AIS Vessel Identification and Techniques

Dave Winkler
CGI Federal
USCG Operations Systems Command
david.winkler@cgifederal.com
david.m.winkler@uscg.mil

2012 Annual Assembly Meeting and Conference of the Radio Technical Commission for Maritime Services

September 23-28, 2012
Orlando, FL USA

Mission

True Maritime Domain Awareness (MDA) - demands continuous, real-time identification of vessels broadcasting AIS

- Which targets are actually maritime vessels?
- Which are land-based (improperly broadcasting with Class A/B AIS)?
- Which are broadcasting:
  - the wrong identity?
  - another vessel's identity?
  - a credential issued to another party for the same vessel?
  - expired identifiers?
  - Radio identifiers issued by a different flag than their registry?
- How many vessels' movements cannot be audited historically by using duplicative MMSI numbers (>1 vessel simultaneously broadcasting same MMSI)?

Even as 100% identification is not possible (typically 50-100 vessels/day out of 8,500-10,000 within the U.S. NAIS System range, or ~1% unverified), allows vessels to be divided into **known** and **unknown** sets, enabling focus on unknown vessels

CGI

Background / Current State

Approximately 50% of AIS Static Data transmissions have errors

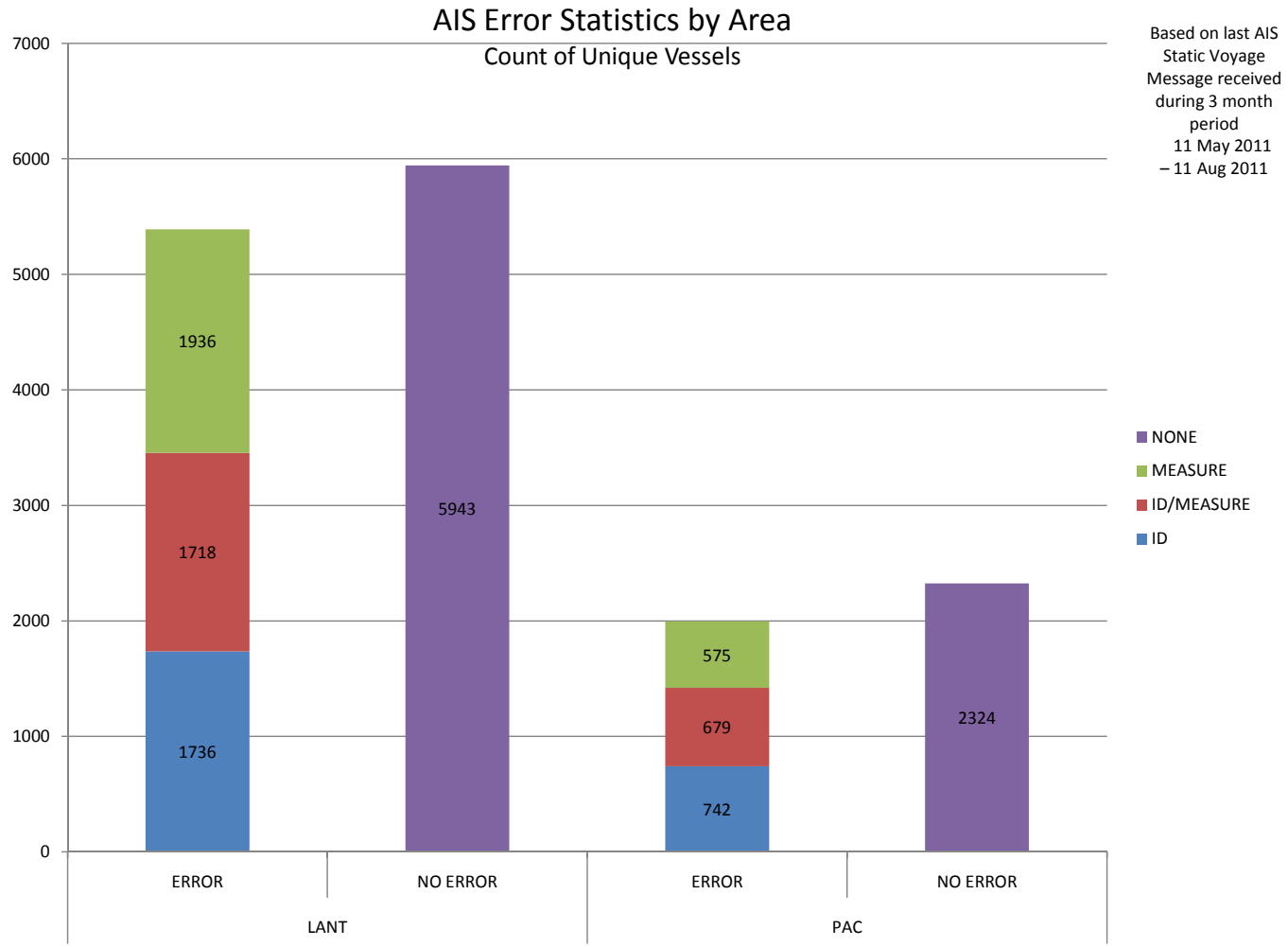
Of those with errors:

