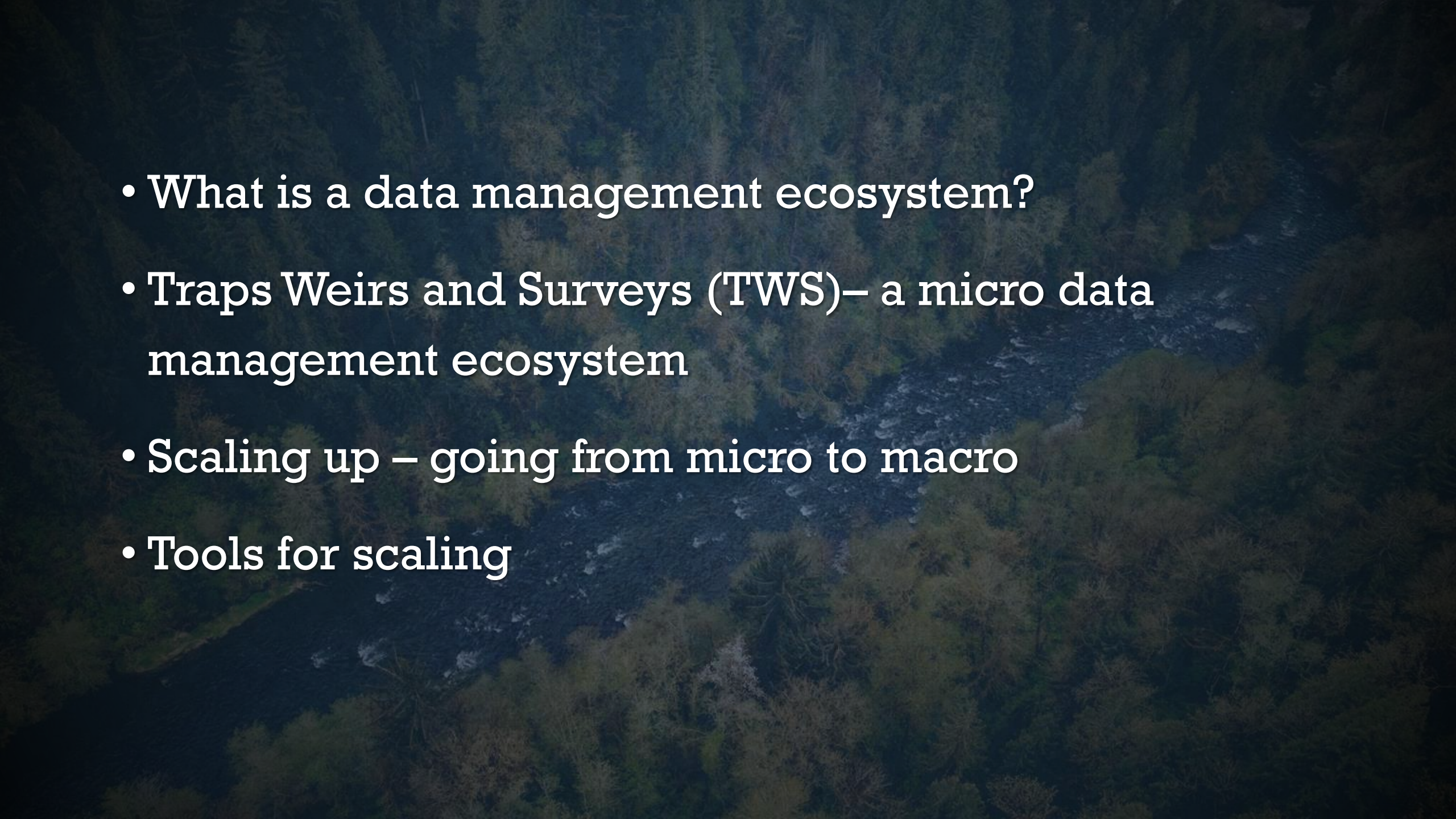


An aerial photograph of a river winding through a dense forest. The river is dark and turbulent, with white rapids visible. The surrounding forest is a mix of green and brown trees, suggesting a transition in seasons or different tree species. The overall tone is dark and naturalistic.

FROM MICRO TO MACRO

A look at scaling-up a data management ecosystem and some tools to make it happen

Benjamin Warren | Danny Warren | Jacob Shapley
WDFW Biological Data Systems Unit

- 
- An aerial photograph of a river winding through a dense, green forest. The river is dark and reflects the surrounding trees. The forest is thick with various types of trees, creating a textured green canopy.
- What is a data management ecosystem?
 - Traps Weirs and Surveys (TWS)– a micro data management ecosystem
 - Scaling up – going from micro to macro
 - Tools for scaling

What is a data management ecosystem?

