

Design goals, technical implementation and practical use of Another data entry system (Andes)

Daniel Ricard, David Fishman, Lindsay Beazley, Brian Boivin, Jamie Emberley, Ryan Martin, Jenni McDermid, Nicolas Rolland, David Sean-Fortin, Quentin Stoyel and Pablo Vergara

Science Branch
Gulf Region
Fisheries and Oceans Canada
Moncton, New Brunswick, E1C 5K4, Canada

2023

Canadian Technical Report of Fisheries and Aquatic Sciences ####



Fisheries and Oceans
Canada

Pêches et Océans
Canada

Canada

Canadian Technical Report of Fisheries and Aquatic Sciences

Technical reports contain scientific and technical information that contributes to existing knowledge but which is not normally appropriate for primary literature. Technical reports are directed primarily toward a worldwide audience and have an international distribution. No restriction is placed on subject matter and the series reflects the broad interests and policies of Fisheries and Oceans Canada, namely, fisheries and aquatic sciences.

Technical reports may be cited as full publications. The correct citation appears above the abstract of each report. Each report is abstracted in the data base *Aquatic Sciences and Fisheries Abstracts*.

Technical reports are produced regionally but are numbered nationally. Requests for individual reports will be filled by the issuing establishment listed on the front cover and title page.

Numbers 1-456 in this series were issued as Technical Reports of the Fisheries Research Board of Canada. Numbers 457-714 were issued as Department of the Environment, Fisheries and Marine Service, Research and Development Directorate Technical Reports. Numbers 715-924 were issued as Department of Fisheries and Environment, Fisheries and Marine Service Technical Reports. The current series name was changed with report number 925.

Rapport technique canadien des sciences halieutiques et aquatiques

Les rapports techniques contiennent des renseignements scientifiques et techniques qui constituent une contribution aux connaissances actuelles, mais qui ne sont pas normalement appropriés pour la publication dans un journal scientifique. Les rapports techniques sont destinés essentiellement à un public international et ils sont distribués à cet échelon. Il n'y a aucune restriction quant au sujet; de fait, la série reflète la vaste gamme des intérêts et des politiques de Pêches et Océans Canada, c'est-à-dire les sciences halieutiques et aquatiques.

Les rapports techniques peuvent être cités comme des publications à part entière. Le titre exact figure au-dessus du résumé de chaque rapport. Les rapports techniques sont résumés dans la base de données *Résumés des sciences aquatiques et halieutiques*.

Les rapports techniques sont produits à l'échelon régional, mais numérotés à l'échelon national. Les demandes de rapports seront satisfaites par l'établissement auteur dont le nom figure sur la couverture et la page du titre.

Les numéros 1 à 456 de cette série ont été publiés à titre de Rapports techniques de l'Office des recherches sur les pêcheries du Canada. Les numéros 457 à 714 sont parus à titre de Rapports techniques de la Direction générale de la recherche et du développement, Service des pêches et de la mer, ministère de l'Environnement. Les numéros 715 à 924 ont été publiés à titre de Rapports techniques du Service des pêches et de la mer, ministère des Pêches et de l'Environnement. Le nom actuel de la série a été établi lors de la parution du numéro 925.

Canadian Technical Report of
Fisheries and Aquatic Sciences nnn

2023

6 DESIGN GOALS, TECHNICAL IMPLEMENTATION AND PRACTICAL USE OF ANOTHER DATA 7 ENTRY SYSTEM (ANDES)

by

Daniel Ricard¹, David Fishman¹, Lindsay Beazley², Brian Boivin³, Jamie Emberley⁴, Ryan Martin⁴, Jenni McDermid¹, Nicolas Rolland¹, David Sean-Fortin³, Quentin Stoyel¹ and Pablo Vergara¹

¹² ¹³ ¹Gulf Fisheries Centre, Fisheries and Oceans Canada, 343 Université Avenue, Moncton, New Brunswick, E1C 9B6

¹⁴ ¹⁵ ²Bedford Institute of Oceanography, Fisheries and Oceans Canada, 1 Challenger Dr., Dartmouth,
Nova Scotia, B2Y 4A2

¹⁶ ³Institut Maurice Lamontagne, Fisheries and Oceans Canada, 850 Rte de la Mer, Mont-Joli,
¹⁷ Québec, QC G5H 3Z4

¹⁸ Saint Andrews Biological Station, Fisheries and Oceans Canada, 125 Marine Science Dr.,
¹⁹ St. Andrews, New Brunswick, E5B 0E4

20 © His Majesty the King in Right of Canada, as represented by the Minister of the
21 Department of Fisheries and Oceans, 2023
22 Cat. No. Fs97-6/nnnE-PDF ISBN XXX ISSN 1488-5379

23 Correct citation for this publication:

24 Ricard, D., Fishman, D., Beazley, L., Boivin, B., Emberley, J., Martin, R., McDermid, J., Rolland,
25 N., Sean-Fortin, D., Stoyel, Q. and Vergara, P. 2023. Design goals, technical implementation
26 and practical use of Another data entry system (Andes). Can. Tech. Rep. Fish. Aquat. Sci.
27 nnn: vi + 41 p.

CONTENTS

29	ABSTRACT	v
30	RÉSUMÉ	vi
31	1 Introduction	1
32	2 Methods	2
33	2.1 Flexibility, scalability and reliability	3
34	2.2 Version control/source control	4
35	2.3 Unit Testing	4
36	2.4 Backup strategy	4
37	2.5 Customizable protocols	4
38	2.6 Quality control	5
39	2.7 User interface	5
40	2.8 Reactivity	5
41	2.9 Multilingualism	6
42	3 Results	6
43	3.1 Architecture	6
44	3.2 Andes Modules	7
45	3.2.1 Bridge	7
46	3.2.2 Set manager	8
47	3.2.3 Ecosystem survey	8
48	3.2.4 Shrimp ecosystem survey app	8
49	3.2.5 Oceanography	9
50	3.2.6 Forecasting utility modules	9
51	3.2.7 Port sampling	10
52	3.3 Technical Implementation of Other Design Goals	11

53	3.3.1 Version control/source control	11
54	3.3.2 Unit testing	11
55	3.3.3 Backup strategy	11
56	3.3.4 Customizable protocols	12
57	3.3.5 Quality control	12
58	3.3.6 Reactivity	13
59	3.3.7 Multilingualism	13
60	4 Discussion	14
61	4.1 Integration of Andes with Existing Data Repositories	15
62	4.2 Future Directions	15
63	4.2.1 Atlantic Zone (Offshelf) Monitoring Program (AZMP/AZOMP)	
64	oceanographic surveys	15
65	4.3 Governance	16
66	5 Acknowledgments	16
67	6 References	17
68	7 Tables	18
69	8 Figures	19
70	9 Acronyms	37
71	10 Glossary	38
72	APPENDICES	41
73	A Tasks for administrative user	41
74	B Further resources about Andes	41

ABSTRACT

76 Ricard, D., Fishman, D., Beazley, L., Boivin, B., Emberley, J., Martin, R., McDermid, J., Rolland,
77 N., Sean-Fortin, D., Stoyel, Q. and Vergara, P. 2023. Design goals, technical implementation
78 and practical use of Another data entry system (Andes). Can. Tech. Rep. Fish. Aquat. Sci.
79 nnn: vi + 41 p.

80 An event logger and data entry system to facilitate the capture of information collected during
81 scientific field work conducted by DFO Gulf Region was developed. The application, called
82 "Another data entry system" (Andes) was implemented as a browser-based environment that
83 facilitates its deployment in a variety of data capture scenarios. Here, we describe the design
84 goals and implementation details of Andes, and present a number of usage cases. Although
85 the system is tailored primarily for use on fishery-related ecosystem surveys, the flexibility of
86 the application in terms of adapting to new gear types and deployment options promotes its use
87 in a wide variety of applications. To promote transparency, longevity, endorsement and use of
88 Andes by other scientific groups, the application utilizes Open Source software and a shared
89 code development platform.

RÉSUMÉ

91 Ricard, D., Fishman, D., Beazley, L., Boivin, B., Emberley, J., Martin, R., McDermid, J., Rolland,
92 N., Sean-Fortin, D., Stoyel, Q. and Vergara, P. 2023. Design goals, technical implementation
93 and practical use of Another data entry system (Andes). Can. Tech. Rep. Fish. Aquat. Sci.
94 nnn: vi + 41 p.

95 Un système d'entrée de données pour faciliter la saisie des informations recueillies lors des
96 activités scientifiques menées par la Région du Golfe du MPO a été développé. L'application, qui
97 s'appelle "Another data entry system" (Andes), a été mise en œuvre en tant qu'environnement
98 basé sur un fureteur, ce qui facilite son déploiement dans une variété de scénarios de saisie de
99 données. Les objectifs de conception et les détails de mise en œuvre d'Andes sont décrits, et
100 un certain nombre de cas d'utilisation sont présentés. Pour promouvoir l'utilisation d'Andes par
101 d'autres groupes scientifiques et assurer la pérennité du projet, l'application s'appuie sur des
102 logiciels code source ouvert et utilise une plate-forme de développement de code partagé.

103

1 Introduction

104 The scientific activities conducted by the Atlantic regions of Fisheries and Oceans Canada (DFO)
105 include ecosystem surveys, oceanographic monitoring surveys, sentinel fisheries programs,
106 coastal surveys, port sampling and many others where field data are collected. These monitoring
107 programs require the collection of a wide variety of information on the marine environment and
108 ecosystems, and are used to inform science-based decision-making. To support the collection
109 of data and metadata during those surveys, paper-based systems were initially used during field
110 work on research vessels, fishing vessels or at commercial ports. With the advent of computing
111 capabilities, paper data sheets were digitized and the resulting data was stored for subsequent
112 analyses. In the 1980s and early 2000s, a number of computer-based data entry systems were
113 developed, in particular to support the activities of scientific ecosystem cruises where a variety
114 of instruments are used to collect data. For scientific trawl surveys, this meant a tool that could
115 support the activities associated with processing the catches from the net, and that could also
116 be used to manage other survey activities while at sea. In the Maritimes and Gulf regions of
117 DFO, the Groundfish Survey Entry (GSE) and later the Ecosystem Survey Entry (ESE) were
118 deployed on workstations in the wet laboratory of research ships. These applications supported
119 the capture of data and information on individual specimens, including the collection of length
120 frequencies and more detailed sampling based on the survey goals and associated protocols. In
121 the Québec Region of DFO, a similar tool called the "Module des Relevés de Recherche" (MRR)
122 was developed and used to directly enter data in digital format on the research vessel.

123 The later versions of those data entry systems were built with Visual Basic for Applications (VBA)
124 and relied on Microsoft Access databases. While highly functional, both the ESE and the MRR
125 have important limitations:

- 126 • Despite being used over a local area network (LAN), these tools are not designed to work
127 in a networked configuration.
- 128 • These tools are unilingual (French or English depending on the DFO region) and their
129 infrastructure does not readily support multiple languages in their user interface.
- 130 • Pre-.NET Visual Basic (VB), the main development framework used to build the
131 applications, are dated programming languages. To adapt the software so that modern
132 libraries can be used and to develop an Application Programming Interface (API) in VB will
133 become increasingly difficult over time.
- 134 • The use of a Version Control System (VCS) is incompatible with the Microsoft Access
135 applications that VB applications rely on. The absence of a VCS presents serious
136 challenges for managing, disseminating and troubleshooting versions of an application
137 across multiple platforms and is a major impediment for a collaborative approach to
138 development.
- 139 • While these applications were built by DFO, multiple attempts to obtain support for these
140 tools from Information Management and Technology Services (IMTS) (now Chief Digital
141 Officer Sector, CDOS) were unsuccessful. This lack of support means that technicians
142 often resort to either using a spreadsheet application such as Excel, or use a pen and
143 paper to perform activities unsupported by the applications. Performing data entry using

144 separate and disconnected methods means that the data has to be manually consolidated
145 at a later date which is time-consuming and prone to error.

- 146 • The lack of a true server-side application (e.g. web application) means that agile
147 development is effectively impossible; especially in the context of a scientific mission that is
148 underway.

149 While the status quo was meeting the immediate data capture needs of DFO scientists, a
150 proactive stance towards addressing the above spurred the development of *Another data entry*
151 system (Andes). This project effectively began in the summer of 2018 as an application to be
152 deployed on field tablets to support port sampling data collection, and later evolved into an
153 application with a wider range of uses. The first field deployment of Andes for an at-sea scientific
154 mission took place in the fall of 2019 during the southern Gulf of St. Lawrence September
155 ecosystem survey. In its pilot year, Andes was used in parallel to the ESE on board CCGS
156 *Teleost* and was tested and modified to ensure that it performed as intended in supporting data
157 entry during scientific activities.