- 1/3 have ID errors
- 1/3 have Measurement errors
- 1/3 have **both** ID and Measurement errors

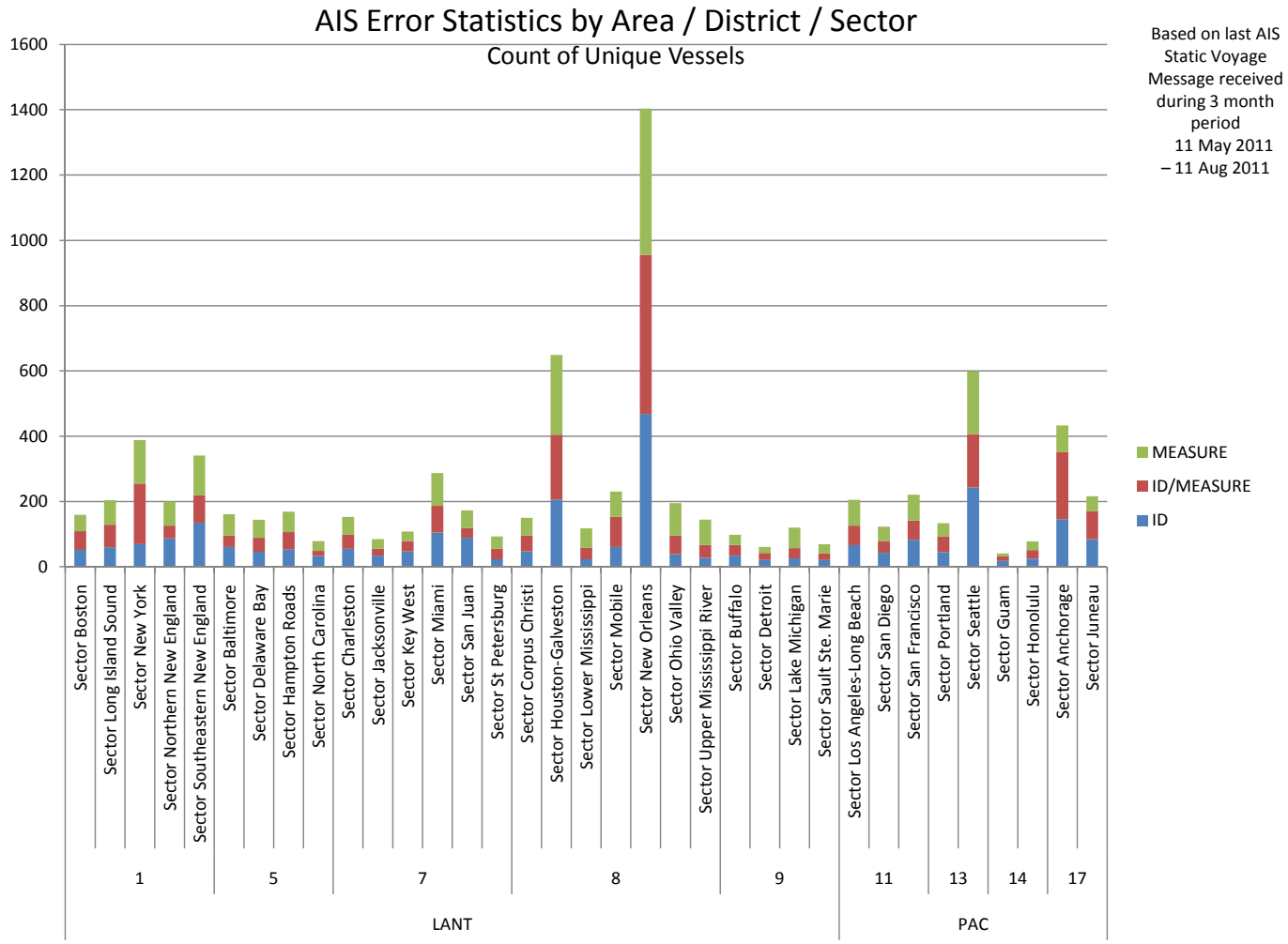
Overall, 1/3 of ALL vessels have at least one incorrect identifier of MMSI number, IMO number, Call Sign and or Ship Name – Maritime Security / Intel

Another 1/3 of ALL vessels have at least one error in measurements or some other non-identifying static data element – Maritime Safety

