



• *Using data to build-in Resiliency & Feedback to improve the business of video.*

• *Diane Strutner, CEO*

• Datazoom - The Video Data Infrastructure Platform

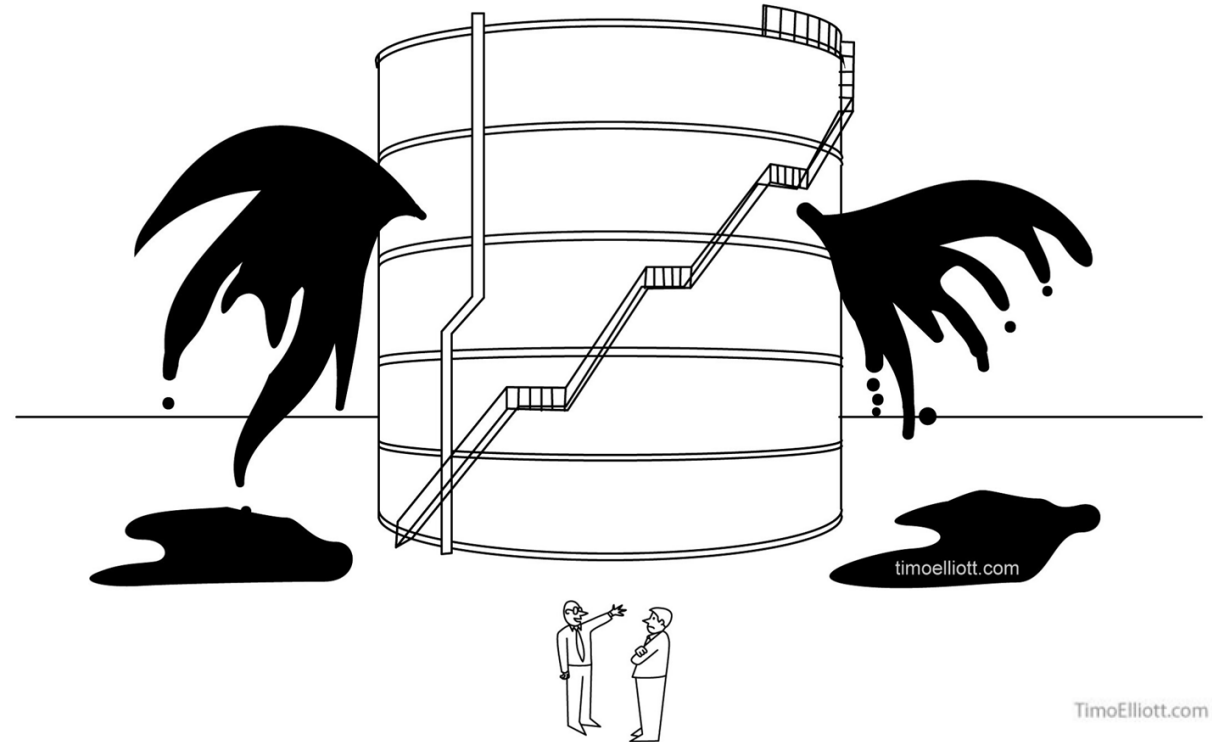
• [www.datazoom.io](http://www.datazoom.io)

“Organisations struggle to corral data into useable and actionable intelligence. **Without clean, relevant, and labeled data**, organisations are **stymied in their efforts to move aggressively on AI**, which CEOs overwhelmingly ‘agree’ will have a significant impact on their business **within the next five years.**”

PwC's 22nd Annual CEO Survey (1,378 CEOs in +90 territories)

## Problem

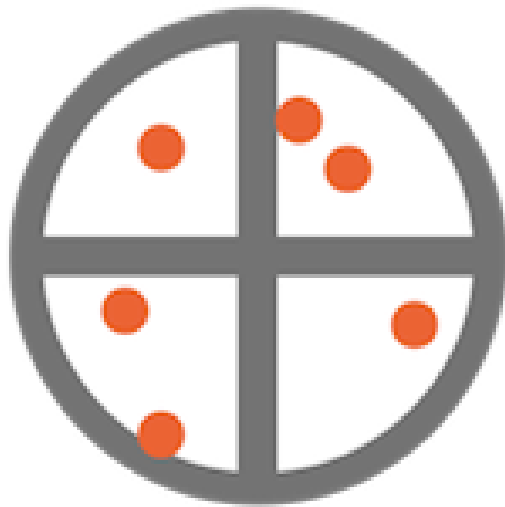
Video Data isn't **standardized, centralized, or organized**. No **easy way exists** for streaming companies, or the vendors and internet infrastructure that support streaming, to analyze and make adjustments to improve video performance.