158 This report documents the design principles that guided the development of Andes, provides
159 technical details about its implementation and details the usage cases encountered thus far. It
160 provides guidelines for assessing the suitability of Andes for novel applications (e.g., other field
161 activities and sampling programs) and discusses the ways in which the application could gain
162 further adoption in supporting a diverse range of data collection activities within DFO.

163 **2 Methods**

164 The overall goals for the redevelopment of the data logging system used in the DFO Gulf Region
165 were to utilize a contemporary software development environment, facilitate transparency
166 through use of a shared code development environment, have the ability to support multiple
167 languages, and to support flexible deployments. The envisioned application would be
168 modularized into different consoles to allow for a diverse range of uses, while leveraging the
169 same core features and infrastructure that the existing application were built on. As the starting
170 point of Andes was the ESE which was used during the annual September survey conducted
171 by DFO Gulf Region, the replacement data entry software had to first replicate all the proven
172 functional capabilities of the ESE. Early versions of Andes achieved those goals and the system
173 was further developed over its usage by scientists in the Gulf Region.

174 The genesis of Andes was to support data collection during commercial port sampling activities
175 (Benoît and Daigle 2007). To provide an alternative to the paper-based forms used, an
176 application that could deployed on field tablets for scientific staff that collect data related to
177 commercial fishing activities was developed. A typical usage case for such an application is
178 to obtain length-frequency samples from commercial fisheries landings, either onboard fishing
179 vessels, on wharves or in fish processing plants. Sampling protocols often include additional
180 requirements such as length-stratified sampling of hard parts such as otoliths, or the collection of
181 whole individuals for later processing in the laboratory. The data structure, protocol flexibility,
182 and deployment requirements of this initial port sampling application provided many of the
183 foundational components for other usage cases.

184 Andes was designed with 3 main usages in mind: A) as the main data entry system for
185 ecosystem survey observations B) to track and capture data and metadata from oceanographic
186 observations, and C) for commercial port sampling activities. For use on the ecosystem surveys,
187 the application must allow for the capture of all information and parameters detailed in the
188 sampling protocols of the survey (e.g. Hurlbut and Clay (1990)). This entails the capture of all
189 information related to fishing activities, plus ancillary variables such as weather, sea state, and
190 water temperature. As catch contents are sorted and identified during the ecosystem surveys,
191 the data capture application must support users in obtaining catch weights and abundance
192 by taxon, individual observations of a specimen's length, weight, maturity status, etc. The
193 application should also be designed to flag users when specimens are to be collected, or when
194 additional sampling requirements were present (e.g. collecting fish otoliths, gonad samples, fin
195 clips, etc.).

196 A large component of the ecosystem surveys conducted by the Gulf, Québec, Maritimes, and
197 Newfoundland and Labrador Regions includes the collection of physical, chemical, and biological
198 oceanographic data in support of the Atlantic Zone Monitoring Program (AZMP) (Therriault et al.
199 1998). A CTD/Rosette system is deployed at a subset of fishing stations where vertical profiles
200 of temperature, conductivity (salinity), dissolved oxygen, chlorophyll and other parameters
201 are collected. Water samples (e.g., nutrients, dissolved oxygen, salinity) are collected at
202 predetermined depths using the Rosette system and attached Niskin bottles, and plankton
203 samples are also collected via vertical ring net tows. These data support annual state of the
204 ocean reporting conducted by the [AZMP](#), and are also used in various stock assessments and
205 ecosystem research projects to provide additional ancillary variables for interpreting the captures
206 data.

207 Historically, the event metadata associated with the deployment of oceanographic equipment
208 on the ecosystem surveys was recorded using the Electronic Logbook ([ELOG](#)) system, a URL-
209 based logging system developed by the Paul Scherrer Institute, Switzerland. However, ELOG
210 was installed and operated completely separately from logging systems used for the biological
211 data, making it difficult to merge the two datasets upon completion of a survey. Ideally, the new
212 application would streamline the capture of oceanographic data and simplify its association with
213 the corresponding biological data.

214 The following section outlines additional user requirements and design principles that also played
215 a role in decision-making during the development of Andes.

216 **2.1 Flexibility, scalability and reliability**

217 The development environment used for the application must be a contemporary programming
218 language with a proven track record for performance, usability and adaptability. Within the
219 context of DFO Science, there are a high number of usage cases under which this application
220 can be deployed. For example, the number of users might range from a single employee alone in
221 the field to several dozen scientists, technicians and vessel personnel participating in a research
222 cruise. Similarly, the application might need to be deployed on a stand-alone device or accessed
223 from within a LAN or even over a wide area network (WAN). The application must be able to
224 accommodate a wide range of practical scenarios.

225 **2.2 Version control/source control**

226 The utilization of a VCS is an indispensable component of a sustainable development workflow.
227 Version management is especially important in the context of having concurrent instances of
228 the application in production at any given time. Knowing the version of a production instance
229 is necessary to resolve any issues that might arise. Similarly, databases and backup files are
230 intimately linked to an application's version number. In order to successfully re-instantiate a
231 backup file, the precise version under which it was produced must be known. The VCS will
232 also provide an indispensable framework for the coordination, examination and integration of
233 contributions from collaborators.

234 **2.3 Unit Testing**

235 The application performance needs to be reliable, especially considering its potential to be
236 deployed in remote field environments. The implementation of unit tests is a practical way to
237 ensure the maintenance of core functionality over time. At one extreme, the addition of any code
238 can be preceded by the creation of unit tests (i.e., test-driven development). This approach will
239 maximize the stability of an application, but can hinder the momentum of a project, especially in
240 its early stages. At the other extreme, application development in the complete absence of unit
241 testing occurs at a relatively fast pace but will result in a project that is vulnerable to breaking in
242 unexpected ways and one that is difficult to maintain and pass on to other developers.

243 **2.4 Backup strategy**

244 In all deployment scenarios, data of high business value will be captured and there is little
245 to no tolerance for data loss. Accordingly, the application must have a way to facilitate the
246 implementation of a robust data-backup strategy. Capturing numerous snapshots of the
247 application (and database) is ideal since doing so provides redundancy as well as the ability
248 to revert to a specific point in time. If possible, the backup files and snapshots should be stored
249 on storage volumes that have some form of redundancy.

250 **2.5 Customizable protocols**

251 The flexibility of scientific project leads to design, modify and report on their sampling protocols,
252 without depending on developers, is very important. The application should allow users to
253 provide a variety of detailed information based on their particular sampling protocols, without
254 the need to change the application source code. Similarly, different protocols utilize different
255 code conventions for the identification of biological catches. The application should offer the
256 flexibility of users to utilize their preferred system of catch codes.

257 **2.6 Quality control**

258 The implementation of quality control checks in a data entry application is of paramount
259 importance. For the usage case that supports a scientific fisheries survey, the following quality
260 control checks must be part of the application:

261 **Fishing sets** The application should ensure all the required information about a fishing
262 location have been filled in. Users should be warned if the fishing station start and/or end
263 coordinates fall outside the expected sampling stratum (if applicable).

264 **Catches** The application should verify the validity of catches that do not have any specimen-
265 level data entry associated with them. This validation is important to help identify catches
266 that might have been entered accidentally.

267 **Specimens** The application should flag specimens whose length falls outside an acceptable
268 range. Similarly, a validation of the specimen's length-to-weight ratio should be performed
269 to warn users if a recorded weight falls outside the expected range for the measured
270 length.

271 **Observations** Individual observations are characterized by an observation type. Observation
272 types should have predefined data types such as integer, float, string or categorical.
273 The application should ensure that inputted observation values should never be left null
274 and that they respect the data type of the corresponding observation type. In the case
275 where an observation type (e.g., sex) has a set of defined categories (e.g., male, female,
276 unknown), the application should ensure that any entered values fall within the set of
277 available options.

278 **2.7 User interface**

279 The user interface of the application can have a significant impact on user experience and
280 on-boarding. By ensuring the application has a modern and intuitive interface, the barriers
281 related to on-boarding new users are significantly reduced. Furthermore, an intuitive interface
282 will reduce the need for extensive help documentation. Wherever extra annotation is required,
283 documentation should be inserted directly in the application in the form of tool tips and help
284 bubbles. By appealing to the end users' intuitions and by providing in-situ help documentation,
285 we reduce the likelihood that fields and features get used incorrectly.

286 **2.8 Reactivity**

287 In the context of being on a research vessel survey, data entry happens at a very high rate and
288 on numerous devices; often with multiple transactions per second. Accordingly, it is imperative
289 that the application does not create a bottleneck for data entry and is able to keep pace with
290 experienced technicians. The usage of a reactive Javascript library in conjunction with an API
291 would allow data entry to occur without webpages having to constantly refresh.

292 **2.9 Multilingualism**

293 The ability for users to choose the language of their choice in the application is of considerable
294 importance to this project. Previous tools that have been used were unilingual, and this by itself
295 would have limited the scope of their use in a Canadian national context.

296

3 Results

297 **3.1 Architecture**

298 The architecture of Andes differs significantly from that of its predecessor (see Figure 1).
299 The application and its associated services are centralized on one or several servers and
300 include: 1) a web-service for handling HTTP requests and responses; 2) a database service
301 for storing data associated with the application; 3) a file-sharing service for handling the storage
302 of backups and related files; 4) a printing service used for printing out specimen labels; and 5)
303 a message-brokering service for handling asynchronous tasks. Client devices, such as data-
304 entry workstations, no longer require the installation of anything more than a modern web
305 browser; i.e., one that is capable of supporting HTML5 and ECMAScript 2016 (Javascript).
306 Accordingly, this increases the range of devices and operating systems that may be used for
307 accessing the application. For instance, the switch to using Andes has allowed for the integration
308 of mobile phones, tablets and linux workstations into the data entry workflow. Finally, the new
309 configuration means multiple stations can simultaneously receive and enter data into the same
310 Andes instance.

311 The [Django Web Framework](#) was selected for the backend of this application due to its
312 modularized nature; virtually all aspects of the programming framework can be decoupled.
313 Furthermore, the Django framework is written in pure Python language; an open-source,
314 generalized object-oriented programming language that is popular for use in data-heavy
315 applications. In addition to standard Django templates used to render HTML webpages to
316 end-users, the web framework also contains an elaborate REST-API component. Django uses
317 an Object Relational Model (ORM) to handle the data layer, and includes APIs for a variety of
318 modern relational database management systems (i.e., PostgreSQL, MariaDB, MySQL, Oracle,
319 SQLite, etc.). Several of the application's frontend templates contain reactive components
320 implemented in Vue.js.

321 The frontend of the application is built in HTML5, JavaScript and Cascading Style Sheet. Most
322 users will be familiar with the flow and functionality of a web browser and will be comfortable
323 navigating and entering data into a website. Facilitated by the Django model and form classes,
324 all controls (i.e., fields) on the website contain verbose descriptions and help text. The [Bootstrap](#)
325 [v5.0](#) CSS and JavaScript libraries were utilized in order to give the application a sleek, modern
326 look and to ensure compatibility with different types of devices (e.g., personal computers, tablets
327 and mobile devices). The Bootstrap library also provides palatable styles for displaying help text
328 in the form of popups and tool tips.

329 While web applications are most often used over a network, the Django library comes with
330 a development web-server that permits users to serve and use the application locally. In

331 this scenario, a single computer acts simultaneously as server and client. While there are
332 important limitations to the use of the Django development web-server in a full-scale production
333 environment, the option to run and use Andes on a single device is a advantage for stand-alone
334 use-cases (e.g. field tablet used for port sampling).

335 The architecture used by Andes creates networking requirements that were not previously
336 present in the ESE and MRR. The server and the client devices must be connected to the same
337 network. The network does not need to have access to the WAN (i.e., the internet) connection.
338 As long as they are connected to the same LAN, they can be configured for work together. This
339 suits the networking environment on board remote vessels that can have sporadic connection
340 failures with the wide area network.

341 The Andes application provides access to its various components based on an internal system of
342 authentication and authorization. The credentials of a given user will affect what action they are
343 able to do. For example, while the chief scientist is able to modify sampling requirements for the
344 mission, other users are not.

345 **3.2 Andes Modules**

346 The main index page of the Andes user interface (Figure 2) allows users to access several
347 different modules that loosely correspond to the main use cases defined in the above Methods
348 section. Access to the different modules is determined by a system of authentication and
349 authorization. Consequently, the index page will appear different to users depending on the
350 permissions they have been granted.

