Name: _____

Date: _____   Section: _____

Google's Algorithm Description
Reading Guide

Directions: Read this to understand features of Google's **algorithm**. In other words, what **rules** does Google follow to decide how to order search results?

https://www.google.com/search/howsearchworks/how-search-works/ranking-results/

How results are automatically generated
With the vast amount of information available, finding what you need would be nearly impossible without some help sorting through it. Google's **ranking systems** are designed to do just that: sort through hundreds of billions of webpages and other content…to present the most relevant, useful results in a fraction of a second.

Key factors in your results
To give you the most useful information, **Search algorithms** look at many factors and signals, including the words of your query, relevance and usability of pages, expertise of sources, and your location and settings. The **weight** applied to each factor varies depending on the nature of your query. For example, the freshness of the content plays a bigger role in answering queries about current news topics than it does about dictionary definitions.

Meaning of your query
To return relevant results, we first need to establish what you're looking for — the intent behind your query. To do this, we build language models to try to decipher how the relatively few words you enter into the search box match up to the most useful content available.
This involves steps as seemingly simple as recognizing and correcting spelling mistakes, and extends to trying to our sophisticated **synonym system** that allows us to find relevant documents even if they don't contain the exact words you used. For example, you might have searched for "change laptop brightness" but the manufacturer has written "adjust laptop brightness. Our systems understand the words and intend are related and so connect you with the right content.
….Our systems also try to understand what **type of information** you are looking for. If you used words in your query like "cooking" or "pictures," our systems figure out that showing recipes or images may best match your intent.

Relevance of content
Next, our systems analyze the content to assess whether it contains information that might be relevant to what you are looking for.
The most basic signal that information is relevant is when content contains the **same keywords** as your search query. For example, with webpages, if those keywords appear on the page, or if they appear in the headings or body of the text, the information might be more relevant.

Quality of content

After identifying relevant content, our systems aim to prioritize those that seem most helpful. To do this, they identify signals that can help determine which content demonstrates expertise, authoritativeness, and trustworthiness.

For example, one of several factors we use to help determine this is understanding if **other prominent websites link or refer to the content.** This has often proven to be a good sign that the information is well trusted. Aggregated feedback from our Search quality evaluation process is used to further refine how our systems discern the quality of information.

Usability of webpages

Our systems also consider the usability of content. When all things are relatively equal, content that people will find more accessible may perform better.

For example, our systems would look at page experience aspects, such as if content is mobile-friendly, so that those on mobile devices can easily view it. Similarly, they look to see if content loads quickly, also important to mobile users.

Context and settings

Information such **as your location, past Search history, and Search settings** all help us to ensure your results are what is most useful and relevant for you in that moment.

We use your country and location to deliver content relevant for your area. For instance, if you're in Chicago and you search "football", Google will most likely show you results about American football and the Chicago Bears first. Whereas if you search "football" in London, Google will show results about soccer and the Premier League.

Search also includes some features that personalize results based on the activity in your Google account. For example, if you search for "events near me" Google may tailor some recommendations to event categories we think you may be interested in.

These systems are designed to match your interests, but they are not designed to **infer sensitive characteristics** like your race, religion, or political party.

You can control what Search activity is used to improve your Search experience, including adjusting what data is saved to your Google account, at myaccount.google.com.

Name: _____

Date: _____   Section: _____

Google's Algorithm Post-Reading Questions:

Directions: Explain each factor used in Google's search algorithm, in your own words. Then, imagine you Googled: "how to teach your dog to fetch." How would each algorithm factor apply to or influence the ranking of your search results?

