Practice update: Building a data integration and visualisation platform for resilience research in New Zealand

Joanne R. Stevenson [1]
Elora Kay [1]
John Vargo [1]

[1] Resilient Organisations Ltd.

Author correspondence:
Resilient Organisations
Unit 2, 188 Durham Street South
Christchurch 8011
New Zealand
Email: joanne.stevenson@resorgs.org.nz

URL: http://trauma.massey.ac.nz/issues/2018-1/AJDTS_22_1_Stevenson.pdf

Abstract

This article summarises the process and progress toward developing the New Zealand Resilience Data Integration and Visualisation En Masse Platform, otherwise referred to as DIVE. The DIVE Platform is a prototype for online data cataloguing, sharing, and collaboration. It is being developed to enable integrated and engaged research that will enhance New Zealand's resilience to hazards. The development of this platform is intended to interface with, and supplement, other efforts to integrate data sharing across New Zealand and beyond. Development of a beta prototype for the DIVE Platform has been completed, allowing users to upload relevant metadata into the system through a data entry form. The prototype features the ability to effectively categorise data. It allows for user friendly data searching and the creation of virtual organisations to facilitate collaborative research. Testing of the beta prototype is currently underway, meaning that end users are already interacting with the web-based DIVE prototype. This stage allows them to provide valuable feedback and showcase examples of resilience research emerging within New Zealand. The beta version also enables users to contribute to the ongoing development of the DIVE Platform itself. The current paper concludes with a discussion of challenges surrounding the development of DIVE, alongside plans for the future development of this platform.

Keywords: *Resilience, data integration platform, metadata*

New Zealand is exposed to a wide range of natural hazards, in no small part because the country straddles the boundary between two tectonic plates with its attendant risk of earthquakes and volcanic eruptions. The New Zealand Government has committed across several platforms (for example: The Resilience to Nature's Challenges National Science Challenge; the New Zealand Centre for Earthquake Resilience, or QuakeCoRE; the Ministry of Civil Defence and Emergency Management's National Resilience Strategy; the Sendai Framework for Disaster Risk Reduction) to building the resilience of its people, places, and economy to ensure safety, stability, and prosperity in the face of significant exposure to disruption. Understanding the current state of the nation's resilience and achieving systemic improvements requires cross-institutional and transdisciplinary collaboration and research innovation. These aspects of systematic improvement, however, present challenges to the status quo of data sharing and management (Medyckyj-Scott et al., 2016). Information inefficiencies and gaps hinder the progress of those tackling the most complex and important issues in disaster risk reduction (DRR) and resilience building. Meaningful progress toward positive DRR outcomes requires collaboration across institutions and disciplines, and requires effective information management. This means creating spaces where data can be captured, safely shared, and managed to ensure quality, appropriate use, and ongoing development.

The Data Integration and Visualisation En Masse (DIVE) Platform aims to create such a space. This Platform is a research and development programme designed to examine how information could be managed to enhance the impact of the research emerging from QuakeCoRE and the Resilience to Nature's Challenges (RNC) - National Science Challenge. DIVE was developed between 2016 and 2018 through a series of consultations and iterative development phases. We have released several tools associated with DIVE on a web-based platform. The web-based DIVE Platform includes: a metadata[1] catalogue and data repository, capacity to search and access resilience data, and

1 Data about data, including keywords, data collection dates and other information.

a series of data literacy resources to upskill the user community. The current paper provides an overview of the DIVE Platform's development, it's current web-based offerings, and the platform's future. We also draw on the experiences of others working in similar areas, to frame how we will promote uptake and guide future developments of the DIVE Platform.

Background

In 2011, the New Zealand government approved a set of principles, asserting that the data and information it holds should be open, readily available, well managed, reasonably priced, and re-usable unless there is a reasonable expectation of information protection (Internal Affairs Te Tari Taiwhenua, 2011). This is underscored by the Open Government Data Programme which champions an *open by default* cross-government data ecosystem. Policy recommendations from the Ministry of Business Innovation and Employment (MBIE) extend this ethos to publicly-funded science data, with the intention of managing and sharing information for better collaborations, more efficient and powerful science, and greater connectedness to end-users (Ministry of Research Science and Technology, 2010). New Zealand has also become a signatory to the Sendai Framework for Disaster Risk Reduction, which promotes "real time access to reliable data, the use of space and in situ information, including geographic information systems (GIS), and use information and communications technology innovations to enhance measurement tools and the collection, analysis and dissemination of data," (UNSIDR, 2015, p.14).

