



Systemes d'information pour la gestion de l'education

4 ÉDUCATION DE QUALITÉ



Efficiency and effectiveness in the choice and use of an MIS

Recommendations concerning the management of data and the functionalities of Information Systems for the Management of Education (MIS)

Chris van Wyk et Luis Crouch

Systemes d'information pour la gestion de l'education

Efficiency and effectiveness in the choice and use of an MIS

Recommendations concerning the management of data and the functionalities of Information Systems for the management of education (MIS)

Chris van Wyk et Luis Crouch



UNESCO

L'Acte constitutif de l'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO) a été adopté par 20 pays lors de la Conférence de Londres en novembre 1945 et est entré en vigueur le 4 novembre 1946. L'UNESCO compte actuellement 195 États membres et 11 membres associés.

L'objectif premier de l'UNESCO est de contribuer au maintien de la paix et de la sécurité dans le monde en resserrant, par l'éducation, la science et la culture, la collaboration entre nations afin d'assurer le respect universel de la justice, de la loi, des droits de l'homme et des libertés fondamentales pour tous, sans distinction de race, de sexe, de langue ou de religion, que la Charte des Nations Unies reconnaît à tous les peuples.

L'UNESCO a cinq fonctions principales inscrites dans son mandat : 1) des études prospectives sur l'éducation, la science, la culture et la communication pour le monde de demain ; 2) le progrès, le transfert et le partage des connaissances par des activités de recherche, de formation et d'enseignement ; 3) des actions normatives en vue de la présentation et l'adoption d'instruments internes et de recommandations réglementaires ; 4) l'expertise par le biais de la coopération technique aux États membres en faveur de leurs projets et politiques de développement ; 5) l'échange d'informations spécialisées.

Institut de statistique de l'UNESCO

L'Institut de statistique de l'UNESCO (ISU) est le bureau de la statistique de l'UNESCO et le dépositaire des Nations Unies pour les statistiques mondiales en matière d'éducation, de science et de technologie, de culture et de communication.

L'ISU est la source officielle des données comparables au plan international, utilisées pour effectuer le suivi des progrès accomplis vers la réalisation de l'Objectif de développement durable sur l'éducation, et des cibles principales liées à la science, la culture, la communication et l'égalité des sexes.

Publié en 2020 par :

Institut de statistique de l'UNESCO
P.O. Box 6128, Succursale Centre-Ville
Montréal, Québec H3C 3J7 Canada

Tél. : +1 514-343-6880
Courriel : uis.publications@unesco.org
<http://www.uis.unesco.org>

ISBN : 978-92-9189-248-8
Réf. : UIS/2020/ED/TD/2

© UNESCO-UIS 2020

Cette publication est disponible en accès libre sous la licence Attribution-ShareAlike 3.0 IGO (CC-BY-SA 3.0 IGO) (<http://creativecommons.org/licenses/by-sa/3.0/igo/>). En utilisant le contenu de cette publication, les utilisateurs acceptent de se conformer aux conditions d'utilisation du service d'archive des publications en accès libre de l'UNESCO (<http://fr.unesco.org/open-access/terms-use-ccbysa-fr>).

Les appellations employées dans cette publication et la présentation des données qui y figurent n'impliquent de la part de l'UNESCO aucune prise de position quant au statut juridique de tout pays, territoire, ville ou zone ou de leurs autorités, ni quant au tracé de leurs frontières ou limites.

Les idées et opinions exprimées dans cette publication sont celles des auteurs. Elles ne représentent pas nécessairement les vues de l'UNESCO et n'engagent à rien l'Organisation.

Table des matières

ACRONYMES & ABRÉVIATIONS	5
LISTE DES FIGURES, ENCADRÉS ET TABLEAUX	6
1. INTRODUCTION	7
2. À PROPOS DU « GUIDE DE L'ACHETEUR » ET DU « GUIDE DE L'UTILISATEUR »	9
3. DÉFINITION DU SIGE	11
4. ARCHITECTURE DU SIGE	13
5. GUIDE DE L'ACHETEUR	15
5.1 Périmètre et contexte du SIGE	15
5.1.1 Les composantes clés du SIGE	15
5.1.2 Comment les indicateurs de l'ODD 4 ont modifié le périmètre des SIGE	17
5.1.3 L'importance d'une politique de données ou SIGE.	19
5.2 Comment se procurer un SIGE	20
5.3 Un cadre de qualité totale des données.	22
5.4 Fonctionnalité et normes du SIGE – Aspects du Guide de l'acheteur	24
5.5 Conditions et normes minimales supplémentaires concernant la technologie	34
5.5.1 Exigences en matière de logiciels	35
5.5.2 Exigences en matière de bases de données	36
5.5.3 Exigences en matière de télécommunications et de réseaux de données (y compris les réseaux et Internet)	38
5.6 Considérations relatives aux futures technologies	39
6. GUIDE DE L'UTILISATEUR	41
6.1 Besoins en personnel	41
6.2 Cycle de vie de production du SIGE	42
6.2.1 Élaboration du questionnaire d'enquête (Conception et développement)	42
6.2.2 Attribution d'identifiants uniques	46
6.2.3 Maintenance du répertoire (registre) des établissements scolaires	47
6.2.4 Saisie des données	48
6.2.5 Stockage de données.	49

6.2.6 Interopérabilité des données	50
6.2.7 Analyse des données et rapports	51
6.2.8 Publication, diffusion et exploitation des données	52
BIBLIOGRAPHIE.	55
ANNEXE A. LISTE DES CONTRIBUTEURS.	57
ANNEXE B. LISTE DES INDICATEURS DE L'ODD 4 ET CONSIDÉRATIONS RELATIVES AU SIGE.	59
ANNEXE C. EXEMPLE DE CALENDRIER DE COLLECTE SIGE	73

Acronymes & abréviations

ADEA	Association pour le développement de l'éducation en Afrique
DHIS	Logiciel d'information sanitaire de district
HISP	Programme des systèmes d'information sur la santé
ISU	Institut de statistique de l'UNESCO
ME	Ministère de l'Éducation
ODD	Objectif de développement durable
OLAP	Traitement analytique en ligne
PME	Partenariat mondial pour l'éducation
Saas	Logiciel en tant que service
SGBD	Système de gestion de base de données
SGBDR	Système de gestion de base de données relationnelle
SIGE	Système d'information pour la gestion de l'éducation
STT	Système de traitement transactionnel
UA	Union africaine
UNESCO	Organisation des Nations Unies pour l'éducation, la science et la culture
URS	Cahier des charges

Liste des figures, encadrés et tableaux

LISTE DES FIGURES

Figure 1. Composantes d'une architecture d'entreprise	13
Figure 2. Architecture type des données d'un SIGE	14
Figure 3. Cycle de vie de production du SIGE	42

LISTE DES ENCADRÉS

Encadré 1. Les « normes »	10
Encadré 2. Données à collecter en urgence en période de COVID-19	18
Encadré 3. Synthèse des principes des Nations Unies en matière d'établissement de rapports sur les indicateurs des ODD.	24
Encadré 4. Compétences techniques attendues du personnel SIGE, au niveau national, infranational et scolaire	41
Encadré 5. Méthode du recensement scolaire : saisie manuelle ou via le web ?	43
Encadré 6. Écarts entre la collecte manuelle de données et les écrans de saisie SIGE	44
Encadré 7. Stratégie d'interopérabilité	50

LISTE DES TABLEAUX

Tableau 1. Dimensions de la qualité des données	23
Tableau 2. Normes de fonctionnalité du SIGE, par niveau de la collecte de données	25
Tableau 3. Voies d'évolution de la technologie.	40
Tableau B.1. Indicateurs de l'ODD 4 et leur traitement possible par le SIGE	60

1. Introduction

L'Organisation des Nations Unies pour l'éducation, la science et la culture (UNESCO) et le Partenariat mondial pour l'éducation (PME) ont uni leurs efforts et tenu la première Conférence internationale sur les systèmes d'information pour la gestion de l'éducation (SIGE) au Siège de l'UNESCO du 11 au 13 avril 2018.¹ Les représentants des unités nationales SIGE ont fréquemment exprimé leur besoin d'orientation de la part de la communauté internationale concernant i) les modalités de prise de décisions pour la mise à niveau de leur SIGE ; ii) une base technique sur laquelle négocier avec les prestataires ou les fournisseurs et iii) les moyens de discuter au mieux de leurs besoins avec les partenaires de développement qui souhaitent contribuer à l'élaboration de leurs données sur l'éducation, une préoccupation motivée par l'absence d'informations comparatives sur ce qu'un SIGE moderne devrait être capable de faire.² Contrairement au secteur de la santé où un système standard par défaut a émergé sous la forme du *District Health Information Software* (DHIS - logiciel d'information sanitaire de district), on a assisté dans le secteur de l'éducation à la prolifération de systèmes proposés par différents donateurs et fournisseurs. Par ailleurs, au niveau des ministères de l'Éducation, il n'y a pas de vision claire sur les interactions nécessaires entre le SIGE et les autres systèmes de données (ex. évaluations, RH enseignants). De ce point de vue, un besoin s'est également manifesté de disposer

de moyens plus unifiés de diagnostic pour le SIGE, ainsi que d'une plate-forme SIGE standard, plus normalisée (par exemple pour assurer l'éducation dans les situations d'urgence). Le présent guide SIGE – composé essentiellement d'un « Guide de l'acheteur » et d'un « Guide de l'utilisateur » – constitue une tentative de la part de l'Institut de statistique de l'UNESCO (ISU) et du PME de répondre aux besoins exprimés par les pays. Il est le fruit de nombreuses discussions au sein de diverses instances qui ont précédé et inclus une réunion de la Table ronde sur les solutions aux problèmes de données dans l'éducation de septembre 2019, en marge de l'Assemblée générale des Nations Unies.³

La dernière décennie a vu l'apparition d'une pléthore de systèmes de collecte, de gestion et de diffusion des données sur l'éducation. Le choix du système à utiliser représente donc une tâche complexe et ardue car les capacités du système doivent être alignées sur les objectifs et les priorités du pays. L'objectif de ce rapport est donc d'aider les pays à faire un choix plus éclairé pour s'équiper de solutions logicielles SIGE spécifiques et pertinentes. Comme son nom l'indique, le présent rapport est un guide (ou un manuel) – son objectif n'est pas de théoriser (sauf si cela est absolument nécessaire), ni de rendre compte de résultats de recherches ou de s'engager dans un plaidoyer. Il s'adresse à un public technicien averti (données, informatique, planification) dans les pays en développement ou dans les agences internationales et les établissements universitaires qui travaillent avec eux. Néanmoins, les décideurs

1 Voir <https://fr.unesco.org/themes/politique-planification-education/emis-2018>

2 Voir <https://en.unesco.org/news/unesco-gpe-launch-first-international-conference-education-management-information-systems-0>. On a également constaté un besoin d'harmonisation des différents cadres d'évaluation du SIGE utilisés par les agences internationales (Banque mondiale, UNESCO, UNICEF, ADEA). Même s'il s'agit d'un besoin valable et important, le présent guide met l'accent sur les systèmes eux-mêmes.

3 <https://www.globalpartnership.org/fr/content/resultats-de-la-table-ronde-sur-les-solutions-aux-problemes-de-donnees-dans-leducation> – voir la Recommandation 4.

non techniciens d'un ministère pourront aussi trouver des informations utiles dans certaines des premières sections en particulier.

Pour être efficace, la prise de décision doit se fonder sur des données de qualité gérées par des systèmes d'information efficaces. L'information est une ressource nécessaire, le produit des systèmes d'information, et c'est un élément essentiel de la gestion et de la prise de décision dans le secteur de l'éducation. Un SIGE n'est pas seulement une solution technologique limitée aux processus opérationnels : en tant que système d'information, il doit faciliter la prise de décisions stratégiques, la formulation de politiques et la budgétisation, et, si cela est possible et pertinent, la gestion quotidienne au niveau des districts pour aider les écoles. En règle générale toutefois, le présent rapport ne traite pas des systèmes d'enregistrement des données au niveau des établissements scolaires, qui sont destinés à la gestion interne de ceux-ci.

Un système d'information pour la gestion de l'éducation (SIGE) constitue généralement une

fonction et une unité au sein du ministère de l'Éducation (ME) d'un pays, chargées de l'acquisition et de la tenue à jour d'un système intégré d'information sur l'éducation, pour la gestion de l'éducation. Une définition plus complète du SIGE est fournie à la Section 3.

Le SIGE a la responsabilité de recueillir, de traiter, d'analyser et de diffuser des données et des informations sur l'éducation dans le but d'appuyer le suivi et l'évaluation des performances du système éducatif. Que la responsabilité de la phase d'analyse relève du SIGE (ou de l'unité SIGE ?) ou que cela se fasse dans les bureaux de la planification, du budget ou des politiques par exemple, cela différera d'un pays à l'autre, mais le SIGE devrait dans l'idéal être implanté là où se trouve le bon profil de compétences (voir la Section 6.1 pour plus de détails sur les compétences techniques requises du personnel SIGE). Au minimum, l'unité SIGE peut offrir une valeur ajoutée en tant que service rendu à d'autres unités, même si ce n'est que pour produire des indicateurs qui leur sont pertinents.

2. À propos du « Guide de l'acheteur » et du « Guide de l'utilisateur »

Ce guide élaboré par un consultant SIGE désigné par l'ISU est le fruit d'un effort collectif⁴. Le processus d'élaboration s'est inspiré d'une revue documentaire et des contributions d'acteurs clés de plates-formes SIGE existantes et en développement. Il se divise en deux grandes parties :

1. le « Guide de l'acheteur » se réfère aux normes de fonctionnalité qu'un système SIGE devrait présenter (Section 5) ; et
2. le « Guide de l'utilisateur » se réfère à la manière d'exploiter au mieux le SIGE, lorsqu'un pays s'en est doté (Section 6).

Bien qu'un ME « n'achète » jamais vraiment un SIGE – car le SIGE s'apparente plutôt à un écosystème au sein duquel on effectue des opérations et qui relie de nombreuses unités du ministère de l'Éducation, d'autres ministères, du monde universitaire et de la société civile – bon nombre de décisions à prendre lors de l'acquisition ou de la mise à niveau d'un

SIGE ressemblent à un processus d'achat. Le « Guide de l'acheteur » - un terme que l'on utilisera ici dans son acception la plus large - définit les normes de fonctionnalité du SIGE, quelle que soit l'approche effectivement adoptée, acquisition ou développement. (Voir l'Encadré 1 pour le sens que nous conférons au terme « normes », quelque chose qui n'est pas prescriptif.) Selon notre hypothèse, les ministères achètent pour la plupart (ou sont équipés par un partenaire de développement) de nouveaux éléments ajoutés à « une épine dorsale SIGE »⁵ existante, ou bien ils se procurent des outils prêts à l'emploi (produits commerciaux) pour remplacer ou ajouter de nouvelles fonctionnalités. Il pourra arriver aussi qu'un ministère souhaite refondre entièrement son SIGE. Par ailleurs, étant donné que *certain*s aspects au moins de la mise à niveau d'une plate-forme existante (ou très occasionnellement du développement d'une plate-forme entièrement nouvelle) peuvent être achetés par un ministère ou achetés et fournis par un partenaire de développement, cette section inclut des conseils concernant les marchés publics. Après avoir « acheté » un système, l'acheteur le reçoit et s'engage ensuite dans un processus d'utilisation - d'où le « Guide de l'utilisateur » qui vient ensuite. Encore une fois, cette désignation est très générale et répond plutôt à un objectif narratif.

4 Chris van Wyk (Université de Stellenbosch, Afrique du Sud), chrisvanwyk@sun.ac.za. L'élaboration de ce rapport n'aurait pas été possible sans la coopération, la consultation et la participation des représentants d'organisations de plates-formes SIGE existantes et en développement. La collaboration des représentants de ces organisations a été extrêmement utile et nous leur adressons nos plus vifs remerciements pour leur aide. Luis Crouch (RTI International, consultant individuel auprès du PME) a collaboré au projet de rédaction. Les contributeurs qui ont rédigé certaines petites sections ou qui nous ont aidés de toute autre manière sont cités à l'Annexe A.

5 Pour la définition et une description, voir la Section 3.

Encadré 1. Les « normes »

Dans ce guide, il est question à plusieurs reprises de « normes ». Il ne s'agit pas d'imposer une solution aux besoins des pays. Au contraire, ces normes sont censées être utiles aux unités SIGE en leur indiquant ce que peuvent faire (ou aspirent à faire) les unités SIGE d'autres pays, et ce qui semble être de l'avis des experts une bonne idée. Nous ne suggérons cependant pas qu'un système doive être conforme à toutes ces normes pour être considéré comme un système de qualité. Bien des adaptations sont possibles, et même nécessaires, au niveau local. En voici deux exemples :

Niveau de décentralisation du pays. Dans un État véritablement fédéral, le rôle du niveau central dans la définition des normes peut être assez différent de celui d'un État unitaire, ou même d'un État extrêmement décentralisé. Par exemple, un État fédéral peut disposer de pouvoirs plus limités pour demander aux États fédérés ou aux provinces de mettre en œuvre une plate-forme logicielle commune. Dans tous les cas, cette solution risque de ne pas être optimale puisque les besoins peuvent varier. Les choix qui s'offrent aux États fédéraux sont exposés tout au long du document pour illustrer ces différences et la façon dont ces normes peuvent varier. Dans un État unitaire décentralisé, la structure de gouvernance du SIGE peut prévoir que des décisions conjointes soient prises dans le cadre d'un comité de responsables provinciaux du SIGE. Une approche de ce type peut être moins nécessaire dans un État centralisé.

Situation initiale de l'éducation dans le pays. Chaque pays peut avoir ses propres priorités quant à l'intensité des efforts à déployer pour les différents indicateurs. Par exemple, dans un pays où l'accès demeure un défi majeur, il faudra peut-être que les indicateurs d'accès soient plus granulaires ou affinés et établir plus de liens avec des bases de données susceptibles d'aider à déterminer les raisons de la faiblesse de cet accès.

Dans le cadre de la préparation de ce document, différentes plates-formes logicielles SIGE et connexes ont été examinées afin de dresser une liste des normes de fonctionnalité, en partant d'un raisonnement fondé sur les exigences habituelles en termes de rapports, notamment pour les indicateurs de l'Objectif de développement durable (ODD) 4, ainsi que pour les variantes régionales telles que celles qui ont été choisies pour le suivi de la Stratégie continentale de l'éducation pour l'Afrique de l'Union africaine (CESA 16-25) – mais sans accorder la primauté à ces dernières, car les besoins des pays doivent être primordiaux. L'objectif

a été d'élaborer un ensemble de normes portant sur les opérations que le SIGE doit être en mesure de réaliser pour fournir des informations précises et utiles aux décideurs du secteur de l'éducation, aux directeurs d'école et aux organisations internationales à qui la plupart des pays doivent rendre compte. La description des liens entre les différentes plates-formes existantes et les normes énoncées dans le « Guide de l'utilisateur » ci-dessous (Section 6) est provisoirement contenue dans un document *séparé* (et non une annexe) intitulé « Systèmes d'information pour la gestion de l'éducation : plates-formes logicielles ».

3. Définition du SIGE

Le présent guide s'appuie sur la définition suivante fournie par l'UNESCO :⁶

« Un SIGE peut se définir comme « un système de collecte, d'intégration, de traitement et de diffusion de données et d'informations pour aider la prise de décisions, l'analyse politique ainsi que la formulation, la planification, le suivi et la gestion du système éducatif à tous les niveaux. C'est un système dans lequel les individus, les technologies, les modèles, les méthodes, les processus, les procédures, les règles et les régulations fonctionnent ensemble pour fournir aux dirigeants du secteur de l'éducation, aux décideurs et aux gestionnaires de tous niveaux un ensemble global d'informations et de données intégrées pertinentes, fiables, non ambiguës et en temps voulu pour les aider à assumer leurs responsabilités. » (UNESCO, 2019)

Cette définition stipule, implique ou émet une hypothèse forte, selon laquelle :

- Le SIGE n'est pas une simple solution technique
- L'éducation est de nature intégrée et distribuée
- Tous les niveaux de la structure éducative doivent être pris en compte (institutionnel, infranational et national)
- Tous les systèmes et infrastructures techniques doivent être en place

6 Bien que les experts puissent ne pas tomber d'accord sur la façon de définir un SIGE, cette définition est conforme aux objectifs du présent guide.

Dans ce qui suit, « l'épine dorsale » du SIGE se réfère à l'ensemble des processus et des données qui constituent les décomptes administratifs traditionnels (tels que les effectifs scolarisés, le nombre d'enseignants, les fournitures), créés sur la base de recensements annuels (ou périodiques) et les décomptes des données sur les établissements scolaires, parfois augmentés de données démographiques ou d'autres données similaires, nécessaires pour créer les indicateurs d'accès et de flux d'un système éducatif (tels que le taux brut de scolarisation, le taux net de scolarisation, les estimations des taux d'achèvement pour chaque niveau). L'« entrepôt de données » (utilisé ici de façon interchangeable avec « datamart ou mini-entrepôt de données ») se réfère à l'ensemble des données qui peuvent être ajoutées et reliées aux données traditionnelles de l'épine dorsale du SIGE, pour autoriser d'importantes analyses à valeur ajoutée, que celles-ci soient effectuées par le bureau du SIGE ou par d'autres bureaux utilisateurs tels que la planification, la budgétisation, l'analyse des politiques, la formation des enseignants, la santé ou la pauvreté.⁷ Pour être utiles, ces bases de données doivent être mises en relation au moyen d'identifiants d'établissement scolaire, de district, etc. Que ce soit pour l'épine dorsale du SIGE proprement dite ou pour un entrepôt élargi, il conviendrait d'envisager

7 Il n'existe pas de distinction absolue entre ce qui constitue l'épine dorsale traditionnelle du SIGE et un entrepôt de données plus vaste et lié, qui est utile au ministère. Les décomptes traditionnels des effectifs scolarisés, établissement par établissement, font partie de l'épine dorsale. Les données sur les performances des apprenants fondées sur des évaluations par échantillons aléatoires ou sur des examens publics font typiquement partie d'un entrepôt élargi. Mais il y a des variables, comme la population de la circonscription de l'école, qui peuvent être si souvent utilisées qu'elles pourraient être considérées comme faisant partie de l'épine dorsale.

une collecte automatique des données, par exemple à partir d'un système de gestion de l'apprentissage ou des données administratives recueillies au niveau des établissements scolaires, bien que cela relève sans doute plutôt, dans les pays à revenu le plus faible, d'une orientation pour l'avenir (pour des scénarios plus futuristes, voir la Section 5.6.). Il existe enfin de nombreux indicateurs importants, souvent binaires ou qualitatifs, qui ne font pas partie à proprement parler de l'épine dorsale du SIGE, ou

même de l'entrepôt de données ; c'est le cas par exemple d'un pays qui a un système de financement scolaire pro-pauvres ou une politique curriculaire sur X ou Y. Même si de nombreuses parties prenantes sont intéressées à connaître ces éléments, et même s'ils font souvent partie des déclarations obligatoires du ministère (y compris sur l'ODD 4 – voir la Section 5.1.2), ils ne sont pas maintenus ou recueillis de manière optimale par un SIGE, mais plutôt par un bureau de planification ou de politique du ministère.