Factor	Explain in Your Own Words	Example of How Each Factor Might Apply to Your Search Results
1. Meaning of your query		*E.g., Google may change the word "teach," to "train." Google may search for, "how to" guides.*
2. Relevance of Content		
3. Quality of Content		
4. Usability of webpages		
5. Context and settings		

Name: _____

Date: _____   Section: _____

YouTube's Algorithm Description
Reading Guide

Directions: Read this to understand features of YouTube's **algorithm**. In other words, what **rules** does YouTube follow when recommending videos to you?

https://blog.youtube/inside-youtube/on-youtubes-recommendation-system/

When YouTube's recommendations are at their best, they connect billions of people around the world to content that uniquely inspires, teaches, and entertains. For me, that means diving into lectures exploring the ethical questions facing technology today or watching highlights from the University of Southern California football games I remember seeing as a kid…

…There's an audience for almost every video, and the job of our recommendation system is to find that audience. **Think about how hard it would be to navigate all of the books in a massive library without the help of librarians.** Recommendations drive a significant amount of the overall viewership on YouTube, even more than channel subscriptions or search.

What is a recommendation system?
Our recommendation system is built on the simple principle of **helping people find the videos they want to watch and that will give them value.** You can find recommendations at work in two main places: your homepage and the "Up Next" panel. Your homepage is what you see when you first open YouTube—it displays a mixture of personalized recommendations, subscriptions, and the latest news and information. The Up Next panel appears when you're watching a video and suggests additional content based on what you're currently watching, alongside other videos that we think you may be interested in.

To do this, we start with the knowledge that everyone has unique viewing habits. Our system then compares your viewing habits with **those that are similar to you** and uses that information to suggest other content you may want to watch. So if you like tennis videos and our system notices that others who like the same tennis videos as you also enjoy jazz videos, you may be recommended jazz videos, even if you've never watched a single one before. Without recommendations, [you] would never have known these videos were available. Unlike other platforms, we don't connect viewers to content through their social network. Instead, the success of YouTube's recommendations depends on accurately predicting the videos you want to watch.
But of course, we also know not everyone wants to always share this information with us. So we've built controls that help you decide how much data you want to provide. You can pause, edit, or delete your YouTube search and watch history whenever you want.

How we personalize recommendations
 A number of signals build on each other to help inform our system about what you find **satisfying**: clicks, watchtime, survey responses, sharing, likes, and dislikes.

Clicks: Clicking on a video provides a strong indication that you will also find it satisfying. After all, you wouldn't click on something you don't want to watch.

Watchtime: Your watchtime—which videos you watched and for how long—provides personalized signals to our system about what you most likely want to watch. So if our tennis fan watched 20 minutes of Wimbledon highlight clips, and only a few seconds of match analysis video, we can safely assume they found watching those highlights more valuable.

Survey Responses: To really make sure viewers are satisfied with the content they're watching, we measure what we call **"valued watchtime"**—the time spent watching a video that you consider valuable. We measure valued watchtime through user surveys that ask you to rate the video you watched from one to five stars, giving us a metric to determine how satisfying you found the content. If you rate a video one to two stars, we ask why you gave such a low rating. Similarly, if you give the video four to five stars, we ask why—was it inspirational or meaningful? Only videos that you rate highly with four or five stars are counted as valued watchtime.

Sharing, Likes, Dislikes: On average, people are more likely to be satisfied by videos that they share or like. Our system uses this information to try to predict the likelihood that you will share or like further videos. If you dislike a video, that's a signal that it probably wasn't something you enjoyed watching.

A Focus on Reliable Recommendations

The rise of misinformation in recent years led us to further expand the ways we use our recommendation system to include problematic misinformation and borderline content—that is content that comes close to, but doesn't quite violate our Community Guidelines. This includes **conspiracy theory videos** ("the moon landing was faked") or other content that spreads **misinformation** ("orange juice can cure cancer").

We're able to do this by using **classifiers** to identify whether a video is "authoritative" or "borderline". These classifications rely on **human evaluators** who assess the quality of information in each channel or video. These evaluators hail from around the world and are trained through a set of detailed, publicly available rating guidelines. We also rely on **certified experts**, such as medical doctors when content involves health information.