There are several ongoing projects designed to enhance the visibility and usability of data across New Zealand. For example, significant effort and resources have been invested across Government to construct the Integrated Data Infrastructure (IDI) (Statistics New Zealand, 2017a) and the Longitudinal Business Database microdata repositories (Statistics New Zealand, 2017b). Additionally, DigitalNZ (The National Library of New Zealand, 2018) and data.govt.nz (2018) are data cataloguing and search platforms designed to make New Zealand community and Government data easier to find.

There have also been numerous information management portals specifically designed in New Zealand and abroad, for hazards data. The New Zealand Geotechnical Database (NZGD, 2018) and the post-earthquake data clearinghouse system hosted by The Earthquake Engineering Research Institute (EERI) are largely driven by the earthquake engineering community in private industry, academia, and crown research institutes. The EERI established the Kaikoura Earthquake Virtual Clearinghouse website for those wishing to publish information relevant to the public and international researchers (Kaikoura, NZ Earthquake Clearinghouse, 2018). EERI data clearinghouses are geared towards initial data capture after an earthquake and are not updated with recovery data.

The Canterbury Earthquake Digital Archive (CEISMIC) mainly catalogues documentary resources, with an emphasis on images, news media, video and audio files, and cultural heritage collections associated with the 2010/2011 Canterbury earthquake sequence. Other aggregation portals for hazards data in New Zealand include NIWA's Historic Weather Events Catalogue, GeoNet's Geohazards applications and data, and portals for data captured by the public, such as GeoNet's Felt Reports and NIWA's community air quality observation network (GeoNet, 2018; NIWA, 2018a, 2018b). Internationally, programmes like the AGORA project, DataONE, and the Sahana Foundation provide interdisciplinary research and information networks to enhance environmental and community outcomes in the face of disruption (AGORA, 2018; DataONE, 2018; Sahana Foundation, 2018).

Despite these efforts, New Zealand does not yet have a unified space where researchers and research stakeholders can share and locate the information they are collecting across multiple hazards, in a way that is curated and archived. As a result, in 2016, the NZ Centre for Earthquake Resilience (QuakeCoRE) and the Resilience to Nature's Challenges (RNC) – National Science Challenge, funded a small team of researchers to investigate how to best enable teams of researchers to access and share data and information to enhance resilience outcomes for New Zealand.

Procedures

The DIVE research team used a design-thinking approach to scope and design the first prototype. Design thinking is a solution-focused process which incorporates the in-depth insights of end-users into iterative prototype development (Brown & Wyatt, 2010). Proponents of design thinking refer to this process as a system of overlapping spaces, as opposed to sequential steps. Brown and Wyatt (2010) identify several spaces in the design thinking process: *Inspiration,* "the problem

or opportunity that motivates the search for solutions" (p. 33); *Ideation*, "the process of generating, developing, and testing ideas" (p. 33); and *implementation,* "the path that leads from the project stage into people's lives" (p. 33). Each of these spaces were explored in the development of the DIVE Platform.

Inspiration

Between March and November 2016, the team initiated a consultation process involving workshops, surveys, interviews, and software prototype design and testing. The inspiration phase begins with a 'brief', a general framework of constraints and goals of the design process and benchmarks against which progress can be measured. For DIVE, the brief was to: Create data management systems that enable teams of researchers to address complex social problems that make New Zealand more resilient to hazards and disasters. The standard against which this system was benchmarked is whether it is useful, usable, and used. These principles were loosely defined in the early phase of the project and can be guided by a series of questions based on an evolving understanding of the system, as shown in Table 1.

Table 1.
Principles for Evaluating the Success of DIVE

Criteria	Description
Useful	Is the data up-to-date? Is the quality of the data being managed? Can others understand and use the data that is being uploaded?
Useable	Is data searchable and accessible for a wide range of users? For example, can it facilitate 'citizen science' or council data collection efforts if that is what the users need?
Used	Are communities of practice being established and self-sustaining?