AIS Error Types – by Area



AIS Error Types – by Sector



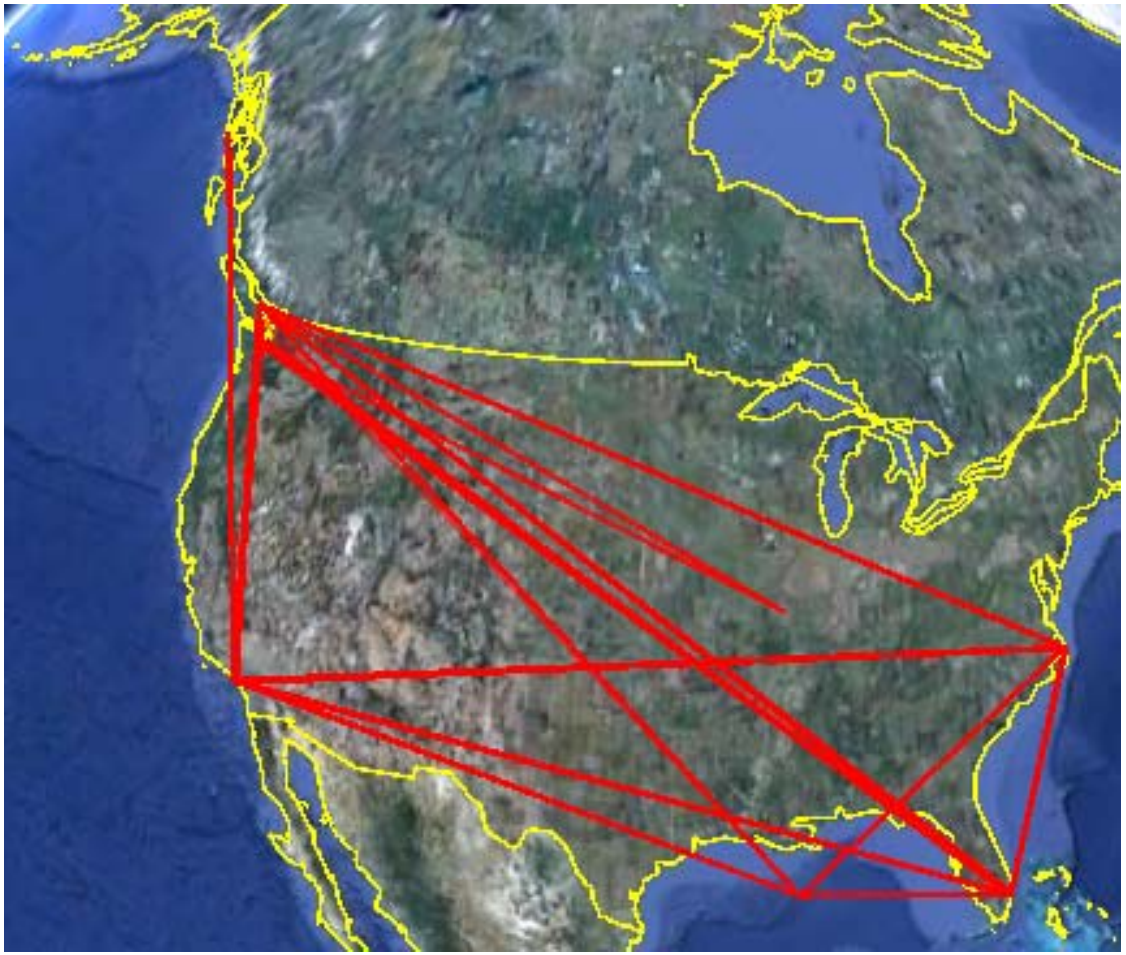
AIS Error – MMSI Duplication

- Largest problem for systems consuming unvalidated AIS data
  - Safety problem when multiple ships use same MMSI in same local region
  - For historical data analysis, often difficult to track history of a vessel which uses a duplicative MMSI
- Limited domain of duplicative MMSIs
  - Only approx. 150 MMSIs over the past 3 years
  - #1 problem: Nauticast X-Pack-US default MMSI 1193046
    - Why? Keeps coming back until operators repair or replace their transponder
    - Typical MMSIs: 111111111, 123456789, 987654321, 1, 5, etc.
  - If another data element is correctly configured its identity can be verified, but often spatial analysis (ports/facilities visited, nearest neighbor vessels) must be used to get “eyes on the target”