- In general, the term is meant to describe:

A system where data are:

Collected
Transformed
Stored
Maintained
Retrieved



A Data Nightmare

Location (circle): Grays / Elochoman / Coweeman / Green / Washougal Date: 9/18/13 Page 1 of 2
 Start time: 7:35 Weather: Showers/Cloudy Samplers (Last name, First initial): [REDACTED]
 Trap Performance: Fishing Number of fish below trap: 200+ Gauge Height: _____ at _____ (time)
 LOP Shape: Square Tag Color: Pink
 Comments:

Line #	Quantity	Species	Tag (Maiden /Recap)	Tag #'s		Fork Length (cm)	Sex (M/F/J)	Fin Clip	Scale Card (1001-10)	I2IG DNA ID #	CWT (+/-)	Disposition	Comments (include weir wash-ups and trap morts)
				Tag 1	Tag 2								
1	1	FCH	Maid	2780	2781	72	M	AD	7875-1	292	~	Upstream	live sampled
2	1	FCH		2782	2783	73	M	AD	7875-2	293	~	Upstream	1
3	1	FCH		2784	2787	65	FM	AD	7875-3		~	Upstream	Flotag gun jammed, note fly to # numbers! Changed sex to Male
4	1	FCH		2788	2789	69	M	AD	7875-4		~	Upstream	live sampled
5	1	FCH		2790	2791	60	F	AD	7875-5		~	Upstream	
6	1	FCH		2792	2793	81	F	AD	7875-6		~	Upstream	
7	1	FCH		2794	2795	67	M	AD	7875-7		~	Upstream	
8	1	FCH		2796	2797	67	M	AD	7875-8		~	Upstream	
9	1	FCH		2798	2799	58	M F	AD	7875-9		~	Upstream	Changed Sex to Match S.S. Recap based on Conc. Card, ton F+P.
10	1	FCH		2910	2917	88	M	NM	7877-1	292	~	Upstream	
11	1	FCH		2915	2914	78	F	NM	7877-2	293	~	Upstream	
12	1	FCH		2913	2912	74	F	NM	7877-3	294	~	Upstream	
13	1	FCH		2851	2852	78	F	NM	7877-4	295	~	Upstream	2851, 2852 fly #
14	1	FCH		2853	2854	79	M	NM	7877-5	296	~	Upstream	
15	1	FCH		2858	2856	70	F	NM	7877-6	297	~	Upstream	
16	1	FCH		2859	28108	88	F	NM	7877-7	298	~	Upstream	
17	1	FCH		2811	2812	70	M	NM	7877-8	299	~	Upstream	
18	1	FCH	✓	2813	2814	70	M	NM	7877-9	300	~	Upstream	
19													
20													

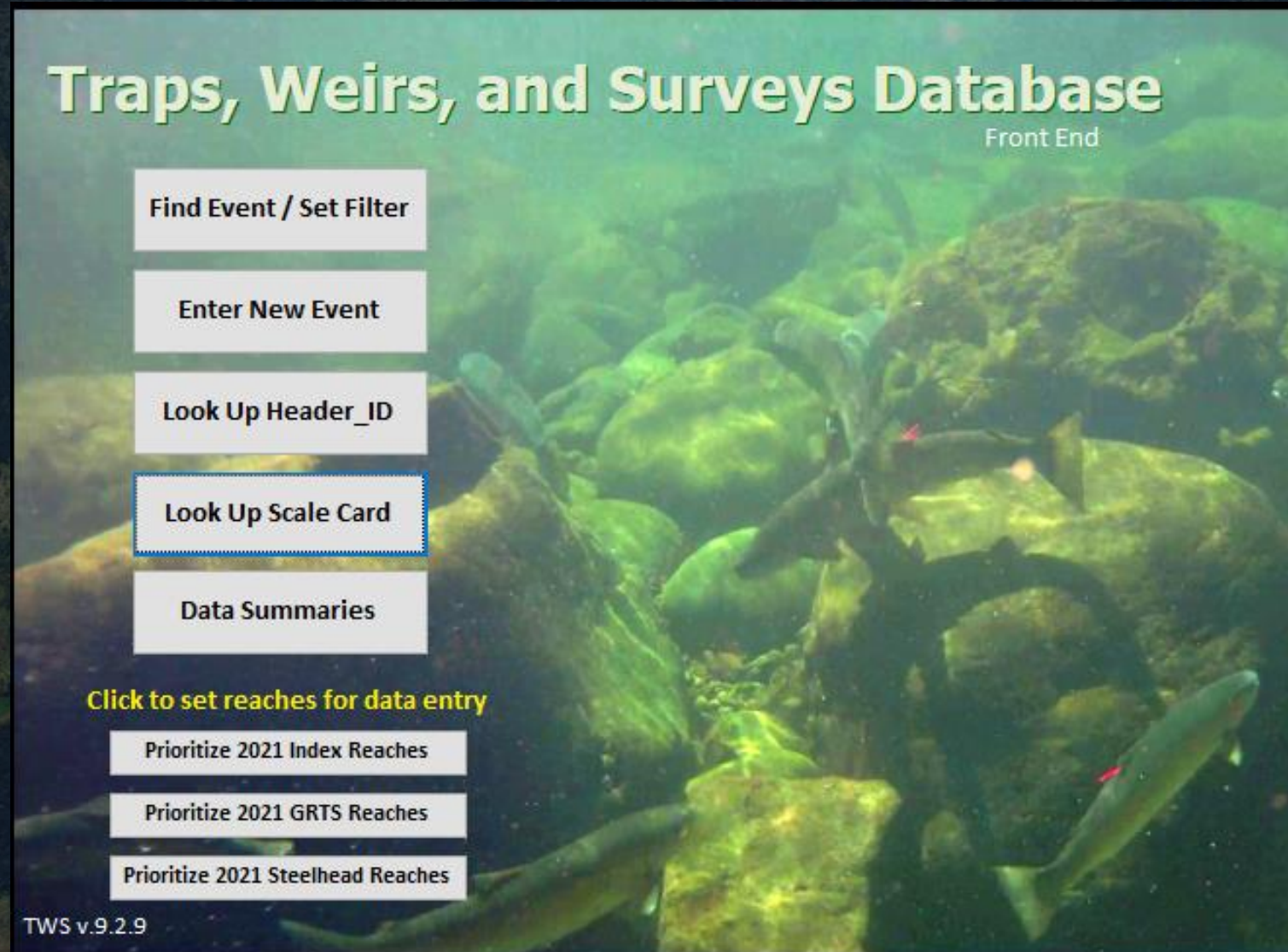
KEY: Fin clips: UM = unmarked, AD = adipose-clipped, LV = left ventral clipped, ADLV = adipose + left ventral clipped
 Disposition = Upstream, Downstream, Surplus-Nutrient enhancement, Surplus-Food bank, Surplus-Unknown, Broodstock-location, Sacrificed for Research, Escaped Upstream, Escaped Downstream