Once the brief is set, the inspiration phase explores the needs of stakeholders through direct consultation and observation. The initial stakeholder group was comprised of QuakeCoRE and RNC researchers and representatives from several key data providing organisations, such as government ministries, local councils, and the Earthquake Commission. We began this phase with a series of workshops and an assessment of the way stakeholders were interacting with other data management systems. These procedures were supplemented by an online survey and several informal interviews with subject matter experts to gather more in-depth information about user-needs and processes that may be useful. Detailed results of this data collection and subsequent analysis are summarised in Stevenson, Vargo, & Brown (2016) and Stevenson et al. (2017).

Ideation

The ideation phase involves synthesizing and translating outcomes of the inspiration phase into visions and choices that can guide the design of a system (Brown & Wyatt, 2010). We developed a series of operational use cases, which are user journeys that describe a flow of operations for interacting with a system, and that can be used to identify the functions, operating systems, boundaries, and constraints that are relevant to potential users (Summers, 2012). The operational use cases focused on important data-related challenges facing QuakeCoRE and RNC researchers. More details concerning these use cases can be found in Stevenson, Vargo, Thomson, and Walsh (2017). In a second stakeholder workshop in mid-2016, we presented the synthesized findings and use cases to collaboratively examine workflow and problem-solving processes. We also received and incorporated additional feedback from participants at this stage.

Implementation

The subsequent, implementation, phase involved developing a pre-alpha prototype[2] of the DIVE software system. In a third workshop the DIVE Platform development team demonstrated the DIVE prototype's current features and facilitated discussion about how such a system could be used to address key issues in resilience research. Approximately 65 stakeholders and subject matter experts contributed to this initial DIVE development process. In addition to informing the design of the DIVE Platform, the workshops were also a helpful starting point for forming relationships that will be central to the successful implementation of a transdisciplinary and cross-institutional collaborative platform.

In December 2017, the DIVE programme received additional funding from QuakeCoRE to develop the beta prototype[3]. The ensuing project's objectives were to:

1) Develop metadata entry and data sharing processes
2) Develop a prototype web-based metadata interface and data catalogue
3) Continue producing educational content for the metadata and geospatial data literacy programme to increase DIVE platform uptake

2 Pre-alpha refers to all software development activities before formal testing.

3 A *beta* prototype is a product that is closer to completion compared to an alpha prototype and is often tested by product end users.

We are hosting the DIVE platform on the Comprehensive Knowledge Archive Network (CKAN), which is an open source data platform. CKAN makes it possible for DIVE users to process register their user information, upload and share their metadata (with an optional upload data function), search metadata, and categorize data and users to better enable collaborations.

The focus was on delivering a minimum viable product that would allow users to catalogue metadata for all resilience-related material and share critical unique datasets. For example, a dataset of heritage churches of the Anglican diocese in New Zealand was entered into the beta prototype. This dataset contains information including the location, year built, construction type, proneness to earthquake damage, notes about vulnerabilities and other useful and vital data. Datasets such as these are often unavailable in the public domain but may enable innovative research in disaster risk reduction and resilience. The front-end layout of DIVE was also updated from the pre-alpha to beta stages, creating an interface that was much more visually engaging (see Figure 1).

The use of metadata has allowed researchers to share information about past and current research endeavours even if they cannot share datasets themselves. The current beta prototype contains a comprehensive data entry form, allowing users to enter high quality metadata, improving the ability for others to interpret and build on past and current resilience research. The *organisations* feature has allowed users to see and search all of the datasets from a particular organisation in one place. Additionally, it has allowed members of that organisation to categorise their data by selecting their respective organisation during the data entry process. Users are now also able to create their own groups and compile pre-existing DIVE datasets into one readily accessible folder. This feature has the potential to facilitate collaboration through the cataloguing of datasets for a project, team, or data-related theme. The development of this minimum viable product has led to a beta version that is now ready to be delivered to QuakeCoRE and RNC researchers for testing. CKAN's ability to link to other CKAN based platforms has also allowed the DIVE team to enable searches across more than one platform.