*“Data is the new oil? Absolutely—**toxic if mishandled!**...”*

What ***data*** do we have today? How good is it?

# Data quality is measured by Accuracy & Precision



**IMPRECISE  
INACCURATE**



**PRECISE  
INACCURATE**



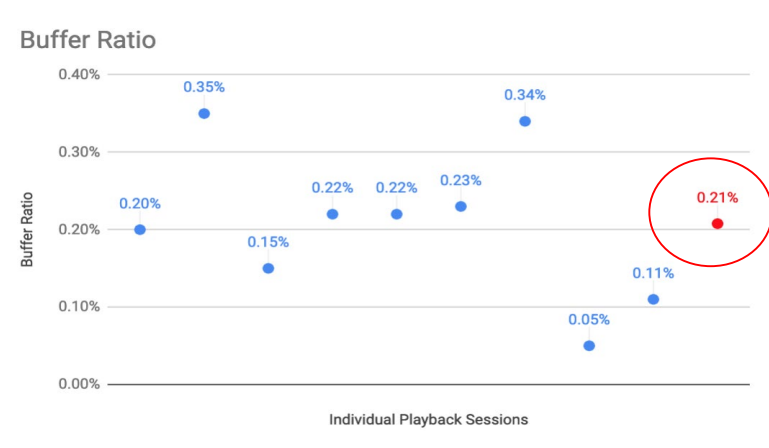
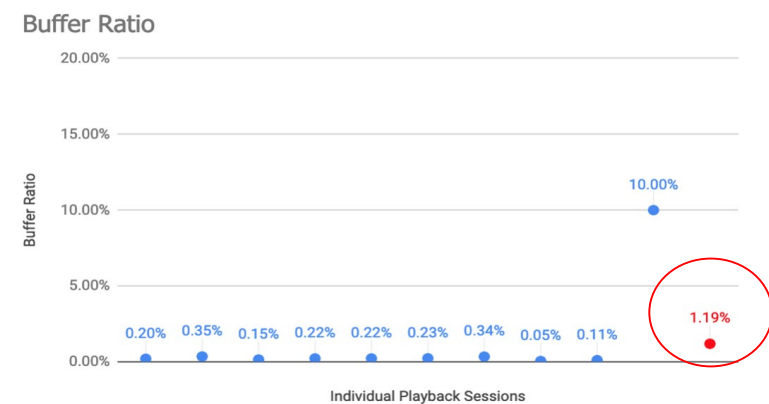
**IMPRECISE  
ACCURATE**