BDS TEAM ASSEMBLE!

- Streamline data collection
- Create efficiencies in field, regional and corporate datasets
- Maintain or increase data integrity and security
- Facilitate timely data availability for analysis and the effective storage of that data in a cloud environment

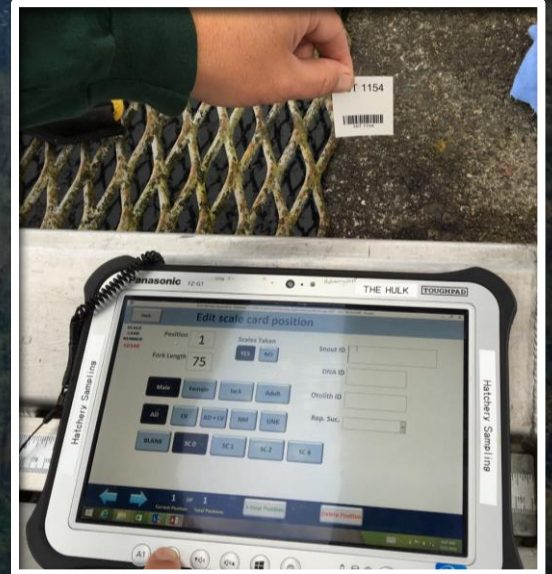


Traps, Weirs and Surveys (TWS) – a micro data management ecosystem



Traps, Weirs and Hatcheries:

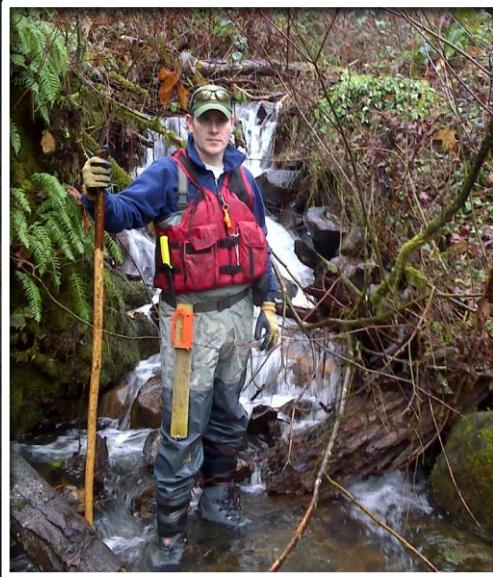
- In use in Southwest Washington since 2012
- Custom MS Access applications running on Panasonic FZ-G1 Toughpads
- Various i/o configurations (barcode, PIT, GPS)



TWS MOBILE: SURVEYS

Surveys:

- In use in Southwest Washington since 2015
- Custom iFormBuilder applications running on Apple iPads
- Cheaper and more mobile than Toughpads
- Some speed limitations (speed of data entry)



Enter Live Fish

Species *

Steelhead	Pacific Lamprey	Coho	Chinook
-----------	-----------------	------	---------

Spawner or Holder? *

Spawner	Holder
---------	--------

Tagged? *

Untagged	Tagged	Unknown
----------	--------	---------

Mark Type *

UM	AD	UNK	AD+LV
----	----	-----	-------

Fish Count *

3

GPS

Latitude: 45.695623, Longitude: -122.667235, Altitude: 54.903931, Speed: 0.000000, Horizontal Accuracy: 10.000000, Vertical Accuracy: 12.000000, Time: 4:42:16 PM PDT

Comments

Done

2016 Winter Steelhead Surveys - Region 5

Survey Method *

Foot	Raft	Snorkel	Jetboat
------	------	---------	---------

Survey Direction

Upstream	Downstream	Meet in middle
----------	------------	----------------

Is the data for ... Split? *

Who will you ... with? *

Clarity (ft) *

5.5

Clarity Code

To Bottom	Actual Measurement	Deeper than Measured
-----------	--------------------	----------------------

Weather *

Sunny	Cloudy	Showers	Raining	Snowing
-------	--------	---------	---------	---------

Conditions *



TWS – Why has it succeeded?