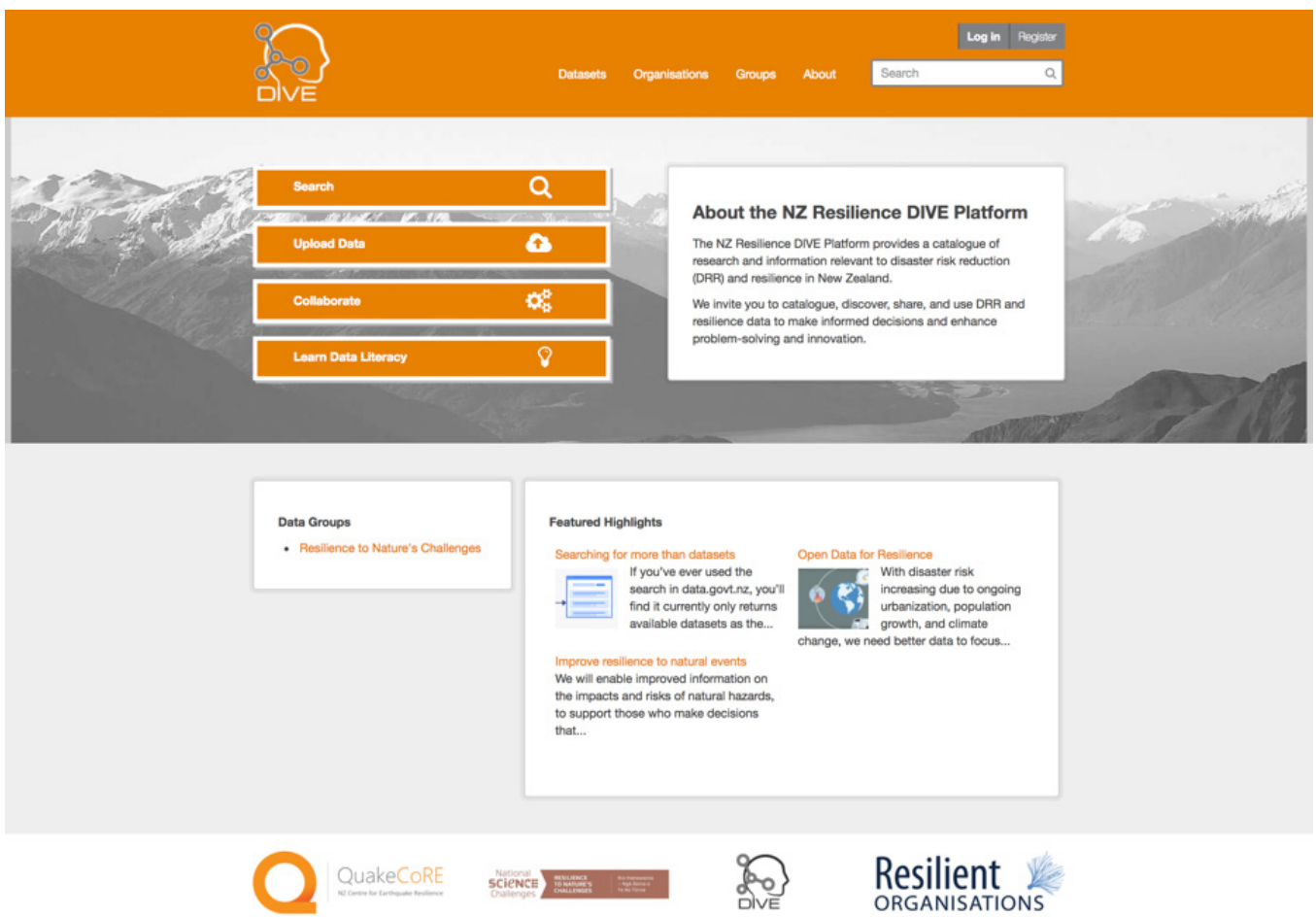

Figure 1. The DIVE homepage (beta version), showing the visual style of the beta prototype.

For example, data.govt.nz is one other platform that also uses CKAN, and which may now be included in multi-platform, or federated, searches.

Ideation: Beta testing phase

During the beta testing phase, researchers from within QuakeCoRE and the RNC will be encouraged to enter their data into DIVE using a step by step guide and assistance from the researchers. This phase will allow us to elicit feedback from users regarding the ease of data entry, and will allow us to showcase some of the exciting and important work being undertaken within New Zealand. If possible, the DIVE prototype will be further refined based on user feedback. Current funding constraints may mean that further software development may not be viable.