351 While Andes has been implemented using a modular design, there is a core set of components
352 that are used across modules. This is true at the level of the database, where in addition to a
353 number of shared tables, sub-modules will have tables that are specific to a given use case. For
354 example, while the Mission table is used across several modules, the Specimen sub-module is
355 specific to the Ecosystem Survey module.

356 Similarly, Andes has modules that are used across multiple user scenarios; in particular, the
357 Bridge module (Figure 3) and the Cruise Dashboard (Figure 4). These modules display high-
358 level information to end users such as queued stations, vessel speed, position and heading and
359 various summaries of science activities that are underway or that have already taken place. All
360 core pages of the application can be toggled to night mode, as desired.

361 **3.2.1 Bridge**

362 The Bridge module (Figure 3) is used by navigation and fishing officers in the wheelhouse
363 (usually deployed on a tablet) to input fishing set metadata. The data entry occurs in two ways:
364 1) by directly editing the set form (a.k.a. the set card); and/or 2) via the Fishing Console, which is
365 displayed in Figure 14. The Fishing Console can capture a number of different events that take
366 place during fishing as well as information coming from the sonar and trawl mensuration system
367 sensors (e.g., Scanmar system).

368 **3.2.2 Set manager**

369 The Set manager module is used to manage upcoming sets and review completed set cards.
370 Typically accessed by the chief scientist, it provides the ability to select upcoming sampling
371 locations and initialize the corresponding set cards. The Set manager module is where sets
372 can be activated and deactivated. Additionally, this module is where any quality assurance flags
373 associated with a set and its catches can be reviewed and accepted.

374 **3.2.3 Ecosystem survey**

375 The Ecosystem Survey module is the main entry point that technicians will use to input survey
376 data. A depiction of the main tables involved in the Ecosystem Survey module, and their
377 relationships are displayed in Figure 5. This component of Andes replicates the capabilities of
378 the ESE for capturing detailed information on length, weight, ageing material, maturity, etc. about
379 fish and invertebrate specimens. This module, typically accessed from the wet laboratory of a
380 survey vessel, is used for all entry of data related to measurements and observations of marine
381 organisms. New catches are first entered into the Active Set page, as portrayed in Figure 6. The
382 Active Set page accepts regional catch codes as a way to input new catches into a set. If a code
383 is not known, a search feature is available. Next, baskets and their corresponding weights and
384 statuses (e.g., sampled vs. not sampled) are then entered into the Catch Card page (Figure 7).
385 Finally, specimens are entered into the Data Entry page (Figure 8). As the data entry progresses,
386 users are dynamically prompted with observation fields that follow the catch-specific sampling
387 protocol. An overview of the sampling protocol is displayed on the right-hand side of the Data
388 Entry page.

389 **3.2.4 Shrimp ecosystem survey app**

390 The Shrimp Ecosystem Survey (“Shrimps”) module resulted from an early adoption of Andes
391 in the Québec Region. The niche workflow requirements of shrimp survey operations justified
392 the creation of its own Andes module which operates as an extension of the ecosystem survey
393 module.

394 Shrimp data is collected by the Quebec region in the course of its yearly ecosystem survey in
395 the estuary and northern gulf of the Saint-Lawrence as well as during other occasional surveys.
396 These data are collected for the stock assessment of the northern shrimp (*Pandalus borealis*) as
397 well as to assess the biodiversity and abundance of shrimp species. In the case of the Quebec
398 region, the biological measures related to shrimps are done onboard in contrast to other regions
399 where shrimps subsamples are frozen and biological data is collected in the laboratory after the
400 end of the mission. The shrimps application was developed to allow the entry of shrimps data
401 as it follows a different workflow compared to the ecosystem survey application. At every set,
402 a subsample of between 2 and 3 kilograms is first collected from the total catch. This sample
403 is then separated by species, then by maturity stage and only then are the biological data
404 (cephalothorax length for all species and occasionally weight for *Pandalus borealis*) taken.

405 **3.2.5 Oceanography**

406 The Oceanography module is a component of the Andes application that is capable of being
407 deployed independently of the Ecosystem Survey module. This module is used to track and
408 record the deployment of CTD/Rosette systems and plankton nets (Figure 9). Basic metadata,
409 such as the location and date/time of deployment, ‘on bottom’ and recovery of gear can
410 be collected. Additionally, the Oceanography module allows for users to enter which water
411 samples/parameters were collected from pre-determined depths after CTD/Rosette deployment
412 (Figure 10). The simplified Entity Relationship Diagram of the Oceanography module of Andes
413 is presented in Figure 11. Upon completion of a survey, Andes is designed to provide a series of
414 oceanographic summary reports that summarize aspects of the data collected:

415 **Mission instrument report** Provides a summary of the oceanographic equipment used on the
416 survey, including component type, model and serial number, and date of last calibration (for
417 CTD sensors).

418 **CTD metadata report** For each fishing set and station where the CTD/Rosette system was
419 deployed, the date, time, position, sounding, and bottle sample IDs are provided.

420 **Hydrolog report** For each fishing set and station number, a summary of the CTD/Rosette
421 and ring net deployments are provided, including the event numbers corresponding to
422 each gear deployment, comments entered into Andes regarding each gear deployment,
423 and summary information related to each set and station, including surface temperature,
424 sounding, day of year, and date/time.

425 **Plankton report** This report provides a summary of the plankton net deployments conducted in
426 relation to each fishing set and stratum. The wire out and wire angle, as well as flow meter
427 start and end are provided.

428 **Bottle report** This report provides a detailed summary of each CTD/Rosette operation (e.g.,
429 altimeter height, bottle height and depth, max. CTD depth), the water samples/parameters
430 collected, and also includes the results of the Winkler titrations for dissolved oxygen
431 samples, if entered into the Oceanographic Activity detail page.

432 **TS report** This report provides a summary of the surface and bottom temperature and salinity
433 data from each CTD cast in relation to each set, station, and stratum.

434 **ELOG report** This report emulates the .log summary report produced by ELOG. This report is
435 required in order to upload the survey data into the Microsoft Access template used by
436 AZMP to load data to BioChem, DFOs national repository for discrete and plankton data.

437 These reports facilitate the post-processing of the data, its integration into existing databases
438 and its distribution and upload to various open data platforms.

439 **3.2.6 Forecasting utility modules**

440 There are a number of task-specific modules designed to assist in mission forecasting
441 and planning. These modules are designed to provide timely information to scientific staff

442 participating in a research cruise, and to also assist the chief scientist and watch leaders in the
443 planning and execution of a mission.

444 **Cruise dashboard**

445 The cruise dashboard module is used to provide an overview of the current status of a mission
446 to all Andes users (Figure 4). It provides a range of real-time statistics that are used for planning
447 purposes, including current status (e.g., fishing, steaming, deploying/retrieving net), the list of
448 upcoming stations, and data from the latest trawl.

449 **Forecasting**

450 The forecasting tool is designed to provide real-time prognostics of survey completion targets
451 based on assumed transit speeds and time spent fishing and processing the trawl catches
452 (Figure 15). This feature provides a useful tool for the chief scientist to evaluate different
453 sampling objectives and to compare different survey route options. As the conditions change
454 regularly during a survey, this tool links with the set manager to provide an estimate of the
455 amount of time required to complete planned stations. This tool can be used for short-term
456 planning (i.e. a day's worth of sampling) or longer-term planning (i.e. a mission's worth of
457 sampling).

458 **Progress Map**

459 Another task-specific utility is the progress map which shows what strata have been completed
460 based on target and minimum number of sets per stratum. This map provides the chief
461 scientist with a clear visual depiction of what has been accomplished and what remains to be
462 accomplished during the survey. The number of sets conducted in each stratum is compared to
463 the minimum and target number of sets per stratum to determine the colour that each stratum will
464 appear in the progress map (Figure 16).

465 **3.2.7 Port sampling**

466 Andes was adapted from a previous port sampling module to support activities where
467 technicians obtain length frequency samples from commercial fishing activities. The port
468 sampling module of Andes is fully stand-alone and is primarily used independently of all
469 other modules. The protocols used in the Port Sampling module are much simpler than those
470 belonging to the Ecosystem Survey module. Through the protocols, users are able to control
471 collection quotas (e.g., “keep two specimens per bin”), the flow of data entry (e.g., which field
472 should be displayed in the sample form) and the layout of the data entry page (e.g., length
473 bins organized in a vertical or horizontal configuration) (see Figure 12). Typically, production
474 instances of this module are deployed on ruggedized field tablets that are suitable for use in wet
475 environments.

476 The design of the user interface is simple and intuitive, where each length bin is a large button on
477 the display. As specimens are tallied, the corresponding buttons on the touchscreen are pushed.
478 A screenshot of the data entry page is presented in Figure 12. Data export abilities from the field
479 tablets to existing database for port sampling data is also part of the module. The various tables
480 of the Port Sampling module and how they relate to one another is shown in Figure 13.

481 **3.3 Technical Implementation of Other Design Goals**

482 **3.3.1 Version control/source control**

483 [Git](#) was selected as the VCS for this project due to the fact it is widespread in use and open-
484 source. The remote repository for this project is currently hosted as a private project on the
485 [Gulf Science organizational GitHub account](#). The project is additionally making use of GitHub
486 infrastructure including, pull requests, issue management (e.g., bugs, feature requests and
487 general enhancement requests), security alerts and version releases. The Andes documentation
488 is also being served using GitHub Pages and can be accessed [here](#).

489 **3.3.2 Unit testing**

490 For Andes development, we use a mixture of test-driven development for critical components
491 of the application, and are also continually adding unit tests for more user-specific components.
492 Using the built-in Python/Django testing framework, this strikes a middle ground between the
493 two approaches described in the above section. While the goal is not to implement test-driven
494 development for the application, the use of unit tests is highly encouraged, especially to back up
495 the core functionality of the application.

496 **3.3.3 Backup strategy**

497 The Andes application has the capability to perform manual and automatic backups. These
498 backups consist of two parts: a JSON export of the entire database and the recording of the
499 current version of the application (i.e., the git hash). The structure of the application models and
500 associated data structure will change over time with development. Accordingly, in order to re-
501 instantiate a particular data snapshot, it is critical to know the precise version from which it was
502 exported. This combination of data export and git version number, gives users the perpetual
503 ability to recreate the exact application environment from the time of the snapshot, no matter how
504 much the application has changed in the interim. In the application, backups are automatically
505 created upon closing sets. Moreover, users also have the ability to manually trigger a backup at
506 anytime.

507 **3.3.4 Customizable protocols**

508 Andes provides project leads the ability to create and modify sampling protocol through the user
509 interface. By doing so, project leads are able to shape the flow and control the behaviour of the
510 application during data entry. This including deciding which fields to display in a form (e.g., set
511 cards), importing stations and other geographical features (e.g., sampling strata, Northwest
512 Atlantic Fisheries Organisation areas, Marine Protected Area (MPA)s, etc.) and the quotas and
513 observation fields associated with different catch items. Examples of catch-specific sampling
514 requirements that can be programmed by project leads can be found in Table 1.

515 Andes also allows users to specify a preferred system of catch codes by attaching the code
516 collection to a sampling protocol. When a given protocol is active, all catch items are referenced
517 by their corresponding entry in the attached catch collection list, an essential feature for making
518 Andes usable in multiple regions.

519 **3.3.5 Quality control**

520 Andes implements a suite of quality control checks. All quality control flags that are raised during
521 data entry are appended to a report and require sign off before a set can be closed.

522 **Sets**

523 The Andes provides clear feedback regarding the completeness status of a given set. Flags
524 are raised if a set's start and/or end coordinates are outside the expected sampling stratum.
525 Additionally, an alarm is sounded in the bridge console module when deviating from the desired
526 stratum in real-time. The bridge console also provides personnel with fishing timers and tow
527 distance displays (see Figure 14).

528 Andes will flag when the tow distance as calculated by the cruise track, differs from the tow
529 distance as calculated by a straight line between the start and end coordinates by more than
530 5%. Andes will flag when the start or stop coordinates of a set are not within the expected NAFO
531 area (if applicable).

532 **Catches**

533 There are numerous flags that can be attributed to individual catches. Flags are generated for
534 catches that do not have any data entry associated with them. In addition to this, Andes will flag
535 when the weights entered for baskets that are considered suspect. This is determined by either
536 the default maximum basket weight (mission level), or the maximum basket weight for a given
537 species. Andes will also flag when the difference of total weight of *sampled* basket differs by
538 more than 25% from the total calculated specimen weight. The total calculated specimen weight
539 is a combination of actual weights (when collected) and those which were estimated from length
540 measurements. The latter is achieved by using regression coefficients estimated from historical
541 length (L in centimeters) and weight (W in grams) observations using the following equation:

$$W = aL^b \quad (1)$$

542 Andes offers the option to specify separate regression coefficients (a and b) for males, females or

543 unspecified individuals (as shown in Figure 17).