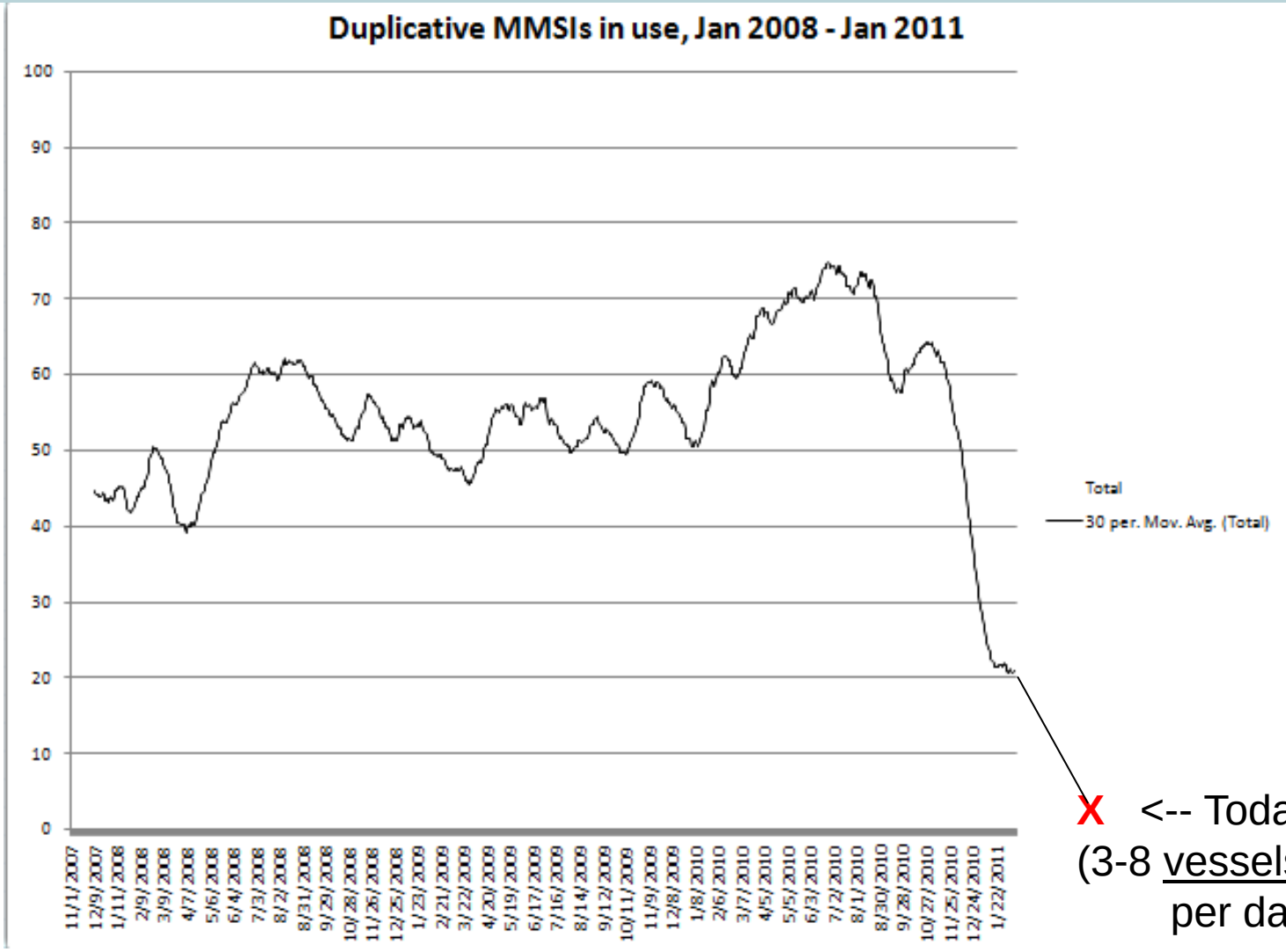
AIS Error – MMSI Duplication

Vessel Track of MMSI 111111111

(30 days, 22 Feb - 24 Mar 2011)



AIS Error – MMSI Duplication



X <-- Today
(3-8 vessels per day)

Registration Error – Multiple MMSI assignments

AIS_UID	MMSI	IMO	CALL SIGN	NAME
338066377||||||||0||||||||||||||||AGAPE	338066377			AGAPE
366813650||||||||0WDA4836|||||||||||||AGAPE	366813650	0	WDA4836	AGAPE
338109284||||||||0||||||||||||AQUATHERAPY	338109284			AQUA THERAPY
366987190||||||||0WDC2063||||||||AQUATHERAPY	366987190	0	WDC2063	AQUA THERAPY
338053633||||98000||19258||||||||||||AUDREYK	338053633	98000	19258	AUDREY K
367172810||||||||0WDD6425||||||||||||AUDREYK	367172810	0	WDD6425	AUDREY K
338055716||||||||0||||||||||||||BAREFOOTN	338055716	0		BAREFOOT'N
366789550||||||||0WDA2377||||||||||BAREFOOTN	366789550	0	WDA2377	BAREFOOT'N
338103485||||||||0||||||||||||CAPTTHANHII	338103485			CAPT THANH II
367156270||8847662WBC7796||||||||CAPTTHANHII	367156270	8847662	WBC7796	CAPT THANH II
338105715||||||||0||||||||||||CAPTAINMORGAN	338105715			CAPTAIN MORGAN
366829590||||||||0WDA6186||||||CAPTAINMORGAN	366829590	0	WDA6186	CAPTAIN MORGAN
338086233||||||||0||||||||||||||CHEKARA	338086233			CHEKARA
366800490||||||||0WDA3632||||||||||||CHEKARA	366800490	0	WDA3632	CHEKARA
338098853||||||||0||||||||||||CHILANGUITATOO	338098853			CHILANGUITA TOO
367448980||||||||0WDF4187||||||CHILANGUITATOO	367448980	0	WDF4187	CHILANGUITA TOO
338104023||||||||0||||||||||||||CHULAMAR	338104023			CHULAMAR
368433000||||||||0WDF7092||||||||||||CHULAMAR	368433000	0	WDF7092	CHULAMAR
338104025||||||||0||||||||||||||CROWDPLEEZER	338104025			CROWD PLEEZER
366937780||||||||0WDB7178|||||||CROWDPLEEZER	366937780	0	WDB7178	CROWD PLEEZER
338090025||||||||0||||||||||||||DAYDREAMS	338090025			DAY DREAMS
366877080||||||||0WDB2543||||||||||DAYDREAMS	366877080	0	WDB2543	DAY DREAMS
338105974||||||||0||||||||DIVINEINTERVENTION	338105974			DIVINE INTERVENTION
367120140||||||||0WDD2806||DIVINEINTERVENTION	367120140	0	WDD2806	DIVINE INTERVENTION
338054427||||||||0||||||||||||||DREAMINON	338054427	0		DREAMIN ON
366722960||||||||0||||||||||||||DREAMINON	366722960	0		DREAMIN ON
338109066||||||||0||||||||||||||FRAIDKNOT	338109066			FRAID KNOT
367478750||||||||0WDF6877||||||||||FRAIDKNOT	367478750	0	WDF6877	FRAID KNOT

