



# GSICS Action Tracker: A Supervised Machine Learning Content Extraction Tool on Google Cloud

Manik Bali<sup>1</sup>, Lawrence E. Flynn<sup>2</sup>, Mitch Goldberg<sup>2</sup>, Lori Brown<sup>3</sup> and Likun Wang<sup>4</sup>

<sup>1</sup>ESSIC/CICS University of Maryland, [Manik.Bali@noaa.gov](mailto:Manik.Bali@noaa.gov), <sup>2</sup>NOAA/NESDIS, <sup>3</sup>Caelum Research, Riverside Technologies



## Introduction

The Global Space based Inter-Calibration System (GSICS, [gsics.wmo.int](http://gsics.wmo.int)) is a consortium of 16 Satellite agencies that have come together under the aegis of the World Meteorological Organization (WMO, [www.wmo.int](http://www.wmo.int)) and the CGMS to monitor their Earth Observation Satellites.

**The GSICS Action Tracker is a Content extraction tool developed on the NOAA/UMD Google Cloud.**

The main goal is to attain high quality Earth observation data in real time by inter-comparing with in-orbit stable references and applying corrections. An second goal is to construct time series of Climate Quality data in retrospect dating back to early 1970s that can be used to extract signals of Climate change.

Currently there are over 268 (Ref: [oscar.wmo.int](http://oscar.wmo.int)) Earth observing satellites from various agencies across (such as NASA, NOAA, EUMETSAT, ESA, KMA, CMA, JMA, ISRO) the world providing crucial weather data that is used Earth System Monitoring.

It is envisaged that Effective Coordination of inter-calibration activities among Satellite agencies would play a key role in successfully monitoring earth observation satellites globally. This would create a platform wherein suitable adjustments are applied to Earth observation measurements (L1B) to remove trends and biases from them which would positively impact downstream services such as development of Climate Data Records and Environmental Variables.



Earth Observation Satellites of GSICS member agencies ( CMA, KMA, NASA, NOAA, ROSHYDROMET, ESA, EUMETSAT) (Image Courtesy CGMS)

## Challenges in Coordination among agencies

- Over 100 Actions Decisions Recommendation Generated from GSICS Related Meetings
- These are recorded in Meeting Minutes documents with varying formats on multiple platforms such Wiki Pages and Zoho Docs.
- Manual transferring of these takes several weeks of full time effort
- Additional resources are required to maintain the status of the Actions Decisions and Recommendations.

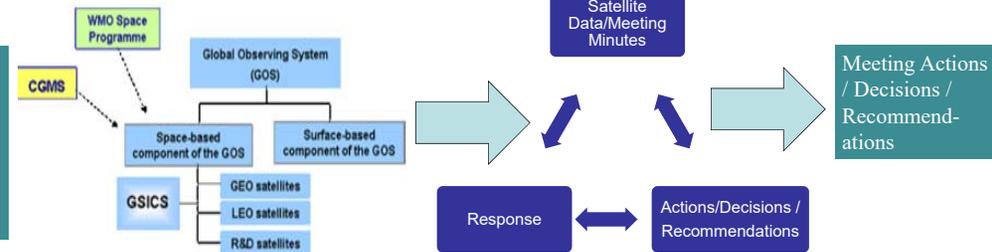


Figure above shows the GSICS Meeting Content data flow from acquisition to its usage

## Proposed Solution

Designed a collaborative paradigm on google cloud for extracting and Managing Meeting minutes and Actions, Decisions and Recommendations by utilizing the features of the Google Cloud Collaboration Platform.



Code a Model (also acts as a classifier) to parse the docs and extract content of interest Actions / Decisions / Recommendations.

Store the Actions / Decisions / Recommendations on a Google Cloud Sheet and publish them on a Webpage

Meeting Minutes prepared collaboratively by satellite agencies across the world.

## Advantages of Google Cloud approach

Reduces overheads as minutes are prepared collaboratively across member countries.

Model is able to parse the google doc and extract Actions/Recommendations/Decisions automatically.

Ease of management -- Scalable approach: Can extract content from large documents at low cost.

Fast search on Actions/Decisions/Recommendations



## Content Extraction Model Algorithm

There are two main platforms of collaboration used by the GSICS members. These are the GSICS Wiki and GSICS Google Cloud. The Action Tracker is designed to parse the google docs as well as GSICS wiki and extract Actions/Decisions and Recommendations from it.