- TWS adoption and use in SW Washington:
Fall 2012 – present
- User focused from the beginning
 - Buy-in from those that collect the data, manage the data, and analyze the data
- Outputs and analytic tools that improve on day-to-day tasks
- Get data in and out efficiently

Current Statistics:

Number of distinct sampling events: >100k
about 10k a year

Number of detail records: >800k
adult fish, redds, scale samples etc.

Number of users annually: Around 100

How do we scale from a micro to macro?



Technical considerations

- Base level of technical know-how
- A sufficient level of infrastructure
- Access to the right tools

Process considerations

- Staff engagement
- Manager support
- Introspective look at what didn't work in the past

BDS TEAM ASSEMBLE...AGAIN!





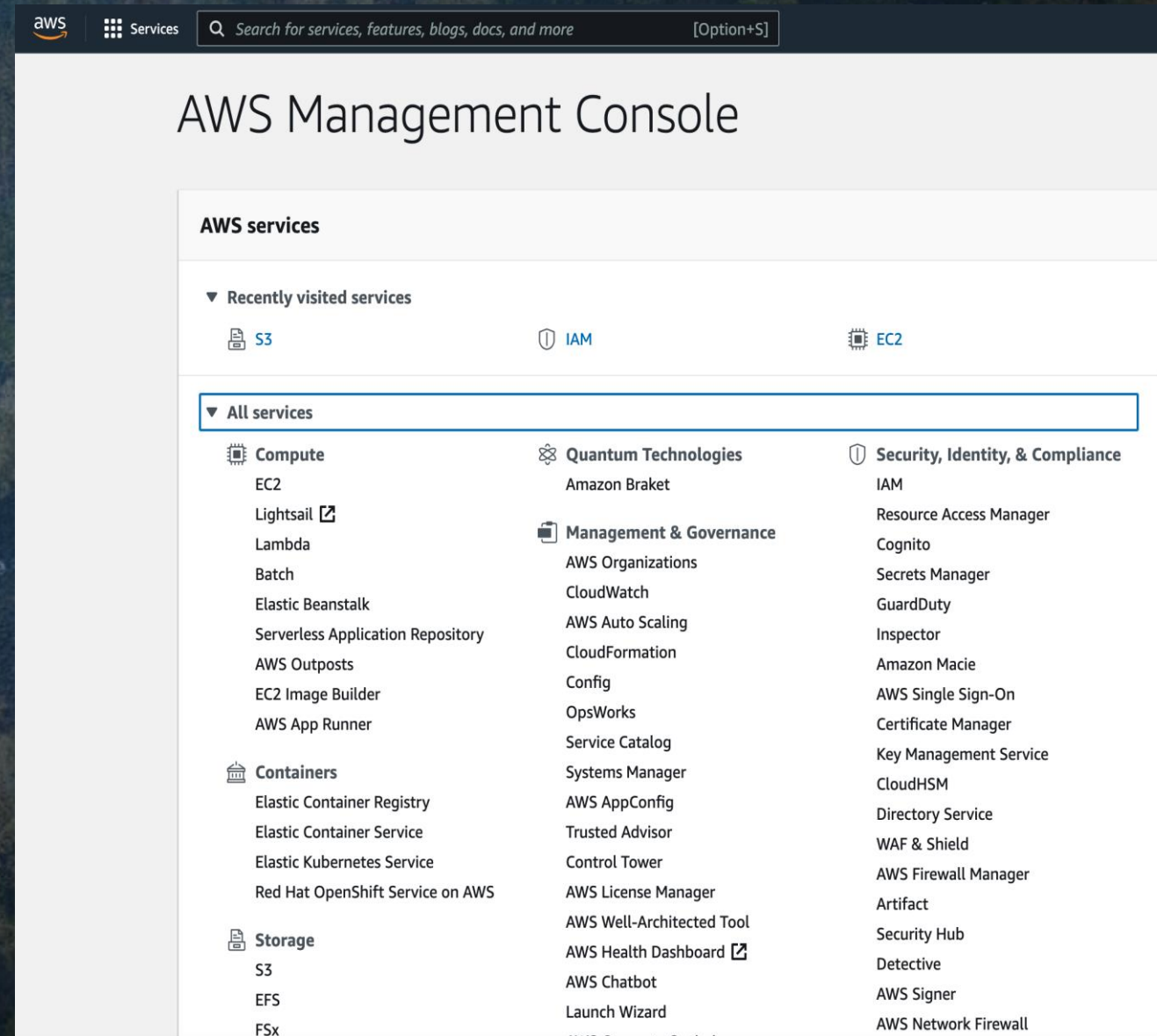
AWS: a powerful tool for authentication and storage

Amazon Web Services (AWS) – Infrastructure as a Service (IaaS)

- Linux Web servers
- Linux ArcGIS servers
- Linux PostgreSQL database servers (EC2) -> Aurora Relational Data Service (RDS)
- Load balancers and bastions
- File Storage and Static Website hosting (S3)
- Messaging (SNS)
- Unauthenticated resource access (Cognito)
- Networking and Security (IAM, ACL, WAF, etc.)