DIVE is currently available in a live production environment and is accessible to anyone through the website www.resiliencedata.org.nz. The feedback gained from the beta testing phase of the production environment will help us develop future use cases which can be used to justify ongoing funding for continued DIVE development. We also stress to our testing users that no effort will be lost, even if future development funding is not secured. DIVE has deliberately been created in a CKAN platform to ensure that the metadata catalogue can be transferred to another host such as data.govt.nz in the eventuality that the original DIVE Platform is decommissioned.

Observations, Conclusions & Recommendations

The outcomes of a system intended to enable better resilience research are likely to be enhanced by a problem-focus, rather than being divided by funding or disciplinary boundaries. A problem-focused system will enhance the visibility of the work going on to improve the resilience of New Zealand. It will be a place where communities of researchers, decision makers, data holders, private industry, and citizen scientists can view, upload, and download data. Such a system should facilitate the creative collision of secondary and primary research data, local narratives, real-time hazard monitoring, *Mātauranga Māori* indigenous knowledge, and multi-media information.

There are, of course, challenges to promoting and maintaining a system like DIVE. There has been extensive research on the factors that influence technology acceptance among end-user populations, including the system's perceived usefulness, perceived ease of use (Schweik et al., 2005), social influences (Venkatesh et al. 2003), the way the project is communicated to potential users, and their belief in the system (Seymour et al., 2007). DIVE uptake management will draw on lessons from previous work such as research by Prasanna and Huggins (2016), whose work on information system adoption in emergency operations centres found that there were several factors mediating and moderating technology adoption. They found, for example, that performance expectancy significantly impacted information system acceptance. As a result, they recommend providing clear guidance for system implementation and investing in efforts that will boost performance expectancy (Prasanna & Huggins, 2016).

Although many researchers within the RNC, QuakeCoRE, and other research programmes have indicated that a system like DIVE is needed, one anticipated challenge is that potential users may not believe they have time to input their data into the platform. This is a common issue faced by metadata collection initiatives (Schweik et al., 2005; Jones & Vines, 2016). To capture good quality metadata, multiple available input fields should be populated within the data entry form. This will allow future users to fully understand each dataset in line with Dublin Core standards. Dublin Core is a well-established, set of internationally recognised metadata standards that recommends fifteen general properties for resource descriptions that are broad enough for almost any type of resource (Dublin Core Metadata Initiative, 2012). It is possible that the length of the data entry form may deter some researchers, many of whom may already be experiencing time pressure in their current projects. For DIVE, this has meant making only the most essential metadata items mandatory. Users are given the option of adding additional detail.

The development of the beta prototype of the DIVE Platform was funded by QuakeCoRE. Our current development budget has been fully allocated, and useful features that could be developed in the future cannot be enabled until additional funding is available. We propose the continued funding and development of DIVE, towards an interactive online space for researchers and practitioners to share information relevant to their ongoing research and the information gathered from disaster events as they unfold. Capturing this data in a federated portal that is curated, properly archived, and strategically shared will facilitate future research, aid

response and recovery actions and decision making, and may become a resilience building tool when broader communities are able to contribute data on the hazards they are experiencing or relevant trends that they are seeing unfold in their communities.

Acknowledgements

The work discussed in this practice update was funded by the Tertiary Education Commission through the QuakeCoRE and by the MBIE through the Resilience to Nature's Challenges - National Science Challenge. The authors would like to extend sincere thanks to everyone who attended workshops, filled out surveys, and shared their expertise throughout the development of the DIVE Platform.

References

AGORA Project. (2018). *A geospatial open collaborative architecture for building resilience against disasters and extreme events*. Retrieved from www.agora.icmc.usp.br/site/agora/.

Brown, T., & Wyatt, J. (2007). Design thinking for social innovation. *Stanford Social Innovation Review*, *12*, 29–43. doi: 10.1596/1020-797X_12_1_29

Data.govt.nz. (2018). *Discovery and use data*. Retrieved from https://data.govt.nz/.