Analogous to obtaining a drivers license from multiple states

Problem Statement

Unique tracking over time - what should be used as the fixed variable when identifying vessels?

No single legal identifier (Official Number, IMO Number, State Registration Number, Call Sign, MMSI, etc.) meets the criteria that it is available / issued to every vessel for the purpose of unique tracking

For U.S., the USCG's Maritime Information for Safety and Law Enforcement (MISLE) and Vessel Documentation System (VDS) serves as the nation's vessel registry

Problem Statement (cont'd)

The MISLE / VDS systems represent any vessel with a unique VESSEL_ID sequence number

This VESSEL_ID number used within the USCG's System of Record is transparent - publicly available through the USCG's CG-MIX Port State Information Exchange (PSIX) search pages and web services.

Correlation of vessels to a permanent, immutable number allows for consistent tracking of vessels over time even as other legally assigned numbers change

Data sharing between IT systems and different organizations is vastly improved as multiple systems, including but not limited to AIS, rely upon the same fixed variable for vessel identification

Techniques for vessel identification

- Correlation – from what?
  - Single identifying elements of MMSI, IMO, Call Sign and Name cannot be relied upon
    - Individual data elements may be wrong or missing
  - Composite of 4 discrete identifiers are turned into an "**AIS Vessel Signature**"
  - AIS Signature is tokenized into a 45-character string

    positions 1-9=MMSI positions 10-18=IMO

    positions 19-25=Call Sign positions 26-45=Name
    - Use vertical pipe character to prefix (front pad) any identifiers which do not use all allocated characters
    - selected because it is not part of the NMEA 6-bit dictionary
    - Remove all non-alphanumeric characters
      - Usually variations in punctuation, spacing, or even inclusion of emoticons :-(
      - Can be regarded as non-significant differences
      - Only affects strings (Call Sign and Name)