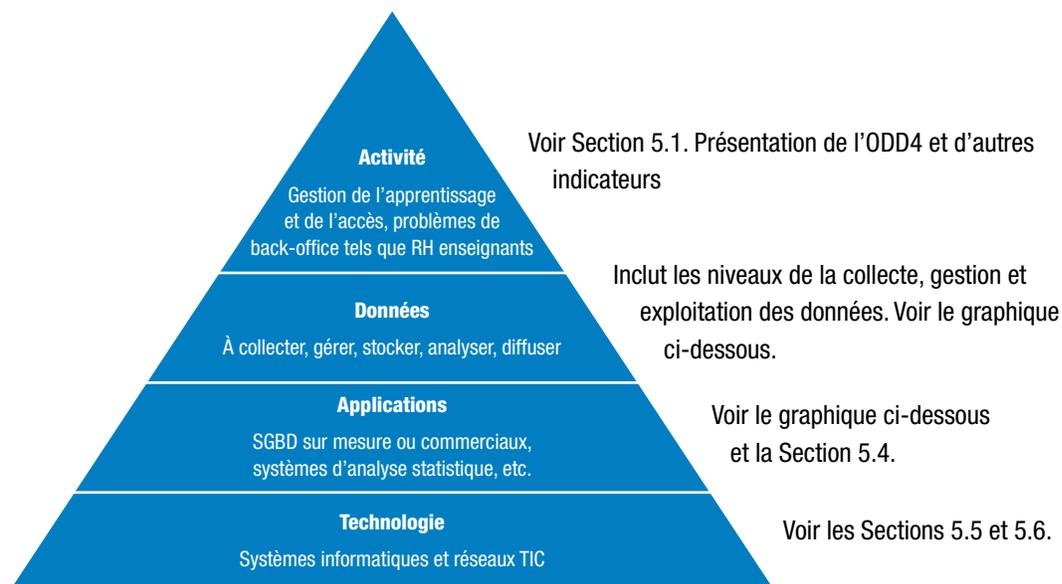
4. Architecture du SIGE

En termes simples, l'architecture du SIGE sert de fondement aux modes de collecte, de stockage et de rapport des données et elle est au cœur des Guides de l'acheteur et de l'utilisateur. Les composantes de toute solution logicielle incluent la *collecte de données*, la *gestion et l'exploitation d'une base de données* (qui inclut des *outils d'analyse* et de *rapport*). Les Guides de l'acheteur et de l'utilisateur s'appuient sur la présentation de l'architecture SIGE contenue dans cette section, où l'on trouvera le contenu principal et la structure de ces parties. En règle générale, l'architecture du SIGE (or de tout système d'information ou de toute organisation) doit refléter *la structure générale d'une entreprise* – dans ce cas

précis, celle d'un ministère. La Figure 1 illustre quatre composantes d'une architecture générale d'entreprise à envisager, à savoir :

- 1) *Architecture de travail* : les besoins/exigences nationaux et internationaux du ministère de l'Éducation – brièvement traités à la Section 5.1 (voir la discussion sur l'ODD 4).
- 2) *Architecture des données* – traitée ci-dessous.
- 3) *Architecture des applications* – traitée ci-dessous, ainsi qu'à la Section 5.4.
- 4) *Architecture technologique* : les technologies seront choisies en fonction des choix/besoins

Figure 1. Composantes d'une architecture d'entreprise



Source : Auteurs.

des niveaux ci-dessus – et traitées en partie à la Section 5.4 et décrites brièvement à la Section 5.5.

Voici en résumé les composantes les plus courantes de toute *architecture de données et des guides SIGE* :

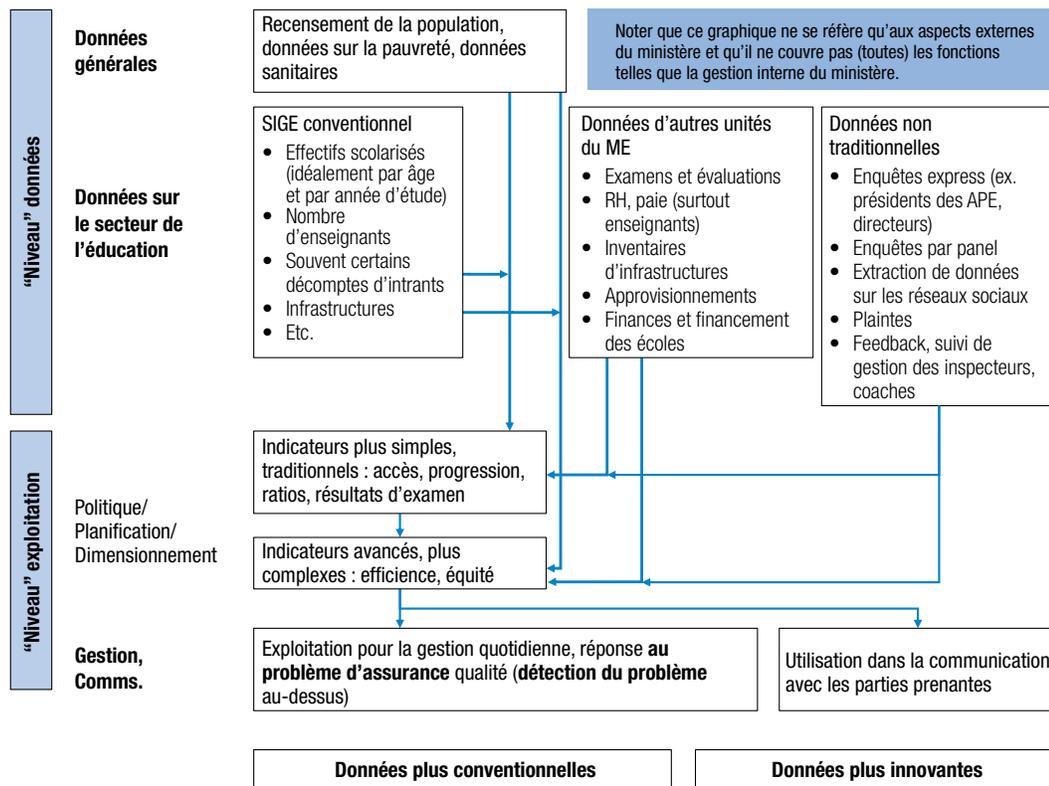
- *Niveau de la collecte de données* : les modalités de collecte des données
- *Niveau de la gestion et du stockage des données* : les modalités de gestion et de stockage des données dans le datamart/ entrepôt de données et leur transfert d'un niveau à l'autre du système s'il n'y a pas de véritables saisie et gestion réparties.

- *Niveau de l'exploitation des données* : les modalités d'analyse, de rapport et de diffusion des données.

La Figure 2 montre les relations qui les lient au sein de l'architecture de données, et les sources de données habituelles.

Les sections suivantes offrent un aperçu plus détaillé des composantes les plus courantes d'une architecture SIGE et de ses fonctions, activités et processus connexes, depuis la collecte de données, leur traitement et leur stockage, jusqu'à l'analyse, les rapports, la diffusion et l'exploitation des données.

Figure 2. Architecture type des données d'un SIGE



Source : Auteurs.

5. Guide de l'acheteur

5.1 Périmètre et contexte du SIGE

Par nature, les systèmes éducatifs sont extrêmement complexes et multiformes, ce qui signifie que le besoin de données, la planification, le suivi et l'évaluation, la prise de décision, la budgétisation, la formulation de politiques, l'affectation et la gestion des ressources, le tout fondé sur des données probantes, peuvent être tout aussi complexes et multiformes. Par ailleurs aucun système éducatif ne ressemble à aucun autre, même lorsqu'ils ont de nombreuses caractéristiques communes (par exemple, il est difficile d'imaginer, au moins pour le moment, un système éducatif sans enseignants, sans élèves, sans écoles, sans administrateurs ou sans manuels, ou sans un besoin de financement). Il est donc possible de conceptualiser une architecture de données relativement commune, mais on n'échappe jamais à la nécessité de s'adapter aux objectifs et aux exigences du ministère en particulier.

Le SIGE n'est pas conçu de manière optimale et mandaté pour produire *tous* les ensembles de données nécessaires pour répondre aux exigences du secteur de l'éducation. Il doit toutefois offrir à l'analyste des « *crochets* » (points d'entrée vers des actions) permettant d'intégrer d'autres données (ex. pauvreté, population), même si ce n'est qu'a posteriori et dans le cadre d'un processus à valeur ajoutée qui ne fait pas partie de l'architecture même du SIGE. L'interopérabilité des données (la capacité des différentes bases de données de se relier les unes aux autres, par exemple via des identifiants d'établissement communs) est un principe majeur de la production de données de qualité et un élément clé de l'architecture globale

du SIGE – c'est réellement la principale source de valeur ajoutée du SIGE aux échelons supérieurs du ministère. En outre, le SIGE doit être un système capable de répondre aux demandes actuelles, mais aussi d'évoluer pour suivre les demandes et les possibilités techniques émergentes et futures. Dans certains cas, il peut s'agir d'un service fourni par l'unité SIGE dans la création d'un entrepôt de données. De plus, le SIGE ne doit pas être considéré comme une solution technologique isolée qui se limite aux seuls processus opérationnels. Il ne doit donc pas tenter d'englober toutes les formes possibles de processus de collecte de données, car cela pourrait vite devenir ingérable et risquerait d'alourdir la charge de travail du personnel. Le SIGE ne peut tout simplement pas être le guichet unique de tous les processus de collecte de données requis dans le secteur de l'éducation ; c'est-à-dire qu'il ne peut pas être tout pour tous. Néanmoins, un SIGE solide doit être capable de se relier à tout le reste.

5.1.1 Les composantes clés du SIGE

Cette section décrit le contexte et le périmètre du SIGE, ainsi que les processus, les procédures et les défis liés à la collecte de données administratives dans le domaine de l'éducation, à prendre en compte lors de l'élaboration d'un SIGE intégré. Par « données administratives » nous entendons les données recueillies auprès des établissements scolaires par le bureau national chargé de la tenue des dossiers, de la planification, du suivi, de l'évaluation et de la formulation des politiques. À la base, le SIGE doit inclure les données concernant les établissements scolaires, les apprenants, les enseignants et les installations physiques. Un SIGE

complet devrait dans l'idéal inclure des données sur les effectifs scolarisés, la fréquentation, les taux d'achèvement, les évaluations de l'apprentissage, la santé des élèves, les finances, les caractéristiques des enseignants et certaines statistiques administratives.

Aux fins du présent guide et en fonction des systèmes agrégés de collecte de données (c'est-à-dire des ensembles de données généralement limités aux bases de données centralisées au niveau de l'école), le SIGE devrait inclure, entre autres :

Décomptes de base, effectifs et intrants :

- Effectifs scolarisés (par année d'étude, sexe, pauvreté, âge, matière, handicap, groupes linguistiques, etc.)
- Redoublement (par année d'étude, sexe), décrochage, achèvement, flux généraux
- Données démographiques sur les apprenants
- Enseignants (expérience, qualification, sexe, placement, nomination, etc.)
- Établissements scolaires (installations physiques, informations générales, implantation, salles de classe, etc.)

Autres composantes du secteur de l'éducation :

- Programme de repas scolaire (le cas échéant)
- Services de bibliothèque et d'archives
- Pensionnats (le cas échéant)
- Santé des apprenants
- Matériels d'apprentissage (manuels scolaires au minimum, en général)
- Données sur les performances des élèves (examens et évaluations des apprenants – dans l'idéal, cela devrait faire partie intégrante du SIGE ; au minimum, le SIGE doit pouvoir se connecter en toute transparence aux évaluations)
- Qualification et salaire des enseignants (liens vers les RH et la paie)

- Financement des établissements scolaires (dans de nombreux pays, les écoles reçoivent des financements directs pour certaines dépenses) et budget

Autres ministères fonctionnels :

- Enseignement supérieur (dans le cas où il relève d'un ministère séparé, au lieu d'être la responsabilité d'un niveau du ministère de l'Éducation)
- Travail/emploi (généralement les programmes de formation au niveau de la formation des adultes ou de l'enseignement supérieur)
- Affaires intérieures ou affaires étrangères (flux d'étudiants internationaux, problèmes de visa, etc.)
- Santé (statistiques sur la santé des élèves)
- Services sociaux (subventions accordées aux familles pour les études et données sur la pauvreté, le cas échéant)
- Financement ou planification (données sur la pauvreté en vue d'un ciblage, par exemple)

Autres organismes (n'entrant probablement pas dans le cadre du SIGE lui-même, mais autorisant des liens, si possible)

- Organisations non gouvernementales (ONG)
- Données démographiques (ex. Institut national de statistique)
- Municipalités (lorsqu'elles sont pertinentes pour l'éducation)

Tous ces éléments devraient être disponibles par niveau (à partir du préscolaire et du développement de la petite enfance/DPE) et par circonscription administrative (école, district, province et national, selon le cas, pour le pays en question).

D'autres ensembles de données recueillis par d'autres composantes du secteur de l'éducation et d'autres ministères fonctionnels ont tout autant

d'importance dans la prise de décisions fondées sur des données probantes. Bien qu'il soit impossible d'en dresser une liste complète – puisque cela dépend des priorités de chaque pays – ils peuvent généralement inclure des éléments tels que les variables liées à l'ODD 4 que le SIGE d'un pays peut vouloir produire ; d'ailleurs, l'exemple de l'ODD 4 peut fournir une liste par défaut. Une autre référence utile est la liste des cibles définies dans le cadre de la Stratégie continentale de l'éducation pour l'Afrique (CESA) et harmonisées avec l'ODD 4 (UA et ADEA, 2018). Certaines questions intéressantes soulevées par ces cadres sont traitées dans la section suivante. En outre, dans l'*Annexe B*, nous fournissons la liste des indicateurs que la plupart des pays sont, ou devraient être, engagés à fournir dans le cadre de l'ODD 4, car il s'agit à la fois d'un engagement international pris par les gouvernements de tous les pays de déclarer leurs données, mais aussi d'une liste mûrement réfléchie et logique qui devrait convenir en grande partie à de nombreux pays. Il est impératif de souligner qu'il ne faut pas interpréter le présent Guide comme un outil d'orientation sur la façon de répondre aux exigences des donateurs et des organisations internationales. Voir l'Encadré 3 sur les rapports sur les ODD où nous réaffirmons la primauté des besoins des pays.

5.1.2 Comment les indicateurs de l'ODD 4 ont modifié le périmètre des SIGE

Le principal défi posé par les indicateurs de l'ODD 4 a trait au fait que bon nombre d'entre eux (en réalité à peu près tous) nécessitent le recours à des bases de données qui débordent de l'architecture de données du SIGE. C'était déjà vrai dans une large mesure pendant la période précédente, lorsque les Objectifs du Millénaire pour le développement (OMD) étaient au cœur des préoccupations. Mais le contraste entre l'espoir et la réalité est un peu plus frappant aujourd'hui. L'importance de créer des entrepôts

pour conserver des ensembles de données qui dépassent l'architecture conventionnelle des données SIGE devient de plus en plus évidente, et pour pouvoir répondre à ces demandes croissantes, le SIGE doit fonctionner en tant que système intégré et interconnecté avec des ensembles de données venant de sources diverses. Pour être réactif, il doit travailler sur les données entrantes, les processus et les extrants des différents niveaux et modalités de l'offre éducative, avec un degré élevé de granularité, c'est-à-dire en partant des données agrégées au niveau national jusqu'au détail d'une école, d'une salle de classe ou même d'un élève individuel (tout en tenant compte des préoccupations liées au respect de la vie privée, suivant les normes transversales et éducatives nationales et internationales⁸). Par ailleurs, les ODD exigent une ventilation en fonction de facteurs qui ne sont pas toujours présents dans l'architecture SIGE, comme la situation de richesse. Dans certains cas, il peut être nécessaire par exemple de relier les données de l'architecture SIGE à une carte détaillée de la pauvreté dans le pays. Il est également assez clair que si des rapports simples présentent déjà des difficultés, les véritables défis en matière de données ont trait à leur *exploitation* pour améliorer les résultats dans le domaine des ODD.

L'*Annexe B* donne la liste des indicateurs que la plupart des pays se sont, ou devraient s'être, engagés à fournir dans le cadre de l'ODD 4, car il s'agit d'un engagement international pris par le gouvernement de chaque pays, et elle indique comment le SIGE pourrait les traiter. Il est toutefois impératif de souligner que ce guide ne doit pas être interprété comme un outil de réponse aux demandes

8 Les normes de respect de la vie privée et de sécurité constituent un minimum pour les ministères des TIC (ou leur équivalent) des pays. Le ministère de l'Éducation peut souhaiter aller encore plus loin, en vérifiant l'application de ces normes à des données dont seul dispose généralement le secteur de l'éducation, comme le développement cognitif de chaque enfant.

des donateurs et des organisations internationales. Au contraire, les *besoins des pays* sont primordiaux.

Voici trois exemples de la façon dont l'ODD 4 pourrait influencer la collecte et la production de rapports de données :

Résultats d'apprentissage : la nécessité et la possibilité de suivre les acquis de l'apprentissage (ODD 4.1.1, 4.2.1, 4.2.3, 4.4.2, etc.). Traditionnellement, les unités SIGE ne recueillent pas de données sur l'apprentissage, mais ces données pourraient être entreposées à proximité du SIGE pour voir, par exemple, le niveau de corrélation entre les données et les mesures de l'apprentissage que le SIGE peut habituellement produire, comme les profils de progression d'année en année au sein de l'école, pour disposer d'une alerte précoce en cas de problème dans les résultats d'apprentissage.

Pour ce faire, il faudrait établir des relations de meilleure qualité et plus automatiques entre les ensembles de données.

Population (presque tous les sous-indicateurs d'accès de l'ODD, tels que 4.1.3, 4.1.5, 4.2.2, 4.2.4). Il est nécessaire de mieux comprendre les données démographiques et leur interaction avec les données sur les effectifs scolarisés. Par exemple, des comparaisons pourraient attirer l'attention sur des inefficacités de flux et des blocages de taux à des niveaux très faibles (ex. 65 %). Ceci pourrait à son tour être un indicateur indirect d'une qualité d'apprentissage faible, et mettrait en évidence le fait qu'il peut y avoir beaucoup plus de redoublement que de décrochage dans le cycle du primaire ou pourrait servir à la triangulation des faiblesses dans les données de redoublement. Pour ce faire, le personnel SIGE devrait pouvoir accéder

Encadré 2. Données à collecter en urgence en période de COVID-19

En juin 2020, les écoles ont commencé à rouvrir après les fermetures décidées à l'échelon national en raison de la pandémie engendrée par le nouveau coronavirus SRAS-CoV-2, certaines d'entre elles restant toutefois fermées. Au vu de la situation, l'ISU a publié des directives sur la nécessité de collecter les données essentielles dans ces circonstances⁹, portant au minimum sur :

- La participation des élèves à toutes les plates-formes électroniques d'enseignement (ou à l'apprentissage à distance en général), ventilée selon les caractéristiques individuelles des élèves comme le genre et la pauvreté.
- La participation des enseignants à toutes les plates-formes électroniques d'enseignement (ou à l'enseignement à distance en général), ventilée selon les caractéristiques individuelles des enseignants comme le genre et le statut contractuel.
- Le recours à des tests rapides et courts pour mesurer fréquemment l'apprentissage des élèves.

Les enseignements dégagés de cette situation pourraient s'appliquer à des questions plus vastes concernant d'autres situations d'urgence, les réfugiés et les personnes déplacées à l'intérieur de leur pays.

9 Voir, par exemple, le bulletin d'information « Collecter les données essentielles de l'éducation durant la crise du COVID-19 : une nécessité » à l'adresse <http://uis.unesco.org/sites/default/files/documents/fs58-need-for-essential-education-data-fr.pdf>

à des données démographiques aussi détaillées que possible, sans oublier toutefois que plus l'unité d'analyse est petite, moins les données démographiques deviennent fiables lorsqu'elles sont comparées aux données sur les effectifs scolarisés. Ici, la possibilité de se relier à d'autres ensembles de données tels qu'une carte détaillée de la pauvreté est essentielle car il ne sera pas possible d'estimer dans la plupart des pays l'accès par niveau de pauvreté, par exemple, sans avoir accès à un tel ensemble de données.

Langue et questions relatives à la diversité linguistique (ODD 4.5.2). Il peut être nécessaire de s'intéresser aux écoles présentant une grande diversité linguistique, en particulier dans le cycle du primaire et dans les contextes où aucun des enfants ne parle la langue d'instruction à la maison. Toutefois le SIGE ne recueille traditionnellement pas ce type de données. Dans ce cas, le problème n'est pas tant de pouvoir relier ces données à d'autres variables (comme les résultats d'apprentissage), aussi intéressant que cela puisse être, mais tout simplement de produire les données. Les différentes possibilités incluent de tester un petit changement à l'instrument qui permettrait de recueillir ces données, en tant que contribution à l'amélioration du SIGE, ou, dans une approximation moins précise, mais aussi beaucoup moins coûteuse, de se relier aux ensembles de données provenant d'un recensement contenant des informations linguistiques.