544 A recent version of Andes implemented a new optional layer of quality control. This feature
545 allows project leads to assign allow-lists and restrict-lists, commonly known as “whitelists” and
546 “blacklists”, to a mission or to a specific geographic feature (e.g., a stratum). In this way, the
547 validity of each catch entered into a set can be assessed. For example, if a catch being recorded
548 is *not* on the that set’s associated allow-lists, the end user will be notified that this is an unusual
549 observation and will be prompted to collect documentation. Similarly, if a project lead adds a
550 catch to the mission’s restrict-list, users who enter this catch will receive a warning message,
551 asking them to double-check the assignment. This is useful when project leads want to limit
552 the usage of certain taxa during data entry, e.g., *Alosa* sp. is preferred over the use of *Alosa*
553 *pseudoharengus*.

554 **Specimens**

555 Specimen lengths and length-to-weight ratios are validated against the parameters entered
556 in the sampling requirements for that catch. Acceptable length-to-weight ratios are assessed
557 by comparing the actual weight to the estimated weight, as described in the section above.
558 Additionally, Andes will flag when there is a mismatch between fish maturity and somatic length.
559 Project leads can specify mature length thresholds for males, females or unspecified individuals
560 in the sampling requirement of a given species (as shown in Figure 18).

561 **Observations**

562 When entering observations, Andes enforces the data type of the corresponding observation
563 type; invalid entries are not accepted. When entering an observation for an observation type
564 that has categories, Andes will display the list of options to the end user and inhibit users from
565 entering invalid selections. Andes will also ensure observation types are not left blank. However
566 NaN entries are permitted when a particular observation is meant to be skipped. Certain
567 observation types are meant to be unique, e.g., unique tag number. If this is flagged at the level
568 of the observation type, end users will be notified if there is a violation of this structure.

569 **3.3.6 Reactivity**

570 The [Django REST framework](#) was used to construct the WebAPI component of the application.
571 The project takes a hybrid approach, combining the use of standard Django views and Vue.js
572 frontend applications embedded in the templates. The latter were used to avoid the need for
573 constantly reloading webpages and to optimize the flow of traffic across the network. Reactive
574 javascript frontend applications also provide a better experience from the point of view of an
575 end-user.

576 **3.3.7 Multilingualism**

577 The Django framework has excellent support for internationalization and localization, including
578 the translation of text and the formatting of dates, times and numbers. It achieves this using a
579 system of ‘hooks’ used by developers to indicate which parts of the code should be localized.

580 See [Django - Internationalization and localization](#) for more details on this process. In our
581 application, an end-user can toggle between English and French by simply clicking on a button.
582 In this way, each client can view the application in the language of their choice.

583

4 Discussion

584 Since its initial deployment during the 2019 September ecosystem survey in the southern Gulf
585 of St. Lawrence, the capabilities and performance of Andes have significantly improved. The
586 software has also been adopted as the main logging software used during ecosystem surveys
587 conducted by DFO Gulf, Maritimes and Québec Regions. The current capabilities of Andes have
588 evolved over the last five years where the system has been used to support many field activities.

589 The forecasting tool and progress map presented in this report were not in the original
590 deployment of the application and were added during the cruise based on the needs of scientific
591 and Canadian Coast Guard (CCG) personnel. The reporting facilities provided by Andes were
592 developed to support the many users of the data collected during scientific activities. Two types
593 of reports are available for Andes: 1) reports meant to be used during field activities and 2)
594 reports meant to be used after field activities are completed.

595 If a vessel is located within range of cellular phone signals, it is possible to make Andes
596 accessible via the internet. While this has some important advantages, exposing an IP to the
597 internet comes with security risks that have to be mitigated.

598 Andes is designed so that data backups are immune to the potential pitfalls of relying on earlier
599 versions of required packages. Having backups that are associated with a specific git version,
600 which includes a comprehensive list of all dependencies and an exact match of the application
601 data model, allows users to readily bring back Andes to where it was at the time of the backup.
602 Users can then access the reporting features that will allow them to create a version of the
603 required data that suits their needs and facilitates inclusion in existing relational database
604 management systems or other types of data solutions.

605 When present, the client-server separation lends itself nicely to be adaptable to new
606 requirements. One can potentially expand upon the *system* with minimal changes to Andes
607 by interacting directly with its exposed REST-API. For example, an imaging system can be aware
608 of current GPS coordinates or current Set/Station. Photos or videos acquired by such an imaging
609 system could automatically append these as metadata, or even add media identifiers directly to
610 the database as they are captured.

611 Most modern vessels will already provide the minimum networking infrastructure necessary for
612 system deployment where connectivity between a dry laboratory, a wet laboratory and the ship's
613 wheelhouse is necessary. In the case where static networking options are not available, this
614 requirement can be met using inexpensive consumer-grade, portable networking equipment
615 (wireless router, switch, etc.). However, as installing an ad-hoc network for scientific needs in a
616 vessel may lead to code violations, we urge Andes users to collaborate with vessel management
617 body like the CCG or chartering authority.

618 The addition of allow-lists and restrict-lists was an important addition to Andes that came out of a

619 Regionalization workshop that took place in December 2022 at the Gulf Fisheries Center. While
620 these features will help improve the quality of data collected, it is important that if implemented,
621 they should be detailed explicitly in the survey sampling programme. For example, if project
622 leads add certain catch items to the restrict-list, this should be mirrored in the survey sampling
623 programme.

624 Andes also now supports the use of various electronic tools such as calipers, electronic
625 measuring boards and scale. Drivers for the measuring boards and scale were developed by
626 the Gulf and Quebec regions and are available on Github.

627 While there is always an effort to make the user interface intuitive and self-explanatory, there is
628 still the need to maintain documentation for the project. The Andes docs are maintained through
629 GitHub Pages and are publicly available here: <https://dfo-gulf-science.github.io/andes/>

630 **4.1 Integration of Andes with Existing Data Repositories**

631 Andes is a standalone application with an underlying relational database management system,
632 and the data collected during research activities must be extracted from the application
633 and integrated into existing databases. This step has proven to be challenging but has also
634 encouraged close inter-regional collaboration since the challenges associated with integration
635 were shared by the different groups involved.

636 **4.2 Future Directions**

637 Onboarding efforts for additional coastal surveys (scallop, sea-cucumber, whelk) and regions is
638 an ongoing process driven by Andes' success.

639 **4.2.1 Atlantic Zone (Offshelf) Monitoring Program (AZMP/AZOMP) oceanographic 640 surveys**

641 Andes was first trialed in parallel to ELOG on a dedicated oceanographic survey led by the
642 Maritimes Region Atlantic Zone Monitoring Program (AZMP) in the spring 2022, and was
643 recently used as the sole event logger during the 2023 Atlantic Zone Offshelf Monitoring
644 Program (AZOMP) survey. During its initial trial on the 2022 spring AZMP survey, a number of
645 limitations were noted. ELOG assigns each gear deployment a sequential 3-digit 'event' number.
646 In contrast, Andes groups all gear deployments in relation to a Set number, which is a primarily
647 fisheries-related term that is not applicable to oceanographic survey data collection. However,
648 Andes could easily be adapted to the AZMP's normal metadata naming convention by renaming
649 Set to 'station', given that multiple sampling events normally occur at the same AZMP station.

650 The initial version of Andes included only two oceanographic gear options: CTD/Rosette and
651 ring net deployments. However, Argo floats, oceanographic buoys, moorings, multinet, and other
652 gear are routinely deployed during AZMP and AZOMP surveys, and the order of operations must
653 be flexible. In previous versions of Andes, the order of operations under each Set could not be

654 modified between stations. With developer support, Andes was recently modified to allow for
655 the inclusion of additional gear options and the ability to toggle and select between gear types
656 should the order of operations change between stations.

657 ELOG allows users to create various ‘logbooks’ that allow for custom recording of metadata. On
658 AZMP surveys, a custom ELOG logbook was created to capture metadata related to samples
659 collected from underway or flow-through thermosalinograph (TSG) systems used on these
660 surveys. A sub-module could be developed in the future to allow for a similar level of data
661 capture using Andes.

662 Andes shows great promise for enhanced data recording and oceanographic sample tracking
663 relative to ELOG. Future modifications of Andes could incorporate information on water budget
664 requirement for each CTD cast, and could allow for the tracking of laboratory post-processing.
665 However, making changes to Andes on the fly would not be possible on local installations (i.e.,
666 not on the custom servers permanently installed on the ecosystem survey vessels) of Andes
667 without on-board developer support. With this, Andes could be adopted as the full-time event
668 metadata logger on AZMP and AZOMP’s oceanographic surveys in the future.

669 Andes has been successfully used by a number of DFO regions and has proved to be a useful
670 data entry application that fulfills its intended design goals. The infrastructure used in the
671 development of the application also means that novel modules and functionalities can be
672 implemented to support new users requirements. For example, adding functionality to add
673 barcodes and QR codes to physical samples such as otoliths and tissue samples, or the addition
674 of imaging capabilities to link digital images to specimens would be possible.

675 **4.3 Governance**

676 Andes does not currently have a well-defined governing body which makes the project vulnerable.
677 To alleviate this situation, we ask that this technical document, as well as the developer’s guide,
678 be considered when outsourcing development efforts towards Andes. This will ensure that the
679 core vision is maintained and that software sustainability principles are upheld. We also hope
680 that this report can stimulate the establishment of a governance structure for the development of
681 the application, one that ensures regional participation and that facilitates collaborative efforts.

682 Andes reached a critical point in its lifecycle. It has been used as the main data entry strategy for
683 numerous missions over the span of a few years. Although new features can always be added,
684 its core functional requirements can now be locked-in. These boundaries are now well-defined
685 and described in this report.

686 **5 Acknowledgments**

687 We thank all DFO personnel who were involved in the early testing and deployment of Andes
688 prior to and during the 2020 southern Gulf of St. Lawrence ecosystem survey. We thank the Gulf
689 Region publications coordinator for his/her assistance with getting this document published.

6 References

- 691 Benoît, H.P., and Daigle, D. 2007. Procedures for commercial catch sampling of finfish and
692 shrimp in the southern Gulf of St. Lawrence. Can. Manuscr. Rep. Fish. Aquat. Sci. 2833:
693 iv + 63 p.
- 694 Hurlbut, T., and Clay, D. 1990. [Protocols for research vessel cruises within the Gulf Region](#)
695 [\(demersal fish\) \(1970-1987\)](#). Can. Manuscr. Rep. of Fish. Aquat. Sci. 2082: 143 p.
- 696 Therriault, J.-C., Petrie, B., Pepin, P., Gagnon, J., Gregory, D., Helbig, J., Herman, A., Lefavre, D.,
697 Mitchell, M., Pelchat, B., Runge, J., and Sameoto, D. 1998. [Proposal for a Northwest Atlantic](#)
698 [Zonal Monitoring Program..](#) Can. Tech. Rep. Fish. Hydrogr. Ocean Sci. 194: vii + 57 p.

7 Tables

Table 1. Example of sampling requirements for five species that are regularly captured in the annual southern Gulf of St. Lawrence September survey.

Species	Example of sampling requirement
Atlantic Cod (<i>Gadus morhua</i>)	<p>Collect length, sex, weight, maturity.</p> <p>Collect otoliths from:</p> <ul style="list-style-type: none"> - one specimen per cm per set for specimens 25 cm and under; - two specimens per cm per set for specimens between 26-45 cm; - and three specimens per cm per set for specimens 46 cm and over. <p>Collect 10 specimens from the Banc des Américains MPA that are between 40-70 cm.</p>
Atlantic Herring (<i>Clupea harengus</i>)	Collect length from every specimen; preserve 2 specimens per every 5 mm length bin per set.
Atlantic Halibut (<i>Hippoglossus hippoglossus</i>)	<p>Collect length, sex, weight, maturity, stomach weight and otoliths from every specimen.</p> <p>Collect 300 fin clips per mission.</p> <p>Collect 25 female gonads per mission.</p>
American Lobster (<i>Homarus americanus</i>)	<p>Length and sex from one specimen per sex per 3 mm per set.</p> <p>If female, check for lobster eggs.</p> <p>If female size is greater than 70 mm, check molt stage and shell disease index.</p> <p>If female size is greater than 70 mm AND has eggs, check abdominal width, egg stage and clutch fullness rate.</p>
Winter Skate (<i>Leucoraja ocellata</i>)	Collect length, sex, weight, maturity, disk width and tail description from every specimen.