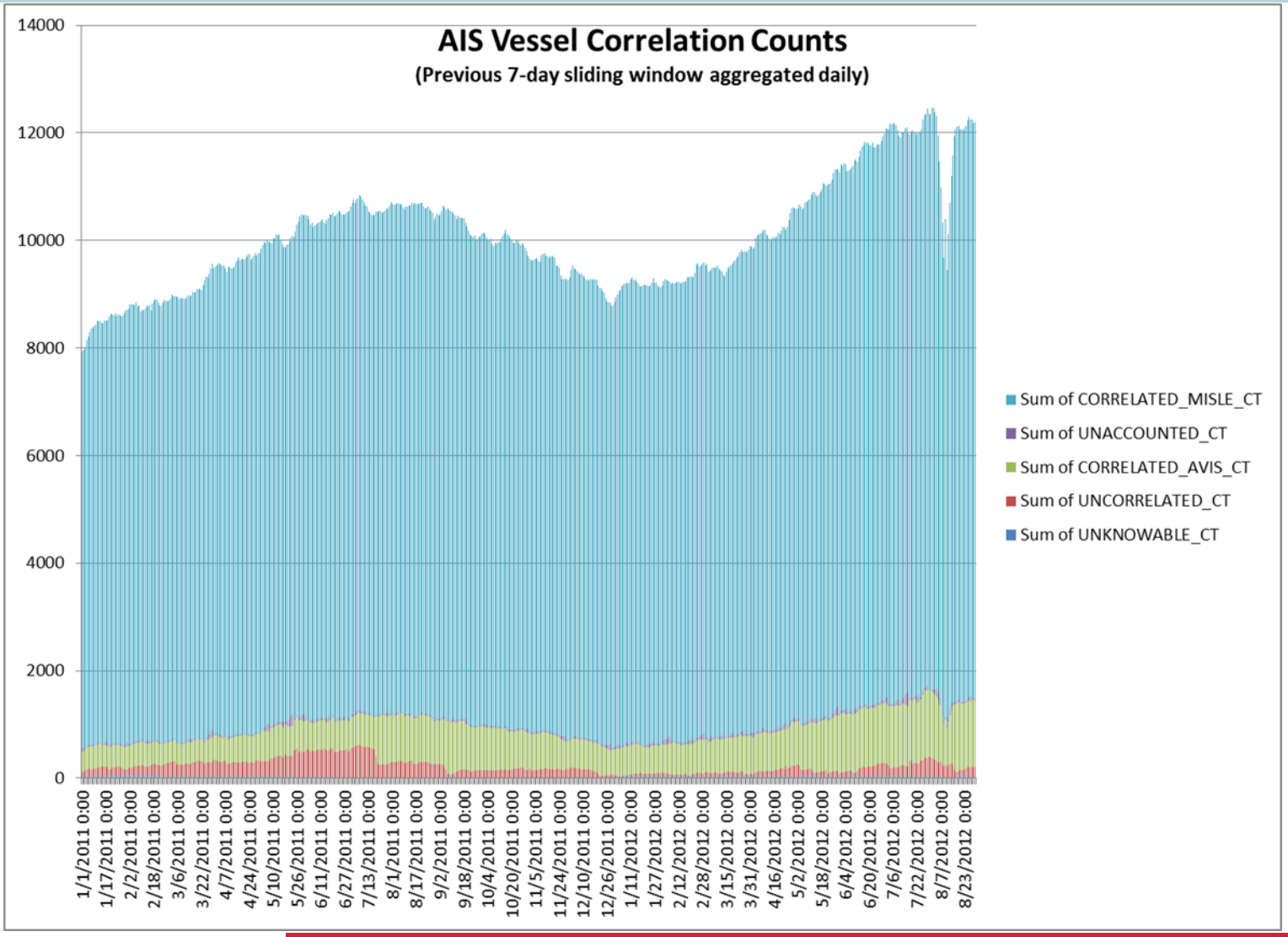
CGI

Vessel Correlation Example

AIS_UID	MMSI	IMO_NUMBER	CALL_SIGN	NAME	RECORDS	MIN_ONPLOT_DT	MAX_ONPLOT_DT	MISLE_VESSEL_ID
367051230||||||||0WDC6095||||LTSAMUELSCOURSEN	367051230	0	WDC6095	LT SAMUEL S COURSEN				609844
367051230||||||||1WDB3537|||||||||||||COURSEN	367051230	1	WDB3537	COURSEN	130,827	2007-11-08 22:40:16.000	2008-07-26 23:57:19.000	609844
367051230||||||||1WDC6095|||||||||||||COURSEN	367051230	1	WDC6095	COURSEN	370,739	2007-11-08 22:43:00.000	2009-05-10 13:12:24.000	609844
||1193046303174162WDC6095|||||||||||||COURSEN	1193046	303174162	WDC6095	COURSEN	15	2008-08-13 12:31:11.000	2008-08-13 12:52:30.000	609844
|36705123||||||||1WDC6095|||||||||||||COURSEN	36705123	1	WDC6095	COURSEN	7	2008-08-13 12:54:58.000	2008-08-13 12:55:06.000	609844
||1193046303174162WDB3537||||||SAMUELSCOURSEN	1193046	303174162	WDB3537	SAMUEL S COURSEN	57	2009-06-02 14:22:01.000	2009-12-28 18:52:58.000	609844
367051230||||||||1WDB3537||||||SAMUELSCOURSEN	367051230	1	WDB3537	SAMUEL S COURSEN	873,748	2009-06-02 14:31:00.000	2011-03-19 22:01:21.296	609844
||1193046303174162WDB3537|||||||SAMUELSCURSEN	1193046	303174162	WDB3537	SAMUEL S CURSEN	11	2009-06-02 14:47:26.000	2009-10-28 13:33:21.000	609844
367051230||||||||1WDB3537|||||||SAMUELSCURSEN	367051230	1	WDB3537	SAMUEL S CURSEN	38,391	2009-06-02 15:00:57.000	2009-12-28 15:06:00.000	609844
||||||||1|36701230WDB3537||||||SAMUELSCOURSEN	1	36701230	WDB3537	SAMUEL S COURSEN	684	2009-09-04 11:20:10.000	2009-09-05 13:05:19.000	609844
367051230||||||||1|D11233|||||||||||NAUTICAST	367051230	1	D11233	NAUTICAST	114	2009-10-07 11:00:40.000	2011-03-19 22:21:36.430	609844
367051239||||||||1WDB3537||||||SAMUELSCOURSEN	367051239	1	WDB3537	SAMUEL S COURSEN	22	2009-10-09 11:29:20.000	2009-10-09 11:32:09.000	609844
367051230||||||||1WSB3537||||||SAMUELSCOURSEN	367051230	1	WSB3537	SAMUEL S COURSEN	60,665	2009-12-28 19:16:16.000	2010-03-27 22:28:46.176	609844
367051230||||||||1WDC6095||||||SAMUELSCOURSEN	367051230	1	WDC6095	SAMUEL S COURSEN	1,086	2011-03-19 21:11:59.186	2011-03-20 16:25:08.123	609844
367051230||||||||0||||||||||||||||||||	367051230	0			4,498	2011-04-12 15:39:47.156	2011-08-09 14:32:15.810	609844
367051230||||||||0WDC6095||||||SAMUELSCOURSEN	367051230	0	WDC6095	SAMUEL S COURSEN	122,045	2011-04-12 15:42:16.730	2011-09-22 23:58:25.403	609844

CGI

Vessel Correlation Statistics



Vessel Data Validation

In order to measure whether a vessel is properly identified, a standard must exist which can be used to compare the AIS data against

“Official” data sources exist within authoritative systems of record, but that does not guarantee they are correct, complete, current or unique!