Sauf s'ils peuvent être reliés entre eux, les ensembles de données du type de ceux qui sont énumérés ci-dessus resteront des îlots de données fragmentés et isolés. Comme nous l'avons déjà mentionné, l'interopérabilité des données est un principe important de la production de données de qualité et un élément clé de l'architecture

globale d'un SIGE. En fait, l'interopérabilité est le seul moyen de permettre aux utilisateurs finaux - y compris aux décideurs politiques - de comprendre la valeur ajoutée significative qu'offrent les données SIGE dans le traitement des défis clés. La prise de conscience de cette valeur ajoutée potentielle peut à son tour accroître la demande et le soutien vis-à-vis de l'unité SIGE pour la planification, le financement des établissements scolaires et d'autres unités utilisatrices qui s'adressent directement aux niveaux supérieurs du ministère. Le SIGE doit avoir comme fonction première bien plus que la simple « gestion » des données. Lorsqu'elles sont centralisées au niveau national, les données sont souvent difficiles à exploiter pour une vraie gestion au jour le jour et un SIGE offrant une capacité d'interopérabilité des données et ayant des « *crochets* » avec les unités qui ont le pouvoir de créer de la valeur ajoutée est plus susceptible de générer une plus grande demande pour ses services et donc un appui budgétaire du ministère.

5.1.3 L'importance d'une politique de données ou SIGE

Il devrait exister un cadre juridique ou politique imposant le SIGE en tant que dépositaire des données sur l'éducation, ce qui lui permettrait de partager et de coordonner les données de façon adéquate avec d'autres unités de collecte au sein du secteur de l'éducation ainsi qu'avec d'autres processus de collecte extérieurs au SIGE et au secteur de l'éducation. L'intégration des données de tous ces ensembles doit constituer l'une des principales stratégies de développement du SIGE. À propos du « Guide de l'acheteur », par exemple, il convient de préciser que sans un cadre politique directeur, il est difficile de savoir ce que l'on doit « acheter » (ou construire ou recevoir en don). Les aspects spécifiques à envisager pour la politique incluent les deux éléments suivants,

qui ne représentent en aucun cas la plus haute priorité, mais qui sont pertinents pour déterminer les fonctionnalités d'utilisation :

- Au niveau national, le SIGE devrait prendre l'initiative d'élaborer et de mettre en œuvre une politique de partage et d'intégration des données dans le pays et au-delà de ses frontières. Cela peut porter sur des thèmes comme le respect de la vie privée et la sécurité, les obligations, l'éthique, les droits et les responsabilités concernant les données.¹⁰ Une telle politique pourrait spécifier les besoins en termes de budget et de personnel.
- Afin de faciliter l'intégration, le partage et l'échange des données, le bureau national devrait mettre en place officiellement un groupe de travail ou une équipe de travail chargé de nettoyer, lier et maintenir les normes concernant les données. Un tel groupe pourrait être composé de responsables thématiques choisis au niveau national et de responsables SIGE issus des niveaux infranationaux pertinents, pour élaborer des protocoles et des accords pour la mise en relation et le maintien d'identifiants uniques communs pour les ensembles de données en rapport avec l'éducation. Les membres du groupe devraient également s'assurer de la réalisation de ces opérations en supervisant le processus, et se tenir mutuellement responsables de leurs responsabilités et engagements respectifs. Le travail effectif serait généralement réalisé par des techniciens subalternes.

10 Pour un ensemble type de préoccupations dans le système d'un pays à revenu élevé, voir https://pourparlerprofession.oeeo.ca/mars_2008/privacy.asp. Ce site contient aussi des références aux politiques nationales en matière de respect de la vie privée et de sécurité au Royaume-Uni : <https://www.gov.uk/government/publications/esfa-privacy-notice>. Pour une discussion dans le contexte d'une économie émergente (Afrique du Sud), voir <https://www.isasa.org/the-protection-of-personal-information-act-and-south-african-schools/>.

5.2 Comment se procurer un SIGE

Se procurer une technologie et des services informatiques signifie acquérir la technologie et les outils techniques qui correspondent aux besoins de l'unité SIGE. Il s'agit du processus d'acquisition de l'infrastructure (l'actif) du système d'information qui pilotera les opérations du SIGE. Noter, ainsi que nous l'avons expliqué à la Section 2, qu'un ministère ne va pratiquement jamais acheter un système complet flambant neuf sur le marché. Au contraire, il peut acquérir certains éléments, ou recevoir en don des éléments fournis par les partenaires de développement. Même dans ce dernier cas, il est essentiel de savoir précisément ce que recherche le ministère, pour quelle raison et à quelles conditions, à la fois à des fins d'efficacité et de budget (car les dons ont eux aussi des coûts d'opportunité).

La vision présidant à l'acquisition doit être en ligne avec le reste du Guide de l'acheteur ainsi qu'avec les recommandations ou normes contenues dans le Guide de l'utilisateur. On trouvera ci-dessous, après un préambule plus général, les pistes de départ du cahier des charges.

L'acquisition d'un système d'information est à envisager dans une perspective de long terme, de cycle de vie. Les systèmes modernes, conçus sous la forme d'applications web progressives et destinés à des terminaux web et mobiles, sont exécutables sur toute une gamme de matériels (qui sont les équipements informatiques servant à effectuer les activités d'entrée, de traitement et de sortie – voir plus de détails dans la Section 5.5). Les applications fonctionnent de façon optimale avec les navigateurs les plus récents et des systèmes d'exploitation mobiles tenus à jour, sans nécessiter forcément un matériel haut de gamme, puisque la plupart des calculs lourds s'effectue généralement sur un serveur. Il est sans doute plus important

d'investir dans une connectivité de qualité que dans un matériel ultra-moderne.

Un SIGE national est un système complexe, accueillant potentiellement des dizaines voire des centaines de milliers d'utilisateurs sur le long terme, si l'on tient compte non seulement des administrateurs et des planificateurs, mais aussi des enseignants, et même des communautés et des parents qui pourraient avoir besoin d'accéder à une partie des informations. Même si la technologie progresse rapidement, les attentes des différentes parties prenantes vis-à-vis du système, en termes de complexité, ne manqueront pas de croître aussi. Il est donc essentiel de prendre en compte la structure du support au sein de l'organisation. C'est dans une certaine mesure un aspect technique (gestion des serveurs et des terminaux), mais aussi et surtout organisationnel, en ce sens que l'information est l'épine dorsale du secteur de l'éducation. Pour réaliser le plein potentiel des investissements, il faut non seulement envisager le coût élevé de l'achat de milliers de terminaux et le prix de la connectivité Internet, mais aussi la formation et le renforcement des compétences au sein du secteur pour une exploitation des données qui débouche sur une action véritable. Par conséquent, lors de l'acquisition d'une solution logicielle, ce qu'il est également important de prendre en compte, c'est la communauté des praticiens du système à long terme, c'est-à-dire des personnes qui se consacrent à l'amélioration de la collecte et de l'exploitation des données éducatives pour atteindre une série d'objectifs. L'acquisition d'un système national ne représente pas simplement, par exemple, l'offre d'un contrat de 2 000 ordinateurs portables afin que toutes les écoles disposent d'un accès (bien qu'un tel achat soit probablement nécessaire), mais plutôt la définition d'un plan portant à la fois sur la formation aux fonctionnalités de base et sur l'expansion future du système, au fur et à mesure que l'organisation découvrira ce qui est

possible et ce qui offre les plus grands bénéfices en alliant sa propre expérience à celle des autres membres de la communauté internationale SIGE. La vision de l'acquisition doit donc être holistique : le processus général doit être considéré comme étant à long terme, tandis que les contrats de matériels spécifiques peuvent être traités selon les mêmes procédures que les autres marchés publics, par des appels d'offres publics. Il peut être également logique pour le ministère de l'Éducation de coordonner son action avec celle des ministères d'autres secteurs comme la communication, la santé et l'agriculture. De même, le support du système peut souvent bénéficier de la gestion d'un centre de données intra-gouvernemental ou d'un nuage national, qui impliquera le secteur privé comme les entreprises de télécommunications ou d'autres prestataires de services et fournisseurs, car la gestion appropriée et sécurisée des serveurs met souvent à rude épreuve les capacités des ministères fonctionnels.

Les facteurs à prendre en considération dans les marchés publics sont les suivants :

- **Mise en œuvre alignée sur le processus de marchés publics des systèmes internes**
 - S'informer sur les directives et parfois obtenir l'approbation du département des marchés publics du ministère, ou des marchés publics à l'échelle de l'État, si nécessaire.
 - Définir le cahier des charges informatique (c'est-à-dire que l'informatique doit pouvoir gérer les demandes contenues dans le présent guide – noter également que certaines options informatiques peuvent ne pas fonctionner sur un système ancien).
 - Instaurer une discussion avec les experts des marchés publics au sein du ministère ou plus généralement au sein du gouvernement, sitôt qu'on se fixe l'objectif d'une passation de marché public

informatique. Cela devrait être basé sur un diagnostic des besoins des utilisateurs ou d'un cahier des charges.

- Gestion des fournisseurs – il convient de s'interroger pour savoir dans quelle mesure il est possible et nécessaire d'impliquer les fournisseurs dans l'analyse des besoins dès le début (ils peuvent être utiles, mais ils peuvent alors acquérir des informations auxquelles d'autres n'ont pas accès ; se conformer aux directives se rapportant aux marchés publics).
- Vérifier la qualité des produits et services devant être fournis, en demandant des références, en s'informant sur les performances passées, tant en termes de processus que de satisfaction finale des autres consommateurs.

■ Aspects différenciés de l'acquisition en fonction de ce qui est acheté ; on peut s'adresser à différents fournisseurs pour différents éléments

- Achat de matériels : le matériel inclut les serveurs, les routeurs, les ordinateurs, les imprimantes, etc.
- Achat de logiciels : les logiciels incluent les logiciels sous licence, les logiciels développés en interne, etc.
- Services : l'infrastructure informatique ne se limite pas aux ordinateurs et aux applications, elle comprend aussi l'infrastructure réseau qui relie les différents ordinateurs, imprimantes et serveurs en interne et sur le web.

■ Plans et coûts continus de support

- Coût total « sur la durée de vie du produit » et hypothèses relatives aux exigences techniques que le ministère devrait être prêt à assumer. Cela devrait également couvrir les besoins et les plans de dotation

en personnel. La formation du personnel pourrait faire partie de la fourniture informatique, à la fois sur le court et le long terme. Noter que si les logiciels *open source* peuvent offrir un faible coût initial, le coût final peut être élevé sur la durée de vie du produit et qu'il peut y avoir d'autres frais indirects.

■ Risques associés à l'acquisition

- Délais d'acquisition – il est important d'être réaliste, c'est-à-dire ne pas tenter de se procurer quelque chose de très coûteux en 10 jours ; s'informer sur les directives générales du ministère ou des marchés publics en général.
- Financement/budget – le budget ne doit pas être irréaliste (insuffisant), car cela pourrait encourager les fournisseurs à faire des promesses irréalistes.
- Le calendrier de réalisation doit être réaliste, sinon les fournisseurs risquent de faire des promesses irréalistes pour s'y conformer ; consulter des experts en TI.
- Il faut anticiper les coûts de la « durée de vie du produit » et les exigences techniques du ministère.

5.3 Un cadre de qualité totale des données

Toutes les normes de fonctionnalité doivent finalement prendre en charge certains aspects de la qualité totale des données. Pour ce faire, le SIGE national doit instaurer *un système de qualité des données* basé sur une approche qualité pour normaliser les données, fournir des définitions et instaurer des conventions de dénomination des données. Ceci est très utile lorsque les données sont transférées d'un niveau à l'autre, par exemple, du niveau national au niveau infranational et jusqu'aux écoles, ou du niveau infranational au niveau national.

Un tel système doit identifier les exigences et critères de qualité, les processus clés du SIGE, leur séquençement et leur interaction, les politiques et enfin les critères et méthodes pour la définition, la détection, la correction et la prévention de la non-conformité.

La qualité des données peut être définie en termes d'*exhaustivité*, de *pertinence*, de *précision*, d'*actualité* et d'*accessibilité*. Ces dimensions de la qualité sont décrites au Tableau 1.

Une autre réflexion utile à propos des principes déterminants pour la qualité (ainsi que d'autres questions comme la facilité d'utilisation, la durabilité, etc.) ressort du document des Nations Unies intitulé « Principes concernant les plates-formes d'établissement de rapports et de diffusion des indicateurs des ODD et orientations pour leur mise en œuvre » (voir la synthèse dans l'Encadré 3).¹¹

11 Pour une explication complète, voir <https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3a-Principles-guidelines-SDG-Monitoring-Reporting-Platforms-E.pdf> (en anglais uniquement)

Tableau 1. Dimensions de la qualité des données

Dimension	Description
Exhaustivité	Les données relatives à tous les éléments de données, pour toutes les entités déclarantes (ex. établissements scolaires, apprenants) ont dû être soumises. Les données sont exhaustives lorsque des valeurs de données sont fournies pour tous les enregistrements, les occurrences ou les entités logiques de l'enseignement préscolaire, primaire et secondaire, y compris les apprenants ayant des besoins éducatifs spéciaux. ¹²
Pertinence	Il ne sert à rien de recueillir des données si celles-ci ne peuvent pas être exploitées de façon significative dans les décisions portant sur l'éducation. Il convient d'éviter les questionnaires trop longs.
Précision	La précision doit être priorisée dans les méthodes de capture et de validation des données, ainsi que dans les processus de vérification des données, quelles qu'en soient éventuellement les implications en termes de temps ou de coût. Les données doivent se situer dans la plage normale des données collectées pour l'élément de données en question et pour une entité spécifique. Rechercher les valeurs aberrantes – la capacité à détecter automatiquement la plupart des valeurs aberrantes lors de la saisie doit faire partie du cahier des charges utilisateur.
Actualité	Les données doivent être actuelles ou à jour. Elles doivent être fournies rapidement et disponibles en temps voulu, sous peine de réduire la crédibilité du système d'information. Les données de toutes les institutions déclarantes doivent être soumises dans les délais impartis.
Accessibilité	Les données doivent être mises à la disposition des utilisateurs, à tous les niveaux du système éducatif. Les utilisateurs doivent savoir quelles sont les données, où les trouver et les récupérer. Les métadonnées et les dictionnaires de données sont importants pour améliorer l'accessibilité des données au sein d'une organisation. Les données doivent être accessibles en temps voulu, selon les modalités requises, par toute personne qui en a besoin, aussi librement que cela est autorisé par les règles plus générales de respect de la vie privée et de la sécurité, ainsi que par la politique du pays en matière de données multisectorielles. (Voir la note de bas de page 11.)

Source : van Wyk 2006, modifié par les auteurs.

12 Notez que ces descriptions sont plutôt axées sur le présent et elles conviennent le mieux aux situations type que l'on rencontre actuellement dans les pays à revenu faible. Une orientation plus futuriste pourrait voir des descripteurs légèrement différents – voir Section 5.6.

Encadré 3. Synthèse des principes des Nations Unies en matière d'établissement de rapports sur les indicateurs des ODD

- Clarté des mécanismes institutionnels et de gestion des données
- Adéquation par rapport aux objectifs
- Plates-formes de diffusion conformes aux principes fondamentaux de la statistique officielle
- Durabilité
- Interopérabilité et normes statistiques
- Appropriation nationale
- Collaboration
- Conception axée sur l'utilisateur
- Accent mis sur la communication des données
- Ventilation des données
- Modularité et extensibilité
- Interfaces normalisées
- Disponibilité de métadonnées
- Données ouvertes
- Données liées

5.4 Fonctionnalité et normes du SIGE – Aspects du Guide de l'acheteur

Le Tableau 2 présente des normes de fonctionnalité, c'est-à-dire les caractéristiques que doit posséder un système ; elles sont établies sur la base de l'expérience des auteurs et des réviseurs en matière de SIGE et de littérature sur les SIGE ; d'autres normes sont des choix plus ouverts, avec leurs propres avantages et inconvénients. La technologie d'un système, ou celle sur laquelle il doit être basé, constitue également une considération essentielle qui est brièvement décrite à la Section 5.5.¹³ Le tableau fait référence à l'architecture de données type d'un SIGE, ainsi que mentionné à la Section 4.

Les trois conventions suivantes sont utilisées dans cette section pour mettre en évidence des informations importantes (c.-à-d. non accessoires) :

- **Note** : des informations plus détaillées et complémentaires explicitant la fonctionnalité à prendre en considération.
- **Conseil** : un conseil utile, tel que la façon d'effectuer plus efficacement une tâche particulière.
- **Attention** : un avertissement à propos d'un choix ou d'une décision à laquelle l'« acheteur » doit réfléchir.

¹³ Pour une analyse des nouveaux défis créés par les indicateurs de l'ODD 4, voir la Section 5.1.2.

Tableau 2. Normes de fonctionnalité du SIGE, par niveau de la collecte de données

NIVEAU DE LA COLLECTE DE DONNÉES (DONNÉES AGRÉGÉES)		
Fonction	Description	Points essentiels à envisager
NOTE :	<p>Données agrégées vs. données individuelles : les données agrégées ou sommaires se rapportent aux données recueillies au niveau des établissements scolaires (méthode du recensement scolaire). Leur collecte fournit des informations sur des questions d'ordre institutionnel, telles que les effectifs scolarisés par année d'étude, âge ou sexe. Les données (individuelles) de l'unité d'enregistrement se rapportent aux données recueillies sur chaque apprenant par le biais du système d'administration et de gestion de l'école ; elles sont décrites après cette section.</p>	
Répertoire	Le système dispose d'une fonctionnalité pour gérer un répertoire d'établissements .	
	Le système dispose d'une fonctionnalité pour ajouter et supprimer des établissements (c'est-à-dire pour tenir le répertoire à jour).	
	Le système a la capacité d'importer des informations institutionnelles à partir d'autres systèmes (ex. Excel, csv).	
Identifiant unique	Attribution d'un identifiant unique aux établissements , selon des algorithmes spécifiques.	<p>Conseil : le système de fichier maître des établissements est un processus clé qui attribue un identifiant unique à chaque établissement et doit être supervisé au niveau national, ce qui est plus ou moins facile selon que le système est réellement fédéral et non pas simplement décentralisé.</p> <p>Conseil : afin de faciliter le lien avec les bases de données qui ne « descendent » pas au niveau de l'école (ex. pauvreté), l'identifiant unique peut contenir les codes géographiques des organes de gouvernance situés à un niveau supérieur. Ou bien le fichier contiendra des champs désignant l'unité organisationnelle dont relève l'école (ex. district, province).</p> <p>Conseil : dans l'idéal, un système de qualité doit offrir la possibilité de générer (éventuellement avec une assistance manuelle) de nouveaux codes à partir d'une base de données de consultation et d'un ensemble de règles. Par exemple, il peut affecter à une école un code d'identification basé sur le district et la province dans lequel ou laquelle elle se trouve.</p> <p>Attention : éviter d'utiliser un numéro séquentiel ; il est préférable de recourir à un algorithme spécifique qui inclut des codes représentant les niveaux infranationaux.</p>
Conception du questionnaire	Le système permet aux pays de concevoir des formulaires de questionnaire personnalisés à utiliser au niveau des établissements (méthode du recensement scolaire).	Le système est personnalisable et flexible pour autoriser l'ajout de champs en fonction du cas. Par exemple, il devrait permettre l'ajout d'un nouveau champ pour rendre compte d'une nouvelle catégorie d'enseignants qui n'avait pas été envisagée.
Transfert des données de base	Le système permet aux utilisateurs de transférer les données existantes des années antérieures dans une nouvelle base de données couvrant tous les établissements.	<p>Note : le système a la capacité de stocker des données sur plusieurs années, permettant ainsi le transfert des informations en toute transparence d'une année sur l'autre.</p> <p>Conseil : cette capacité est essentielle pour générer des tendances chronologiques et d'autres formes de données à valeur ajoutée.</p> <p>Conseil : le cas échéant, le système devrait offrir la possibilité de pré-imprimer (ou de pré-remplir si cela se fait en ligne) les données de base dans le questionnaire physique avant de le diffuser aux établissements.</p> <p>Attention : la pré-impression (contrairement au pré-remplissage en ligne, lorsque cela est possible) accroît considérablement les besoins logistiques car il faut envoyer exactement le bon questionnaire à la bonne école.</p>
Codes-barres	Le système dispose d'une fonction de codage à barres pour gérer la diffusion et la collecte des questionnaires.	Note : cela ne s'applique que lorsque les questionnaires sont imprimés sur papier avant d'être diffusés dans les écoles.