8 Figures

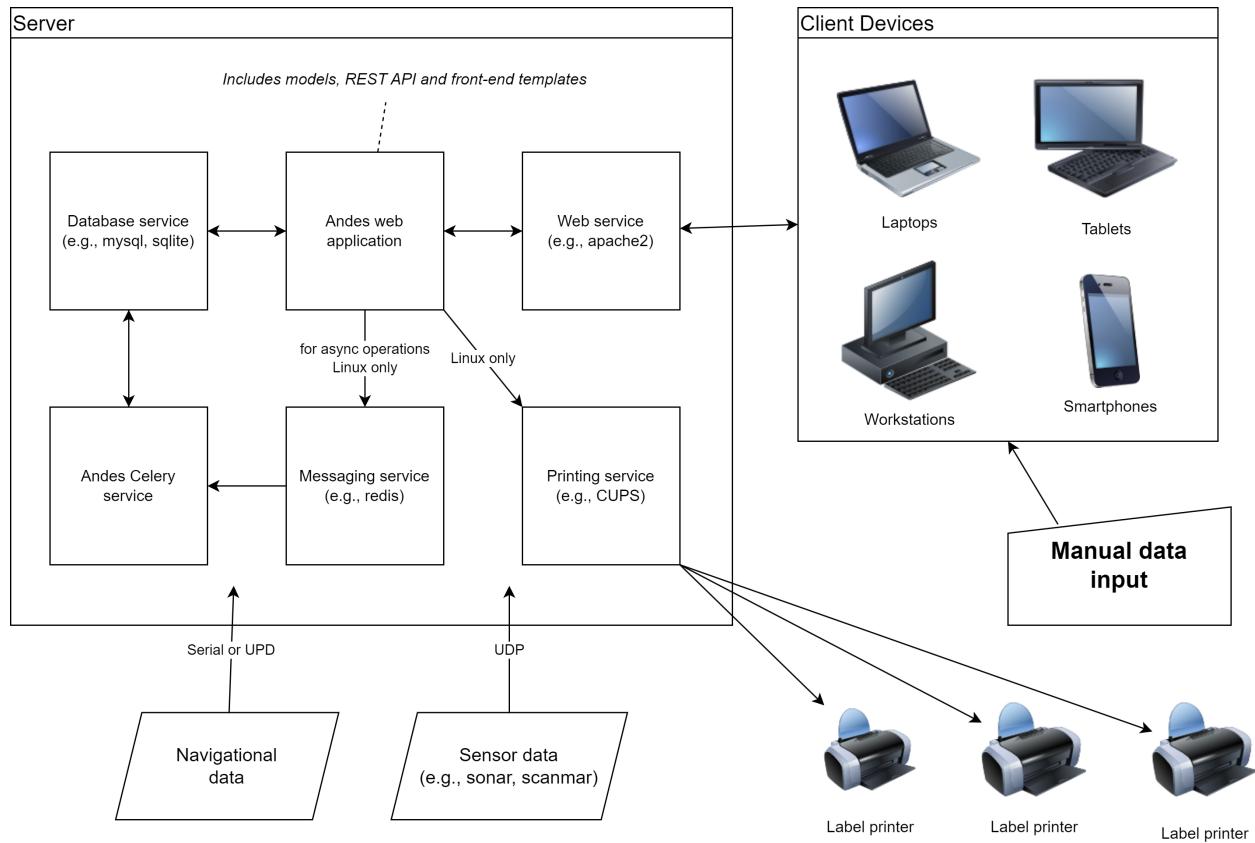


Figure 1. The generalized system architecture diagram of the Andes application and its associated services.

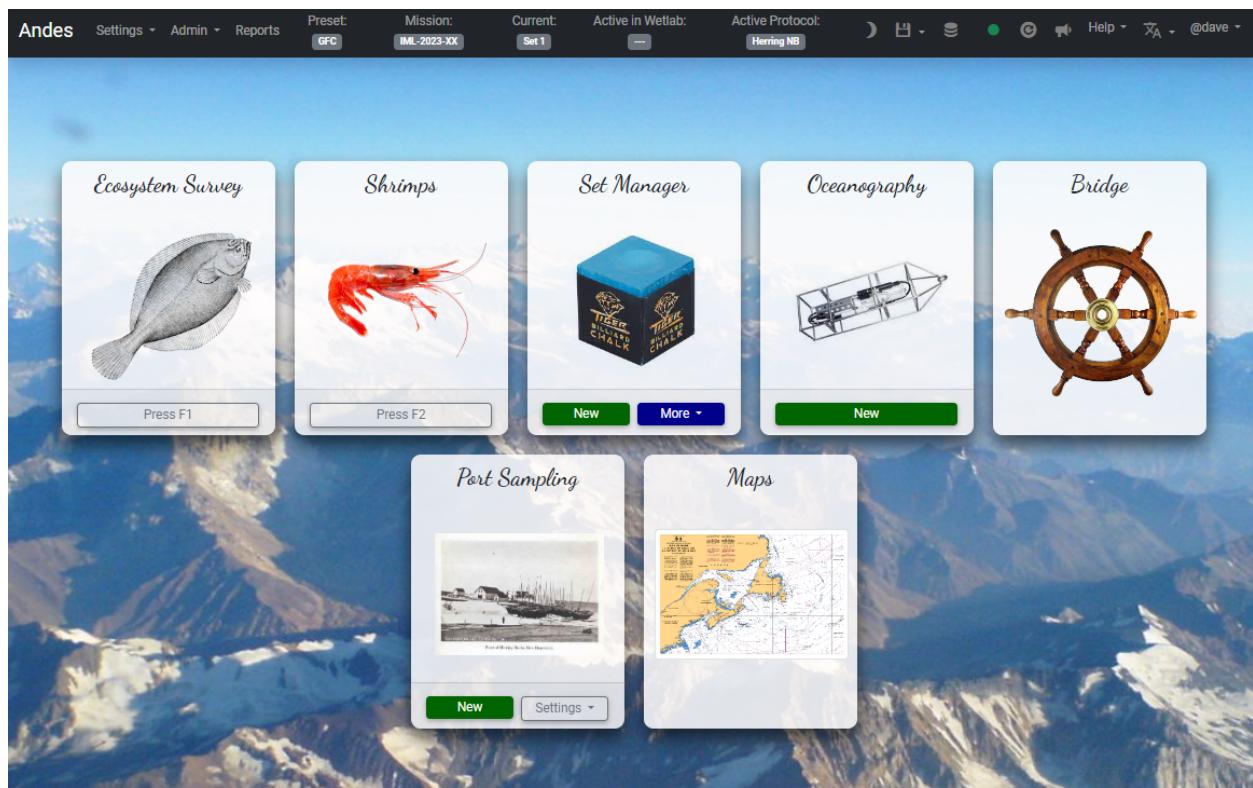


Figure 2. A screenshot of the main index page of Another data entry system (Andes) showing the principal modules of the application. Which are displayed on the index page will depend both the users permissions and profile.

+
Fishing

✍
Set Card

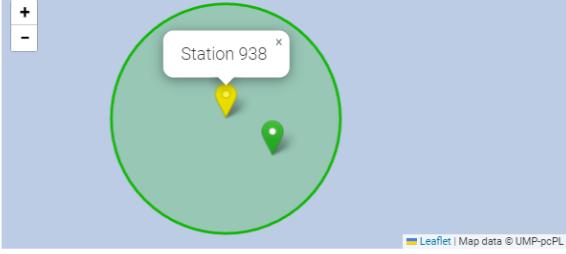
📍 **Current Set**

Set Number	169
Station / Stratum	Station 938 (Alternate) Stratum 472
Operations	Fish, ctd
Current Speed Over Ground	—
Current Course Over Ground	29.4°
Current Depth (true)	22.52m
Distance to Station (w/in radius)	—
ETA	Arrived on station.
Fishing Status	Fishing is underway (2 months ago)
Is Active in Wetlab?	No
Is Set Card Complete?	No
Last Specimen Entered in Wetlab	AMERICAN PLAICE (3 minutes ago)
Set Start Time (Canada/Atlantic)	2023-03-29 16:01:54 (2 months ago)
Starting Coordinates	43° 3.918000' / -63° 38.149000'
Set End Time (Canada/Atlantic)	n/a
Ending Coordinates	n/a
Calculated duration	n/a
Calculated distance	n/a
Calculated speed	3.08 kts
Metadata	<small><u>Created:</u> 2023-03-29 13:04:14 ADT by ryan_martin</small> <small><u>Updated:</u> 2023-06-12 09:17:13 ADT by dave</small>

⌚ **Stations**

Set	Station	Stratum	Depth Int.	Latitude	Longitude	Operations	Dist. to Next
167 ✓	203	498	367-732	42° 45.8916'	-63° 42.0180'	Fish, ctd	8 nm
168 ✗	950	478	184-366	42° 48.1044'	-63° 32.1960'	Fish, ctd	17 nm
169 ✓	938	472	91-183	43° 4.2358'	-63° 38.7053'	Fish, ctd	28 nm
170 ✓	940	473	<91	43° 20.2131'	-64° 9.9992'	Fish, ctd	7 nm
171 ✓	939	472	91-183	43° 21.9618'	-64° 1.3979'	Fish	62 nm
172 O	1001	460	91-183	44° 16.3300'	-63° 19.9300'	ctd, zoo	—

📍 **Current Set / Station**



Leaflet | Map data © UMP-pcPL

weathermap **Forecast Map**



Figure 3. A screenshot of the Andes Bridge console. This app is used by bridge personnel across multiple user scenarios. This dashboard is a critical link in the communication between the scientist and crew personnel. An example of this is the station list; as a chief scientist plans out his or her route, queue stations and the associated activities will appear in the Bridge console. The Bridge console is also used by bridge officers to trigger the start and ends of operations and to input set metadata.

21

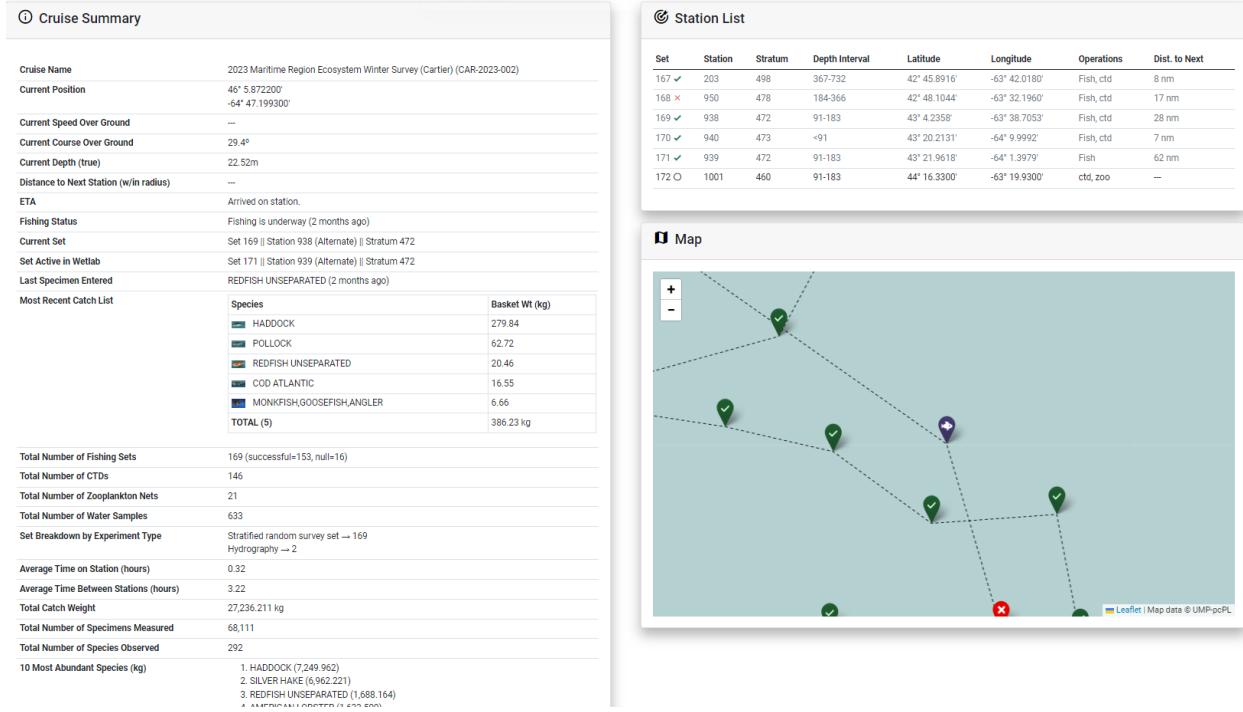


Figure 4. A screenshot of the Andes Cruise Dashboard. This dashboard helps to communicate queued stations, fishing activities underway and estimated times of arrival to science staff and crew personnel. The dashboard also displays high level summaries of Science activities that were conducted.

