

ISO/IEC TR 20547-2:2018-01 (E)

Information technology - Big data reference architecture - Part 2: Use cases and derived requirements

Contents		Page
Foreword		v
Introduction		vi
1	Scope	1
2	Normative references	1
3	Terms and definitions	1
3.1	Terms defined elsewhere	1
3.2	Terms defined in this document	1
3.3	Abbreviated terms	1
4	Use case properties for survey	6
4.1	Overall description	6
4.2	Current solution	7
4.3	Big data characteristics	7
4.4	Big data science	7
4.5	Overall big data issues	8
4.6	Big data use case Template	8
5	Use cases summaries	9
5.1	Use case development process	9
5.2	Government operation	10
5.2.1	Use case 1: Census 2010 and 2000 -- Title 13 big data	10
5.2.2	Use case 2: NARA Accession, Search, Retrieve, Preservation	10
5.2.3	Use case 3: Statistical survey response improvement	11
5.2.4	Use case 4: Non-Traditional Data in Statistical Survey Response Improvement (Adaptive Design)	11
5.3	Commercial	12
5.3.1	Use case 5: Cloud Eco-System for Financial Industries	12
5.3.2	Use case 6: Mendeley -- An International Network of Research	12
5.3.3	Use case 7: Multi-media streaming service	13
5.3.4	Use case 8: Web Search	13
5.3.5	Use case 9: Big data Business Continuity and Disaster Recovery Within a Cloud Eco-System	14
5.3.6	Use case 10: Cargo Shipping	14
5.3.7	Use case 11: Materials Data for Manufacturing	14
5.3.8	Use case 12: Simulation-Driven Materials Genomics	15
5.4	Defense	16
5.4.1	Use case 13: Cloud Large-Scale Geospatial Analysis and Visualization	16
5.4.2	Use case 14: Object Identification and Tracking from Wide-Area Large Format Imagery or Full Motion Video--Persistent Surveillance	16
5.4.3	Use case 15: Intelligence Data Processing and Analysis	17
5.5	Health care and life sciences	17
5.5.1	Use case 16: Electronic Medical Record Data	17
5.5.2	Use case 17: Pathology Imaging/Digital Pathology	18
5.5.3	Use case 18: Computational Bioimaging	18
5.5.4	Use case 19: Genomic Measurements	19
5.5.5	Use case 20: Comparative Analysis for Metagenomes and Genomes	19
5.5.6	Use case 21: Individualized Diabetes Management	19

5.5.7	Use case 22: Statistical Relational Artificial Intelligence for Health Care	20
5.5.8	Use case 23: World Population-Scale Epidemiological Study	20
5.5.9	Use case 24: Social Contagion Modeling for Planning, Public Health, and Disaster Management	21
5.5.10	Use case 25: Biodiversity and LifeWatch	21
5.6	Deep Learning and Social Media	22
5.6.1	Use case 26: Large-Scale Deep Learning	22
5.6.2	Use case 27: Organizing Large-Scale, Unstructured Collections of Consumer Phot.....	22
5.6.3	Use case 28: Truthy--Information Diffusion Research from Twitter Data	23
5.6.4	Use case 29: Crowd Sourcing in the Humanities as Source for Big and Dynamic Data	23
5.6.5	Use case 30: CINET--Cyberinfrastructure for Network (Graph) Science and Analytics	23
5.6.6	Use case 31: NIST Information Access Division -- Analytic Technology Performance Measurements, Evaluations, and Standards	24
5.7	The Ecosystem for research	24
5.7.1	Use case 32: DataNet Federation Consortium	24
5.7.2	Use case 33: The Discinnet Process	25
5.7.3	Use case 34: Semantic Graph Search on Scientific Chemical and Text- Based Data	25
5.7.4	Use case 35: Light Source Beamlines	26
5.8	Astronomy and physics	26
5.8.1	Use case 36: Catalina Real-Time Transient Survey: A Digital, Panoramic, Synoptic Sky Survey	26
5.8.2	Use case 37: DOE Extreme Data from Cosmological Sky Survey and Simulations	27
5.8.3	Use case 38: Large Survey Data for Cosmology	27
5.8.4	Use case 39: Particle Physics--Analysis of Large Hadron Collider Data: Discovery of Higgs Particle	28
5.8.5	Use case 40: Belle II High Energy Physics Experiment	29
5.9	Earth, environmental, and polar science	29
5.9.1	Use case 41: European Incoherent Scatter Scientific Association 3D Incoherent Scatter Radar System	29
5.9.2	Use case 42: Common Operations of Environmental Research Infrastructure	30
5.9.3	Use case 43: Radar Data Analysis for the Center for Remote Sensing of Ice Sheets	31
5.9.4	Use case 44: Unmanned Air Vehicle Synthetic Aperture Radar (UAVSAR) Data Processing, Data Product Delivery, and Data Services	31
5.9.5	Use case 45: NASA Langley Research Center/ Goddard Space Flight Center iRODS Federation Test Bed	32
5.9.6	Use case 46: MERRA Analytic Services (MERRA/AS)	32
5.9.7	Use case 47: Atmospheric Turbulence - Event Discovery and Predictive Analytics	32
5.9.8	Use case 48: Climate Studies Using the Community Earth System Model at the U.S. Department of Energy (DOE) NERSC Center	33
5.9.9	Use case 49: DOE Biological and Environmental Research (BER) Subsurface Biogeochemistry Scientific Focus Area	33
5.9.10	Use case 50: DOE BER AmeriFlux and FLUXNET Networks	34
5.10	Energy	34
5.10.1	Use case 51: Consumption Forecasting in Smart Grids	34
5.10.2	Use case 52: Home Energy Management System	34
6	Use cases derived technical considerations	35
6.1	Use case specific technical considerations	35
6.2	Summary of requirements analysis	35
6.3	Features of use cases	37
Annex A Submitted use case studies		40
Annex B Summary of Key Properties		197
Annex C Use case technical considerations summary		207
Annex D Use case detail technical considerations		225
Bibliography		252