NIVEAU DE LA COLLECTE DE DONNÉES (DONNÉES AGRÉGÉES)

Fonction	Description	Points essentiels à envisager
Saisie des données	Le système a la capacité de mettre en place la saisie des données à n'importe quel niveau (national, infranational et institutionnel).	Attention : <i>centralisation vs. décentralisation</i> – en pratique, la décentralisation des systèmes implique le transfert de l'autorité du bureau national vers les bureaux infranationaux. Bien qu'il faille reconnaître que les systèmes d'information décentralisés présentent d'importants avantages, il est aussi nécessaire d'en connaître les risques et les inconvénients. Étant donné que le bureau national est plus éloigné du détail des décisions prises aux niveaux infranationaux, dont il n'a généralement pas connaissance, il doit veiller à ne pas décentraliser ses responsabilités et ainsi à ne pas perdre le contrôle des normes et de la supervision des processus de collecte de données.
Saisie des données : technologie	La saisie des données s'effectue localement sur clavier d'ordinateur dans le cadre d'un processus manuel.	Note : cette fonctionnalité existe sur un ordinateur local, sans connectivité.
	La saisie des données s'effectue via un navigateur via des protocoles Internet. Cela peut nécessiter des processus manuels dans le navigateur ou des processus automatisés avec appels d'une interface de programmation applicative (API).	Note : cette solution est particulièrement avantageuse dans un SIGE décentralisé. Attention : veiller à ce que la bande passante et la connectivité soient adaptées aux besoins du système. Attention : pour les régions d'un pays où la bande passante n'est pas entièrement fiable et à haut débit, veiller à ce que le système puisse autoriser la saisie asynchrone des données sur n'importe quel terminal acceptable tel que des tablettes. Note : une API est un logiciel capable de faire des appels de données vers d'autres bases de données programmées par un logiciel différent, par exemple l'application Contacts d'un téléphone cellulaire, capable d'extraire une liste de directeurs d'école depuis le SIGE.
Le système dispose de la fonctionnalité permettant à l'organisation de développer et de livrer du contenu aux terminaux mobiles ; il tire parti des capacités natives des terminaux mobiles (ex. une application de contacts et sa capacité à utiliser le réseau de données cellulaires ou le Wifi).		
Saisie des données : assurance qualité	Le système dispose de règles d'assurance qualité intégrées telles qu'une authentification et des validations.	Note : le système permet à l'utilisateur de définir différentes règles de validation pour identifier les violations. Au point de saisie des données, une vérification doit être effectuée pour contrôler si les données se situent dans les niveaux acceptables de valeur minimale et maximale d'un élément de données particulier. Conseil : le système doit calculer automatiquement certains ratios, pas nécessairement pour les inclure dans le formulaire, mais pour les utiliser pour dépister les erreurs, car les plages d'erreur des ratios (ex. taux d'encadrement) sont plus faciles à configurer que des plages de valeurs absolues. Attention : sachez que les règles de validation et de vérification intégrées ne doivent pas ralentir le processus de saisie des données. En outre, les règles de validation sont là pour alerter l'utilisateur, mais elles ne doivent pas l'empêcher de poursuivre le processus de saisie des données.

NIVEAU DE LA COLLECTE DE DONNÉES (NIVEAU DE L'UNITÉ)

Fonction	Description	Points essentiels à considérer
<p>Remarque : les données d'enregistrement des unités Apprenant renvoient aux données recueillies sur chaque apprenant par le biais du système d'administration et de gestion de l'école. (Cela peut aussi s'appliquer aux enseignants, nous prenons ici simplement l'exemple des apprenants).</p> <p>Attention : faut-il développer un système d'enregistrement d'unité et stocker au niveau national les données d'enregistrement des unités Apprenant ? Cette question soulève des défis spécifiques en matière de <i>protection de la vie privée</i>, de <i>sécurité</i> (voir ci-dessous), ainsi que des <i>défis techniques</i> et de <i>capacité</i>. Par exemple, selon le nombre d'apprenants dans un pays, on arrivera facilement à une table de base de données comportant des millions d'enregistrements Apprenant. Il devient quasiment impossible de traiter des ensembles de données aussi volumineux par les techniques de stockage et les outils d'écriture de requête existants au sein des gouvernements. Le système national ne devra donc pas être une réplique du système de registre des unités opérationnelles qui existe au niveau institutionnel et infranational. Au niveau national, les données d'enregistrement d'unité Apprenant devront être uniquement les données enregistrées dans le système national d'identification des apprenants à des fins d'immatriculation et de suivi des apprenants.</p>		
Répertoire	Le système dispose d'une fonctionnalité pour gérer un répertoire d'apprenants individuels .	
	Le système dispose d'une fonctionnalité pour ajouter et supprimer des apprenants individuels .	Attention : lorsque l'on supprime un apprenant, il faut veiller à ne pas supprimer les données historiques. Il peut être préférable de <i>clôturer</i> l'inscription d'un apprenant dans une école donnée plutôt que de supprimer l'apprenant, afin de conserver intactes les données historiques.
	Le système a la capacité d'importer des informations sur les apprenants à partir d'autres systèmes (ex. Excel, csv).	
Identifiant unique	Attribuer un identifiant unique à un individu , selon les algorithmes prescrits.	<p>Note : un code d'identification unique doit être attribué à chaque apprenant. Il est important que cet identifiant reste cohérent et précis au fil du temps. Un identifiant unique est un numéro unique, non dupliqué, qui est attribué à un apprenant pour la totalité de son parcours d'études, qu'il ou elle change d'école ou non.</p> <p>Conseil : le bureau national doit attribuer à chaque apprenant un identifiant national unique (numéro matricule). Le bureau national élabore des procédures pour s'assurer que deux identifiants du même type ne soient pas attribués au même apprenant.</p> <p>Note : le système utilise un nombre spécifique d'éléments de données relatifs aux caractéristiques d'un apprenant. Ces éléments de données sont utilisés par un algorithme qui attribue un identifiant unique à l'apprenant.</p>
Conception du questionnaire	Le système dispose d'une fonctionnalité pour configurer les écrans en fonction du questionnaire.	
	Le système dispose d'une capacité pour ajouter des champs dans un questionnaire dispensé au niveau des unités. Il est personnalisable et flexible pour permettre d'ajouter des champs en fonction du cas.	
Transfert des données de base	Le système dispose d'une capacité permettant aux utilisateurs de transférer les données existantes de l'année précédente vers une nouvelle base de données couvrant tous les apprenants (report).	Conseil : <i>couverture longitudinale des données</i> – l'identifiant unique de l'apprenant permet de suivre sa progression dans le système grâce à l'identifiant configuré en données longitudinales. (Les données longitudinales sont des données recueillies sur le même apprenant d'une année à l'autre)
Saisie des données	Le système dispose d'une capacité pour la mise en place de la saisie des données à n'importe quel niveau (national, infranational et institutionnel).	

NIVEAU DE LA COLLECTE DE DONNÉES (NIVEAU DE L'UNITÉ)

Fonction	Description	Points essentiels à envisager
Saisie des données : technologie	La saisie des données s'effectue localement sur clavier d'ordinateur dans le cadre d'un processus manuel.	Note : cette fonctionnalité doit exister sur un ordinateur local, sans connectivité.
	La saisie des données s'effectue via un navigateur via des protocoles Internet ou elle est automatisée via une API.	Note : cette solution est particulièrement avantageuse dans un SIGE décentralisé. Attention : veiller à ce que la bande passante et la connectivité soient adaptées aux besoins du système.
Le système dispose d'une fonctionnalité permettant à l'organisation de développer et de fournir du contenu aux terminaux mobiles et d'exploiter les capacités natives des terminaux mobiles.		
Saisie des données : assurance qualité	Le système dispose de règles d'assurance qualité intégrées , telles que l'authentification et des validations.	Note : le système permet à l'utilisateur de définir différentes règles de validation pour identifier les violations. Au point de saisie des données, une vérification doit être effectuée pour contrôler si les données se situent dans les niveaux acceptables de valeur minimale et maximale d'un élément de données particulier. Comme nous l'avons mentionné ailleurs dans le présent guide, il peut être plus facile d'utiliser des ratios contextuels (ex. un taux d'encadrement qui alertera si les données des effectifs scolarisés ne sont pas compatibles avec les données concernant les enseignants). Attention : veiller à ce que les règles de validation et de vérification intégrées ne ralentissent pas le processus de saisie des données. En outre, les règles de validation doivent alerter l'utilisateur, mais sans l'empêcher de poursuivre le processus de saisie des données.

NIVEAU DE LA COLLECTE DE DONNÉES (GÉNÉRAL) (S'applique aux données agrégées ou aux données de niveau d'unité)

Fonction	Description	Points essentiels à considérer
Choix du type de système	<p>Système de traitement transactionnel (STT) : un système de traitement transactionnel (STT) est un système de capture et de traitement d'informations détaillées (individuelles) nécessaires pour actualiser les données sur les opérations fondamentales d'une organisation. Système utilisé pour la gestion quotidienne des écoles.</p>	<p>Note :</p> <p><i>Avantages :</i></p> <ul style="list-style-type: none"> • Disponibilité des données en temps réel • La collecte de données n'entraîne pas de charge de travail supplémentaire car elle fait partie du fonctionnement quotidien de l'école • Amélioration de la qualité des données <p><i>Inconvénients :</i></p> <ul style="list-style-type: none"> • Préoccupations en matière de sécurité • Coûteux à mettre en place
	<p>Système de suivi de l'enregistrement des apprenants : système de stockage des données d'enregistrement d'unité Apprenant et d'unité Enseignant dans une base de données centrale nationale ayant pour fonction de suivre le mouvement des apprenants individuels d'une école à l'école et d'une année à l'autre.</p>	<p>Note : le système n'est pas un système d'administration et de gestion de l'école (opérationnel) au vrai sens du terme. Il est principalement utilisé comme système d'enregistrement et de suivi des apprenants. Le système est développé et géré au niveau central par le bureau national. Sa principale fonctionnalité est d'enregistrer les apprenants, de suivre et de surveiller les apprenants individuels dans le pays, par le biais de leur enregistrement, des transferts d'apprenants entre les écoles, etc.</p> <p><i>Avantages :</i></p> <ul style="list-style-type: none"> • Il est plus facile de suivre le mouvement des apprenants d'une école à l'autre et d'une année à l'autre, tout au long de leur parcours d'études • Cela permet de créer un ensemble de données longitudinales et de déterminer exactement le nombre d'apprenants d'une cohorte donnée qui ont décroché du système, le nombre qui a progressé dans le système sans redoubler et le nombre qui est encore dans le système après un ou plusieurs redoublements <p><i>Inconvénients :</i></p> <ul style="list-style-type: none"> • Préoccupations en termes de sécurité, en raison des données personnelles des individus • Le volume des données peut poser un problème pour les méthodes de stockage traditionnelles. Un autre défi concerne la vitesse de traitement requise pour traiter et gérer ces grands volumes de données.
	<p>Système de collecte de données agrégées : le système recueille des données sommaires (agrégées) auprès des établissements à des fins de rapport et de statistique.</p> <p>La collecte des données s'effectue au moyen d'un questionnaire de recensement annuel (méthode du recensement scolaire) envoyé aux écoles pour recueillir les renseignements sur l'éducation dont le gouvernement a besoin pour suivre le système d'éducation.</p>	<p>Note : la collecte de données agrégées ou sommaires fournit des informations sur des questions d'ordre institutionnel, telles que les effectifs scolarisés par année d'étude et par âge, et les effectifs par sexe.</p> <p>Note :</p> <p><i>Avantages :</i></p> <ul style="list-style-type: none"> • Aucun souci quant au respect de la vie privée parce qu'il ne contient aucune identification personnelle • Plus facile d'identifier les schémas et les tendances des données. <p><i>Inconvénients :</i></p> <ul style="list-style-type: none"> • Les données recueillies initialement au niveau institutionnel ne peuvent pas être exploitées pour les niveaux d'agrégation inférieurs, par exemple, pour le suivi d'élèves individuels au fil du temps et entre les établissements • La qualité des données est toujours problématique

NIVEAU DE LA COLLECTE DE DONNÉES (GÉNÉRAL) (S'applique aux données agrégées ou aux données de niveau d'unité)

Fonction	Description	Points essentiels à envisager
Choix du logiciel applicatif	Le système est réalisé sur mesure et auto-développé à des fins spécifiques, il appartient à l'organisation.	Note : auparavant, la seule option qui s'offrait aux ministères et aux gouvernements était de développer une solution SIGE complète, soit en interne soit réalisée sur mesure par un fournisseur de logiciels.
	Il s'agit d'un système standard (produit commercial).	Note : le système est un progiciel d'information de gestion, disponible sur le marché et capable d'effectuer bon nombre des tâches requises par un SIGE. De tels systèmes ne nécessitent souvent que des modifications mineures à apporter au produit commercialisé. Note : il existe une distinction entre l'offre commerciale prête à l'emploi (COTS) et l'offre destinée aux administrations publiques, également prête à l'emploi (GOTS). Cette dernière doit souvent répondre à des exigences ou des attentes différentes en termes d'acquisition et elle aura certainement un cycle de vie différent en termes de mises à niveau.
Systèmes d'exploitation	Le système d'exploitation sur lequel s'exécute l'architecture SIGE sert de plate-forme à un environnement prenant en charge l'interaction utilisateur via des interfaces (ex. écran, souris, clavier, imprimante). Exemples : Windows, macOS, Linux, Android, iOS et Chrome OS.	Note : <i>facilité d'utilisation, attrait visuel et intégration des flux de travail</i> : se rapportent à la facilité d'administration et de déploiement de la plate-forme, de la création et de la consommation de contenu, de l'interaction avec le contenu, ainsi qu'au niveau d'attraction visuelle du produit.
Formation	Le système fournit un soutien à la formation , c'est-à-dire qu'il intègre un minimum de fonctions d'« aide » de qualité, qu'il est relié à des conseils de formation ou, mieux encore, à des manuels.	Note : les managers devraient mettre en place des processus afin que le ministère de l'Éducation dispose de personnels SIGE correctement déployés, possédant les compétences nécessaires pour atteindre les objectifs du SIGE. Note : le système doit comporter un ensemble complet de menus de recherche d'aide en ligne, FAQ, etc., ainsi que d'un guide autonome bien indexé. Conseil : veiller à ce que le personnel SIGE possède des compétences minimales en bases de données et soit capable d'écrire des requêtes simples pour extraire des données d'un système de base de données. Les requêtes peuvent être automatisées, mais s'assurer que les membres (au moins certains) du personnel SIGE sont capables d'écrire des requêtes (ou même du code) constitue un bon moyen de définir un niveau de compétences à atteindre et d'assurer une certaine flexibilité au lieu de dépendre intégralement de l'automatisation.

NIVEAU DE LA GESTION ET DU STOCKAGE DES DONNÉES

Fonction	Description	Points essentiels à considérer
Choix du type de base de données	<p>Système de fichiers à deux dimensions (plats) : le système fonctionne comme un fichier plat (pas de relation entre les enregistrements).</p>	<p>Note : le système repose sur une structure simple (une seule table de données) et il ne peut pas contenir plusieurs tables, comparé à une base de données structurée avec un modèle relationnel.</p> <p>Attention : éviter ce type de structure de base de données. Plus on ajoute de données et plus ce système de fichiers devient inefficace, car il n'y a pas de relations entre les différents ensembles de données. Bien que les données soient plus faciles à comprendre et à visualiser dans une structure de fichier à deux dimensions, l'accès aux données et la recherche dans les données peuvent être extrêmement lents et pesants.</p>
	<p>Système mono-utilisateur : la base de données du système installé est destinée à un seul utilisateur (application autonome). Les applications autonomes sont des logiciels conventionnels installés sur l'ordinateur de chaque utilisateur.</p>	
	<p>Systèmes multi-utilisateurs : le système de gestion de base de données (SGBD) est destiné à partager des informations via un réseau dans l'ensemble de l'organisation. De nombreux utilisateurs peuvent accéder simultanément à la base de données. Naturellement, c'est l'option privilégiée dans presque tous les cas. Certains systèmes anciens peuvent ne pas posséder cette fonctionnalité.</p>	
Gestion des bases de données	Le système de gestion de base de données (SGBD) a une structure relationnelle .	<p>Note : un système de gestion de base de données relationnelle (SGBDR) est un système de gestion de base de données qui permet la jointure temporaire ou permanente de tables de données par le biais d'un champ commun (clé primaire et clé étrangère). Chaque ligne, chaque enregistrement ou chaque instance d'une base de données est associée à un ensemble fixe d'attributs ou de champs. Chaque table possède une clé primaire qui identifie chaque enregistrement de façon unique. La table peut également contenir une clé étrangère, qui est identique à la clé primaire d'une table externe. Une jointure relationnelle est réalisée en faisant correspondre les valeurs de la clé étrangère aux valeurs correspondantes de la clé primaire de la table externe. Un exemple serait une « table » d'écoles telle que (i) pour chaque école de cette table il existe un lien vers toutes les salles de classe de cette école et (ii) ces salles de classe elles-mêmes sont des tables de données caractérisant la salle de classe, sans avoir à faire apparaître dans la table des écoles, les salles de classe imbriquées dans l'école.</p>
Choix du type de logiciel	Le système est développé avec un logiciel open source .	<p>Note : le logiciel « open source » est distribué gratuitement (selon le cas) avec son code source accessible et disponible pour les utilisateurs.</p> <p>Avantage : très faible coût et une flexibilité bien supérieure de l'infrastructure.</p> <p>Attention : le coût du logiciel ne représente qu'une faible part du coût total. Le support logiciel, la maintenance logicielle et les compétences techniques en interne sont des exigences essentielles et importantes. Une solution logicielle open source exige souvent de la part de l'utilisateur une meilleure compréhension technique, ce qui pourrait accroître incidemment les coûts.</p>
	Le système est construit avec un logiciel propriétaire .	<p>Note : les logiciels propriétaires sont la propriété d'un individu ou d'une entreprise. Cette situation fait que les utilisateurs n'ont généralement pas accès au code source. Un avantage clé est que le logiciel propriétaire est livré avec des garanties et un support que l'open source n'offre généralement pas.</p>
Système accessible sur le web	L'application de base de données est conçue pour être gérée et accessible via Internet .	<p>Note : un système accessible sur le web offre également la possibilité d'être déployé sur un réseau local ou étendu.</p> <p>Conseil : veiller à ce que le système dispose d'une fonctionnalité hors connexion qui autorisera la saisie des données sans connectivité Internet. Cela signifie que si l'Internet tombe en panne au cours d'une session de travail, l'utilisateur pourra poursuivre la saisie des données dans le système, et il pourra ensuite télécharger une fois la connectivité rétablie.</p>

NIVEAU DE LA GESTION ET DU STOCKAGE DES DONNÉES

Fonction	Description	Points essentiels à envisager
Stockage des données	Le système a la capacité de stocker les données de plusieurs années dans la même base de données.	Note : le stockage dans la base de données de grandes quantités de données correspondant à plusieurs années permet de créer un ensemble de données longitudinal. La couverture longitudinale des données est un besoin essentiel pour suivre les apprenants individuels dans le système éducatif. Elle facilite également le suivi et la représentation des tendances au niveau des établissements scolaires, caractéristique d'analyse à valeur ajoutée essentielle que le SIGE devrait permettre. Il s'agit de créer un ensemble de données qui inclut les informations des mêmes apprenants d'une année sur l'autre.
Entrepôt des données	Le système utilise un entrepôt de données ou un mini-entrepôt (datamart) permettant l'exploration des données.	Note : les données transactionnelles sont extraites des systèmes opérationnels et transformées en une ou plusieurs bases de données dans l'entrepôt de données, en vue de l'exploration et de l'analyse des données par les protocoles et selon les principes de l'entrepôt de données. Pour de plus amples détails sur les entrepôts de données, voir la Section 6.2.5.3.
Sécurité et confidentialité	Le contrôle d'accès strict, la sécurité des données, le respect de la vie privée, la protection des données, la confidentialité, l'authentification et le chiffrement (encryptage) sont des fonctionnalités clés de la solution logicielle.	Note : pour maintenir la sécurité, des capacités sont requises pour sécuriser la plate-forme, telles que le transfert crypté des données (ex. HTTPS, HSTS) et le chiffrement des données stockées ; l'administration des utilisateurs ; et l'audit d'accès et d'utilisation des plates-formes. Une architecture de sécurité flexible devrait permettre de gérer l'accès au système, à différents niveaux, par différents types d'utilisateurs et de rôles personnalisés. Elle sera définie en fonction des exigences du pays.
	Authentification : les utilisateurs sont identifiés et authentifiés de façon unique dans le système.	Note : attribue un identifiant unique à tous les utilisateurs du système, qui sert à les authentifier dans le système.
	Contrôle d'accès par mot de passe : chaque utilisateur peut avoir un nom d'utilisateur et un mot de passe uniques liés à des rôles et privilèges individuels ou de groupe.	Note : le système dispose d'une fonctionnalité permettant d'attribuer un nom d'utilisateur et un mot de passe uniques à chaque individu du système. Chaque utilisateur unique se voit attribuer un rôle au sein d'un groupe. Note : différents types de mots de passe (voir ci-dessous « Hiérarchie des droits d'accès ») peuvent conférer différents droits à l'utilisateur (ex. lecture uniquement, lecture et écriture).
	Hiérarchie des droits d'accès : le système garantit que chaque utilisateur n'a accès qu'aux données dont il a besoin.	Note : ceci est le fruit d'une combinaison d'utilisateurs, de rôles, de groupes et de privilèges. Les « rôles » renvoient aux fonctionnalités auxquelles les utilisateurs ont accès et les « groupes » renvoient au champ des données accessibles. Les utilisateurs sont créés dans le système avec un accès au niveau de la hiérarchie où ils travaillent. L'utilisateur peut capturer ou accéder à des données uniquement liées à ce niveau hiérarchique en fonction des rôles utilisateur et des groupes d'utilisateurs.
Interopérabilité et intégration des données	Le système est compatible avec d'autres systèmes existants.	Note : l'interopérabilité renvoie au protocole de communication entre les systèmes. C'est la capacité pour différents systèmes d'information d'échanger, de partager et d'exploiter des données en toute cohérence. Note : l'intégration est le processus qui permet de créer ou d'assurer l'interopérabilité. L'interopérabilité est le processus de combinaison ou d'intégration de données hétérogènes résidant dans différentes sources et offrant à l'utilisateur une vue unifiée de ces données.
	La fonctionnalité de pouvoir exporter et importer des données à partir d'autres sources existe dans le système.	
	Les données provenant de sources multiples peuvent être liées, intégrées ou fusionnées à l'aide d'un champ commun, sur tout un ensemble de sources de données.	