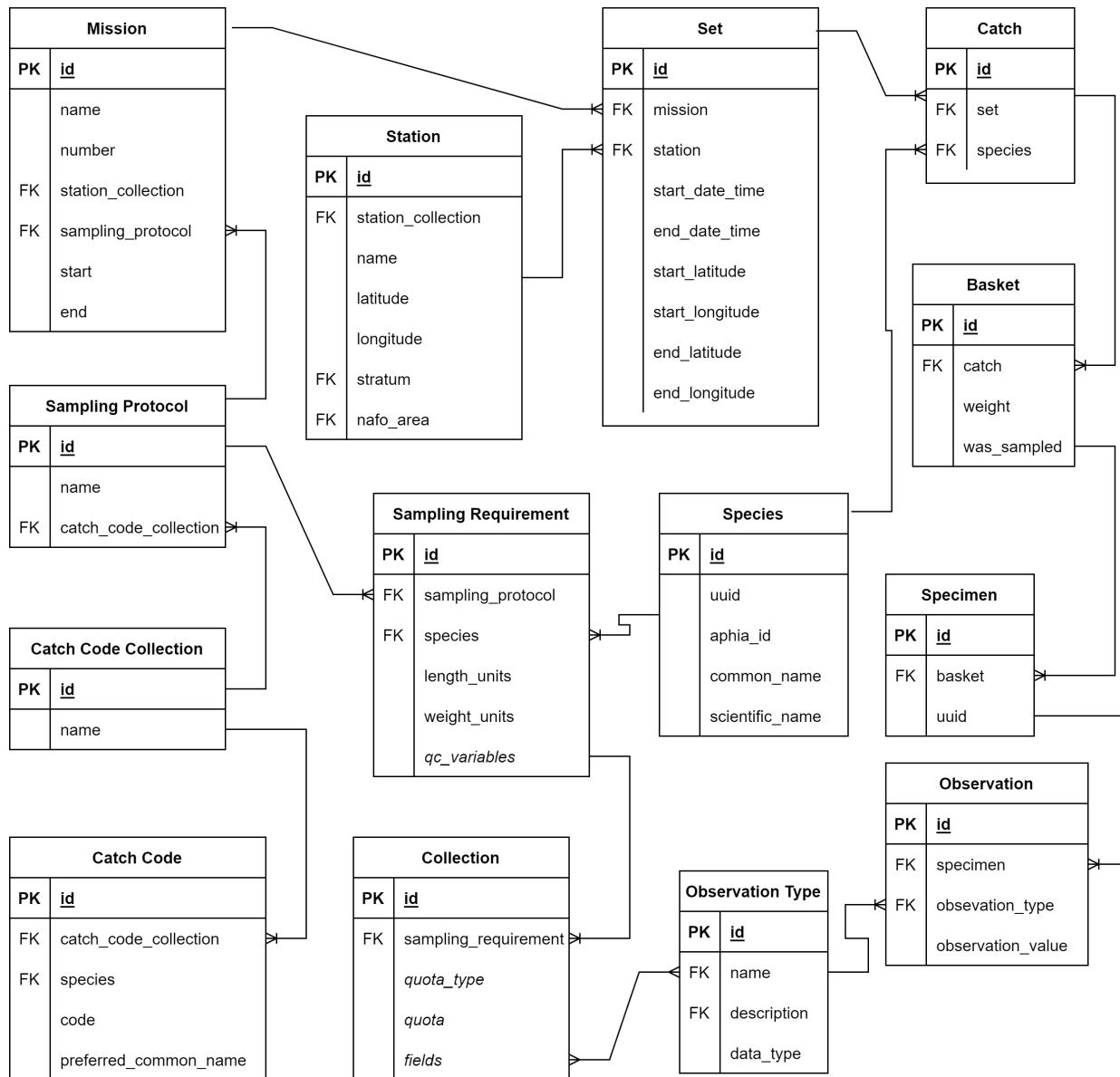


Figure 5. A simplified entity relationship diagram of the data model used by the Ecosystem Survey module of Andes.

Now Processing Set # 170

Station 940 (Alternate) || Stratum 473

Last Species Entered:
MAILED SCULPIN (304)

Enter the species code:

14

Code	Common name	Total basket weight (kg)	Specimens Expected?	Counts			
				Specimens (measured)	Specimens (unmeasured)	Baskets (weighed)	Baskets (unweighted)
10	COD ATLANTIC	15.380	Yes	9	---	1	---
11	HADDOCK	331.681	Yes	306	---	5	---
14	SILVER HAKE	0.062	Yes	4	---	2	---
16	POLLOCK	4.700	Yes	3	---	1	---
23	REDFISH UNSEPARATED	43.130	Yes	259	---	2	---
30	HALIBUT ATLANTIC	0.960	Yes	3	---	1	---
40	AMERICAN PLAICE	0.690	Yes	6	---	1	---
50	STRIPED ATLANTIC WOLFFISH	0.026	Yes	1	---	1	---
60	HERRING ATLANTIC	7.810	Yes	59	---	1	---
62	ALEWIFE	0.148	Yes	1	---	1	---

SILVER HAKE – Merlu argenté – *Merluccius bilinearis*
(already added)

Figure 6. A screenshot of the Active Set page. This page is a component of the Ecosystem Survey module. Users in the wet laboratory enter new catches into this page as they are identified.

Active Set / Catch Card

Stats F4 Refresh F5 Comment F6 More Operations... F8 Toggle Fullscreen F11 Print F12 Delete Ctrl-Backspace Back Esc

COD ATLANTIC – *Gadus morhua* (10)

Set 170 || Station 940 || Stratum 473



View Species Identification Card (Ctrl-H)

New Basket

Basket weight (kg)	Size class	Sample → F3 to toggle yes	Specimens
15.3800	1	Yes	9
TOTAL	15.3800 (Total) 15.3800 (Sampled)	--	9

Baskets

Basket #	Weight (kg)	Size class	Sampled?	Specimens
1	15.3800	1	Yes	9
TOTAL	15.3800 (Total) 15.3800 (Sampled)	--	--	9

Catch Summary:

Total basket weight	15.38 kg
Total sampled basket weight	15.38 kg
Total number of unweighted baskets	0
Total specimen count, unmeasured	0
Total specimen count, measured	9
Total specimen weight, measured	15312 g
Calculated basket weight	15312 g
Calculated specimen weight / sampled basket weight	100%
Comments	---

Sampling Requirement for COD ATLANTIC

Length: → Fork length, centimeters (cm)
 Weight: → Total, grams (g)
 Minimum length: → n/a
 Maximum length: → 140 cm
 Mature length: →
 Rounding rule: → round up to the nearest whole
 Always collect sex? → No

There is 6 collections under this requirement.
 Press F3 to view the details.

Figure 7. A screenshot of the Catch Card page of the Ecosystem Survey module of Andes. This is where new baskets and their associated weights are entered. This example shows the catch card of Atlantic Cod from a Maritimes Region survey of the George Bank.

Active Set / AMERICAN PLAICE (40) / Observations

Previous Stats Comment Print Delete New List Refresh Back

AMERICAN PLAICE – *Hippoglossoides platessoides* (40)

Set 170 || Station 940 || Stratum 473

Specimen #7 (Size Class 1) → Andes ID 72277

Ctrl+Delete → delete an observation

Length	2	cm – round up to the nearest whole – Fork length	
Sex	1	0-Undetermined 1-Male 2-Female	
Weight	2	g – Total	

Collections:

Standard

Fields: Length, Sex, Weight
Quota: 1 per sex, per cm, per set.
Is applicable for specimen? Yes

Stomach Sampling

Fields: Stomach Fullness, Collect Stomach
Quota: 2 per 5cm, per set.
Is applicable for specimen? Yes

Genetic Sampling 4X (NAFO areas 4XL, 4XM, 4XN, 4XO...)

Fields: Fin Clip
Quota: 100 per mission. Up to a maximum of 100 per mission.
Is applicable for specimen? No

Sample Vial Number (WHEN Fin Clip = [...)

Fields: Sample Vial Number
Quota: unlimited
Is applicable for specimen? No

Figure 8. A screenshot of the specimen data entry interface of the Ecosystem Survey module of Andes. The specimen observations are edited on the left-hand side while the sampling collections are displayed on the right-hand side.

Oceanographic Sample for Set 166 (CAR-2023-002)

Detail

Set	Set 166 (CAR-2023-002)
Station	Station 205 (Primary) Stratum 498
Duration	104.77 min
Operator	kevin pauley
Metadata	<small>Created: 2023-03-29 00:14:59 ADT Updated: 2023-03-29 06:49:34 ADT by kevin</small>
Completion Status	Complete

Depth

22.52 m

2 months ago

Activities

[Add Activity](#)

Event number	Instrument	Duration	Number of bottles	Number of actions	Actions				Completed?
					Deploy	Bottom	Recovery	Abort	
164	Yellow Belly	57.9 min	8	3	✓	✓	✓	---	Yes
165	Plankton net (202µm)	34.42 min	0	3	✓	✓	✓	---	Yes

Station Map

Figure 9. A screenshot of the Oceanographic Sample detail page from the Andes Oceanography module.

164 - Yellow Belly

Detail			
Instrument	Yellow Belly	For CTDs Only:	
Event number	164	CTD filename	23002166.hex
Duration	57.9 min	Min. altimeter height from bottom (m)	2
Wire out (m)	688	Min. bottle height from bottom (m)	3
Wire angle (degrees)	0	Max depth of CTD (m)	683
Tow down speed (m/min)	30		
Tow up speed (m/min)	30		
Metadata	Created: 2023-03-29 00:14:59 ADT Updated: 2023-03-29 06:52:55 ADT by kevin	Temperature (°C)	4.59 5.69
Completion Status	Complete	Salinity	32.23 35.01

Bottles		Saved							
		Add Bottle	5m						
Unique ID	Depth (m)	TIC/pH	Chlorophyll	Nutrients	Oxygen	Salinity	Phytoplankton	Comment	Complete?
496740	5	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Yes X
496739	25	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Yes X
496738	50	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Yes X
496737	100	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Yes X
496736	200	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Yes X
496735	300	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Yes X
496734	400	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		Yes X
496733	deepest bottle 683	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		Yes X

Oxygen Subsample									
Bottle	Operator	Winklers #1 (ml/L)	Winklers #2 (ml/L)	Winklers #3 (ml/L)	Comment	Complete?			
496740 @ 5	kevin	7.569							Yes
496735 @ 300	kevin	3.29	3.302						Yes
496733 @ 683	kevin	5.07	5.05						Yes

Figure 10. A screenshot of the Oceanographic Activity detail page from the Andes Oceanography module. Note that 'Yellow Belly' refers to the colloquial name of the CTD/Rosette system.

