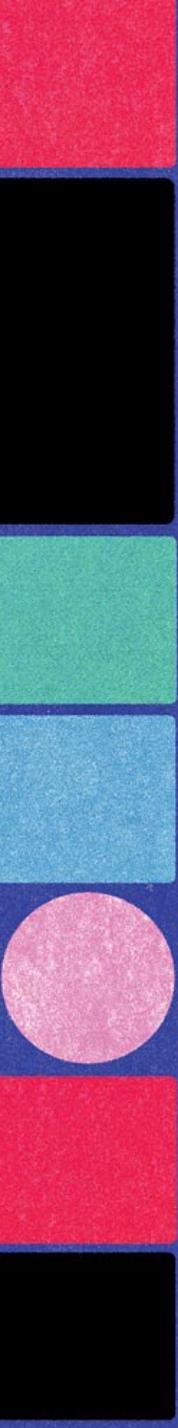
# DESIGN BEYOND DEVICES:

CREATING MULTIMODAL,  
CROSS-DEVICE EXPERIENCES

Available from  
[rosenfeldmedia.com](http://rosenfeldmedia.com) or Amazon.

All paper books on the Rosenfeld  
site come with the e-book  
version!

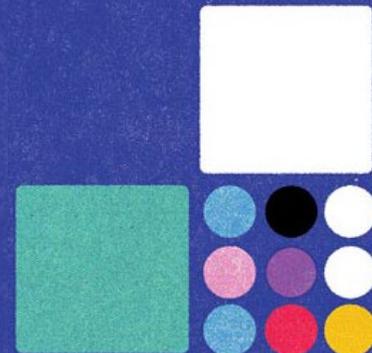




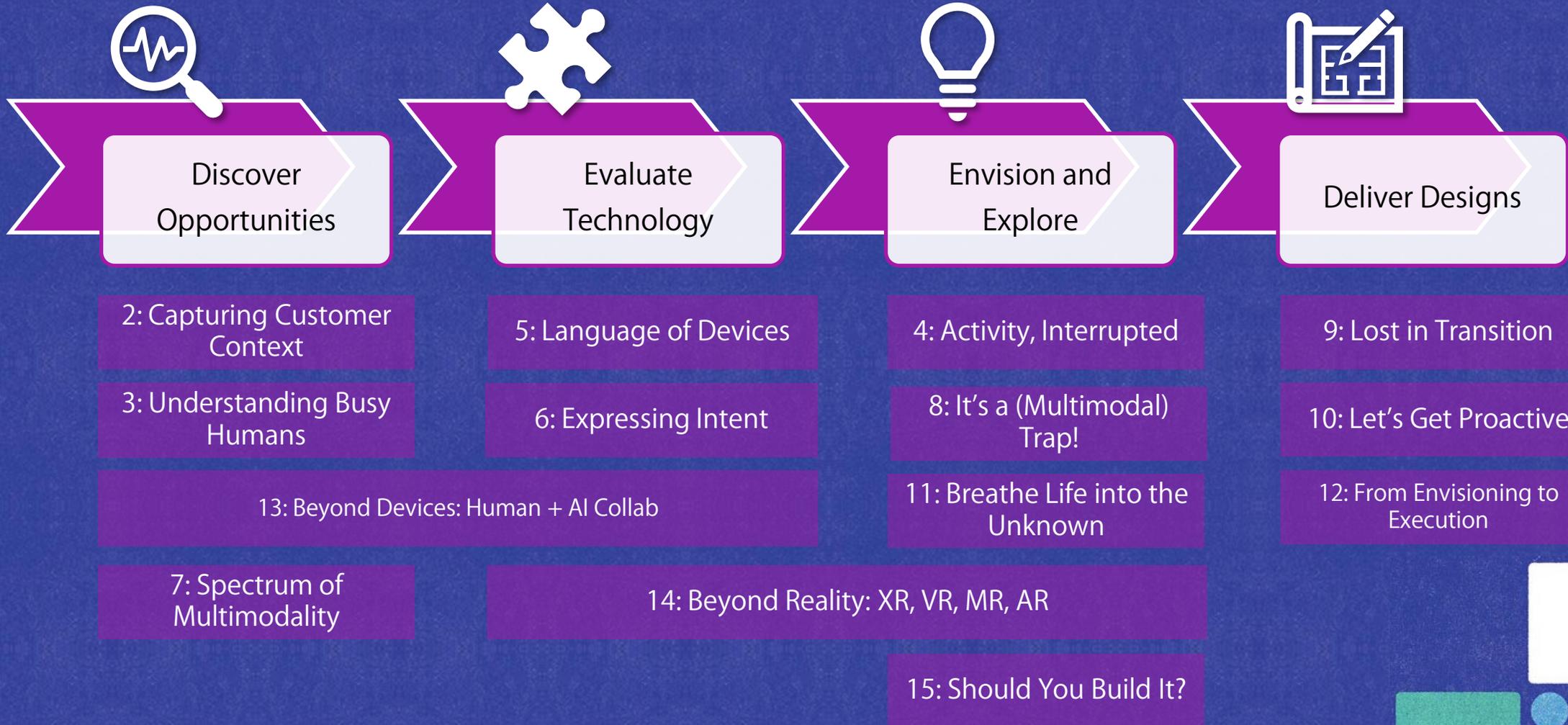
## THE BOOK CAN BE DIVIDED INTO FOUR THEMES.