To determine authoritativeness, evaluators answer a few key questions. Does the content deliver on its promise or achieve its goal? What kind of expertise is needed to achieve the video goal? What's the reputation of the speaker in the video and the channel it's on? What's the main topic of the video (eg. News, Sports, History, Science, etc)? Is the content primarily meant to be satire? These answers and more determine how **authoritative** a video is. The higher the score, the more the video is promoted when it comes to news and information content.

Name: _____

Date: _____   Section: _____

Directions: Explain each factor used in YouTube's recommendation algorithm, in your own words. Then, imagine you watched a YouTube video of how to train your dog to fetch. How would each algorithm factor influence the next video recommended to you?

Factor	Explain in Your Own Words	Example of How Each Factor Might Influence the Next Video Recommended to you?
1. Compares viewing habits to those similar to you		*E.g., Google may recommend funny videos because these are videos watched by others, who also watched dog training videos*
2. Videos you find **satisfying**		
3. Watch time		
4. Valued Watch Time		
5. Video Authoritativeness		

Name: _____

Date: _____   Section: _____

Directions: Read this to understand features of Instagram's **algorithm**. In other words, what **rules** does Instagram follow when deciding how to order posts in your feed?

https://about.instagram.com/blog/announcements/shedding-more-light-on-how-instagram-works

---

Shedding More Light on How Instagram Works

What is "the algorithm"?
One of the main misconceptions we want to clear up is the existence of "The Algorithm."

Instagram doesn't have one algorithm that oversees what people do and don't see on the app. We use a variety of algorithms, classifiers, and processes, each with its own purpose. We want to make the most of your time, and we believe that using technology to personalize your experience is the best way to do that.

When we first launched in 2010, Instagram was a single stream of photos in **chronological order**. But as more people joined and more was shared, it became impossible for most people to see everything, let alone all the posts they cared about. By 2016, people were missing 70% of all their posts in Feed, including almost half of posts from their close connections. So we developed and introduced a Feed that ranked posts based on what you care about most.

How we rank Feed and Stories
Over the years we've learned that Feed and Stories are places where people want to see content from their friends, family, and those they are closest to. With any ranking algorithm, how it works can be broken down into steps.

We start by defining the set of things we plan to rank in the first place. With Feed and with Stories this is relatively simple; **it's all the recent posts shared by the people you follow.** There are a few exceptions, like ads, but the vast majority of what you see is shared by those you follow.

Next we take all the information we have about what was posted, the people who made those posts, and your preferences. We call these "signals", and there are thousands of them. They include everything from what time a post was shared to whether you're using a phone or the web to how often you like videos. The most important signals across Feed and Stories, roughly in order of importance, are:

- **Information about the post.** These are signals both about how popular a post is – think how many people have liked it – and more mundane information about the content itself, like when it was posted, how long it is if it's a video, and what location, if any, was attached to it.

- **Information about the person who posted.** This helps us get a sense for how interesting the person might be to you, and includes signals like how many times people have interacted with that person in the past few weeks.
- **Your activity.** This helps us understand what you might be interested in and includes signals such as how many posts you've liked.
- **Your history of interacting with someone.** This gives us a sense of how interested you are generally in seeing posts from a particular person. An example is whether or not you comment on each other's posts.

From there we make a set of **predictions**. These are educated guesses at how likely you are to interact with a post in different ways. There are roughly a dozen of these. In Feed, the five **interactions** we look at most closely are how likely you are to spend a few seconds on a post, comment on it, like it, reshare it, and tap on the profile photo. The more likely you are to take an action, and the more heavily we **weigh** that action, the higher up you'll see the post. We add and remove signals and predictions over time, working to get better at surfacing what you're interested in.

Another important case to call out is **misinformation**. If you post something that **third-party fact checkers** label as misinformation, we don't take it down, but we do apply a label and show the post lower in Feed and Stories. If you've posted misinformation multiple times, we may make all of your content harder to find.

Name: _____

Date: _____   Section: _____

Directions: Explain each factor used in Instagram's algorithm, in your own words. How does each algorithm factor influence the first post you are likely to see in your feed?