- Consider for which data elements a system of record serves as the **data steward**
  - FCC is the steward of the Call Sign and MMSI, but not the ship name
  - USCG is the steward of a *documented* vessel’s name, but not call sign
  - Lloyds is the steward (for the IMO) of the IMO #, but not the owner

A proper standard should incorporate the authoritative and verified data elements from each authoritative information source in order for comparisons with raw AIS data to yield proper decisions as to whether a vessel is properly identified

CGI

Vessel Catalog – Data Sources

U.S. Radio licenses
- FCC / Boat U.S. / SeaTow / Shine Micro / U.S. Power Squadrons

International radio registrations
- ITU MARS Database

Official Vessel Registration
- U.S. Certificate of Documentation (CG-MIX PSIX) / International Flag State registries

Lloyds Register / IHS
- Equasis

Classification Society records
- IACS member societies' data is regarded as legal record by many flag states

Notice of Arrivals, Fishing Treaty Organization databases, etc

CGI

Vessel Catalog Example

An effective vessel catalog will maintain only those basic data elements which form the basis for a Common Recognition Context for a vessel

- Each data element should be maintained / verified for completeness / correctness / uniqueness

```
<Vessel>
    <Identification>
        <Vessel_ID>1179</Vessel_ID>
        <MMSI>367440780</MMSI>
        <IMO>8968715</IMO>
        <Call_Sign>WDF3513</Call_Sign>
        <Name>FAST SPIRIT</Name>
        <Official_No>1092094</Official_No>
        <State_Reg_No/>
    </Identification>
    <Authority>
        <Vessel_Flag>US</Vessel_Flag>
        <Class_Soc>ABS</Class_Soc>
    </Authority>
    <Category>
        <Vessel_Service>Passenger (Inspected)
        </Vessel_Service>
        <Statcode5>B21A2OC</Statcode5>
        <Statcode5_Desc>Crew/Supply Vessel
        </Statcode5_Desc>
    </Category>
    <Manufacture>
        <Build_Year>2000</Build_Year>
        <Mfgr_Name>Breaux Brothers</Mfgr_Name>
        <Mfgr_Hull_No>542-001</Mfgr_Hull_No>
    </Manufacture>
    <HomePort Home_Port_Code="1000943">
        <Home_Port_City>GALLIANO</Home_Port_City>
        <Home_Port_State>LA</Home_Port_State>
        <Home_Port_Country>US</Home_Port_Country>
    </HomePort>
    <Measurement>
        <Length>50.29</Length>
        <Beam>9.75</Beam>
        <Draft>2.45</Draft>
        <Horsepower>19520</Horsepower>
        <GT>378</GT>
        <NT>113</NT>
    </Measurement>
    <Auditing>
        <Remarks/>
        <Created_Dttm>2010-02-17T09:06Z</Created_Dttm>
        <Validated_Dttm>2011-02-01T00:00Z</Validated_Dttm>
        <Validated_By>WINKLER,DAVID M</Validated_By>
    </Auditing>
</Vessel>
```

CGI

AIS Enforcement

While AIS signal is visible, publicly available information, only flag states have the authority to enforce correct AIS configuration

The **cost** of misconfigured AIS **is far greater than the cost of enforcement**, but:

- costs are not well defined – not easy to quantity ($$$)
- spread across multiple organizations
    - Multiple government agencies, commercial and academic entities rely upon AIS information for security, safety, economic and environmental analysis

For the U.S., 33 CFR 164.46(b) and 46 USC 70114 allows for commercial vessels with an improperly configured AIS to be issued penalties of up to $25,000/day and $50,000 maximum as defined in 46 USC 70119

Estimates for 70% / 30% compliance / non-compliance with a 3-month enforcement program would:

- Correct >95% of all known AIS misprogramming in the U.S. within 3 months
- Collect approximately **$4,000,000** in fines after initial warnings ignored