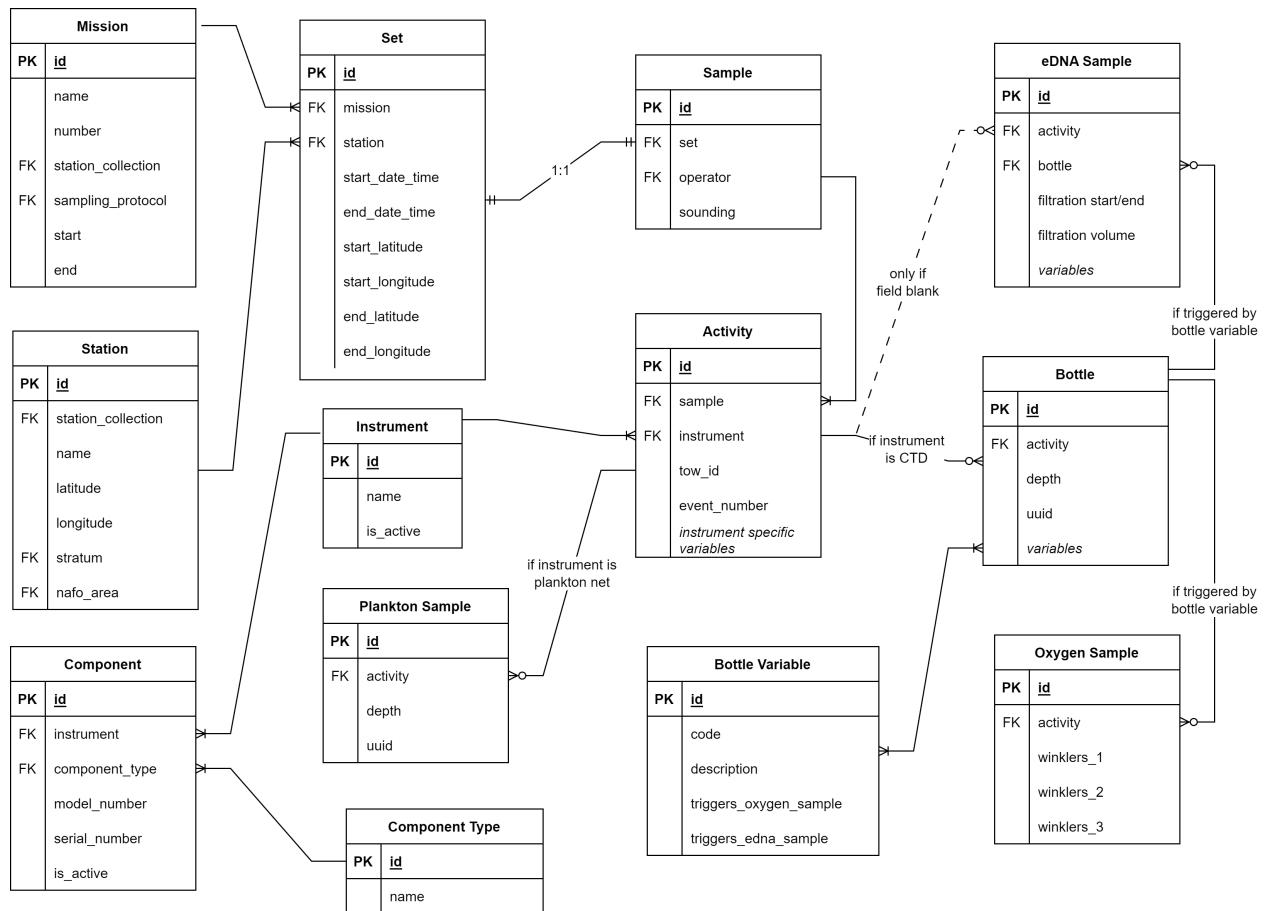


Figure 11. A simplified entity relationship diagram of the data model used by the Oceanography module of Andes.

Andes Settings Admin Preset: GFC Active Protocol: Herring NB Help français @dave

Select a size (cm)

Delete Mode View Summary Manual Mode

5	5.5	6	6.5	7	7.5	8	8.5	9	9.5 1
10	10.5 1	11	11.5	12	12.5 1	13	13.5	14	14.5
15	15.5	16 1	16.5 1	17 2	17.5 1	18 1	18.5 1	19 1	19.5 1
20 1	20.5 2	21 5	21.5 5	22 6	22.5 3	23 4	23.5 4	24 4	24.5 3
25 3	25.5 3	26 2	26.5 3	27 2	27.5 2	28 1	28.5 1	29 1	29.5 1
30 1	30.5 1	31 1	31.5	32	32.5	33	33.5	34	34.5
35	35.5	36	36.5	37	37.5	38	38.5 2	39	39.5
40	40.5	41	41.5	42	42.5	43	43.5	44	44.5
45	45.5	46	46.5	47	47.5	48	48.5	49	49.5
50									

MEASURED: 73 / KEPT: 52

[Back to Sample](#)

Figure 12. A screenshot displaying the Port Sampling module protocol form. The selections made in this form will affect the behaviour of the module during data entry.

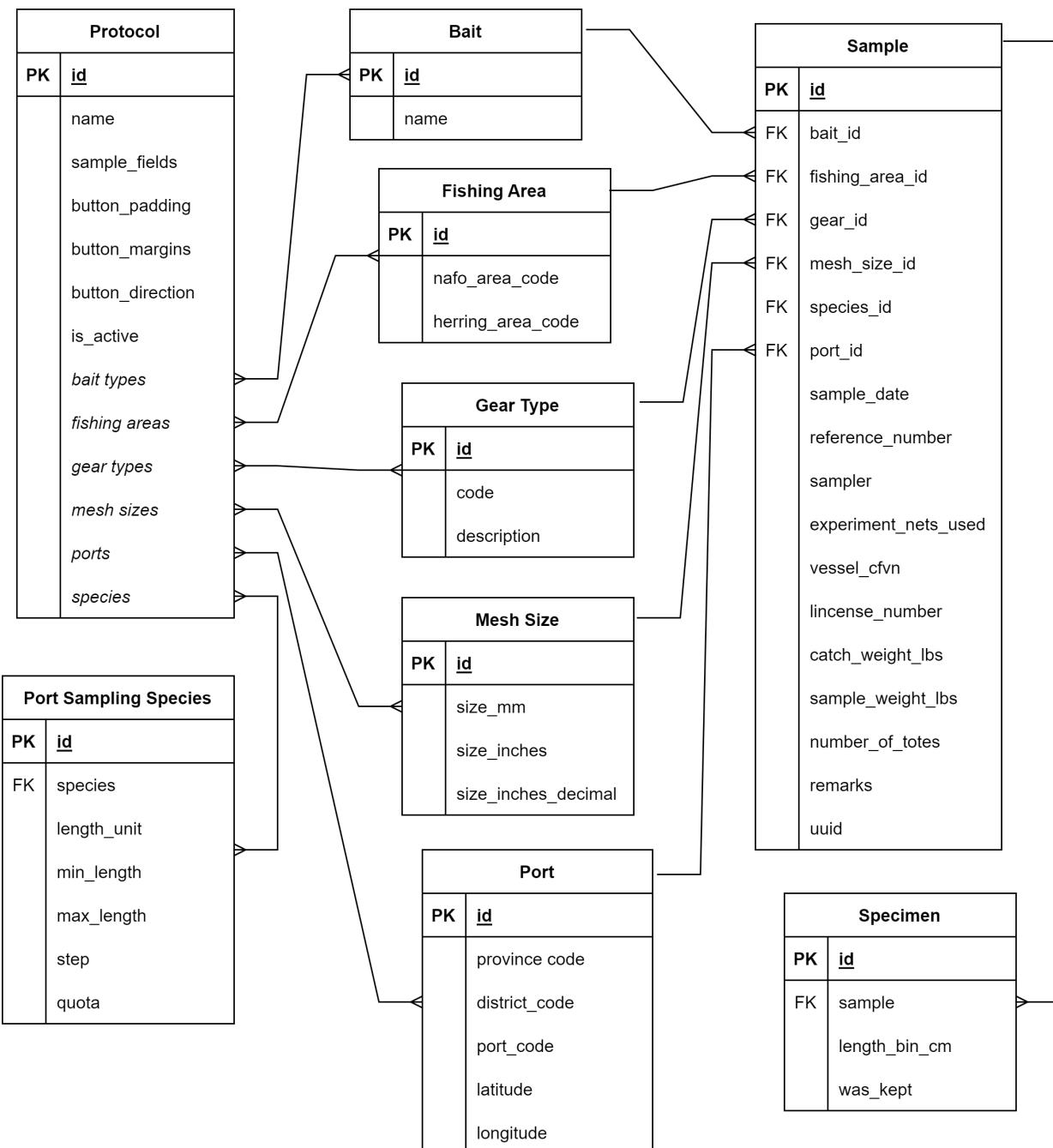


Figure 13. A simplified entity relationship diagram of the data model used by the Port Sampling module of Andes.

Home / CCG Dashboard / Fishing Console

WARNING: Your current location seems to be outside the desired stratum!! ✖

Net Deployed
Doors Deployed
Warp Deployed
Net On Bottom
Haul Back
Net Off Bottom
Door Recovered
Net Recovered

← Back
Abort
Set Card

Start Scanmar = net on bottom
 Stop Scanmar = net off bottom

Net Sensor Data

	Height	Opening	Clearance
Trawl Eye	55.19 m 2 months ago	39.19 m 2 months ago	13.88 m 2 months ago
Depth, Door #1	93.13 m	2 months ago	
Depth, Door #2	84.81 m	2 months ago	
Depth, Head Rope	95.88 m	2 months ago	
Pitch, Scanfactor (Port)	3.5 °	2 months ago	
Pitch, Scanfactor (Starboard)	3.5 °	2 months ago	
Pitch, Door (Port)	7.5 °	2 months ago	
Pitch, Door (Starboard)	3.38 °	2 months ago	
Roll, Scanfactor (Port)	1.94 °	2 months ago	
Roll, Scanfactor (Starboard)	1.19 °	2 months ago	
Roll, Door (Port)	-8.94 °	2 months ago	
Roll, Door (Starboard)	-11.63 °	2 months ago	
Speed Over Ground	0.0 kts	2 months ago	

Fishing Status:

Type	Latitude	Longitude	Date/time	
net on bottom	43° 3.9180'	-63° 38.1490'	2023-03-29 16:01:54 ADT	2 months ago ✓ ✗
warp deployed	43° 3.5160'	-63° 37.8790'	2023-03-29 15:53:20 ADT	2 months ago ✓ ✗
doors deployed	43° 3.1270'	-63° 37.6460'	2023-03-29 15:49:43 ADT	2 months ago ✓ ✗
net deployed	43° 2.9010'	-63° 37.5130'	2023-03-29 15:45:36 ADT	2 months ago ✓ ✗

Timer

0 min 48 sec

Tow Distance

188.815 nm

Figure 14. A screenshot of Andes Fishing Console within the Bridge module. This screenshot shows some quality control validation features including: the fishing timer, the distance display and an alarm message for when the vessel is travelling outside the target stratum. On the bottom left of the screenshot, output from the net monitoring system is displayed. The large buttons in the center of the display are used to trigger fishing events.

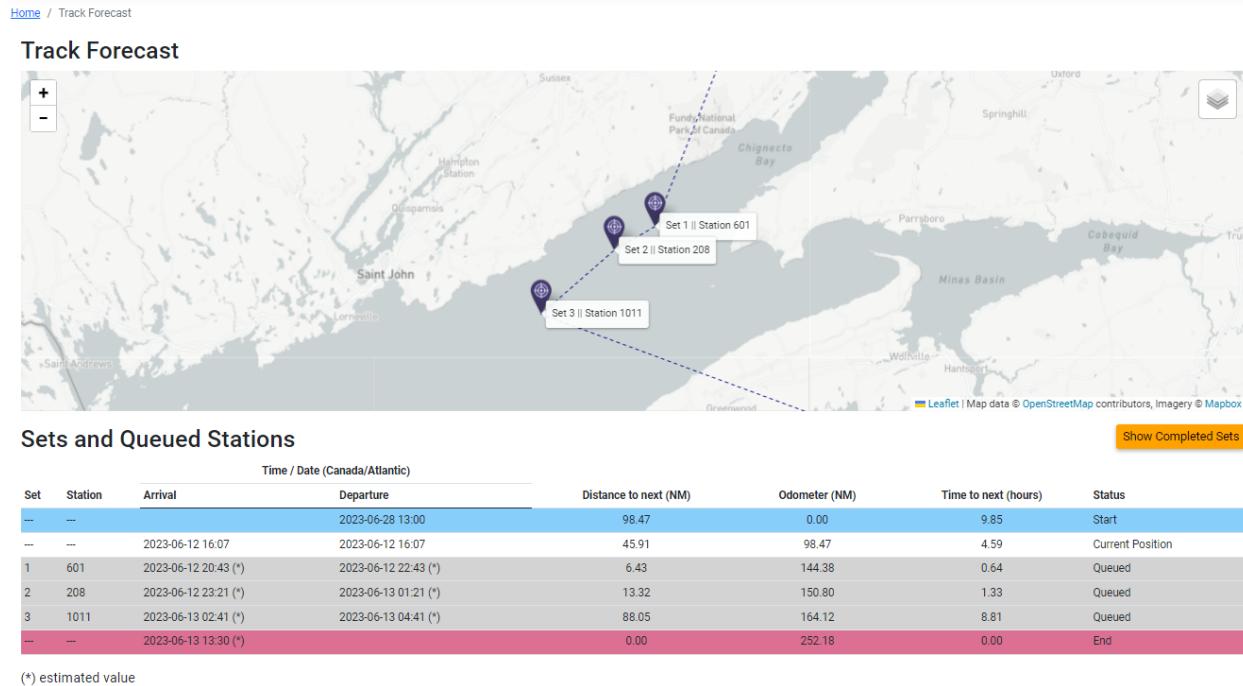


Figure 15. A screenshot of the Track Forecast tool.