Factor	Explain in Your Own Words	Example of How Each Factor Might Influence the Next Post You Would See In Your Feed:
1. Information about the post		*E.g., if more people had liked the post*
2. Information about the person who posted		
3. Your activity		
4. Your history interacting with someone		
5. Misinformation		

Name: _____

Date: _____   Section: _____

TikTok's Algorithm Description
Viewing Guide

Directions: View this video this to understand features of TikTok's **algorithm**. In other words, what **rules** does TikTok follow when deciding which videos to recommend?

https://www.wsj.com/video/series/inside-tiktoks-highly-secretive-algorithm/investigation-how-tiktok-algorithm-figures-out-your-deepest-desires/6C0C2040-FF25-4827-8528-2BD6612E3796?

Pause Time	Question
1:06	What video features play a role in what information TikTok shows users? a. b. c. d.
1:23	*The video says, "through this one powerful signal TikTok learns your most secret interests and emptions."* What is this powerful signal?
1:59	What are two characteristics of the videos you're introduced when you first open the app? a. b.
3:52	What is algorithm transparency?
4:32	How does TikTok's algorithm compare to YouTube's?
6:25	Think about the example of the "sad" video. What may be a limitation or danger of TikTok recommending more and more "sad" videos to a user?
7:14	The video says, *"Whether or not it keeps showing you that type of video, depends on your response to it."* What does this statement mean?

9:21	What is "disruptive content"? Why does TikTok show viewers disruptive content?
10:12	The video says: "*The algorithm is detecting that this depressing content is useful to create engagement.*" What does **engagement** mean in this case? Explain this statement in your own words.
10:39	The video says: *"The algorithm is able to find the piece of content that you're **vulnerable** to."* What does the word vulnerable mean? Why is the word vulnerable used here?
11:44	What are two characteristics of **harmful** content: a. b.
12:44	The video says: *"We are training them and they are training us."* What does this statement mean? Please explain in your won words.

Name: _____

Date: _____

Section: _____

Algorithm Comparison

Directions: Classify each algorithm you learned about according to its function, goals, and other characteristics.

Algorithm Website	**Function:** What does the algorithm do?	What information about **you as a user** does the algorithm consider?	What information about **other people** does the algorithm consider?	How does the algorithm consider <u>information quality</u>?	**Goals:** What are the goals of the algorithm? How do you think the algorithm is improved or evaluated?
Google					
YouTube					
Instagram					
TikTok					

What are two patterns that you notice in the algorithms you classified:

(1)

(2)

Name: _____

Date: _____   Section: _____

Algorithm Comparison
Discussion Questions

Directions: Use your algorithm comparison chart to answer each questions below.

1. What is the function or the main job of algorithms on websites like Google, YouTube, and TikTok? Why is this job important?

2. Why do websites like Google, YouTube, and TikTok design algorithms for presenting content (e.g., instead of presenting content chronologically or in alphabetical order)? What are companies' goals in designing algorithms?

3. How do companies use the information about you (and other users) that they collect? Are there any drawbacks or dangers to how this information is used?

4. Why do companies provide information about their algorithms? In other words, why were the documents you read created? Why might companies not want to provide information about how their algorithms work?

Name: _____

Date: _____   Section: _____

Algorithm Evaluation

Directions: Use what you know about how algorithms work, to identify at least one benefit and at least one drawback of algorithms, for companies and for users.

	Benefits	Drawbacks
Companies (e.g., Google, TikTok)		
Individual Users		

Name: _____

Date: _____   Section: _____

Exit Ticket: Design an Algorithm
You have a website that recommends content (e.g., news stories, jokes, videos) to users. You want to build an <u>algorithm</u> to determine what content to recommend to users in what order. **What are three features that you would consider in developing your algorithm:** **(a)**   **(b)**   **(c)**
Which feature did you consider to be most important for your algorithm? Why?