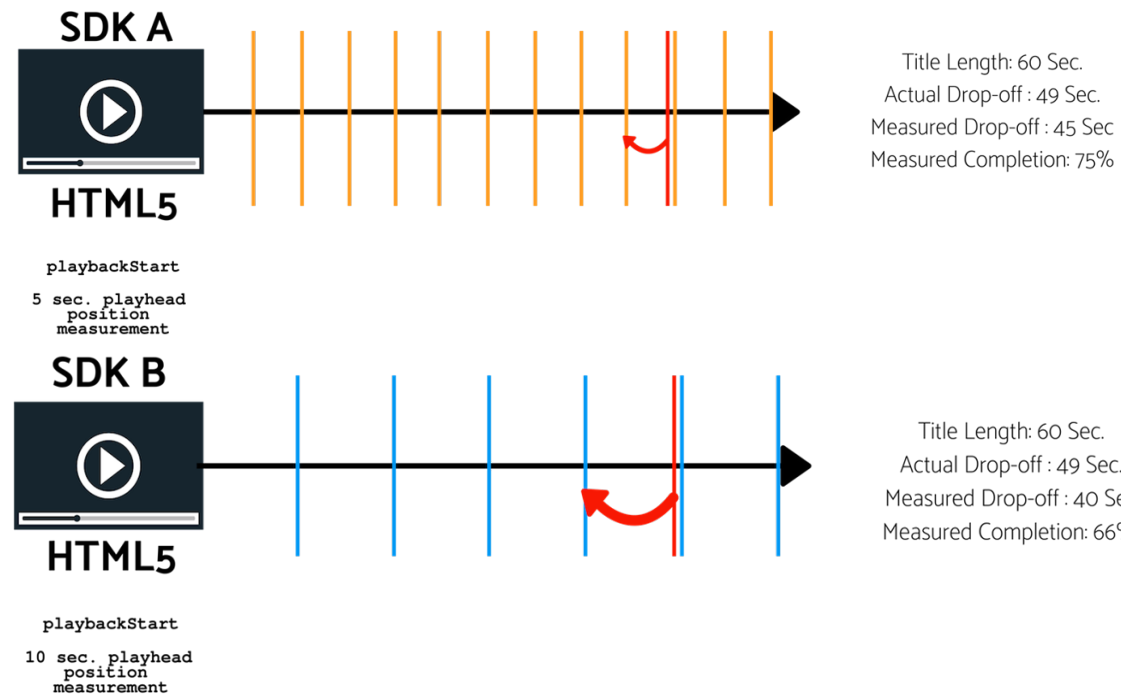
**PRECISE  
ACCURATE**

# Metrics Obfuscate Accuracy & Precision

Different data **calculations** are used.  
*Is buffer ratio 1.19% or .21%?*



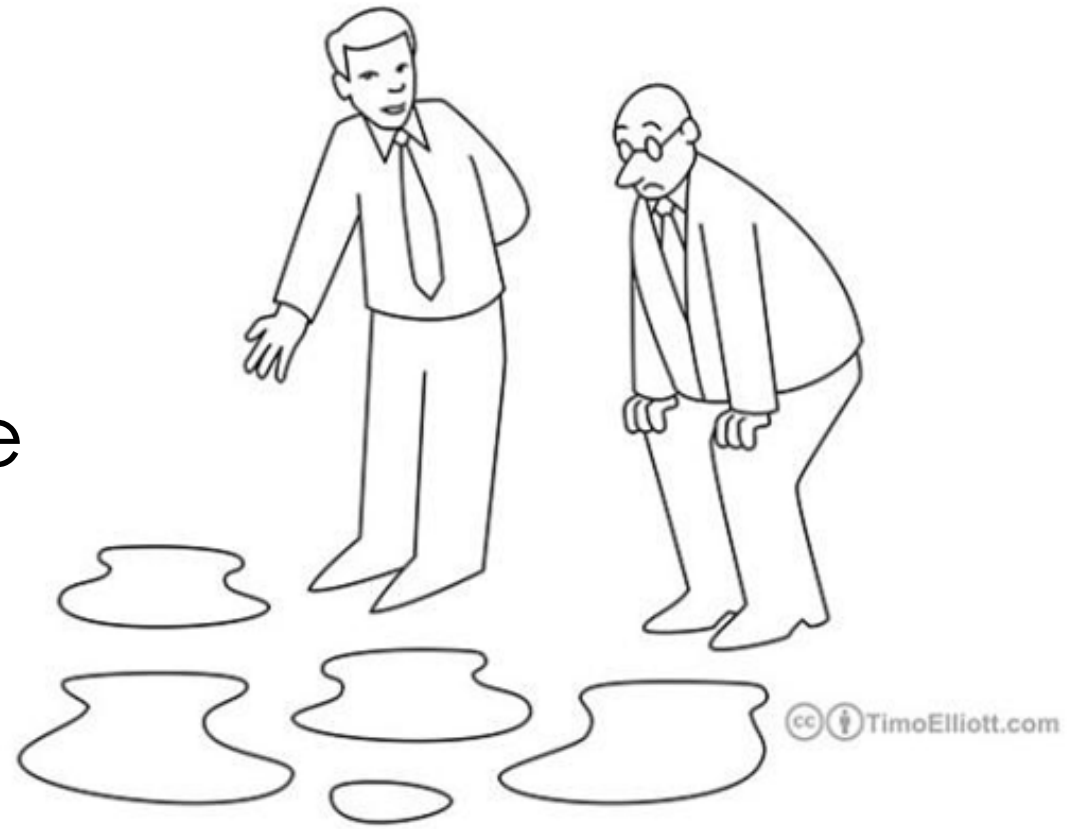
Different data **measurement** is taken.  
*Is completion rate 75% or 66%?*