Custom-built Security Token Service (STS) – Used for every web and mobile app requiring a login

- Leverages public-facing and internal identity providers (IDPs), including Secure Access Washington (SAW) and Active Directory (AD)
- Implements role and object-based access control (ROBAC)
- User management portal for maintaining roles, users and clients (apps)





PostgreSQL

PostgreSQL: a powerful open-source relational database

- Leverages the SQL language
- Modern architecture
- Robust access-control system
- Reliable
- Scalable
- Free!

pgAdmin

File Object Tools Help

Browser

- Scenarios (19)
 - age_and_scales
 - ais_monitoring
 - bdo
 - creel
 - device_inventory
 - fish_id
 - forage_fish
 - logbook
 - marine_fish_common
 - mfbds
 - ms_commercial
 - Aggregates
 - Collations
 - Domains
 - FTS Configurations
 - FTS Dictionaries
 - FTS Parsers
 - FTS Templates
 - Foreign Tables
 - Functions
 - Materialized Views
 - Operators
 - Procedures
 - Sequences
 - Tables
 - Trigger Functions
 - Types
 - Views (16)
 - vw_all_odfw_samplecategory
 - vw_all_mss_data
 - vw_all_nomark_fish_counts
 - vw_all_odfw_commercial_detail
 - vw_all_odfw_commercial_headers
 - vw_all_plus_counts
 - vw_cwtf_output
 - vw_cwtf_recovery
 - vw_cwtf_samplingheaders
 - vw_cwtf_tripheader_groups
 - vw_cwtf_tripheader_groups_assign
 - vw_cwtf_tripheaders
 - vw_odfw_all_detail_to_commercial_id
 - vw_odfw_headergroup_all
 - vw_odfw_headergroup_all_printform
 - vw_wdfrw_snout_lab_export
- pikeminnow
- pit_harvest
- public
- spawning_ground
- species_comp
- str_sturgeon
- tws
- watix

Query Editor

```
-- View: ms_commercial.vw_cwtf_recovery
-- DROP VIEW ms_commercial.vw_cwtf_recovery;

CREATE OR REPLACE VIEW ms_commercial.vw_cwtf_recovery
AS
SELECT concat('R_', ms_commercial.uuid_to_bigint(d.detailid, 15)::numeric) AS recoveryid,
sh.samplingheaderid AS samplingheaderreferenceid,
d.snfd_barcode::character varying(25) AS snouttagcode,
CASE
WHEN d.snfd_barcode::character varying(25) IS NULL THEN false
ELSE true
END AS snouttaken,
CASE
WHEN s.species_id::text = '12'::text THEN '51'::character varying(64)
ELSE s.species_id
END AS species,
sh.run,
to_char(h.sampledate, 'YYYYMMDD'::text) AS recoverydate,
'R'::text AS recoverydatatype,
sh.primaryarea AS recoverylocationcode,
CASE
WHEN d.fin_mark_1::text = '2'::text AND d.fin_mark_2 IS NULL THEN 'AD'::text
WHEN d.fin_mark_1::text = '2'::text AND d.fin_mark_2::text = '4'::text THEN 'AD-LV'::text
WHEN d.fin_mark_1::text = '2'::text AND d.fin_mark_2::text = '5'::text THEN 'AD-RV'::text
WHEN d.fin_mark_1::text = '2'::text AND d.fin_mark_2::text = '6'::text THEN 'AD-LP'::text
WHEN d.fin_mark_1::text = '2'::text AND d.fin_mark_2::text = '7'::text THEN 'AD-RP'::text
WHEN d.fin_mark_1 IS NULL AND d.fin_mark_2::text = '4'::text THEN 'LV'::text
WHEN d.fin_mark_1 IS NULL AND d.fin_mark_2::text = '5'::text THEN 'RV'::text
WHEN d.fin_mark_1 IS NULL AND d.fin_mark_2::text = '6'::text THEN 'LP'::text
WHEN d.fin_mark_1 IS NULL AND d.fin_mark_2::text = '7'::text THEN 'RP'::text
WHEN d.fin_mark_1 IS NULL AND d.fin_mark_2 IS NULL THEN '0000'::text
END AS fin_mark
```