NIVEAU DE LA GESTION ET DU STOCKAGE DES DONNÉES

Fonction	Description	Points essentiels à envisager
Imputation	Un bon système d'information a la capacité d'imputer les données lorsque les données pour certaines écoles sont manquantes .	<p>Note : les méthodes d'édition et d'imputation appropriées sont utilisées et revues régulièrement ou actualisées si besoin est. La méthode d'imputation utilisée peut déterminer l'exactitude de la valeur imputée. Il convient de fournir des informations (détails) sur les raisons pour lesquelles telle ou telle méthode particulière a été choisie et sur la date de son réexamen.</p>
Hébergement sur le web	<p>Hébergement dans le « cloud » : l'accès aux logiciels et aux données sur un serveur Internet, détenu et géré par un organisme tiers.</p>	<p>Note : l'accès à des logiciels hébergés dans le cloud peut moins prêter à controverse que le stockage, par exemple, des données des élèves dans le cloud.</p> <p>Avantage : l'accès aux données est possible de partout, dès lors que l'on dispose d'une connexion Internet. Avec le stockage dans le cloud, toute modification des données est actualisée sur tous les terminaux ayant accès au cloud.</p> <p>Inconvénient : le stockage dans le cloud est dépendant de la connexion Internet. La protection de la vie privée et de la sécurité pourrait être compromise et plus difficile à gérer car les données ne sont plus dans un environnement de stockage physique.</p> <p>Attention : veiller à ce que les serveurs du cloud soient fiables et hautement sécurisés, et que toutes les données soient cryptées. Un fournisseur de cloud d'entreprise approprié doit être en mesure de présenter la documentation ou la certification de sécurité appropriée.</p>
	<p>Auto-hébergement : installe et accède aux logiciels à partir de son propre serveur. L'auto-hébergement est aussi souvent appelé hébergement « sur place » ou « local ».</p>	
	<p>Logiciel en tant que service (SaaS) : Modèle de licence et de fourniture de logiciels selon lequel le logiciel est délivré sous licence, sur la base d'un abonnement et hébergé en central. On l'appelle parfois « logiciel à la demande » et Microsoft s'y référerait précédemment comme « logiciel plus services ». Les applications SaaS sont également connues sous le nom de logiciels basés sur le web, logiciels à la demande et logiciels hébergés. La licence peut être payée sur la base d'une facturation mensuelle du service ou de frais en fonction de l'utilisation.</p>	

NIVEAU DU RAPPORT ET DE L'ANALYSE DES DONNÉES

Fonction	Description	Points essentiels à considérer
Traitement analytique en ligne (OLAP)	Il existe un portail web pour le traitement analytique en ligne (OLAP). Le système permet aux utilisateurs d'analyser simultanément les informations de plusieurs systèmes de base de données.	Note : un portail web est un site web qui sert de point d'accès unique aux informations contenues dans l'entrepôt national de données et d'autres sources pertinentes. Son rôle est de faciliter l'accès à ces sources de données, de façon structurée et systématique. Il peut se connecter à l'interface web de l'entrepôt de données et communiquer avec les ressources pertinentes : cartes, graphiques, rapports, tableaux et documents statiques.
	OLAP offre la fonctionnalité d'analyser jusqu'au niveau de données le plus détaillé.	
	OLAP offre une fonctionnalité appelée « <i>slice and dice</i> ».	Note : OLAP est un processus informatique qui permet aux utilisateurs de sélectionner et d'extraire les données de différents points de vue. L'opération « <i>slice and dice</i> » signifie que les données sont subdivisées en éléments plus petits (ex. ventilation par école, puis par enseignants de l'école, puis par élèves de ces enseignants) ce qui permet d'examiner les données de différents points de vue et d'offrir à l'utilisateur des informations supplémentaires.
	OLAP dispose d'une fonctionnalité de table pivot , c'est-à-dire que l'on peut créer des données tabulaires à partir d'une base de données réelle. Dans une base de données réelle, chaque ligne pourrait préciser le sexe de l'élève. La mise en place d'une table pivot permettrait à l'utilisateur de créer une table récapitulative où le sexe serait une colonne ou une ligne.	
Tableau de bord	Le système a la capacité de créer des tableaux de bord et du contenu hautement interactifs , et de suivre, analyser et afficher les données sous une forme visuelle : les données sont visualisées dans un tableau de bord sous la forme de tableaux, de graphiques linéaires, de diagrammes à barres, etc. Les affichages d'un tableau de bord peuvent être extraits d'un seul ensemble de données ou de nombreux ensembles de données sous-jacents.	
Exécution des requêtes et rapports	Rapports opérationnels : les rapports sont exécutés périodiquement et directement distribués aux utilisateurs concernés.	
	Rapports en libre-service : l'utilisateur exécute les rapports depuis une interface simple pilotée par des menus.	
	Rapports paramétriques : les paramètres commandés par l'utilisateur aident celui-ci à adapter le rapport à des conditions spécifiques lors de l'exécution. Ces paramètres permettent à l'utilisateur de modifier le contenu et ils offrent une certaine flexibilité, moyennant un effort limité et en dépit de connaissances techniques réduites de la part de l'utilisateur.	
	Requêtes ad hoc : les requêtes ad hoc sont rédigées par l'utilisateur, puis stockées en vue d'une utilisation future.	

5.5 Conditions et normes minimales supplémentaires concernant la technologie

En nous fondant sur la section précédente, nous décrivons ici certaines conditions minimales concernant la technologie sous-jacente du SIGE¹⁴

¹⁴ Il y a donc une certaine répétition des concepts clés de la section précédente, en fonction de ce qui est jugé nécessaire.

- le terme « technologie » se rapportant aux *matériels, aux logiciels, aux bases de données et à la mise en réseau*, ainsi qu'aux *télécommunications et aux réseaux*. Dans les pays où le revenu par habitant est très faible ou l'infrastructure informatique globale du gouvernement peu développée, ces conditions peuvent paraître trop ambitieuses et elles constituent plutôt une trajectoire de croissance. Dans d'autres pays, la

technologie déjà utilisée peut avoir dépassé ce qui est indiqué ici.

Les matériels incluent les ordinateurs et les équipements d'information et de communication connexes utilisés pour les activités de saisie, de traitement et de sortie. Ils représentent la partie la plus visible, celle qui joue un rôle critique dans les fonctionnalités et la réussite du système d'information, mais ces fonctionnalités matérielles sont en constante évolution et les options matérielles susceptibles de sous-tendre les conditions logicielles énumérées ci-dessous sont beaucoup trop nombreuses et variées. Dans la présente section, nous allons donc nous concentrer sur les conditions minimales concernant les autres composantes technologiques.

5.5.1 Exigences en matière de logiciels

Les logiciels sont des ensembles d'instructions qui indiquent au matériel ce qu'il doit faire. Ils se répartissent en deux grandes catégories :

logiciels du système d'exploitation et **logiciels d'application**. Pour s'y retrouver dans le nombre incalculable de composantes logicielles, il faut d'abord faire des choix quant aux modules nécessaires : l'architecture du système et son niveau de complexité. Ce qui suit porte sur les principaux composants d'un solide logiciel système.

Logiciel du système d'exploitation : ainsi que cela a été décrit à la Section 5.4, le système d'exploitation est la plate-forme qui fournit un environnement de prise en charge des interactions utilisateur par le biais d'interfaces (ex. écran, souris, clavier, imprimante). Des exemples de systèmes d'exploitation sont Windows, macOS, Linux, Android, iOS et Chrome OS. Le système d'exploitation doit faciliter l'utilisation du SIGE, le rendre visuellement attrayant et permettre l'intégration des flux de travail.

Logiciels d'application : de nombreux pays sont peut-être déjà dotés de systèmes, auquel cas il ne sera pas indispensable de « choisir » un logiciel pour l'ensemble du SIGE. Il faudra peut-être cependant faire des choix concernant d'autres systèmes qui interagissent, manuellement ou non, avec le système existant.

Logiciels sur mesure. Les logiciels sur mesure présentent l'avantage de pouvoir être développés en fonction des besoins spécifiques du pays. Dans la mesure où le ministère détermine le processus et le résultat final, on peut éviter de se doter de fonctions superflues. Cependant, le développement de logiciels sur mesure peut être un processus lent et de longue durée, et le coût initial sera très élevé.

Logiciels standard du marché. Ces systèmes d'exploitation ne requièrent souvent que des modifications mineures au produit commercialisé. Ils peuvent présenter un certain nombre de bénéfices ou avantages, notamment :

- Le processus de modification et de mise en œuvre du logiciel peut être plus rapide. Toutefois, en fonction de l'utilisation, il s'avère parfois difficile ou contestable de modifier un logiciel standard du marché (voir la section précédente).
- La solution logicielle pourrait permettre d'économiser sur les coûts.
- Avec un logiciel standard du marché, il est parfois possible de tester les fonctionnalités avant l'achat.

Les logiciels standard présentent toutefois l'inconvénient d'être souvent fournis équipés de fonctionnalités superflues susceptibles d'ajouter des coûts inutiles.

Logiciels d'application : *fonctionnalités souhaitables*

Partie frontale (ou frontal) : que les données soient saisies électroniquement au niveau local ou manuellement au niveau central, le logiciel doit comporter un module de saisie ayant des fonctionnalités dédiées et intégrant un processus de validation. C'est ce qu'on appelle souvent le « frontal ».

Fonctions de validation des données : la validation des données doit être intégrée aux différents modules à chacune des étapes de la chaîne de production des données : collecte, saisie et diffusion des données.

Logiciels d'application : types

Logiciels open source ou libres.¹⁵ Ces logiciels sont distribués gratuitement, le code source est accessible et mis à la disposition des utilisateurs. Ainsi que nous l'avons déjà indiqué, les principaux avantages sont leur coût très modique et une flexibilité très supérieure au sein de l'infrastructure. Mais le coût du logiciel ne représente qu'une petite partie du coût total. Le support logiciel, la maintenance logicielle et les compétences techniques nécessaires en interne sont des facteurs importants. Une solution logicielle open source exige une meilleure compréhension technique de la part de l'utilisateur, ce qui pourrait augmenter incidemment les coûts.

Logiciel propriétaire. Dans la mesure où la solution logicielle appartient à un particulier ou à une entreprise, les utilisateurs n'ont généralement pas

accès au code source. Un avantage clé est que les logiciels propriétaires peuvent être livrés assortis de garanties et d'un support que le logiciel open source n'offre généralement pas. Noter que si le ministère a choisi de développer son propre logiciel sur mesure, il pourrait en revendiquer la propriété intellectuelle, conformément aux règles en vigueur dans le pays.

5.5.2 Exigences en matière de bases de données

L'approche base de données et les bases de données

Systèmes de gestion de base de données relationnelle (SGBDR). Il s'agit du type de système de gestion de base de données (SGDB) le plus fréquemment utilisé¹⁶, qui autorise la jointure temporaire ou permanente de tables de données par le biais d'un champ commun (clé primaire et clé étrangère). Chaque ligne, chaque enregistrement ou chaque instance d'une base de données est associée à un ensemble fixe d'attributs ou de champs. Chaque table dispose d'une clé primaire qui identifie chaque enregistrement de façon unique. La table peut également contenir une clé étrangère, qui est identique à la clé primaire d'une table externe. Une jointure relationnelle est réalisée en faisant correspondre les valeurs de la clé étrangère aux valeurs correspondantes de la clé primaire de la table externe.

Système de fichiers à deux dimensions (plats)

Ainsi qu'indiqué précédemment, ce système repose sur une structure simple (une seule table de données) et, contrairement à une base de données

15 Il convient d'établir une distinction entre « libre » et « open source ». Il existe beaucoup d'applications propriétaires « libres/freemium », en particulier dans le monde de l'éducation, mais toutes les applications open source ne sont pas gratuites tout le temps. De plus, au sein de l'open source on trouve des licences « copyleft » qui peuvent interdire à des partenaires commerciaux de participer à une activité open source, ainsi que des licences qui sont plus adaptées au développement commercial en open source.

16 Un système de gestion de base de données (SGDB) est un logiciel qui permet aux utilisateurs de créer, extraire, mettre à jour et gérer des données.

structurée avec un modèle relationnel, il ne peut pas contenir plusieurs tables. Il est conseillé d'éviter ce type de structure de base de données car, dans la mesure où il n'y a pas de relation entre les données, plus on ajoute de données et plus le système de fichiers plats devient inefficace. Bien que les données soient plus faciles à comprendre et à visualiser dans une structure de fichiers plats, le processus d'accès aux données et de recherche dans les données peut être lent et lourd.

Traitement des données par un tableur. Si l'on a utilisé des tableurs par le passé, il sera nécessaire de faire migrer rapidement les systèmes. Un tableur peut constituer un choix périlleux lorsqu'il est utilisé pour effectuer des tâches pour lesquelles il n'a pas été conçu, par exemple une base de données ou un logiciel statistique, ou pour exécuter une tâche qui dépasse ses capacités. Comme les informations sont formatées dans la feuille de calcul elle-même, il est compliqué de récupérer les données au moyen des requêtes les plus simples. Bien que le tableur permette d'utiliser les feuilles de calcul pour créer des relations entre les tables, ces relations seront très limitées. En outre, plus une feuille de calcul devient grande, plus la mémoire nécessaire pour la garder ouverte est importante et plus il y a un risque d'erreurs majeures.

Base de données NoSQL (Not-only SQL - pas seulement SQL). Ces bases de données offrent un mécanisme de stockage et d'extraction des données qui est modélisé sous une autre forme que les relations tabulaires utilisées dans les bases de données relationnelles. Il s'agit d'une base de données non relationnelle du type utilisé pour suivre des flux d'événements.

Stockage des données. Le système a la capacité de stocker les données sur plusieurs années, permettant ainsi le transfert en toute transparence

des informations d'une année sur l'autre. S'il n'existe pas d'entrepôt de données disponible, veiller à ce que la base de données soit conçue de manière que le SIGE stocke les données de l'année en cours et des années antérieures. Les données historiques sont importantes pour l'analyse des tendances. Il est nécessaire d'instaurer une politique de conservation des données, car certaines doivent être conservées pendant de nombreuses années, tandis que d'autres peuvent n'avoir d'utilité que pendant quelques jours.

Imputation. Des méthodes appropriées d'édition et d'imputation doivent être utilisées et revues, révisées ou actualisées régulièrement en fonction des besoins. L'imputation est une méthode d'estimation des valeurs manquantes. L'exemple le plus simple est celui d'une école où il n'y a pas de données pour une variable particulière (ex. nombre de manuels fournis à l'école pour une année donnée), mais qui possède des données pour les années antérieures, auquel cas on peut procéder à une extrapolation. Ou si l'on sait que la fourniture des manuels s'effectue en fonction d'une formule basée sur les effectifs, cette formule pourrait être utilisée pour imputer les valeurs. Comme nous l'avons déjà mentionné, un bon système d'information devrait posséder la capacité d'imputer les données lorsque l'on manque de données pour certaines écoles. La méthode d'imputation utilisée peut être déterminante pour la précision de la valeur imputée. Il faut donc fournir des informations (détails) sur les raisons pour lesquelles telle ou telle méthode particulière a été choisie et sur la date de son réexamen.

Système sur le web. Dans la mesure du possible, l'application de base de données devrait être conçue pour être gérée et accessible via Internet. Il peut s'avérer plus facile de déployer une architecture de serveur basé sur Internet et utilisant des navigateurs et des API (selon le cas) qu'un système de serveur client conventionnel basé sur un ensemble structuré

d'interactions entre un ordinateur faisant office de serveur de fichiers et des utilisateurs ayant des besoins particuliers pour accéder à ces fichiers. Comme cela a déjà été décrit, un système basé sur le web peut éventuellement être déployé facilement sur un réseau local ou étendu, ou sur Internet. Veiller à ce que le système dispose d'une fonctionnalité hors connexion autorisant la saisie de données sans connectivité Internet. Cela signifie que si l'Internet tombe en panne pendant une session de travail, l'utilisateur peut poursuivre la saisie des données dans le système, puis les télécharger (automatiquement, dans l'idéal) une fois la connectivité rétablie.

Outils de support

Définitions des termes du domaine de l'éducation (dictionnaire des termes et des concepts). C'est là une condition préalable au fonctionnement efficace de tout système d'information, car cela permet de normaliser les termes et concepts utilisés par une organisation pour ses processus de collecte d'informations et de rapports. Il arrive souvent que différentes terminologies soient utilisées dans les systèmes d'information pour décrire le même concept ou principe dans différentes publications et bases de données. Dans un système de qualité, la terminologie doit être la même pour toutes les institutions afin d'éviter de créer des ambiguïtés. Le dictionnaire pourrait faire partie du système d'aide de la plate-forme logicielle, mais il devrait aussi exister en dehors de la plate-forme.

Dictionnaire de données (métadonnées) et système de métamarquage. Le premier (un référentiel centralisé d'informations à propos des données) explique à l'utilisateur ce que signifient les codes des tables de données (ex. l'âge des élèves peut varier de x à y, ou la plage et les descriptions des réponses valides, par exemple, pour la « langue principale d'instruction »). Le deuxième fournit un

vocabulaire limité ou recommandé pour décrire les données et/ou le contenu d'un système (c'est-à-dire, en un sens, les instructions sur la façon de créer le dictionnaire de données).

5.5.3 Exigences en matière de télécommunications et de réseaux de données (y compris les réseaux et Internet)

Les communications et la connectivité sont des caractéristiques essentielles des systèmes d'informations modernes. La conception d'un SIGE de qualité devrait envisager les aspects suivants :

Bande passante. La bande passante est une mesure de la quantité de données pouvant transiter à tout moment sur le réseau, en différents points du réseau. Une bonne bande passante est importante pour la connectivité Internet ; elle renvoie au débit de transmission maximal et le plus fiable qu'un chemin réseau puisse offrir. Sur certaines plates-formes de réseau comme le haut débit résidentiel, la bande passante effective est limitée. La bande passante est très importante car elle détermine la vitesse à laquelle une page web se charge dans un navigateur. Elle peut être impactée de manière significative selon le site web et la connexion Internet utilisée pour y accéder. Si l'on choisit principalement de stocker les données sur des serveurs dans le cloud, la bande passante a sans doute moins d'importance ; que ce soit pour le transfert des données d'un endroit à un autre ou pour leur saisie à distance, une bande passante adéquate est cruciale.

Technologie mobile. Un service mobile (smartphone et tablettes) est conçu pour permettre une collecte rapide de données mobiles à la fois en mode en ligne et hors ligne. Noter que la couverture de la téléphonie mobile doit être prise en compte.

Niveau de préparation technologique. Ceci est un aspect clé de la mise en œuvre d'une technologie mobile destinée à la collecte de données, et il convient de prendre en compte notamment le coût de propriété des terminaux (prix unitaire, garantie et réparation), de la connectivité Internet et des données. Dans les pays à revenu faible et dans les zones rurales, il peut être fortement conseillé de choisir une technologie avec collecte des données hors ligne.

Transfert des données de base. Le système devrait permettre aux utilisateurs de transférer les données des années antérieures dans une nouvelle base de données couvrant tous les établissements. Ainsi qu'indiqué précédemment, sa capacité est essentielle pour la génération des tendances chronologiques et d'autres formes de données à valeur ajoutée. Le cas échéant, le système devrait offrir la possibilité de pré-imprimer (ou de pré-remplir en ligne) les données de base dans le questionnaire physique avant sa diffusion aux établissements. Noter toutefois que la pré-impression (contrairement au pré-remplissage en ligne) accroît considérablement les exigences logistiques car il faut faire parvenir le bon questionnaire à la bonne école.

5.6 Considérations relatives aux futures technologies

Les cadres conceptuels de prochaine génération portant sur la conception des SIGE ont commencé à se concentrer sur le vaste potentiel des données recueillies principalement auprès des établissements scolaires, des élèves et du personnel, dans le but d'améliorer les résultats d'apprentissage. Pour réaliser ce potentiel dans le contexte d'une architecture SIGE, il faut conduire

une réflexion et une restructuration des systèmes de données dans le secteur de l'éducation. Pour les plates-formes SIGE, la restructuration envisagée placerait les données au cœur des processus décisionnels et améliorerait considérablement le flux de données pour permettre leur exploitation fiable par une grande variété d'intervenants : administrateurs, enseignants, élèves, tuteurs, employeurs.

Les technologies SIGE modernes sont capables d'exploiter les immenses quantités de données éducatives qui sont collectées mais restent en sommeil. En règle générale, il y a bien plus de données collectées que de données analysées pour insuffler des idées aux acteurs de l'éducation. Les architectures SIGE existantes ne sont souvent pas capables d'éliminer de manière adéquate les obstacles à une exploitation rapide des données, par exemple la saisie de données incorrectement alignées se traduit par des lacunes dans les données, la désorganisation et l'incompatibilité des sous-systèmes, une latence dans les rapports et une absence fondamentale de confiance dans la qualité des données. La technologie SIGE moderne peut aider à surmonter ces défis.

La technologie SGBDR/OLAP existante peut maintenant être enrichie de nouvelles approches à l'aspiration, à la transformation, au stockage, à l'analyse et à la visualisation des données. La nouvelle architecture de données SIGE peut fournir des pipelines de données évolutifs et efficaces, offrant un débit élevé et une faible latence.¹⁷ S'il est bien conçu, le pipeline de données SIGE peut offrir une plate-forme complète pour la collecte des données sur l'éducation à partir de diverses sources, la conversion de ces données

¹⁷ La *latence* est le temps que met un paquet à être transféré sur un réseau. Le *débit* est la quantité de données envoyées et reçues par unité de temps.

en informations et leur mise à disposition rapide pour les utilisateurs finaux. Le Tableau 3 tente de résumer ces questions d'avenir, en partie en les

comparant aux défis actuels. Noter que le tableau, loin d'être exhaustif, ne présente que quelques exemples importants.