Singular datapoints  
and measurements  
**don't tell the whole story.**



Understanding datapoints  
and measurements require  
**context.**



*“We call them our data puddles!”*



Placing video data into context means **enabling correlation**

E.g. If we detect a **buffering** problem, looking at only one data source most often will not identify **why** it has occurred.

*List of possible causes for buffering:*

Break in Feed  
Encoding Error  
Transcoding Error  
Packaging Error  
Origin Upload Error  
CDN Asset Error

CDN Congestion  
DNS Error  
DNS Congestion  
Transit Error  
Transit Congestion  
ISP Error

ISP Congestion  
Last Mile Error  
Last Mile Congestion  
Device Error  
Application Error  
Video Player Error

# Review of Current Data Challenges

## Embedded Subjectivity

Can we consistently reproduce our results? Subjectivity is carried into all downstream analysis & functions.

## Data Silos & Restricted Access

Do we have all relevant data in one place?

## Lack of Data Standardization

Is data prepared to be used AI and ML systems to use it?

## Slow Data

Do we have access to data in a relevant time-frame so that real-time Automation can have meaningful results?

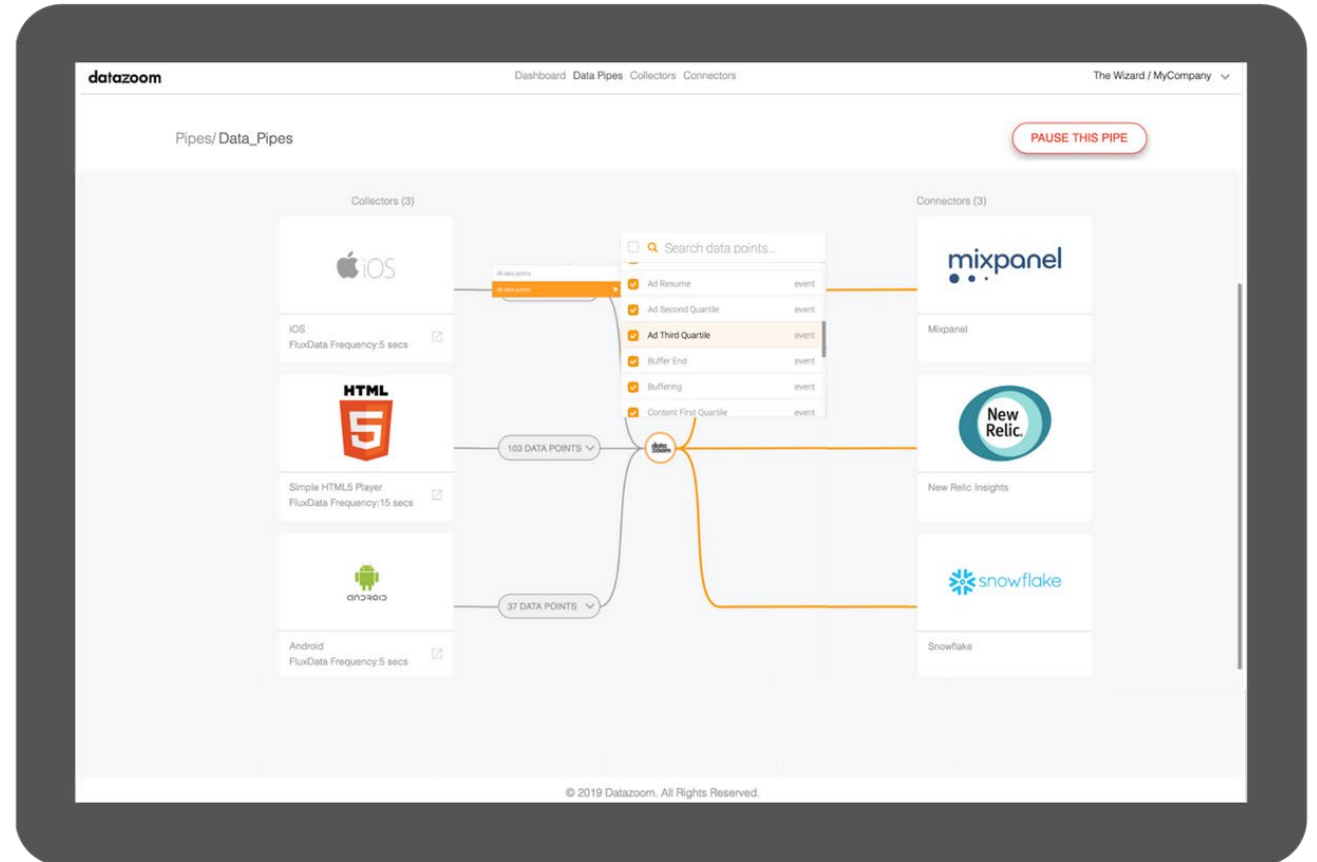
# Datazoom's Video Data Infrastructure Platform