Data Output

recoveryid	samplingheaderreferenceid	snouttagcode	snouttaken	species	run	recoverydate	recoverydatatype	recoverylocationcode	recordedmark	weightamount	weightcode	weighttype	lengthamount	lengthcode	lengthtype	detecto
1	R_170935884232080	SH_2	[null]	false	1	20210426	R	5F33230 R30 15	AD		[null]	[null]	770 0		1	E
2	R_343829714609204	SH_2	[null]	false	1	20210426	R	5F33230 R30 15	AD		[null]	[null]	640 0		1	E
3	R_344837254621484	SH_18	[null]	false	1	20210329	R	5F33220 R20 15	AD		[null]	[null]	[null] 0		1	E
4	R_1950508291902931	SH_18	[null]	false	1	20210329	R	5F33220 R20 15	AD		[null]	[null]	880 0		1	E
5	R_166747325998364	SH_20	[null]	false	1	20210503	R	5F33220 R20 15	AD		[null]	[null]	640 0		1	E
6	R_23996181709580	SH_19	[null]	false	1	20210405	R	5F33220 R20 15	AD		[null]	[null]	680 0		1	E
7	R_123103185168111	SH_19	[null]	false	1	20210405	R	5F33220 R20 15	AD		[null]	[null]	730 0		1	E
8	R_12241542000377	SH_7	21L0600	true	1	20210329	R	5F33220 R20 15	AD		[null]	[null]	670 0		1	E
9	R_278623692236973	SH_1	[null]	false	1	20210419	R	5F33230 R30 15	AD		[null]	[null]	480 0		1	E
10	R_170989151620335	SH_15	21L1503	true	1	20210405	R	5F33220 R20 15	AD		[null]	[null]	640 0		1	E
11	R_175281250448499	SH_16	21L5110	true	1	20210418	R	5F33220 R20 15	AD		[null]	[null]	790 0		1	E
12	R_168412360513245	SH_2	21L0625	true	1	20210426	R	5F33230 R30 15	AD		[null]	[null]	490 0		1	E
13	R_178642560137039	SH_18	[null]	false	1	20210329	R	5F33220 R20 15	AD		[null]	[null]	650 0		1	E
14	R_19646467899769	SH_3	[null]	false	1	20210412	R	5F33230 R30 15	AD		[null]	[null]	830 0		1	E
15	R_176477502054818	SH_13	[null]	false	1	20210425	R	5F33220 R20 15	AD		[null]	[null]	780 0		1	E
16	R_194381250486898	SH_14	[null]	false	1	20210411	R	5F33220 R20 15	AD		[null]	[null]	680 0		1	E
17	R_159526455726010	SH_23	[null]	false	1	20210426	R	5F33230 R30 15	AD		[null]	[null]	830 0		1	E
18	R_147671165703033	SH_23	21L0306	true	1	20210426	R	5F33230 R30 15	AD		[null]	[null]	720 0		1	E
19	R_19680358230869	SH_15	[null]	false	1	20210405	R	5F33220 R20 15	AD		[null]	[null]	800 0		1	E
20	R_182483159373772	SH_2	[null]	false	1	20210426	R	5F33230 R30 15	AD		[null]	[null]	500 0		1	E



Angular: a modern web developer platform

- Web based
- Optimized for speed by using a component framework
- Large library of templates and assets
- Intuitive APIs that minimize code to product ratio

[Angular demo](#)

Washington Department of Fish and Wildlife
Apps Portal

Welcome, Danny! [Sign Out](#)

Chehalis Spawning Ground QAQC

[QAQC](#) [SAVE](#) [DELETE](#)

Survey Date *	Start Time	End Time	Observers	Stream	Reach	Target Species	Survey Type	Survey Method	Survey Direction	Clarity	Clarity Code
11/21/2021	10:21 AM	01:12 PM	keith.bradley	Lucas Creek (LB)	Lucas Creek_3.2-4.2:Blue Car Seat (Sm)	Coho	Index	Foot	Downstream	1.5	Actual depth measur

Weather	Flow For Reach	Visibility	Survey Completion	Reason for incomplete survey
Sunny	Medium-low	Fair	Completed survey	

Lives (not on Redds) +

Species	Run	Spawner/Holder	UNK:SD	UM:SD	AD:SD	UNK:Jack	UM:Male	UM:Female	UM:Jack	UNK:Male	UNK:Female	AD:Male	AD:Female	AD:Jack
Coho salmon	Fall run	Holding	0	0	0	0	0	0	0	1	0	0	0	0

Deads +

Species	Run	Clip Status	Fish Condition	Sex	Fork Length (cm)	CWT Detected	CWT	Scale Card #	Scale Card Position #	DNA #	Otolith #	PIT #
Coho salmon	Fall run	UM	Carcass - Good	Male	56	Coded-wire tag nr		A-74557	2		NWKC021-48	