Agenda Item: 1g GOES-16 calibration and validation – 10:25 (20 minutes)	
Presenter	John Fulbright (NOAA, remote)
Overview	<p>John introduced the two main EO instruments on GOES-16 and -17: ABI and GLM. He reviewed the validation schedule for GOES-16, which is now being followed for the recently launched GOES-17 - but without the extended validation period.</p> <p>John highlighted the benefit of ABI's ability to view the Moon and perform N-S scans to validate the calibration and detector normalization.</p> <p>He explained that the VNIR radiometric calibration was initially unstable, due to ground-processing and identified a bias in Band 2, whereas the IR calibration was found to be stable (within 1K accuracy requirements) and in good agreement with CrIS.</p> <p>John also briefly described the airborne validation field campaigns made with the NASA ER-2, which included overflights of the Sanoran desert site along the line of dedicated ABI N-S scans, as well as flights over the ocean, fires and thunderstorms.</p>
Discussion point, conclusions, Actions, Recommendations, Decisions	
Zhangdong Yang (CMA) asked about the data sharing policy of the GOES-R validation campaign, commenting that it could be useful for TANSAT validation.	

Algorithm utilizes Google properties to decompose document into contextual elements.

Agenda Item: 1g GOES-16 calibration and validation – 10:25 (20 minutes)	
Presenter	John Fulbright (NOAA, remote)
Overview	<p>John introduced the two main EO instruments on GOES-16 and -17: ABI and GLM. He reviewed the validation schedule for GOES-16, which is now being followed for the recently launched GOES-17 - but without the extended validation period.</p> <p>John highlighted the benefit of ABI's ability to view the Moon and perform N-S scans to validate the calibration and detector normalization.</p> <p>He explained that the VNIR radiometric calibration was initially unstable, due to ground-processing and identified a bias in Band 2, whereas the IR calibration was found to be stable (within 1K accuracy requirements) and in good agreement with CrIS.</p> <p>John also briefly described the airborne validation field campaigns made with the NASA ER-2, which included overflights of the Sanoran desert site along the line of dedicated ABI N-S scans, as well as flights over the ocean, fires and thunderstorms.</p>
Discussion point, conclusions, Actions, Recommendations, Decisions	
Zhangdong Yang (CMA) asked about the data sharing policy of the GOES-R validation campaign, commenting that it could be useful for TANSAT validation.	

Identifiers are used to Zero in on the relevant content.

A.GWG.2018.1g.1: Fred Wu (NOAA) to discuss with CMA potential sharing of GOES-R field campaign data for TANSAT validation.

A.GWG.2018.1g.1: Fred Wu (NOAA) to discuss with CMA potential sharing of GOES-R field campaign data for TANSAT validation.

## Action Tracker Implementation in GSICS



Meeting Minutes on Google Doc

Extracted Content on Google Sheet



Action Id	Item	Effort Level	Urgency	Summary	Lead	What to Do	Expected Completion	Actual Completion	Deliverable Usage	Status
GOEW.2016.7.1	Data & Pace of next VIG meetings			Peter to setup a template for the agency report which would help focusing on GSICS activities.	EUM(Peter)	Configuration	2017-annual meeting	Closed by email 2017-02-14 here	Template	Closed
GR.2016.3.1	GSIC-Ring dataset analysis			Rob to consider producing an analysis of GSIC-Ring statistics provided by ECMWF as part of GSICS.	EUM(Rob)	Analysis	2017-annual meeting		Not possible in 2017.	Pending
GR.2016.3.2	GSIC-Ring dataset analysis			SUMMIT to coordinate and/or for GSIC-Ring test dataset from all geostationary satellite operators	EUM(Rob)		9/1/2016		File structure setup. Met 7 & 9 real data available - can upload (not link to GSICS corrections).	Pending
GR.2016.34.1	Selecting GSICS References (R, VIG and reviewing terminology)			Tim revision to consider using terminology used in current GSICS Corrections - going demonstration phase.	EUM(Tim)	Configuration	2017-annual meeting	In revised GSICS submission 2017-02-23	Primary GSICS Corrections now defined and anchor References.	Closed
GR.2016.34.1	Handling Diurnal Cycle in GSIC-IR			Fred to report at next meeting	NOAA(Fred)	Analysis	2017-annual meeting			Pending

<https://www.star.nesdis.noaa.gov/smcd/GCC/MeetingActions.php>

### Next Steps

- Scale the Action Tracker to address the needs of the NOAA/STAR and WMO WIGOS System
- WMO has requested GCC to examine if the algorithm could be scaled to extract satellite specifications imbedded in documents.

### Acknowledgements

This work is supported by the NOAA JPSS Program. Authors would like to thank the GSICS Community for the vital inputs they provided in operationalizing the tool

### Conclusions

An Action Tracker that extracts relevant content from documents has been presented  
 The underlying algorithm uses Supervised machine learning algorithm to zero in onto content of relevance  
 The Algorithm can be scaled to extract content from various Earth Observation Satellite related documents.