**Capture.** Specify data to collect raw data using our SDKs/Libraries, or via 3rd party Logs/APIs.

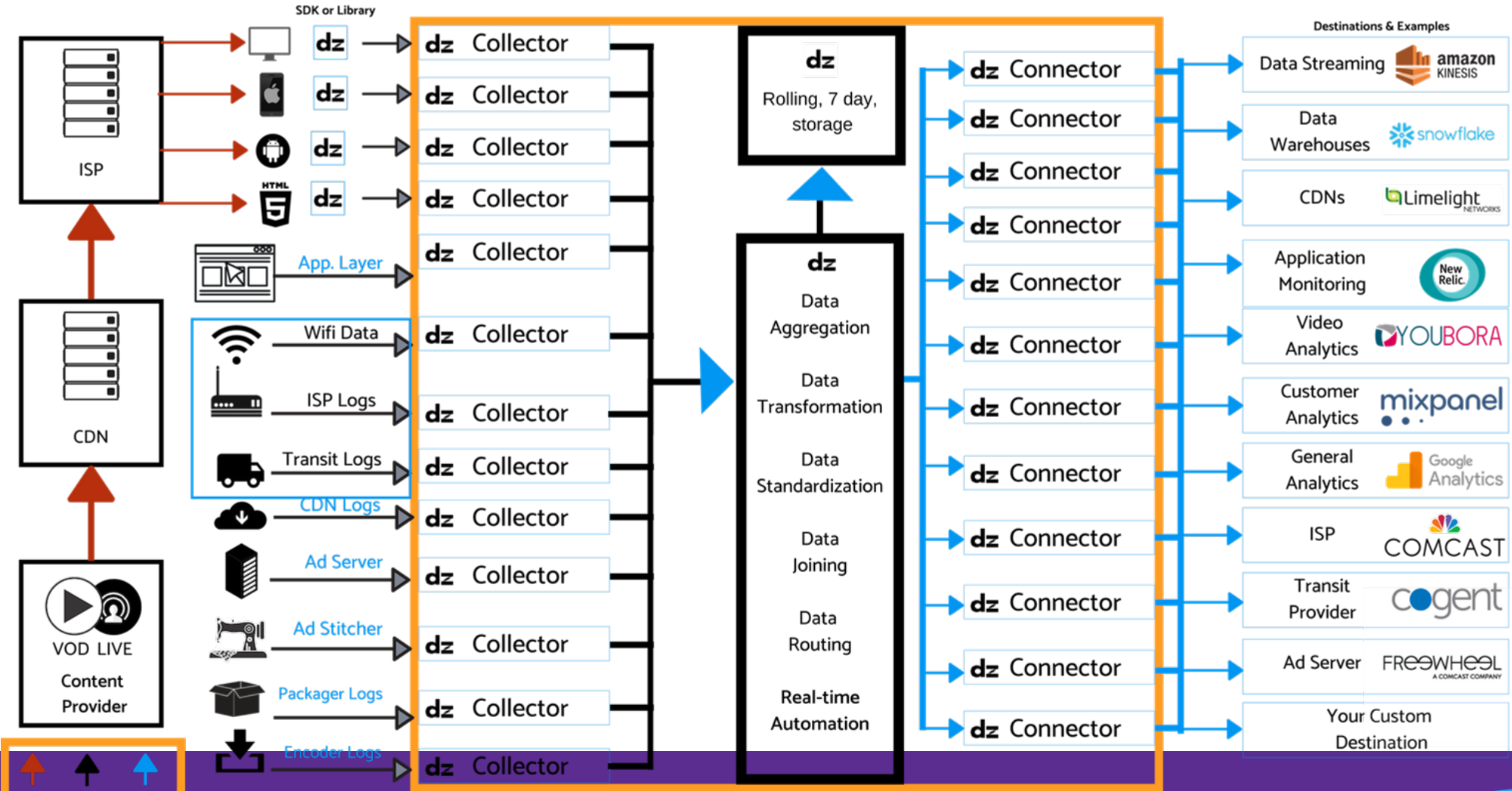
**Classify.** Normalize data according to our Data Dictionary.