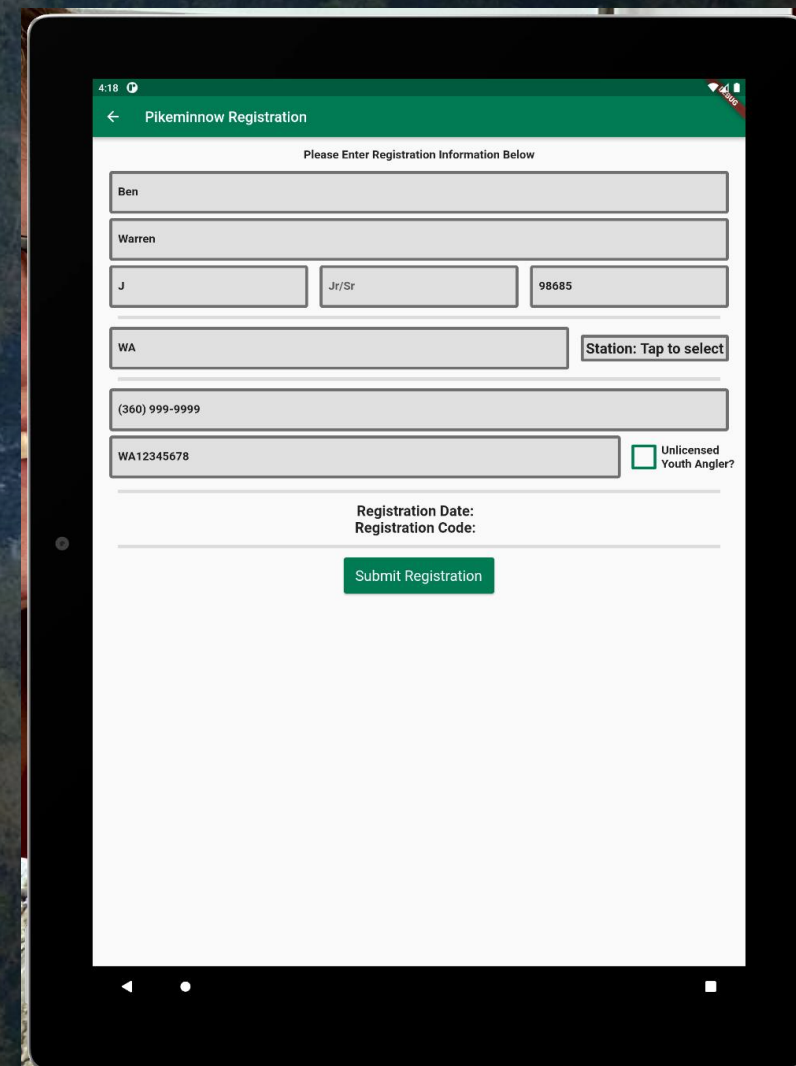
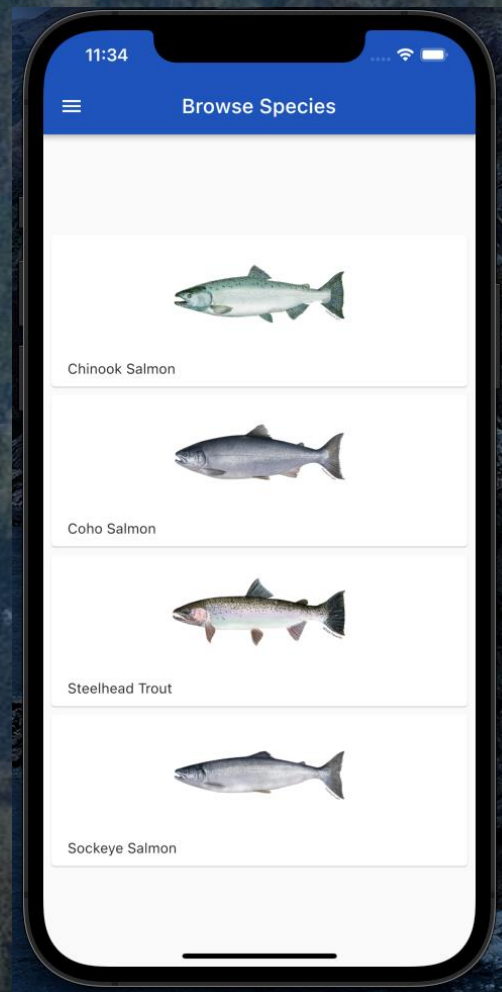
Redds +

Status	Species	Run	Redd Name	Degradation	Redd Comment	TL:
Previous redd, still vi	Coho salmon	Unknown run	6RD5157	50	FT	2
New redd	Coho salmon	Fall run	6RD3748	0		0
New redd	Coho salmon	Fall run	6RD3747	25		0



Flutter and Dart: a new IDE for mobile development

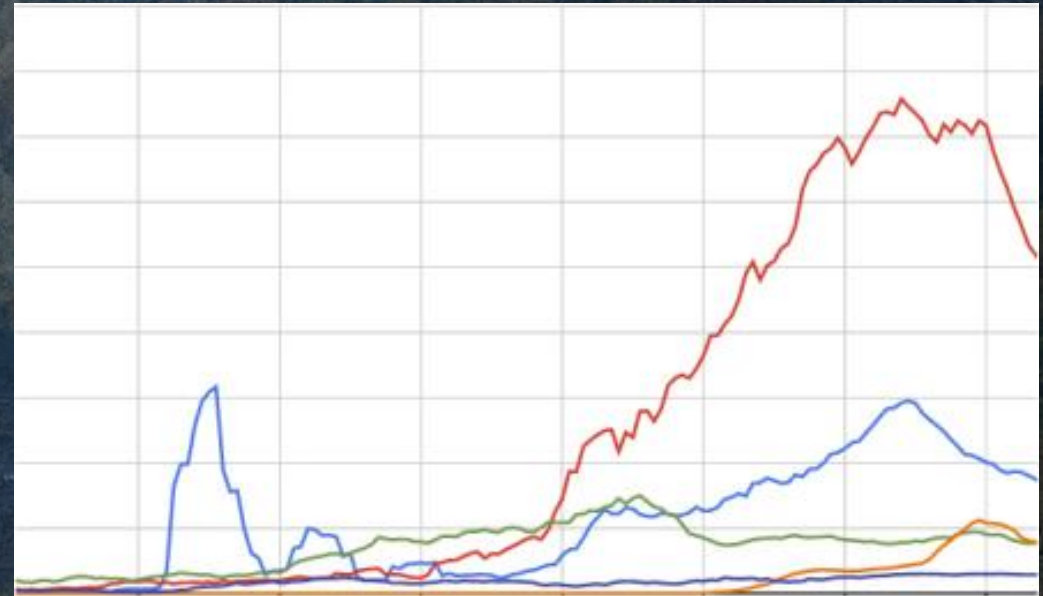
- Open Source (Google)
- Cross-platform (Android, iOS, Desktop, Web) with a single codebase
- Brilliant tooling and low-code entry
- 'Widget' framework makes code repurposing nearly seamless across applications
- Industry leading support and documentation