Each chapter fits into one or two of these core themes, each a critical piece of the puzzle you'll need to complete to become a responsible, resilient multimodal designer.

1. Customer context & ethics
2. Multimodal frameworks
3. Ideation and Execution
4. Emerging technology



# Putting it all together



# DOWNLOAD THE REFERENCE MATERIALS:

- Multimodal reference sheet
- Flow shapes one-pager
- Blank swim lane worksheet

**Ideaplatz** Multimodal Voice Design Quick Reference Guide

**The Spectrum of Multimodality**

**RICH INFORMATION**

<b>QUADRANT 2 Anchored</b> Experiences with rich physical presence where a customer is usually nearby. <i>Fire TV, Xbox One, Cortana on PC</i>	<b>QUADRANT 1 Adaptive</b> Experiences that support both close proximity and long-range interactions. <i>Echo Show, Facebook Portal, Google Nest Hub</i>
<b>QUADRANT 3 Direct</b> Customer and device must be in direct contact or extreme proximity for use. <i>Fitbit, Google Glass, HoloLens, Apple Watch</i>	<b>QUADRANT 4 Intangible</b> Hands-free experience where close proximity to the device is not required. <i>Echo (original), Google Home</i>

**SCOPED INFORMATION**

**Common Transitions**

<b>BETWEEN Modalities</b> 1. Input transitions 2. Output transitions 3. Input/Output Mismatch	<b>BETWEEN Connections</b> Connection loss Unstable connection Proximity loss Insufficient connection	<b>BETWEEN Devices</b> Environmental Device Preference Device Suitability
--	---	--

© 2021 Ideaplatz, LLC | @ideaplatz | http://ideaplatz.com

**Ideaplatz** Multimodal Voice Design Quick Reference Guide

**Flows**

- Simple Flow**  
Use for a single path through an experience, with minimal branching possibilities.
- Triple Flow**  
Depicts a few branching possibilities and makes it clearer when each modality updates.
- Swim Lanes**  
Best for situations where fluidity, implementation, or complexity is high.

**Multimodal flows**

There are no industry standards, but this shorthand has served me well at multiple companies. Adapt and change as needed. Colors are optional.