**Connect.** Join disparate datasets and integrate data across the technology stack, with sub-second delay.

**Automate.** Rules engine to adjust and add resiliency for critical services embedded at the end-point.

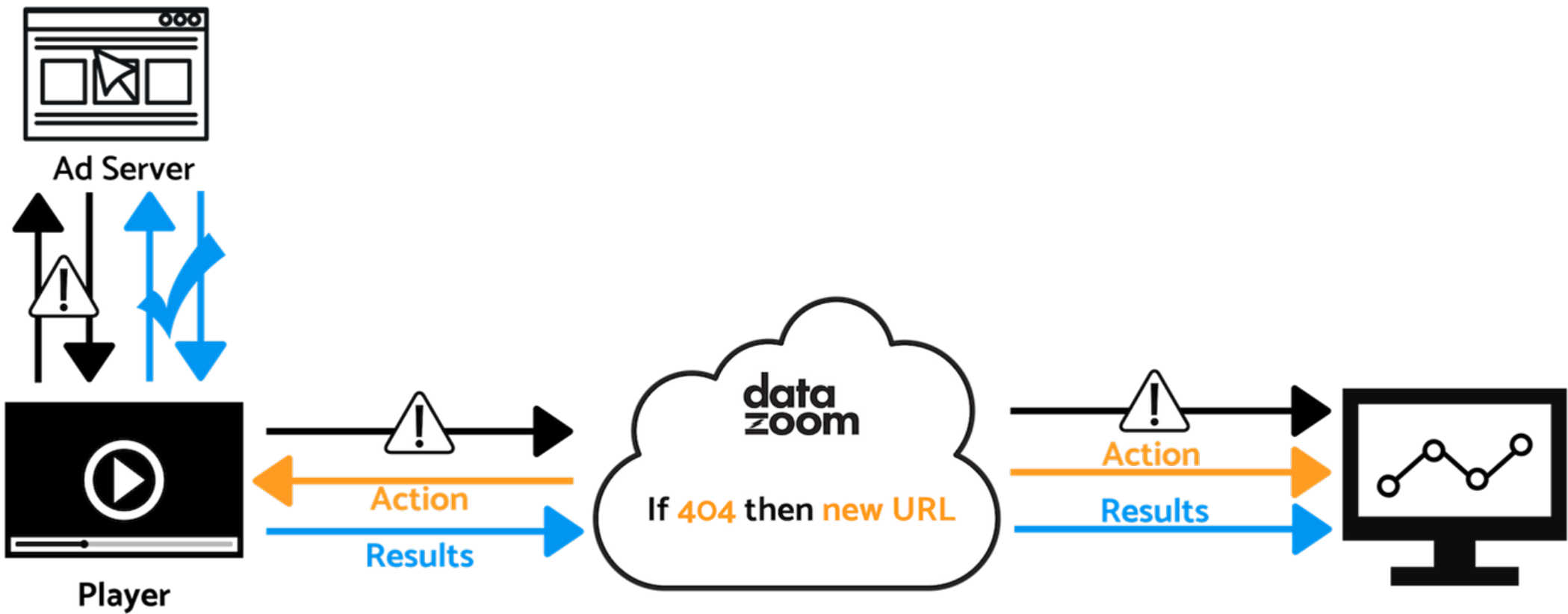


# Goal: holistic 'datatecture' to enable intelligent automation in video

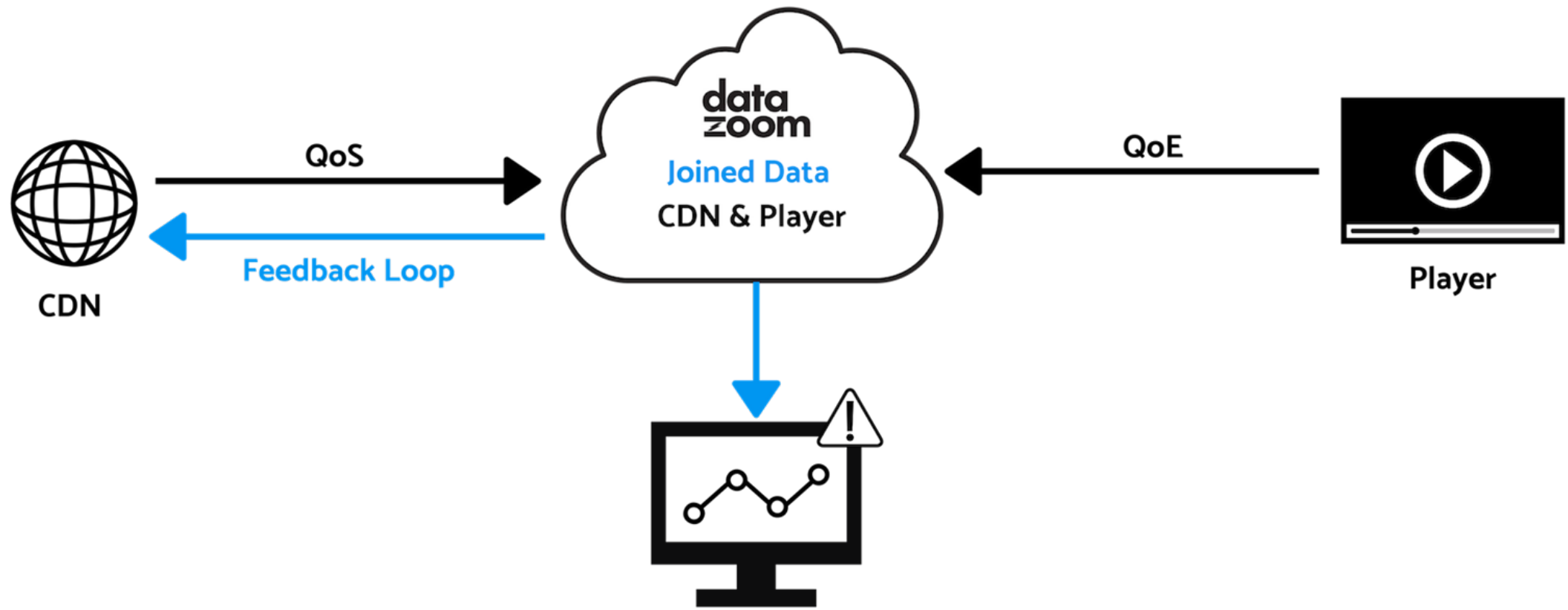


What things can we improve  
using ***intelligent automation***?

Failure recovery to build-in **Resiliency** for the things we can control.



Data sharing to build-in **Feedback** for the things we can't control.



Want a copy? email: [diane@datazoom.io](mailto:diane@datazoom.io)