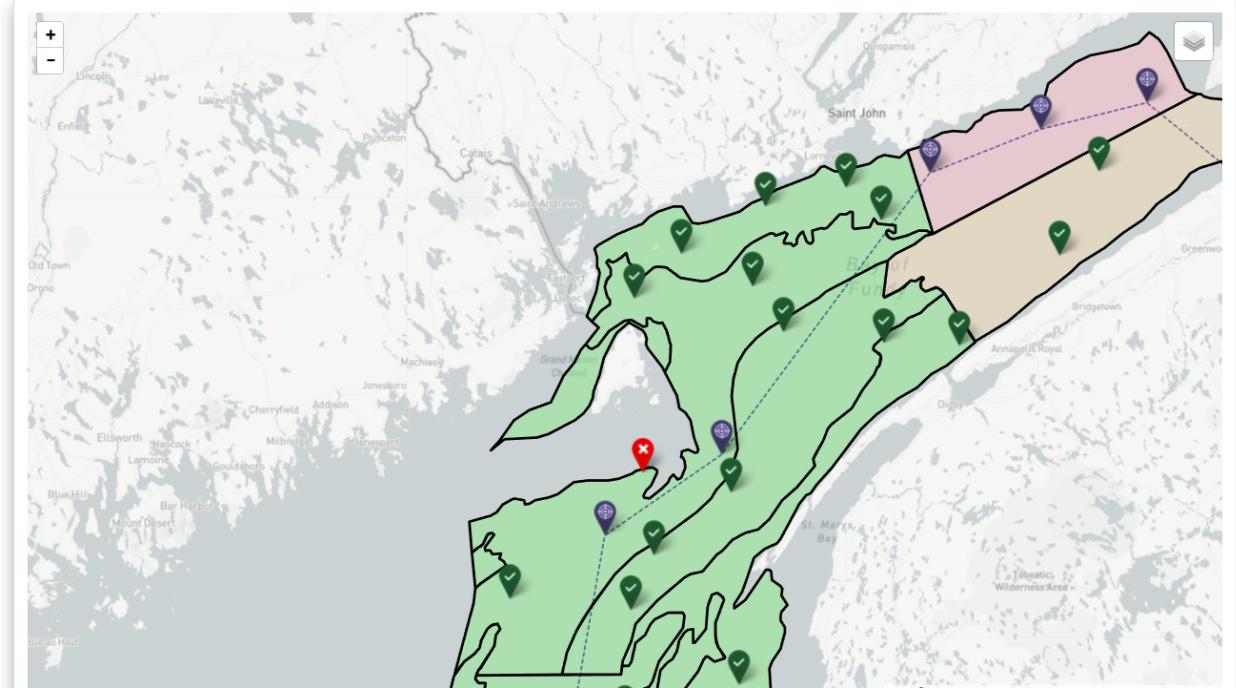
[Stations](#) [Show Track](#) [Hide Strata](#) [Show NAFO Areas](#)

Figure 16. A screenshot of the Andes Progress map. The map colors strata by the status of their set quotas. The green strata are those which have met their specified targets of set conducted; those in yellow have not met the target but have surpassed the minimum; and those in red have not conducted the minimum specified number of sets.

Length-to-Weight Ratio

A+B coefficients (weight = a × length^b)

	A	B
Unspecified	0.00561 The A regression coefficient in the relationship between length and weight for unspecified sex.	3.125999999 The B regression coefficient in the relationship between length and weight for unspecified sex.
Male	0.006227218 The A regression coefficient in the relationship between male length and weight.	3.096902077 The B regression coefficient in the relationship between male length and weight.
Female	0.006141017 The A regression coefficient in the relationship between female length and weight.	3.101193464 The B regression coefficient in the relationship between female length and weight.

Tolerance threshold for length-to-weight ratio (%)

25

The higher the threshold, the more relaxed Andes will be in raising warnings

Wait for sex before testing length-to-weight ratio? If this is set to true, Andes will only check length-to-weight ratio after sex has been observed.

Figure 17. A screenshot of the “Length-to-Weight Ratio” section of the sampling requirement form. Andes offers the option to specify regression coefficients for males, females or unspecified individuals to be used in quality control.

Mature Length

Mature length threshold

Unspecified	30 What is the min length of a mature individual of this species before which a user should receive a warning?
Male	Mature length (male)
Female	Mature length (female)

Maturity observation type

Maturity

Code used to signal an immature specimen

1
Only applicable if a maturity observation type is provided above

This screenshot shows the 'Mature Length' section of a sampling requirement form. At the top, there's a blue header bar with the title 'Mature Length'. Below it, a table is used to specify 'Mature length threshold' for different categories: 'Unspecified', 'Male', and 'Female'. The 'Unspecified' row contains the value '30' and a note asking what the minimum length of a mature individual of the species is before a warning is issued. The 'Male' and 'Female' rows have empty input fields labeled 'Mature length (male)' and 'Mature length (female)' respectively. Below the table, there's a section for 'Maturity observation type' with a dropdown menu set to 'Maturity'. At the bottom, there's a section for 'Code used to signal an immature specimen' with a single digit '1' entered, accompanied by a note that this is only applicable if a maturity observation type is provided.

Figure 18. A screenshot of the “Mature Length” section of the sampling requirement form. Andes offers the option to specify mature length thresholds for males, females or unspecified individuals to be used in quality control.

702

9 Acronyms

703 **Andes** Another data entry system.

704 **API** Application Programming Interface.

705 **AZMP** Atlantic Zone Monitoring Program.

706 **CDOS** Chief Digital Officer Sector.

707 **CSS** Cascading Style Sheet.

708 **DFO** Fisheries and Oceans Canada, formerly the Department of Fisheries and Oceans.

709 **ERD** Entity Relationship Diagram.

710 **ESE** Ecosystem Survey Entry.

711 **GSE** Groundfish Survey Entry.

712 **html5** Hypertext Markup Language.

713 **IMTS** Information Management and Technology Services.

714 **LAN** Local Area Network.

715 **MPA** Marine Protected Area.

716 **MRR** Module des Relevés de Recherche.

717 **NAFO** Northwest Atlantic Fisheries Organisation.

718 **REST** Representational state transfer (REST).

719 **REST-API** Representational state transfer (REST) Application programming interface (API).

720 **VCS** Version Control System.

721 **WAN** Wide Area Network.

10 Glossary

723 **Andes** Another data entry system is an application developed by Fisheries and Oceans Canada
724 to support data collection for a variety of scientific programs.

725 **Andes lead** A person with a functional understanding of the inner workings of Andes. This
726 person has the knowledge and user-rights that permit them to configure a Mission, define
727 a Sampling Protocol, Sampling Requirements, etc.

728 **API** An Application Programming Interface (API) is a particular set of rules and specifications
729 that a software program can follow to access and make use of the services and resources
730 provided by another particular software program that implements that API.

731 **AZMP** The Atlantic Zone Monitoring Program collects and analyses the biological, chemical
732 and physical oceanographic field data in the four Altantic regions of Fisheries and Oceans
733 Canada.

734 **Bridge** The Bridge app is meant to be used by the navigation officer while the fishing officer
735 deploys and retrieves the trawl. The operator inputs fishing events (eg., net deployed,
736 doors deployed, winches locked, net on bottom, haul back, net off bottom, doors recovered,
737 net recovered) or actions (eg., net on/off bottom) directly to Andes via the Bridge app. .

738 **Catch** A table of the Ecosystem Surveys app. A Catch table contains a link to Set and Species
739 entries, as well as an optional: specimen count, unweighted baskets, relative abundance
740 category and an invertabrate catch ratio. Catches can be subdivided into sub-catches
741 having a parent-children relationship.

742 **Catch Card** Catch details recorded prior to detailed sampling.

743 **CCG** The Canadian Coast Guard.

744 **CDOS** The Chief Digital Officer Sector is the IT branch of DFO.

745 **Closed Set** An Open Set Can be closed if the data it contains (eg, from Catch Card) does not
746 trigger a Set Flag.

747 **Configuration Preset** Configuration for a particular deployment scenario. Configuration presets
748 include details on the backup locations, and peripherals like GPS, sonars, label printers,
749 etc. .

750 **Cruise** Same as Mission.

751 **CSS** A Cascading Style Sheet is a style sheet language used for describing the presentation of
752 a document written in a markup language such as HTML or XML.

753 **CTD** Conductivity, salinity and density.

754 **Current Set** The Set that is currently occurring in time. It has begun and is not finished. There
755 can only be one Current Set at a time.

756 **DFO** Fisheries and Oceans Canada is a department of the Government of Canada that is
757 responsible for developing and implementing policies and programs in support of Canada's
758 economic, ecological and scientific interests in oceans and inland waters.

759 **Ecosystem Survey** The Ecosystem Survey app is used to as the main data input interface. .

760 **ESE** The Ecosystem Survey Entry replaced the GSE in the early 2000s.

761 **Fishing Event** An event related to fishing activity of Set. Metadata such as GPS coordinates
762 and time to are associated with the following event types: *net deployed, doors deployed,*
763 *warp deployed, net on bottom, haul back, net off bottom, door recovered, net recovered,*
764 *and aborted*. A Set may contain a list of such event types. The Fishing Events are added
765 to the Set by a crewmember using the Bridge application. .

766 **GSE** The GSE is a data entry tool developed in the 1980s.

767 **html5** The Hypertext Markup Language is the standard markup language for documents
768 designed to be displayed in a web browser.

769 **IMTS** Information Management and Technology Services is the IT branch of DFO.

770 **JavaScript** Javascript is a programming language that is one of the core technologies of the
771 World Wide Web, alongside HTML and CSS.

772 **JSON** JavaScript Object Notation. The JSON format is a popular way to serialize data and is an
773 alternative to the CSV (comma-separated values) and XML (eXtensible Markup Language)
774 formats.

775 **LAN** A series of computers connected to each other and capable of communicating with each
776 other over wired or wireless connections.

777 **Mission** Same as Cruise .

778 **NAFO** Northwest Atlantic Fisheries Organisation.

779 **Open Set** A Set that has not been closed. All new Sets will start in an open state and will
780 remain so until closed. Sets that are activated will automatically be opened. The Current
781 Set is always open.

782 **ORM** Object Relational Mapping are an abstraction of relational entities (database tables) as
783 objects.

784 **Port Sampling** Scientific program collecting samples from commercial fishing activities.

785 **REST-API** Representational state transfer (REST) Application programming interface (API) is
786 an API that follows the REST software architectural style, created to guide the design and
787 development of the architecture for the World Wide Web.

- 788 **Sampling Requirements** The species-specific requirements for samples to be collected during
789 a scientific cruise.
- 790 **Sampling Protocol** The detailed description of what data is to be collected during a scientific
791 cruise.
- 792 **Set** A Set contains all fishing activity and sampling results for a particular Station .
- 793 **Set Flag** A flag used to indicate that the data contained within the set has failed to pass a
794 specific validation tests. Closing an Open Set with active flags can be done by overriding
795 the validation mechanism.
- 796 **Species** An identifiable taxon that can be assigned to a species code.
- 797 **Station** A target location specified by coordinates where a scientific activity is to take place.
- 798 **VB** Pre-.NET Visual Basic for Applications is the early version of Microsoft VBA and is no longer
799 supported or updated by Microsoft.
- 800 **VBA** Visual Basic for Applications is a programming language built into most desktop Microsoft
801 Office applications. More details can be found on the Wikipedia page for this programming
802 language and from Microsoft.
- 803 **VCS** A Version Control System records changes to a file or set of files over time so that specific
804 versions can be recalled later. For example, git is a VSC.
- 805 **Vue.js** A JavaScript framework for building reactive user interfaces. <https://vuejs.org/>.
- 806 **WAN** A telecommunications network that extends over a large geographic area, such as the
807 internet.

808

APPENDIX A Tasks for administrative user

809 Prior to a survey, a number of configurations must be done for Andes to function as intended.
810 At a minimum, the following information must be provided and ingested by Andes in order for
811 a mission to take place:

812 **Strata polygons**

813 **List of stations**

814 **Species list**

815 **Sampling requirements**

816

APPENDIX B Further resources about Andes

817 The main git repository supporting the development of the application is hosted by IMTS. The
818 following resources can be found there:

- 819 • deployment checklist
- 820 • species lists for Gulf, Maritimes and Quebec
- 821 • strata polygons
- 822 • sampling requirements from previous surveys