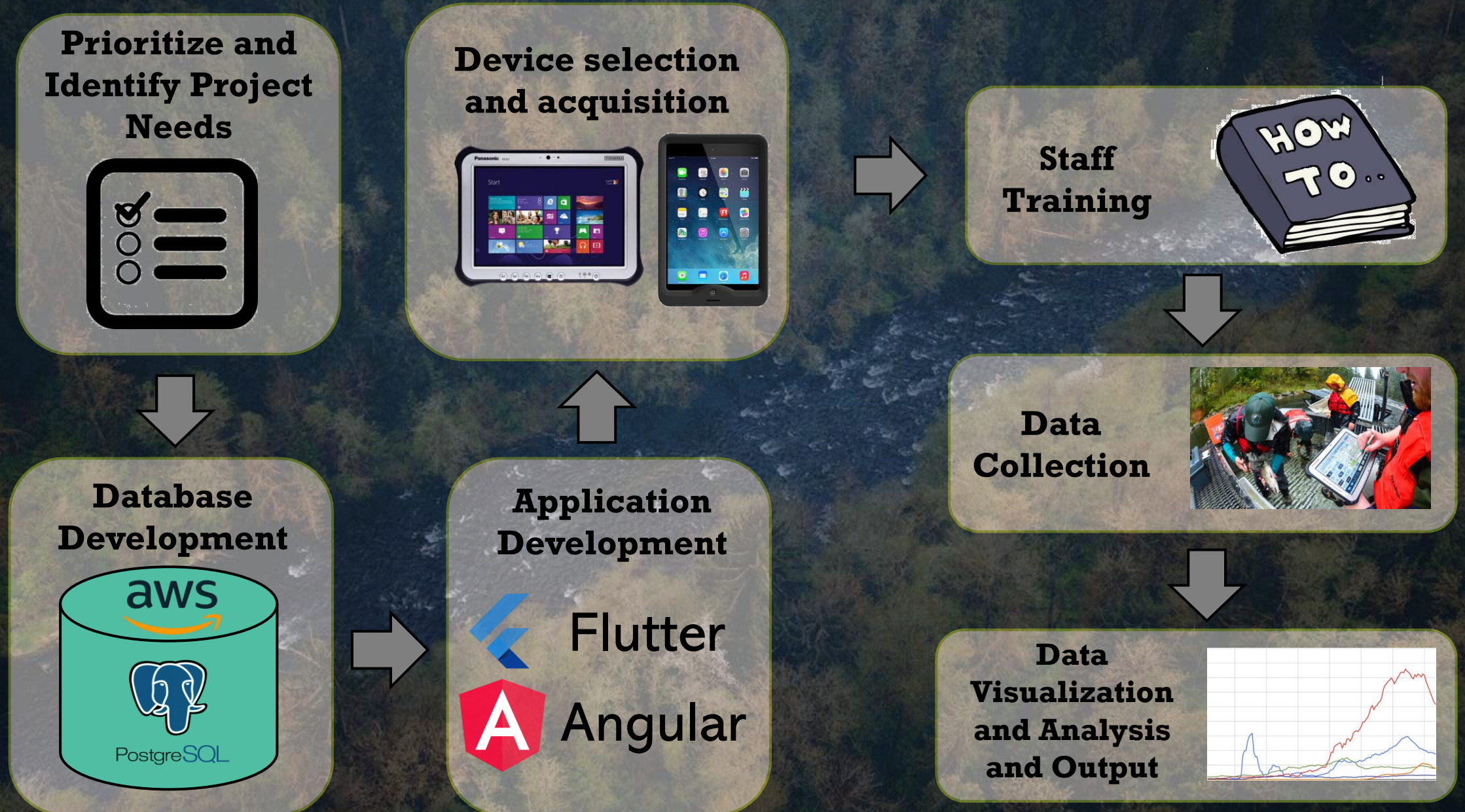


Outputs: tools to generate outputs for clients

- Automated extract, transform, load (ETL) procedures leveraging views, SQL Server Integration Services (SSIS) and Python.
- Transfer CSV and other file formats to third parties
- Update various datasets on Data.wa.gov
- Automated internal fish ticket reporting leveraging views and Microsoft Power BI
- Reports run daily and uploaded to dedicated Microsoft Teams channels
- Custom Reporting RESTful APIs
- Available for third parties to export specific datasets on-demand



Our data management ecosystem: a roadmap





Questions?