CGI

AIS Enforcement - Example

AIS MMSI	AIS Call Sign	AIS IMO	AIS Name	AIS Last Observed	Correct MMSI	Correct Call Sign	Correct IMO	Correct Name	AIS Draft	A Distance from Bow	B Distance from Stern	C Distance from Port	D Distance from Starboard
1	WDE9276	9032824	INTL'RAIDER	9/10/2012	367415510	WDE9276	9032824	INT'L RAIDER	3	12	31	4	6
367077440	WDF5644	918409300	INT'L BRAVE	9/1/2012	367465640	WDF5644	9184093	INT'L BRAVE	3.4	14	30	8	3
367036120	WDD6853	663407	INT'L TRADITION	9/10/2012	367178460	WDD6853	0	INT'L TRADITION	3	12	31	6	4
367046690	WDC5807	149	JIMBO	9/9/2012	367165510	WDD5902	0	JIMBO	3	0	0	0	0
367450560	WDF4330	602952	CAPT BRIAN	9/8/2012	367508080	WDF9716	8978136	CAPT BRIAN	3	10	20	6	3
367184050	WDD7263	0	CAVALIER	9/10/2012	367184050	WDD7263	0	CAVALIER	3	0	0	0	0
367485860	WDF7572	9030773	CLIPPER	9/10/2012	367485860	WDF7572	9030773	CLIPPER	3	10	70	9	1
367163390	WDD5771	641216	GULF SOUTH 1	9/10/2012	367163390	WDD5771	0	GULF SOUTH 1	0	8	26	6	2
367176070	WDD6676	635181	INTL CARRIER	9/9/2012	367176070	WDD6676	0	INT'L CARRIER	3	0	0	0	0
367475790	WDF6587	0	INTL CHARGER	9/10/2012	367475790	WDF6587	8978095	INT'L CHARGER	2.3	0	0	0	0
367464080	WDF5512	8978174	INT'L CHIEF	8/31/2012	367464080	WDF5512	8978174	INT'L CHIEF	2.5	0	0	0	0
367452910	WDF4526	9121716	INT'L COURAGE	9/10/2012	367452910	WDF4526	9121716	INT'L COURAGE	3.2	0	0	0	0
367197840	WDD8333	0	INT L DIAMOND	9/8/2012	367197840	WDD8333	0	INT'L DIAMOND	0	10	20	30	40
367158180	WDD5424	892626	INTERN,L DISCOVERER	9/10/2012	367158180	WDD5424	8926626	INT'L DISCOVERER	0	0	0	0	0
367147750	WDD4664	4	INT,L EXPLORER	9/10/2012	367147750	WDD4664	0	INT'L EXPLORER	3	8	26	8	2
367186230	WDD7424	105000000	INT'L FALCON	9/9/2012	367186230	WDD7424	0	INT'L FALCON	0	7	30	3	6
367483880	WDF7375	8978162	INT'L FLYER	9/7/2012	367483880	WDF7375	8978162	INT'L FLYER	5.5	0	0	0	0
367159140	WDD5493	100000009	INT'L FREEDOM	7/1/2012	367159140	WDD5493	7501065	INT'L FREEDOM	3.5	10	25	6	5
538001293	V7BK9	7703417	INTL FRONTIER	8/3/2012	538001293	V7BK9	8766703	INT'L FRONTIER	18	0	0	0	0
367468990	WDF5947	8978320	INT'L NAVIGATOR	9/10/2012	367468990	WDF5947	8978320	INT'L NAVIGATOR	2.5	6	2	6	26
367191040	WDD7768	591434000	INTL PATRIOT	9/9/2012	367191040	WDD7768	0	INT'L PATRIOT	3.5	12	22	4	5
367165490	WDD5901	641321	INTL PRIDE	7/26/2012	367165490	WDD5901	0	INT'L PRIDE	25.5	0	35	0	8
367464110	WDF5514	0	INT'L QUEEN	8/28/2012	367464110	WDF5514	8978215	INT'L QUEEN	2	10	22	5	4
367152990	WDD5051	608889	INTLRUNNER	9/8/2012	367152990	WDD5051	0	INT'L RUNNER	3	0	0	0	0
367312920	WDD9543	52	INTL SCOUT	9/10/2012	367312920	WDD9543	0	INT'L SCOUT	2.9	0	0	0	0

CGI

Interagency Cooperation

Federal Initiative for Navigation Data Exchange (FINDE)

- A federal working group focused on data sharing and standardization of vessel, port, commodity, owner/operator information
- Partners include USACE (lead), USCG (co-lead), CBP, IRS, NOAA, MARAD
- Achievements include interagency Information Sharing Agreements (ISAs) to share AIS information and reference data sets from the USCG to USACE and sharing of USACE's inland AIS transceiver network data with USCG

(Other similar efforts exist – learned of the UNCLOG working group at RTCM)

Federal-Industry Logistics Standardization (FILS)

- Focused on data standards and information sharing, both industry-industry and industry-government
- Increased automation of reporting to government intended to alleviate reporting requirements to multiple agencies – one-stop reporting is the goal
- USACE is lead agency – Institute for Water Resources (IWR) Navigation Data Center (NDC)

CGI

Take-aways

- Understand the scope of AIS identification and measurement data error
- USCG can share AIS data feeds and corrective analysis with partner government agencies today (perhaps with a broader audience in the future)
- Participate in interagency working groups – engage in data sharing and standing up data services for improved efficiencies between agencies
- White paper and additional resources for the USCG Authoritative Vessel Identification Service (AVIS) may be made available on request
- If your agency manages maritime vessel information, standardize / verify against authoritative sources

david.winkler@cgifederal.com
david.m.winkler@uscg.mil

CGI