Tableau 3. Voies d'évolution de la technologie

Niveau	Actuellement géré en général par ¹⁸	Possibilité future, voire actuelle, à envisager
Niveau de la collecte de données	Papier, Excel, Access	Les applications web progressives sont de plus en plus puissantes et conviviales.
	Papier, basé sur le web, mais toujours par clavier	Pour les opérations de type « bloc », numérisation des codes QR par téléphone ou tablette. La biométrie est aussi de plus en plus utilisable (scan des empreintes digitales, de la main entière, de l'iris).
Niveau de la gestion et du stockage des données	Fichiers locaux, bases de données locales, pièces jointes à des courriels	Utilisation extensive des ressources en ligne avec alternatives hors ligne en cas de connexion de mauvaise qualité (papier, SMS). Services officiels pour les métadonnées clés (liste maîtresse des établissements scolaires, ressources humaines, registre des élèves) et pour les données qui peuvent être stockées en même temps sous la forme de données centralisées dans le cloud (national) et de données conservées localement.
	Données « intentionnelles », c'est-à-dire collectées en tant que données, généralement à des fins administratives ou de politique/planification.	Les données peuvent être « résiduelles » ou des sous-produits de l'extraction de données sur les réseaux sociaux, de formulaires de plainte (même par téléphone), etc.
Niveau de l'analyse et des rapports sur les données	Outils de bureau (mono-utilisateur), principalement Excel, mais aussi outils d'analyse spécialisés tels que Stata, SPSS, ArcGIS. Ceux-ci nécessitent souvent un important traitement préalable des données, sous la forme d'événements ponctuels réalisés par des experts.	Il existe de plus en plus d'applications intégrées pour une analyse interactive ; par exemple, R Shiny et une foule d'outils de visualisation JavaScript peuvent accéder aux données via des API et ils sont mis à la disposition d'un grand nombre d'utilisateurs sur Internet. Les interfaces utilisateur peuvent être adaptées à des besoins spécifiques ; des tableaux de bord permettent de partager l'analyse en ligne et des interprétations peuvent être discutées dans les forums et les chats. Les simulations et les outils d'apprentissage automatique sont aussi de plus en plus disponibles, et particulièrement utiles à des fins de planification.

Source : Auteurs.

18 Cette colonne est quelque peu axée sur les pays à revenu faible, et non sur les pays à revenu intermédiaire de la tranche inférieure et certainement pas sur les pays à revenu intermédiaire de la tranche supérieure ou élevé.

6. Guide de l'utilisateur

Une fois que les responsables et les organismes d'aide auront déterminé les fonctionnalités qu'ils souhaitent avoir dans leur SIGE, la prochaine étape consistera à explorer les modalités de mise en service de ce système. Cette section décrit cette tâche plus en détail. Par ailleurs, l'ISU a rédigé le Guide opérationnel d'utilisation du SIGE pour le suivi de l'ODD 4 qui approfondit plus précisément de quelle manière utiliser le SIGE pour établir les rapports sur l'ODD 4. Il est disponible à l'adresse suivante : http://uis.unesco.org/sites/default/files/documents/operational_guide_to_using_emis_to_monitor_sdg_4_fr.pdf

6.1 Besoins en personnel

Non seulement le Guide de l'utilisateur sert à comprendre les normes de fonctionnalité spécifiques, en expliquant les moyens de les appliquer, mais il doit aussi aider les utilisateurs à déterminer si le personnel possède les compétences nécessaires pour exploiter les systèmes. L'Encadré 4 donne un récapitulatif de certaines de ces compétences, à un niveau assez élevé.

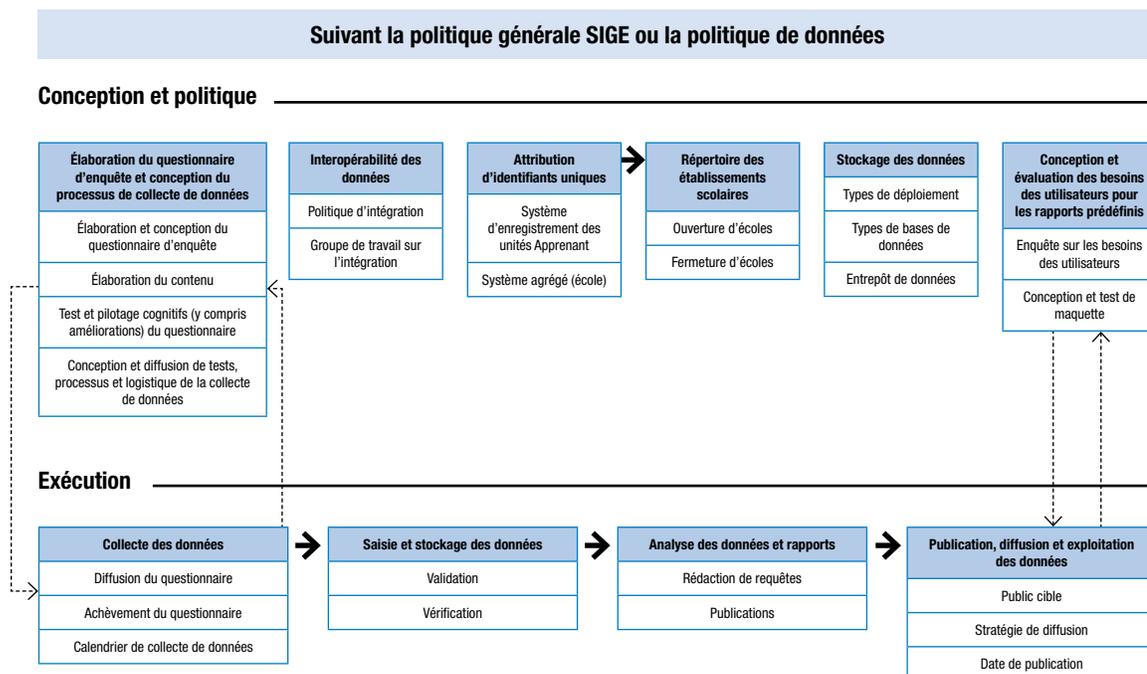
Encadré 4. Compétences techniques attendues du personnel SIGE, au niveau national, infranational et scolaire

La structure organisationnelle du SIGE doit refléter les principaux processus et activités de travail. Le succès d'une initiative SIGE au sein du ministère de l'Éducation dépend de son efficacité à tirer parti des compétences et des connaissances intégrées de son personnel. Il est important d'assurer une présence nationale, infranationale et institutionnelle d'un personnel SIGE possédant les compétences pertinentes, lesquelles doivent refléter les besoins d'informations et l'orientation stratégique du ministère de l'Éducation.

Le profil des compétences du personnel SIGE, à tous les niveaux, devrait inclure au minimum :

- Connaissance des processus et des activités de collecte de données
- Compétences minimales en bases de données (éléments essentiels de la base de données, y compris l'écriture de requêtes complexes)
- Connaissance de la gestion et de l'analyse de données (y compris, si possible, au moins un cours fondamental en statistiques)
- Compétences et connaissances de base en TIC, dans les domaines se rapportant à plusieurs composantes différentes qui permettent aux utilisateurs/personnels d'extraire, stocker, produire et partager des informations, notamment :
 - Matériels – ordinateurs, imprimantes, etc.
 - Logiciels – systèmes d'exploitation et progiciels (ex. tableurs, bases de données, présentations) installés sur des ordinateurs
 - Systèmes de communication – Internet, navigateurs, courriel, collecte d'informations sur téléphone cellulaire, etc.
 - Plates-formes en ligne – visioconférence (par exemple Skype, Zoom, Hangout, Teams)

Figure 3. Cycle de vie de production du SIGE



Source : Auteurs

Note : Bien que le processus de production représenté en partie basse – l'exécution – soit relativement linéaire, la partie haute – conception et politique – est de type plutôt non linéaire, puisque seules quelques-unes des étapes sont séquentielles. Par exemple, il est difficile d'avoir un répertoire approprié des établissements scolaires sans avoir attribué des identifiants scolaires uniques. Dans un souci de simplification du diagramme, les zones qui sont liées de façon linéaire dans la partie haute ne sont pas représentées. Une seule flèche est tracée, à des fins d'illustration. Il y a une interaction entre les deux parties, car la conception est forcément nécessaire à l'exécution et des boucles de rétroaction relient l'exécution à la conception ; dans la réalité, elles sont beaucoup plus nombreuses que celles représentées ici. Noter également que « l'exploitation » se réfère davantage à une utilisation quotidienne ou managériale qu'à l'exploitation à des fins de stratégie et de planification. Dans ce deuxième sens, l'exploitation aura lieu en dehors du cycle représenté ici, au sein d'autres unités du ministère (voir la Section 4 sur l'architecture des données).¹⁹

19 C'est-à-dire plus linéaire que la partie haute, et plus linéaire que l'interaction entre la partie haute et la partie basse.

6.2 Cycle de vie de production du SIGE

Le Guide de l'utilisateur définit les processus, les procédures et les pratiques de collecte, de traitement et de diffusion des données éducatives afin d'assurer la production de données de qualité – c'est-à-dire des données *complètes, pertinentes, précises, actuelles* et *accessibles*.

Dans ce guide, ces processus sont regroupés sous l'appellation *cycle de vie de production du SIGE*, lequel est décrit Figure 3. Le cycle de vie de production du SIGE guide le lecteur pas à pas dans ces processus et pratiques, de la conception initiale et de l'élaboration du questionnaire d'enquête

jusqu'aux processus de collecte de données, d'exploitation et de diffusion des données. Le Guide de l'utilisateur sera basé sur le cycle de vie de production du SIGE en tant qu'objectif pour déterminer toutes les activités essentielles des processus de collecte et d'exploitation des données à tous les niveaux, du niveau national au niveau infranational, jusqu'aux écoles.

6.2.1 Élaboration du questionnaire d'enquête (Conception et développement)

Le questionnaire d'enquête constitue le cœur du SIGE. Dans la plupart des pays en développement, les données requises sont obtenues au moyen

d'enquêtes (méthode du recensement scolaire – voir l'Encadré 5). La collecte de données peut s'effectuer sur papier ou par des moyens technologiques.

Lorsque les données sont recueillies dans le cadre d'une enquête auprès des établissements scolaires, ce qui est la principale source de données, cela s'effectue en général au moyen d'un questionnaire qui sert également d'instrument de développement de la base de données. Dans un système national unitaire, le bureau national (généralement l'unité SIGE au niveau central) est responsable de l'élaboration, de la conception, de l'impression (le cas échéant) et de la diffusion dans les délais du questionnaire d'enquête. Dans un système fédéral, le niveau national pourrait faciliter un accord et développer des normes en collaboration, mais c'est habituellement le premier niveau infranational (province ou État) qui est chargé d'exécuter les opérations ; par ailleurs les normes peuvent varier.

6.2.1.1 Conception du questionnaire

La conception initiale du questionnaire est importante parce que :

- Grâce au questionnaire, on obtient les données d'intrants des établissements scolaires sous une forme physique (sur papier), ou via Internet dans un système web.
- C'est l'élément le plus important pour la construction de la base de données et comme cahier des charges (URS) pour l'élaboration du modèle de base de données²⁰.
- Il est utilisé pour concevoir et développer les écrans de saisie dans le cas d'un système électronique ou basé sur le web (l'Encadré 6 décrit un problème particulier mentionné dans les applications précédentes, qui propose à la fois une saisie sur papier et électronique). Il convient de noter que lorsque la saisie n'est pas uniquement électronique, les versions électronique et sur papier ont exactement la même disposition.

20 Un URS est un document utilisé pour guider les développeurs de la base de données dans la détermination des fonctions qu'il faut inclure dans le système et dans la création du modèle de base de données.

Encadré 5. Méthode du recensement scolaire : saisie manuelle ou via le web ?

La saisie manuelle des données du recensement scolaire présente certains inconvénients. L'impression et la distribution du formulaire d'enquête à toutes les entités concernées sont souvent très gourmandes en temps et elles peuvent s'avérer extrêmement coûteuses en termes de carburant et de papier. L'opération est compliquée au plan logistique et certaines écoles sont presque toujours omises. La saisie et le nettoyage des données prennent souvent aussi beaucoup de temps. Par conséquent, l'obtention de données fiables et de qualité, utilisables pour la planification de l'éducation et la prise de décision, prend trop de temps. Ce système manuel présente néanmoins quelques avantages, en particulier dans les zones très pauvres, où toutes les écoles n'ont pas de connexion Internet et où des terminaux comme des tablettes ou des PC, si elles en sont équipées, peuvent permettre un téléchargement asynchrone (c'est-à-dire qu'après la saisie des données sur le PC local ou une tablette, elles sont téléchargées ultérieurement sur le serveur à un moment où cela est pratique, lorsque le WiFi devient disponible). On peut aussi envisager une saisie manuelle dans certaines régions d'un pays et électronique dans d'autres ; ou manuelle à certains niveaux (par exemple l'école) et électronique à d'autres niveaux qui peuvent avoir accès à Internet (par exemple, les bureaux de district).

Encadré 6. Écarts entre la collecte manuelle de données et les écrans de saisie SIGE²¹

Dans le cas d'une saisie électronique, l'écran de saisie d'un système électronique ou d'un système web reproduit « exactement » le contenu du questionnaire. Pour les enquêtes sur papier, les écrans de saisie de la base de données développés en vue de la capture des données peuvent ne pas représenter l'ensemble du contenu du questionnaire, ce qui peut conduire à un écart « délibéré » entre les données recueillies à travers les questionnaires de recensement scolaire et les données saisies dans la base de données du SIGE. Cette pratique peu répandue a été observée dans certains pays d'Afrique subsaharienne pendant la période de l'EPT et des OMD (en raison de contraintes liées à la capacité du système ou à des ressources limitées).

21 Information fournie avec l'autorisation d'O. Labé (voir l'Annexe A pour une liste de tous les contributeurs).

6.2.1.2 Élaboration du contenu du questionnaire d'enquête

- Le contenu du questionnaire d'enquête est élaboré dans le cadre d'un processus collaboratif, c'est-à-dire que le contenu (les questions) du questionnaire est compilé en collaboration avec les principales parties prenantes.
- Le contenu du questionnaire d'enquête est essentiel, car il permet de s'assurer que toutes les informations pertinentes sont saisies.
- Le contenu du questionnaire d'enquête inclut des catégories telles que des informations générales sur l'école (nom, adresse, secteur, années d'étude) et sur les apprenants (effectifs scolarisés, langue parlée à la maison, âge et matières proposées).
- Des tests cognitifs répétés et un pilotage du questionnaire doivent faire partie du processus, même si l'on a déjà utilisé un questionnaire SIGE ou si des tests cognitifs ont déjà eu lieu antérieurement (afin de l'améliorer au fil du temps).

Le présent guide ne décrit pas en détail les principes généraux qui s'appliquent à la construction de questionnaires de qualité. Pour de plus amples

indications sur le contenu du questionnaire d'enquête, il peut être utile de s'informer sur la façon dont un questionnaire est conçu par les chercheurs qui travaillent sur les enquêtes auprès des ménages ou des établissements, car il leur arrive de piloter les questionnaires à plusieurs reprises, d'utiliser des laboratoires cognitifs pour comprendre comment les individus interprètent les questions, et ainsi de suite. Ces pratiques valent certainement la peine d'être imitées, car ces chercheurs consacrent souvent plus de temps à comprendre comment les déclarants réagissent aux questionnaires que les concepteurs des formulaires administratifs. Par ailleurs, plusieurs institutions disposent de ressources à la fois sur les enquêtes par échantillonnage aléatoire auprès des ménages et des établissements, ainsi que sur la conception des formulaires administratifs.²² Bien

22 Voir, par exemple: https://www.povertyactionlab.org/sites/default/files/documents/Instrument%20Design_Diva_final.pdf; <https://github.com/worldbank/DIME-Resources/blob/master/survey-instruments.pdf>; <http://documents.worldbank.org/curated/en/452741468778781879/Volume-One>; http://surveys.worldbank.org/sites/default/files/files/C4D2-T/C4D2-T2/C4D2-T2_PPTS/designer.pdf; <https://unstats.un.org/unsd/EconStatKB/KnowledgebaseArticle10364.aspx>. Pour des exemples de directives gouvernementales sur les formulaires web, voir <https://www.usability.gov/get-involved/blog/2008/04/usable-online-forms.html> et <https://www.usability.gov/how-to-and-tools/resources/publications/defensive-design-web-how-improve-error-messages-help-forms-and-other-crisis-points.html>. Pour un exemple

que la plupart des SIGE des régions à revenu faible soient peu susceptibles de mettre en œuvre dans un avenir immédiat la saisie directe par les écoles sur ordinateur ou tablette, il est intéressant de noter que la conception de formulaires en ligne de qualité se prête plus ou moins aux mêmes considérations que les formulaires traditionnels avec papier et crayon.²³

6.2.1.3 Diffusion du questionnaire d'enquête

- Dans le cas d'un processus manuel, le questionnaire d'enquête devrait être diffusé en vue d'être renseigné par toutes les écoles concernées, à partir du répertoire des établissements scolaires.
- Dans le cas d'un processus technologique, l'accès au système devrait être fourni à tous les établissements en temps opportun.
- Les deux méthodes peuvent coexister dans un pays, selon le niveau de technologie et le niveau au sein du système (école par rapport au district, par exemple).

6.2.1.4 Remplissage du questionnaire d'enquête

- Tous les établissements scolaires (ex. primaires et secondaires, publics ou privés) doivent remplir le questionnaire d'enquête avant une date butoir précise.
- Les éléments de données du questionnaire sont considérés comme indispensables à la fourniture et à la gestion de l'éducation dans le pays.
- Il convient de décourager l'élaboration de questionnaires trop longs et d'éviter les détails superflus, qui peuvent être sans importance ou sans pertinence.
- Le questionnaire doit être accompagné de directives appropriées (dictionnaire de données

et métadonnées) sur la façon de le remplir et, si nécessaire, d'une formation.

- Il convient de définir clairement dans le questionnaire les termes et éléments techniques et de données utilisés afin que les déclarants puissent comprendre quelles sont les informations à fournir.
- Le remplissage du questionnaire au niveau de l'école doit être effectué en se référant aux documents officiels. Par exemple, pour une question sur l'âge, il convient de consulter le registre de classe contenant le détail des dates de naissance.
- En ce qui concerne la vérification du questionnaire d'enquête (preuve de la qualité des données déclarées), il convient de :
 - Vérifier des informations comme le total des effectifs scolarisés par rapport à celles qui sont contenues dans les documents légaux utilisés pour la tenue des registres par l'école (ex. registre des admissions) ;
 - Vérifier que les tableaux contenant les mêmes informations que le questionnaire, par exemple les effectifs scolarisés par âge, par année d'étude ou par sexe, produisent tous les mêmes totaux. La saisie électronique des données permettrait de signaler ce type de problème.

6.2.1.5 Calendrier de collecte de données

- Un calendrier couvrant toutes les étapes du processus de collecte de données doit être diffusé à tous les acteurs des rôles pertinents. Un exemple est inclus à l'Annexe C.
- Ces étapes doivent être indiquées dans le questionnaire d'enquête.
- Les dates importantes doivent figurer dans le calendrier, telles que :
 - Diffusion des questionnaires d'enquête ;
 - Remplissage des questionnaires d'enquête au niveau des établissements scolaires ;

de directives fondées sur le conseil, voir <https://www.effortmark.co.uk/>. Voir Dillman et coll., 2014 (dans Bibliographie).

23 Cette information est fournie avec la permission de D. Dillman, Washington State University.

- Début de la saisie des données ;
- Achèvement de la saisie des données ;
- Publication des données.

6.2.2 Attribution d'identifiants uniques

Les identifiants uniques sont des numéros d'identification générés au niveau national (ou provincial, mais en suivant un protocole national, y compris dans un pays fédéral) qui sont attribués à toutes les personnes/entités (ex. apprenants – matricule ; enseignants et autres membres du personnel – numéro de paie ; établissements d'enseignement – numéro SIGE).

Lorsque l'on crée un système de données longitudinales, il est nécessaire de relier les différents ensembles de données collectées chaque année sur chaque apprenant ou établissement scolaire, à l'aide d'un champ commun à tous ces ensembles de données. Pour ce faire, des codes d'identification uniques doivent être attribués à chaque apprenant (voir la discussion sur l'« enregistrement des unités Apprenant » au Tableau 2 de la Section 5.4) ou à chaque établissement. Il est important que cet identifiant reste cohérent et précis dans le temps.

Le bureau national du SIGE (ou peut-être le bureau provincial dans un pays véritablement fédéral) devrait élaborer et introduire des normes régissant la création d'identifiants uniques.

6.2.2.1 Système d'identification unique des apprenants

Système d'identification unique des apprenants pour la mise en œuvre d'un SIGE basé sur les enregistrements de niveau unité

Le suivi individuel des apprenants n'est peut-être pas adapté à de nombreux contextes, car il est exigeant au plan technologique et administratif. Dans

ces situations, il convient de procéder à une analyse minutieuse des coûts et des avantages avant de se lancer dans un tel projet. Dans les cas où un système d'identification unique des apprenants est adapté, les considérations suivantes s'appliquent.

Un identifiant unique est un numéro unique, non dupliqué, qui est attribué à un apprenant pour la totalité de son parcours d'études, qu'il change ou non d'école. Aucun apprenant ne peut avoir plus d'un numéro, et deux apprenants ne peuvent pas avoir le même numéro.

L'identifiant unique permet de suivre la progression de l'apprenant dans le système, via un identifiant en données longitudinales (données recueillies sur le même apprenant d'année en année). Le bureau national attribue à chaque apprenant un identifiant national unique (matricule) qui servira à identifier précisément ses enregistrements au fil des ans. (Cet identifiant peut être attribué au niveau provincial dans un pays fédéral, ensuite on fait remonter l'information.)

Le système d'identification unique des apprenants fonctionne de la façon suivante :

- Le bureau national élabore des procédures pour faire en sorte de ne pas attribuer deux identifiants égaux du même type au même apprenant (par exemple, lorsque l'apprenant déménage, il conserve son identifiant) et à ne pas attribuer le même identifiant à deux apprenants. Ces règles doivent être respectées à tous les niveaux.
- Le SIGE national (ou provincial dans un pays fédéral) attribue un identifiant qui suivra chaque apprenant dans toutes ses classes, tout au long de son parcours (ou de ses changements d'école ou de région infranationale) jusqu'à ce qu'il quitte (soit désinscrit) et se réinscrive

(réinscription) dans le système d'éducation publique du pays.