DataONE. (2018) Retrieved from www.dataone.org/data-life-cycle%209

Dublin Core Metadata Initiative. (2012). *Dublin Core Metadata Element Set, version 1.1: Reference description*. Retrieved from http://dublincore.org/documents/dces/

GeoNet. (2018). *GeoNet: Geological hazard information for New Zealand*. Retrieved from www.geonet.org.nz/

Internal Affairs Te Tari Taiwhenua. (2011). *New Zealand Data and Information Management Principles*. Retrieved from www.ict.govt.nz/guidance-and-resources/open-government/new-zealand-data-and-information-management-principles/.

Jones, M., & Vines, R. (2016). Cultivating capability: The socio-technical challenges of knowledge management. *Records Management Journal*, *26*, 242–258. doi: 10.1108/RMJ-11-2015-0035

Kaikoura, New Zealand Earthquake Clearinghouse. (2018). Retrieved from www.eqclearinghouse.org/2016-11-13-kaikoura/.

Medyckyj-Scott, D., Stock, K., Gibb, R., Gahegan, M., Dzierzon, H., Schmidt, J., & Collins, A. (2016). *Our Land and Water National Science Challenge: A data ecosystem for land and water data to achieve the challenge mission.* Wellington, New Zealand: Landcare Research.

Ministry of Research Science and Technology. (2010). *Environmental Data Management Policy Statement (Report SW230201)*. Retrieved from www.mbie.govt.nz/publications-research/publications/science-and-innovation/MBIE_Environmental-Data-Management-Policy-Statement.pdf/view?searchterm=environmental%20data

The National Library of New Zealand. (2018). *DigitalNZ A-Tihi O Aotearoa*. Retrieved from https://digitalnz.org/.

NIWA. (2018a). *Community observation networks for air*. Retrieved from https://www.niwa.co.nz/atmosphere/research-projects/CommunityObservationNetworksforAir

NIWA. (2018b). *NZ Historic Weather Events Catalogue*. Retrieved from https://hwe.niwa.co.nz/

NZGD. (2018). *New Zealand Geotechnical Database*. Retrieved from www.nzgd.org.nz/Registration/Login.asp

Prasanna, R., & Huggins, T.J. (2016). Factors affecting the acceptance of information systems supporting emergency operations centres. *Computers in Human Behavior, 57*, 169-181. doi: 10.1016/j.chb.2015.12.013

Sahana Foundation. (2018). *Sahana Foundation open source disaster management software*. Retrieved from https://sahanafoundation.org/

Schweik, C. M., Stepanov, A., & Grove, J. M. (2005). The open research system: a web-based metadata and data repository for collaborative research. *Computers and Electronics in Agriculture*, *47*, 221–242. doi: 10.1016/j.compag.2004.12.006

Seymour, L., Makanya, W., & Berrange, S. (2007). End-users' acceptance of enterprise resource planning systems: An investigation of antecedents. In *Proceedings of the 6th annual ISOnEworld conference* (p. 26). Las Vegas, NV: ISOnEworld Conference.

Statistics New Zealand. (2017a). *Longitudinal business database (1 January 2017)*. Retrieved from www.stats.govt.nz/integrated-data/longitudinal-business-database/.

Statistics New Zealand. (2017b). *Integrated data infrastructure (16 November 2017)*. Retrieved from www.stats.govt.nz/integrated-data/integrated-data-infrastructure/

Stevenson, J. R., Vargo, J., & Brown, C. (2016). *Data integration and visualisation: Prototyping the QuakeCoRE Data Platform for diverse needs*. Retrieved from https://wiki.canterbury.ac.nz/pages/viewpage.action?pageId=52692301

Stevenson, J.R., Vargo, J., Thomson, C., & Walsh, L. J. (2017). *Resilience & data in New Zealand: The Data Integration and Visualisation En Masse (DIVE) platform 2016 summary*. Retrieved from https://wiki.canterbury.ac.nz/pages/viewpage.action?pageId=52692301.

Summers, B. L. (2012). *Effective methods for software and systems integration*. Boca Raton, FL: CRC Press.

UNISDR. (2015). *Sendai Framework for Disaster Risk Reduction 2015– 2030*. Retrieved from www.unisdr.org/we/coordinate/sendai-framework.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: toward a unified view. *Management Information Systems Quarterly*, *27*, 425-478. doi: 10.2307/30036540