**BASICS** Start/End

**OUTPUT** System Prompt (Text to Speech), Sound or Music

**OTHER** External data, Reusable function

**COLORS** Input, Audio output, Visual output

**Ideaplatz** Multimodal swim lane worksheet

Break up phases with vertical lines

INPUT: Physical	
INPUT: Audio	
SYSTEM Logic	
OUTPUT: Audio	
OUTPUT: Visual	

**Multimodal flows: Example pattern**

**BASICS** input

**OUTPUT** Graphical UI, LED

Decision or conditional behavior

logic Loop

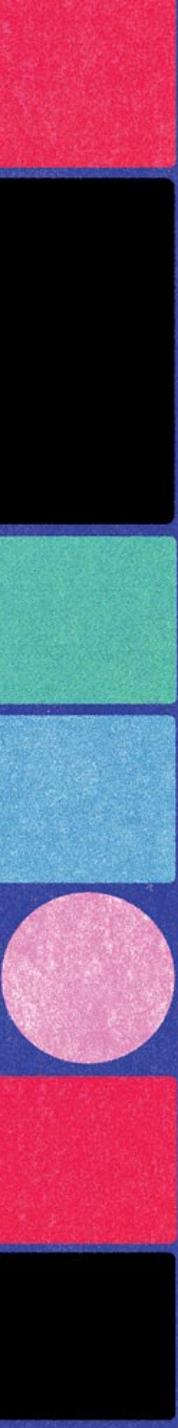
Conditional, Reusable, Backend

Figure 12.2 from *Design Beyond Devices: Creating Multimodal, Cross-Device Experiences*

@IDEAPLATZ

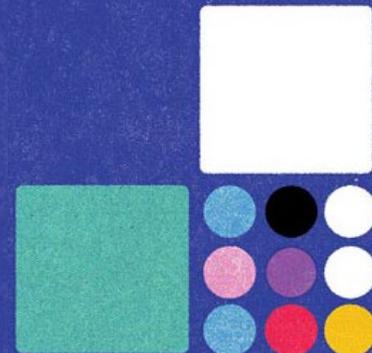
<http://bit.ly/ideaplatz-designbeyonddevices-downloads>





Virtual book signing:  
I'll mail a signed bookplate sticker  
to any Voiceflow attendee!

EMAIL YOUR ADDRESS TO  
[CHERYL@IDEAPLATZ.COM](mailto:CHERYL@IDEAPLATZ.COM).



FOR MORE DESIGN  
EDUCATION  
GOODNESS,  
CHECK OUT  
IDEAPLATZ.COM.

Talks, workshops, articles,  
podcast appearances, and  
downloads.



## HOME

**Ideaplatz: the road to great design.** We offer design inspiration that empowers today's designers to build tomorrow's experiences. Founder Cheryl Platz is an internationally recognized designer, author, speaker, and instructor. Our services are available for conferences, corporate training, and individual coaching.

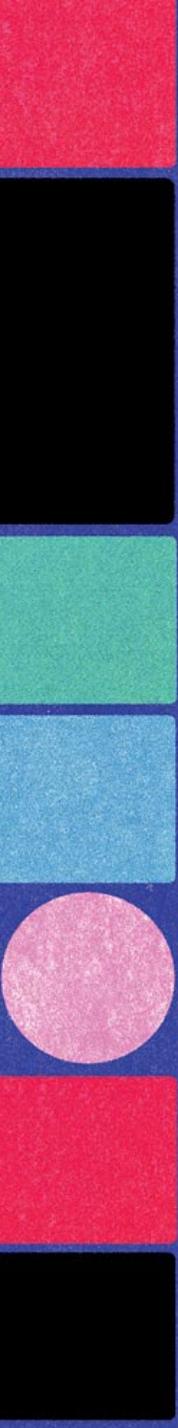
Cheryl is the author of *Design Beyond Devices: Creating Multimodal, Cross-Design Experiences*, published by Rosenfeld Media in December 2020. She is touring (virtually) in support of this release, including **three talks and two new workshops**:

- Opti-Pessimism: Design, AI, and Our Uncertain Future
- Design Beyond Websites: What is Dead May Never Die
- Capturing Customer Context for Cross-Channel Experiences (talk)
- **[NEW for 2021]** Capturing Customer Context (workshop)
  - This workshop version of the content from the book and talk will give you the tools to run your own shared understanding workshop with stakeholders to kick off UX research.
- **[NEW for 2021]** Design Beyond Devices: Multimodal Design
  - A hands-on workshop applying the frameworks introduced in Cheryl's new book.

Previously established talks and workshops are also available. To inquire about bookings, visit our contact form today.

Until COVID-19 vaccines are widely available, we are accepting remote bookings only. Ideaplatz is a proud member of the MURAL Consultants Network, and most of our workshops have been updated for remote delivery and interactivity.





# May the voice be with you!

**CHERYL PLATZ**

Principal Designer & Owner, Ideaplatz

<http://ideaplatz.com>

[cheryl@ideaplatz.com](mailto:cheryl@ideaplatz.com)

[@ideaplatz](#) (Twitter, Medium)

Ideaplatz 