- Le SIGE national :
 - Autorise les niveaux infranationaux à attribuer de nouveaux identifiants aux apprenants qui n'ont pas d'identifiant en cours ;
 - Délivre ou enregistre un identifiant national remonté (dans un système fédéral) qui sera utilisé par les niveaux infranationaux pour rapporter des données de niveau unité au ministère de l'Éducation nationale ;
 - S'appuie sur un nombre spécifique d'éléments de données qui fournissent des caractéristiques sur un apprenant. Ces éléments de données sont stockés dans une base de données sécurisée et confidentielle au niveau national et utilisés pour identifier un apprenant de façon unique. Ces éléments pourraient inclure, entre autres, le nom de l'apprenant, son nom de famille, sa date de naissance, son sexe, etc.

6.2.2.2 Système d'identification unique des établissements (établissements scolaires et autres lieux)

Système d'identification unique des établissements (répertoire des établissements)

Noter que ces identifiants sont généralement plus importants que les identifiants délivrés aux apprenants, car certains pays peuvent ne pas être en mesure d'utiliser (ou ne pas prioriser l'utilisation) d'identifiants uniques pour les apprenants, tandis que les identifiants uniques des établissements et la liste maîtresse sont la clé de tout SIGE fonctionnel.

Le système de fichier maître des établissements est un processus essentiel pour attribuer un identifiant

unique à un établissement ; il doit être géré au niveau central.

Le but d'un tel système est d'attribuer un identifiant unique à chaque établissement du pays. L'identifiant unique de l'établissement permet d'analyser les données longitudinales institutionnelles (données recueillies sur le même établissement d'année en année). On peut également vouloir que l'identifiant de chaque établissement indique automatiquement la place de l'établissement dans la hiérarchie de gouvernance. Autrement dit, l'identifiant peut être formé de codes indiquant les niveaux supérieurs à l'école, bien que cet enchaînement de codes hiérarchiques ne soit pas strictement nécessaire tant que la base de données comporte des champs (par exemple district, province) qui permettent d'identifier l'école.

Le fonctionnement de base du système d'identifiant unique des établissements implique les éléments suivants :

- Le niveau national (ou, en général, le premier niveau infranational dans un système fédéral) attribue à chaque établissement un identifiant institutionnel national unique qui servira à suivre avec précision les enregistrements effectués d'une année sur l'autre.
- Le bureau national élabore des procédures pour veiller à ce que le même identifiant ne soit pas attribué à deux établissements.
- Le bureau national, en conjonction avec le SIGE infranational, élabore des procédures pour la fermeture et l'ouverture des établissements.

6.2.3 Maintenance du répertoire (registre) des établissements scolaires

Un répertoire des établissements scolaires doit être créé, tenu à jour et actualisé à intervalles réguliers (tous les trimestres, tous les semestres ou annuellement). Plus précisément :

- Il faut mettre en place une unité dédiée, responsable du répertoire des établissements scolaires, à un niveau infranational (ou peut-être sous-provincial dans un pays fédéral), responsable de l'ouverture et de la fermeture de tous les établissements privés et publics (notant que certains pays enregistrent très peu d'écoles privées, en particulier dans le préscolaire).
- Cette unité aura pour but, lors de l'enregistrement des établissements scolaires, d'attribuer un identifiant unique à chaque établissement (en collaboration avec le bureau national), de maintenir et d'actualiser le répertoire officiel des établissements et de veiller à ce que le même identifiant ne soit pas attribué à plusieurs établissements scolaires.
- Le répertoire des établissements constitue une étape importante dans l'ensemble du cycle de vie du SIGE, car il facilite la diffusion des questionnaires (dans un système papier) ou l'accès aux questionnaires (dans un système web).
- Le taux de réponse de la collecte de données peut être déterminé avec précision si l'on dispose d'un répertoire complet et à jour des établissements.
- Le répertoire des établissements devrait inclure leurs coordonnées de base, leurs programmes d'enseignement ainsi que des informations linguistiques, et même des informations spatiales requises par l'utilisateur de l'information.

- Un répertoire de tous les établissements devrait être disponible au niveau central et dans chacun des bureaux infranationaux.
- Toutes les entités du ministère qui pourraient recueillir des données indépendamment du SIGE sont tenues (même rétroactivement, c'est-à-dire pour intégrer les identifiants dans des bases de données existantes qui présentent une certaine importance) d'utiliser la liste maîtresse et les identifiants SIGE. Si les entités du ministère souhaitent également utiliser d'autres identifiants à leurs propres fins, elles devront veiller à ne pas créer de confusion et à faire en sorte que les deux identifiants accompagnent toujours les enregistrements de données et figurent dans toutes les bases de données.

6.2.4 Saisie des données

- L'assurance qualité est essentielle à la production de données de qualité (voir la Section 5.3), à des données complètes, pertinentes, précises, actuelles et accessibles.
- La saisie des données doit être vérifiée et validée, en particulier dans le cas d'un processus de saisie manuelle.

6.2.4.1 Vérification

- La vérification assure la prévention des erreurs commises lors de la saisie des données.
- La vérification implique de contrôler que les données ont été saisies correctement dans le système à partir du formulaire d'enquête.
- Les méthodes de vérification pouvant être utilisées sont les suivantes :
 - Vérifier que les données saisies dans le questionnaire d'enquête correspondent bien à celles de la base de données (relecture) ;
 - Utiliser des programmes informatiques (requêtes) pour comparer le total des tableaux de données qui devraient produire

les mêmes totaux, tels que les effectifs scolarisés par âge par rapport aux effectifs scolarisés par sexe par rapport aux effectifs scolarisés par année d'étude, etc.

6.2.4.2 Validation

- La validation est l'opération de contrôle que les données saisies ont une signification réelle ; par exemple, que les données telles que les effectifs scolarisés, l'âge, etc. se situent dans une gamme de valeurs admissibles.

6.2.5 Stockage de données

La base de données pour le stockage des données est un élément clé du cycle de vie du SIGE. Une base de données est organisée de telle manière que les données stockées sont facilement accessibles, selon le type de déploiement.

6.2.5.1 Type de déploiement

- *Déploiement hors connexion* : le déploiement hors connexion signifie que des instances hors connexion autonomes sont installées pour les utilisateurs finaux, généralement au niveau infranational : ce sont principalement les fonctionnaires au niveau infranational qui assurent la maintenance du système.
- *Déploiement en ligne* : on peut décider d'utiliser une architecture de serveur web car l'architecture des serveurs Internet se déploie plus facilement qu'un système de serveur client. Un système web peut aussi être déployé facilement sur un réseau local ou étendu ou sur Internet. Enfin, l'architecture web facilite la fourniture rapide de données.

6.2.5.2 Type de base de données

- La solution optimale pour la saisie des données est l'utilisation d'un système de gestion de base de données relationnelle (SGBDR). C'est

au niveau du cahier des charges URS que l'on optera ou non pour un SGBDR.

- La structure relationnelle facilite l'interrogation de la base de données et l'intégration de grands ensembles de données provenant de sources multiples.
- L'intégration des données signifie généralement que l'on relie différentes sources de données au moyen d'un champ commun sur une série de sources de données.
- Pour ce faire, des codes d'identification uniques doivent être attribués aux ensembles de données utilisés pour l'intégration.

6.2.5.3 Entrepôt de données

- Un entrepôt de données est un ensemble de bases de données interconnectées, conçu et développé principalement à des fins d'analyse. En règle générale, les données sont chargées depuis différents systèmes opérationnels ou transactionnels contenant des bases de données différentes et liées.
- L'entrepôt de données est construit de façon à faciliter l'exploitation des informations de gestion et non à des fins opérationnelles.
- L'entrepôt de données devrait idéalement utiliser les mêmes identifiants uniques que l'épine dorsale du SIGE.
- La question de savoir si l'entrepôt de données doit relever du bureau du SIGE ou d'un autre bureau (analyse ou planification) n'est pas essentielle.
- Les sources de l'entrepôt de données sont les bases de données opérationnelles produites à partir des effectifs scolarisés (SIGE), des évaluations d'apprentissage (examens), de la santé des élèves, des enseignants, des matériels d'apprentissage (manuels scolaires), des installations physiques, etc.
- C'est l'entrepôt de données qui offre finalement une valeur ajoutée significative à l'épine dorsale

Encadré 7. Stratégie d'interopérabilité

Bien souvent, les différentes composantes du système éducatif recueillent et gèrent de grandes bases de données sans les partager entre elles. Par exemple, les données relatives à l'évaluation des élèves (et/ou examens), la qualification et le salaire des enseignants (paie), les effectifs scolarisés (SIGE) et les fournitures de matériels d'apprentissage (manuels) sont stockées dans des bases de données distinctes et non partagées. Si ces ensembles de données ne sont pas reliés entre eux, ils restent fractionnés et existent isolément les uns des autres, empêchant d'effectuer les analyses à valeur ajoutée utiles qui devraient être au cœur de la mission du SIGE et qui sont nécessaires pour assurer sa santé budgétaire. Pour que les données deviennent exploitables dans ce puissant type d'analyse, on peut :

- Élaborer et mettre en œuvre une politique de partage et d'intégration des données ;
- Créer officiellement un groupe (ou une équipe de travail) chargé de créer les liens entre des ensembles de données comme les effectifs scolarisés, l'évaluation, la paie, les matériels d'apprentissage et les enseignants, à partir d'un identifiant unique commun.

du SIGE dans le domaine de l'analyse des politiques et pour des finalités de planification plus complexes. Les données brutes du SIGE sont utiles pour la gestion et une planification plus routinière (ex. ressources scolaires).

- Contrairement aux tables des systèmes opérationnels, les règles de forme normale ne s'appliquent pas et toute dénormalisation de la conception facilitant le processus de collecte d'information est acceptable. Par exemple, le nom d'un enfant ne sera pas reproduit chaque fois qu'un événement de la vie de cet enfant sera enregistré ; c'est son identifiant qui sera utilisé et son nom sera retrouvé si nécessaire.²⁴

6.2.6 Interopérabilité des données

Quand l'information est intégrée, elle offre une valeur ajoutée bien supérieure. La mise en place de définitions standard et de systèmes de codage dans l'ensemble du système éducatif se traduira

par une intégration plus rapide des données entre les unités. Cette intégration signifie généralement que l'on relie différentes sources de données par le biais d'un champ commun sur une série de sources de données (voir l'Encadré 7). Pour ce faire, des codes d'identification uniques doivent être attribués à tous les niveaux du système éducatif pour lesquels des données sont recueillies. C'est l'absence de tels identifiants uniques communs qui autorisent les liens entre les systèmes de données, qui fait que l'on ne dispose pas de systèmes d'information intégrés.

L'interopérabilité permet à différents systèmes (niveaux) de « se parler » et de partager des données. Le partage des données et de l'information est grandement facilité par le recours à un ensemble commun de normes et de métadonnées. Les normes et directives techniques doivent décrire les moyens d'améliorer les données des services gouvernementaux. Elles offrent aux administrations publiques (aux échelons national, régional et local) l'occasion de se joindre entre elles et de partager leurs ensembles de données, autorisant ainsi un

²⁴ Voir [https://fr.wikipedia.org/wiki/Forme_normale_\(bases_de_donn%C3%A9es_relationnelles\)](https://fr.wikipedia.org/wiki/Forme_normale_(bases_de_donn%C3%A9es_relationnelles))

flux transparent d'informations entre les différents départements et niveaux de l'administration.

L'interopérabilité revêt une importance particulière pour la communication des indicateurs de l'ODD 4, car bon nombre d'entre eux requièrent des données qui ne se trouvent pas dans les principales bases de données conventionnelles d'un SIGE et pourraient se rapporter à ce que ce guide appelle un « entrepôt de données » (voir l'explication de la Section 5.1.2).

6.2.7 Analyse des données et rapports

Les ensembles de données générés par le SIGE (et dans certains cas, par les entrepôts connexes) sont souvent des sources de données sous-exploitées. Il faut donc que le SIGE maximise les ensembles de données générés, en faisant la promotion des analyses de données et des pratiques de recherche. Son objectif ne vise pas seulement à recueillir, stocker et traiter l'information, il est aussi d'aider, par la fourniture d'informations pertinentes et accessibles, à l'élaboration et à la planification des politiques en matière d'éducation. Outre la production des rapports statistiques habituels faits de tableaux et de graphiques, le SIGE doit donc envisager des outils permettant de mettre au jour les tendances et les relations dans les ensembles de données qui sont produits. Les outils d'intelligence commerciale supportent ce type d'analyse de données et fournissent la fonctionnalité nécessaire.

Une façon d'accroître à la fois la rationalité de la prise de décision et la demande de données serait qu'une politique SIGE (ou une politique de données sur l'éducation en général) énonce les types et la gamme de décisions qui doivent être fondés sur la présentation d'éléments probants. Par ailleurs, les unités SIGE pourraient encourager la remise en question ou la défense de différentes positions adoptées par les groupes de réflexion et

la société civile à propos des bases de données. Un bon exemple serait de savoir si le décrochage ou le redoublement et les sureffectifs représentent le plus gros problème dans les petites classes, ou tout simplement quelle est la véritable ampleur du décrochage à l'école primaire (par rapport à un faible apprentissage ou au redoublement, ou aux deux, par exemple).

6.2.7.1 Écriture de requêtes et rapports

Les types de rapports suivants sont recommandés :

- **Rapports opérationnels** : les rapports sont exécutés périodiquement et directement distribués aux utilisateurs concernés.
- **Rapports en libre-service** : l'utilisateur exécute les rapports depuis une interface simple pilotée par des menus.
- **Rapports paramétriques** : les paramètres commandés par l'utilisateur aident celui-ci à adapter le rapport à des conditions spécifiques lors de l'exécution. Ces paramètres permettent à l'utilisateur de modifier le contenu et ils offrent une certaine flexibilité, moyennant un effort limité et malgré des connaissances techniques réduites de la part de l'utilisateur.
- **Requêtes ad hoc** : les requêtes ad hoc sont rédigées par l'utilisateur, puis stockées en vue d'une utilisation future.
- **Traitement analytique en ligne (OLAP)** : les utilisateurs peuvent analyser et subdiviser les données en éléments plus petits, dans plusieurs dimensions (district local, sexe, situation de richesse, etc.) afin de mieux comprendre les causes apparentes de problèmes tels qu'un faible taux d'achèvement. Grâce à cette fonctionnalité, les utilisateurs peuvent effectuer des calculs très complexes sur de grandes quantités de données. Il est également possible de télécharger et de traiter les données avec

des outils aussi simples qu'Excel ou aussi avancés que Stata ou R Shiny.

6.2.7.2 Publications

Le SIGE devrait optimiser les ensembles de données grâce à la promotion de rapports sur les données, l'interrogation des données, l'analyse de données et les pratiques de recherche. Le rapport annuel du SIGE sur les statistiques de l'éducation devrait être largement diffusé dans les 12 mois suivant la collecte des données, en suivant les recommandations suivantes :

- Les données du SIGE doivent être accessibles au niveau infranational.
- Le SIGE devrait publier un rapport au niveau infranational, dans le même format que son rapport national.
- Un personnel possédant des compétences et des aptitudes en écriture de requêtes doit être disponible au sein de la fonction SIGE, au niveau national et infranational, car on ne peut pas se fier uniquement à des « requêtes courantes » « toutes faites » ou prêtes à l'emploi ; des requêtes ad hoc sont souvent nécessaires.
- Veiller à ce que le personnel SIGE dispose des compétences fondamentales en bases de données et de la capacité d'extraire des données du système de base de données.

Il faut mettre à disposition les données du SIGE sur support électronique (sur Internet) dans un format de fichier convivial, tel qu'une base de données, une feuille de calcul, voire un fichier texte au format délimité par des virgules, à télécharger dans le respect des règles strictes de confidentialité et des procédures de « droit d'exploitation ».

6.2.8 Publication, diffusion et exploitation des données

Les données produites par le SIGE doivent être diffusées et partagées avec toutes les parties prenantes concernées, à tous les niveaux du système éducatif, afin d'appuyer la prise de décisions, l'analyse et la formulation de politiques, la planification, le suivi et la gestion.²⁵

6.2.8.1 Publication et diffusion des données

Conformément au calendrier discuté à la Section 6.2.1.5, le SIGE devrait :

- Diffuser et rendre accessibles (en particulier au niveau infranational) des statistiques pertinentes, de haute qualité, actuelles et accessibles répondant aux besoins des utilisateurs au niveau du gouvernement.
- Les données non diffusées régulièrement (micro-données et dispositifs alternatifs de publication) devraient être mises à disposition sur demande, conformément à un cadre juridique comprenant notamment des dispositions relatives à la confidentialité.
- Étant donné que la plupart des écoles ne sont pas intéressées par les données détaillées de l'ensemble du pays, une stratégie de diffusion imprimée pourrait être d'envoyer simplement à chaque district un document de trois pages « standard » (ex. un PDF) à distribuer à chaque école et comprenant les données essentielles (en particulier les données qui permettent de comparer les intrants et les extrants et de produire certains ratios clés) de l'école, le contexte de son district et le contexte national. Le district peut alors imprimer 300 pages (3 pages pour 100 écoles – ce n'est qu'un exemple) et les

25 Pour un exemple de plan de diffusion de données (du Royaume-Uni), voir <https://www.gov.uk/government/organisations/department-for-education/about/statistics#ad-hoc-statistics-and-data-releases> ou <https://tinyurl.com/y9cudbvm>.

distribuer. Ceci peut être particulièrement utile lorsque la connectivité Internet est faible ou que les écoles ont une faible capacité même pour exécuter des requêtes simples.

- Une stratégie de diffusion des données devrait exister pour :
 - Identifier les besoins en matière d'information du public cible ;
 - Déterminer le support de diffusion des différents types de documents (papier ou électronique) ;
 - Proposer un calendrier de publication des données sur papier et en version électronique ;
 - Identifier une personne ou une unité responsable de la publication des données ;
 - Déterminer le thème de la publication (ex. effectifs scolarisés, redoublement, toute question pertinente).²⁶

6.2.8.2 Exploitation des données

On a pu constater ces dernières années un accroissement considérable de la disponibilité, et dans certains cas aussi, de l'amélioration de la qualité des données disponibles pour la prise de décisions politiques et la recherche. Ce guide contient quelques suggestions concernant l'accessibilité des données sous une forme susceptible d'améliorer leur exploitation pour les politiques, l'analyse et la planification. Le SIGE devrait promouvoir activement l'exploitation des statistiques et consulter régulièrement les utilisateurs.

²⁶ Pour quelques idées et exemples, consultez les suggestions de la Banque mondiale dans l'article : <http://documents.worldbank.org/curated/en/155601468205458626/pdf/466170NWP0Box31Information0System111.pdf> ou <https://tinyurl.com/y9kmtapd>.

Six recommandations sont pertinentes dans ce contexte :

- Fournir des « requêtes communes » prêtes à l'emploi que l'utilisateur pourra adapter à ses propres besoins (ex. une école ou un district particulier).
- Accroître l'accessibilité des données annuelles de recensement scolaire en les mettant à disposition dans un format de fichier convivial, tel qu'une base de données, des feuilles de calcul ou même un fichier texte délimité par des virgules, à télécharger.
- Ne pas fournir aux utilisateurs plus de données que celles dont ils ont besoin ou qui les intéressent, pour éviter qu'ils ne se perdent dans trop de détails.
- Permettre à l'utilisateur de faire au moins des choix fondamentaux sur ce qu'il pourra télécharger en termes d'enregistrements et de regroupements d'enregistrements (écoles ou groupes scolaires, individuellement ou agrégés, par exemple toutes les écoles d'un district individuellement, ou le total ou la moyenne de toutes les écoles d'un district, autant que possible dans un tableau pivot Excel) ainsi que les champs (variables), là encore disponibles autant que possible dans un tableau pivot Excel.
- La diffusion des données et les rapports annuels de recensement scolaire devraient être accessibles en ligne, ce qui permettrait d'accroître l'exploitation des données si la majorité des utilisateurs est connectée à Internet.
- Mettre à disposition les données SIGE accompagnées d'une documentation telle que le Guide de l'utilisateur, les livres-codes et les fichiers de données.

Bibliographie

Abdul-Hamid, H. (2017). *Data for Learning: Building a Smart Education Data System*. Washington, DC : Groupe de la Banque mondiale. Ed-DQAF (UIS), UNESCO

Abdul-Hamid, H. (2014). « What matters most for Education Management Information Systems: a framework paper. « SABER Working Paper Series No. 7. Washington, DC : Banque mondiale

Union africaine et ADEA (2018). Stratégie continentale de l'éducation pour Afrique : Manuel d'indicateurs (en anglais). Disponible sur http://www.education2030-africa.org/images/systSIGE_edu/docs/Final_Indicators_Manual_compressed.pdf.

Community Systems Foundation (www.communitysystemsfoundation.org) et le personnel d'OpenEMIS ont fourni des informations complètes sur les fonctionnalités d'OpenEMIS

Dillman, D., J. Smyth et L.M. Christian (2014). *Internet, Phone, Mail and Mixed-Mode Surveys: The Tailored Design Method*. New Jersey : John Wiley & Sons Inc.

DHIS2. Documentation DHIS2. Extrait de : <https://www.dhis2.org/documentation>. Consulté en novembre 2019

DHIS2. Guide d'utilisation DHIS2. Extrait de : https://docs.dhis2.org/2.33/fr/user/dhis2_user_manual_en.pdf. Version anglaise consultée en novembre 2019

DHIS2. Plate-forme technologique DHIS2. Extrait de : <https://www.dhis2.org/technology> Consulté en novembre 2019

Gartner (2019). (Analystes) Howson C., J. Richardson, R. Sallam et A. Kronz. Magic quadrant for analytics and business intelligence platforms. Gartner. 11 février 2019

Gartner (2019). (Analystes) Ronthal A., R. Edjlali et R. Greenwald. Magic quadrant for data management solutions for analytics. Gartner. 21 janvier 2019

Becta (2006). Functional requirements for information management including MIS. Extrait de : http://archive.teachfind.com/becta/industry.becta.org.uk/content_files/industry/resources/Specification_key_docs/IMP_RequireSIGEnts_final_v1.doc?resID=20860&download_url=/content_files/industry/resources/Specification_key_docs/IMP_RequireSIGEnts_final_v1.doc. (Consulté en novembre 2019)

Heywood, A. et D. Boone (2015). *Guidelines for Data Management Standards in Routine Health Information Systems*. Measure Evaluation, Carolina Population Center, Chapel Hill, NC.

FHI 360. Notice d'information : <https://www.fhi360.org/sites/default/files/media/documents/resource-m360-factsheet.pdf>

Le personnel de FHI 360 a fourni des informations complètes sur l'architecture mobile FHI 360.

HISP-Afrique du Sud. Le personnel du Health Information Systems Program (HISP) a fourni des informations complètes sur les fonctionnalités du système DHIS2. <http://www.hisp.org>

New York State Student Identification System (NYSSIS) (2008). Disponible sur : <https://www.baruch.cuny.edu/step/documents/nyssisguide.pdf>

OCDE (2007). Glossaire des termes statistiques. OCDE : Paris

OpenEMIS. <https://www.openemis.org/>

OpenEMIS. GPE Open EMIS Questionnaire Response. Obtenu sur OpenEMIS. UNESCO.

OpenEMIS. Lesotho Openemis Pilot. Obtenu de l'équipe technique de la Community System Foundation (CSF)

OpenEMIS. Overview : Obtenu de l'équipe technique de la Community System Foundation (CSF)

Stair R. et G. Reynolds, (2018). *Principles of Information Systems*. Course Technology, Cengage Learning.

Subosa, M. et M. West (2018). « Re-orienting Education Management Information Systems (EMIS) towards inclusive and equitable quality education and lifelong learning » Working Paper on Education Policy, No. 5, UNESCO.

UNESCO (2019). Extrait de : <https://learningportal.iiep.unesco.org/en/glossary/educational-managSIGEnt-information-system-emis> (Consulté en novembre 2019)

Institut de statistique de l'UNESCO (UIS) (2019). Extrait de : <http://uis.unesco.org/en/capacity-development-tools> (Consulté en novembre 2019)

Van Wyk, C (2015). "An overview of key datasets in education in South Africa." South African Journal of Childhood Education, 5:2. Extrait de : <https://sajce.co.za/index.php/sajce/article/view/394>

Van Wyk, C. (2006). "The Development Of An Education Management Information System From A Sensemaking Perspective And The Application Of Quantitative Methods To Analyse Education Data Sets", Thèse de doctorat, non publiée. Université de Stellenbosch.

Wikipedia. « Logiciel en tant que service » Disponible sur : https://fr.wikipedia.org/wiki/Software_as_a_service (Consulté en décembre 2019)

Annexe A. Liste des contributeurs

Nous remercions toutes les personnes figurant dans cette liste qui ont contribué à la rédaction de sous-sections du présent rapport, à un examen par les pairs minutieux par écrit ou participé à des entretiens approfondis.

Plate-forme	Organisation	Nom
DHIS2	Université d'Oslo	Braa, Kristin
DHIS2	Université d'Oslo	Russpatrick, Scott
DHIS2	Université d'Oslo	Sanner, Terje Aksel
DHIS2	Université d'Oslo	Staring, Knut
DHIS2	HISP/Afrique du Sud	Claasen, Elmarie
N/A	Groupe de travail SIGE, ADEA	Maiga, Youssouf
N/A	PME	Irving, Margaret
N/A	PME	Martinez, Raphaelle
N/A	Microsoft	Joyce, Alexa
N/A	ISU	Boade, Georges
N/A	ISU	Labé, Olivier
N/A	ISU	Montoya, Silvia
N/A	ISU	Sigdel, Shailendra
N/A	Yet Analytics	Blake-Plock, Shelly
OpenEMIS	Community Systems Foundation	Kapp, Jon
OpenEMIS	Community Systems Foundation	Oswalt, Kris
OpenEMIS	OpenEMIS	Chapelete, Pierre
OpenEMIS	OpenEMIS	Turnbull, Karl
OpenEMIS	UNESCO	Bin Mahfooz, Sara
OpenEMIS	UNESCO	Chang, Gwang-Chol
OpenEMIS	UNESCO	Yano, Satoko
SABER-SIGE	Banque mondiale	Azevedo, Joao Pedro
StatEduc	ISU	Fall, Yacine

Par ailleurs, les collègues mentionnés ci-dessous ont participé à la revue par les pairs de ce guide oralement et/ ou en fournissant des conseils. Seules les personnes qui ne sont pas déjà citées ci-dessus sont nommées ici.

Nom	Organisation
Bah, Alpha	Gambie - SIGE
Chakroun, Borhene	UNESCO
Dobey, Edward	Ghana – SIGE
Memom, Asif	UNICEF
Myrick, Neal	Tableau Foundation
Ndao, Makha	ADEA
Palacios, Alex	PME
Reuge, Nicolas	UNICEF
Thompson-Quaye, Anna	PME

Annexe B. Liste des indicateurs de l'ODD 4 et considérations relatives au SIGE

Cette section met l'accent sur les façons d'améliorer et de modifier un SIGE au niveau de la collecte de données pour assurer le suivi de la mise en œuvre de l'ODD 4. Pour pouvoir répondre à l'accroissement des demandes, le SIGE doit fonctionner en tant que système intégré et interconnecté avec des ensembles de données provenant de sources diverses. Le [Tableau B.1](#) met en évidence les interactions de ces sources avec les données de l'épine dorsale du SIGE, données qui pourraient se trouver dans un entrepôt qui inclut l'épine dorsale et/ou les rapports généralement produits manuellement, et les indicateurs de l'ODD 4 qui ne sont pas réellement des « données ». La relation spécifique entre l'accent mis sur l'équité et l'égalité dans l'ODD 4, beaucoup plus forte que dans les OMD, est mise en évidence dans la ligne relative à l'indicateur ODD 4.5.1.

Tableau 4. Indicateurs de l'ODD 4 et leur traitement possible par le SIGE

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
<p>Cible 4.1 : D'ici à 2030, faire en sorte que toutes les filles et tous les garçons suivent, sur un pied d'égalité, un cycle complet d'enseignement primaire et secondaire gratuit et de qualité, qui débouche sur un apprentissage véritablement utile</p>					
4.1.1	<p>Pourcentage d'enfants et de jeunes : (a) en 2e ou 3e année d'études ; (b) en fin de cycle primaire ; et (c) en fin de premier cycle du secondaire qui maîtrisent au moins les normes d'aptitudes minimales en (i) lecture et (ii) mathématiques, par sexe</p>	Non	<p>Interopérabilité et intégration des données : cet indicateur exige que le SIGE soit compatible avec les systèmes d'évaluation. Données d'évaluation intégrées ou fusionnées avec les données SIGE, par le biais d'un champ commun (identifiant d'établissement scolaire).</p>	<p>Les données sur les résultats d'apprentissage provenant des évaluations nationales ou transnationales doivent utiliser le même identifiant d'établissement scolaire que le SIGE.</p>	Peut-être
4.1.2	<p>Organisation d'une évaluation de l'apprentissage représentative à l'échelle nationale (i) en 2e ou 3e année d'étude ; (ii) à la fin du cycle primaire et (iii) à la fin du premier cycle du secondaire</p>	Non		<p>Les données demandées sont sans rapport avec le SIGE : il s'agit de données qualitatives sur l'existence de telles évaluations et peut-être aussi sur leurs paramètres ou leurs métadonnées. Données sur l'administration d'une évaluation à grande échelle à partir d'un échantillon national représentatif émanant des bureaux nationaux d'évaluation de l'apprentissage, des ministères de l'éducation ou d'autres organismes chargés des évaluations de l'apprentissage, y compris les organisations régionales ou internationales qui réalisent des évaluations de l'apprentissage (ex : CONFEMEN, EQAP, IEA, OCDE, SACMEQ et LLECE).</p>	N/A

1 Ces données peuvent être numériques ou qualitatives. Si elles sont numériques, elles ne seront pas incluses dans une véritable base de données pouvant être reliée au SIGE ; elles peuvent se trouver dans une simple feuille de calcul Excel, ou même dans un fichier Word, par exemple.

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.1.3 Taux brut d'admission en dernière année d'études (niveau primaire, premier cycle du secondaire)	<p>Oui. Typique dans un SIGE agrégé (approche du recensement scolaire), mais il faudrait obtenir des données externes sur la population à partir des recensements ou d'autres sources.</p> <p>Doit être maintenu de façon longitudinale. La qualité des données sur les redoublants peut constituer une contrainte.</p> <p>Le suivi individuel des apprenants pourrait être utile ici.</p>	<p>Oui. Peut inclure des données sur la population pour le dénominateur.</p>	<p>Dans un système de collecte de données agrégées, les nouveaux entrants sont calculés en soustrayant du total des effectifs de la dernière année, le nombre d'élèves redoublant la dernière année.</p> <p>Dans un système qui capture les apprenants individuels, l'identifiant unique permet de suivre la progression d'un apprenant dans le système. Des chiffres précis sur le nombre total de redoublants et de décrocheurs peuvent être obtenus par les données longitudinales.</p>		<p>Oui. Difficile à automatiser entièrement.</p>
4.1.4 Taux d'achèvement (niveau primaire, premier cycle du secondaire, deuxième cycle du secondaire)	<p>Oui. Pour le numérateur.</p>	<p>Oui. Les données démographiques pour le dénominateur devraient provenir de données de recensement ou similaires.</p>			<p>Oui. Difficile à automatiser.</p>
4.1.5 Taux d'enfants non scolarisés (niveau primaire, premier cycle du secondaire, deuxième cycle du secondaire)	<p>Oui. Pour une partie du numérateur.</p>	<p>Oui. Les données démographiques pour le dénominateur proviennent généralement de données du recensement ou similaires. Les données de l'enquête auprès des ménages pourraient permettre la triangulation des données SIGE + recensement.</p>	<p>Interopérabilité : Données du SIGE à lier aux données de population par le biais d'un champ commun.</p>	<p>En rapport avec le SIGE : effectifs scolarisés pour chaque année d'âge dans chaque niveau d'éducation.</p> <p>Sans rapport avec le SIGE : estimations de la population par année d'âge.</p>	<p>Oui. Difficile à automatiser.</p>
4.1.6 Pourcentage des enfants ayant dépassé l'âge normal d'un niveau (cycle primaire, premier cycle du secondaire)	<p>Oui. Entièrement SIGE.</p>	<p>Oui. Pourrait contenir des données d'enquête auprès des ménages pour vérifier les schémas par âge.</p>			<p>Oui. Difficile à automatiser.</p>

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.1.7 Nombre d'années d'enseignement primaire et secondaire a) gratuit et b) obligatoire garanti dans les cadres juridiques	Non	Non		Il ne s'agit que d'une simple variable politique.	N/A
Cible 4.2 : D'ici à 2030, faire en sorte que toutes les filles et tous les garçons aient accès à des activités de développement et de soins de la petite enfance et à une éducation préscolaire de qualité qui les préparent à suivre un enseignement primaire					
4.2.1 Proportion d'enfants de moins de 5 ans dont le développement est en bonne voie en matière de santé, d'apprentissage et de bien-être psychosocial, par sexe	Non	Oui. Il faudrait disposer des données de l'enquête auprès des ménages ou d'autres données de l'entrepôt de données.	L'UNICEF travaille sur des cadres pour cette mesure et des méthodes devraient commencer à apparaître dans les enquêtes auprès des ménages à partir de 2021.		Oui
4.2.2 Taux de participation à des activités organisées d'apprentissage (un an avant l'âge officiel de scolarisation dans le primaire), par sexe	Oui	Oui. On peut avoir besoin de données de l'enquête auprès des ménages pour amplifier les données du SIGE. Données démographiques nécessaires pour le dénominateur.	Tout dépend de la définition des activités « organisées » d'apprentissage que le pays met en œuvre et qui peuvent être pleinement en accord ou non avec la conception de l'ODD. Cela pourrait également être amplifié via des rapports d'enquête auprès des ménages inspirés par l'UNICEF.		Oui
4.2.3 Pourcentage des enfants de moins de 5 ans vivant dans des environnements d'apprentissage familiaux positifs et stimulants	Non	Oui. Exigerait des données de l'enquête auprès des ménages.			Peut nécessiter des rapports manuels.

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.2.4 Taux brut de scolarisation de l'éducation de la petite enfance a) dans l'enseignement préscolaire et b) dans le développement de la petite enfance	Oui	Oui. Exigerait des données supplémentaires sur la population pour le dénominateur, et serait utile pour une triangulation avec les données de l'enquête auprès des ménages.			Peut nécessiter des rapports manuels.
4.2.5 Nombre d'années d'éducation préscolaire i) gratuite et ii) obligatoire garantie dans les cadres juridiques	Non	Non		Il ne s'agit que d'une simple variable politique.	N/A
Cible 4.3 : D'ici à 2030, faire en sorte que les femmes et les hommes aient tous accès dans des conditions d'égalité à un enseignement technique, professionnel ou tertiaire, y compris universitaire, de qualité et d'un coût abordable					
4.3.1 Taux de participation des jeunes et des adultes à un programme d'éducation et de formation scolaire ou non scolaire au cours des 12 mois précédents, par sexe	Oui	Oui, car le SIGE peut avoir du mal à recueillir ces données. Il serait peut-être nécessaire de les augmenter par les données des enquêtes auprès des ménages, du ministère du Travail, et d'autres sources de données moins habituelles.			Peut nécessiter des rapports manuels.
4.3.2 Taux brut de scolarisation de l'enseignement supérieur, par sexe	Oui	Dans de nombreux pays, ce type de données peut ne pas se situer dans l'épine dorsale du SIGE ; l'éducation tertiaire peut être dans un ministère différent possédant son propre système d'information (HEMIS). Les établissements d'enseignement supérieur ont tendance à être sur la défensive quant à leur autonomie. Il serait peut-être nécessaire de faire une triangulation avec les données de l'enquête auprès des ménages.			Peut nécessiter des rapports manuels.

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.3.3 Taux de participation aux programmes d'enseignement technique et professionnel (15 à 24 ans), par sexe	Oui	Oui, car le SIGE peut avoir du mal à recueillir ces données. Il serait peut-être nécessaire de l'augmenter par les données des enquêtes auprès des ménages, du ministère du Travail, et d'autres sources de données moins habituelles.			Peut nécessiter des rapports manuels.
Cible 4.4 : D'ici à 2030, augmenter considérablement le nombre de jeunes et d'adultes disposant des compétences, notamment techniques et professionnelles, nécessaires à l'emploi, à l'obtention d'un travail décent et à l'entrepreneuriat					
4.4.1 Proportion de jeunes et d'adultes ayant des compétences en informatique et en communication, par type de compétence	Non	Oui. Il faudrait l'augmenter par des données d'évaluation. Ce sera un domaine difficile à mesurer. L'UNESCO travaille sur des méthodes.	Interopérabilité et intégration des données : Le SIGE est compatible avec les systèmes d'évaluation. Les données d'évaluation intégrées ou fusionnées avec les données SIGE utilisent le cas échéant un champ commun (identifiant d'établissement scolaire) (lorsque les enquêtes sont réalisées au niveau établissement plutôt que sur le lieu de travail).	Les données sur les résultats d'apprentissage des évaluations nationales ou transnationales doivent utiliser le même identifiant d'établissement scolaire que le SIGE. Toutefois, les données peuvent provenir d'enquêtes auprès des ménages ou du lieu de travail.	Presque certainement
4.4.2 Pourcentage des jeunes/adultes qui maîtrisent au moins le niveau minimal d'aptitude en culture numérique	Non	Oui. Il faudrait augmenter par des données d'évaluation. Ce sera un domaine difficile à mesurer. L'UNESCO travaille sur des méthodes.	Interopérabilité et intégration des données : Le SIGE est compatible avec les systèmes d'évaluation. Les données d'évaluation intégrées ou fusionnées avec les données SIGE utilisent le cas échéant un champ commun (identifiant d'établissement scolaire) (c'est-à-dire lorsque les enquêtes sont réalisées au niveau établissement plutôt que sur le lieu de travail).	Les données sur les résultats d'apprentissage des évaluations nationales ou transnationales doivent utiliser le même identifiant d'établissement scolaire que le SIGE. Toutefois, les données peuvent provenir d'enquêtes auprès des ménages ou du lieu de travail.	Presque certainement

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.4.3 Taux de réussite scolaire des jeunes/adultes par tranches d'âge, situation au regard de l'activité économique, niveaux d'études et orientation du programme	Non	Oui. Il faudrait l'augmenter par les données de la population (recensement) ou des enquêtes auprès des ménages, généralement recueillies par les instituts nationaux de statistique ou les ministères du travail.			Oui
Cible 4.5 : D'ici à 2030, éliminer les inégalités entre les sexes dans le domaine de l'éducation et assurer l'égalité d'accès des personnes vulnérables, y compris les personnes handicapées, les autochtones et les enfants en situation vulnérable, à tous les niveaux d'enseignement et de formation professionnelle					
4.5.1 Indices de parité (femmes/hommes, urbain/rural, quintile inférieur/supérieur de richesse et autres paramètres tels que le handicap, le statut d'autochtone et les situations de conflit, à mesure que les données deviennent disponibles) pour tous les indicateurs dans le domaine de l'éducation de cette liste pouvant être ventilés	Non. Les indices eux-mêmes ne seraient pas produits normalement par le SIGE ou ne feraient pas partie de son épine dorsale.	Oui. Ce serait bon à avoir dans un entrepôt. Il faudrait des données externes comme des données du recensement, des enquêtes auprès des ménages ou de sources de données innovantes telles que les données sur les citoyens.	Interopérabilité et intégration des données : noter que c'est la grande différence entre les ODD et les OMD. Cela mis à part, les pays eux-mêmes doivent s'intéresser aux inégalités. C'est la nécessité de mettre en corrélation des questions telles que l'accès, l'achèvement, etc., la situation géographique urbain/rural, la richesse, etc. qui fait toute l'importance des fonctions d'entrepôt et d'interopérabilité d'un SIGE moderne.		Exigera très certainement beaucoup de saisie manuelle.
4.5.2 Pourcentage d'élèves dans l'enseignement primaire dont la première langue ou la langue parlée à la maison est la langue d'instruction	Non	Oui. Enquêtes spéciales auprès des ménages ou des écoles.			Oui

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
Mesure dans laquelle les stratégies explicites basées sur une formule réaffectent les ressources de l'éducation aux populations défavorisées	Non	Non	Il s'agit d'une variable de politique complexe qui nécessite des informations qualitatives.		Exigerait un rapport manuel
4.5.3					
Dépenses d'éducation par élève, par niveau d'enseignement et source de financement	Non	Non	Il s'agit d'une variable de politique complexe qui nécessite des informations fournies par le ministère de l'Éducation ou le ministère des Finances, ou les deux.		Exigerait un rapport manuel
4.5.4					
Pourcentage de l'aide totale à l'éducation allouée aux pays les moins avancés	Non	Non	Il s'agit d'une variable complexe applicable uniquement au niveau mondial.		
4.5.5					
Cible 4.6 : D'ici à 2030, veiller à ce que tous les jeunes et une proportion considérable d'adultes, hommes et femmes, sachent lire, écrire et compter					
Pourcentage de la population d'un groupe d'âge donné ayant les compétences voulues à au moins un niveau d'aptitude fixé a) en alphabétisme et b) numératie fonctionnels, par sexe	Non	Oui. Il faudrait augmenter les données par les données d'évaluation et des enquêtes auprès des ménages ou sur le lieu de travail. Ce sera un domaine difficile à mesurer. L'UNESCO travaille sur des méthodes.			Presque certainement
4.6.1					
Taux d'alphabétisme des adultes et des jeunes	Non	Oui. Il faudrait augmenter les données par les données d'évaluation et des enquêtes auprès des ménages ou sur le lieu de travail. Ce sera un domaine difficile à mesurer. L'UNESCO travaille sur des méthodes.			Presque certainement
4.6.2					

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données!	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.6.3 Taux de participation des jeunes et des adultes aux programmes d'alphabétisation	Peut-être, cela dépend dans quelle mesure le SIGE enquête sur les centres d'apprentissage des adultes ou des jeunes. Ceux-ci sont parfois non formels et proposés dans ou par des organisations confessionnelles, des clubs communautaires, etc., de sorte que cela peut être difficile.	Oui. En raison des difficultés rencontrées dans la collecte de données par des moyens administratifs traditionnels, il peut être nécessaire d'utiliser les enquêtes auprès des ménages.		Très probablement	
Cible 4.7 : D'ici à 2030, faire en sorte que tous les élèves acquièrent les connaissances et compétences nécessaires pour promouvoir le développement durable					
4.7.1 Degré d'intégration de i) l'éducation à la citoyenneté mondiale et ii) l'éducation au développement durable, y compris l'égalité des sexes et le respect des droits de l'homme, dans a) les politiques nationales d'éducation, b) les programmes d'enseignement, c) la formation des enseignants et d) l'évaluation des étudiants	Non	Non, il s'agit d'une variable de politique à mesurer par le biais d'informateurs clés.		Oui, mais le rapport est surtout descriptif, moins numérique.	
4.7.2 Pourcentage d'écoles qui dispensent une éducation au VIH et à la sexualité fondée sur les compétences utiles pour la vie	Ce n'est pas traditionnel, mais pourrait être inclus dans l'enquête administrative.	Peut-être, mais nécessiterait une enquête auprès des établissements scolaires.		Oui. Difficile à automatiser.	

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.7.3	Degré auquel le cadre du Programme mondial en faveur de l'éducation aux droits de l'homme est mis en œuvre au niveau national (conformément à la résolution 59/113 de l'Assemblée générale des Nations Unies)	Non	Non, il s'agit d'une variable de politique à mesurer par le biais d'informateurs clés.		Oui, mais le rapport est surtout descriptif, moins numérique.
4.7.4	Pourcentage d'étudiants par groupe d'âge (ou d'un niveau d'études donné) montrant une compréhension adéquate des questions relatives à la citoyenneté mondiale et à la durabilité	Non	Oui. Il faudrait augmenter les données par les données d'évaluation et des enquêtes auprès des ménages ou sur le lieu de travail. Ce sera un domaine difficile à mesurer. L'UNESCO travaille sur des méthodes.		Presque certainement
4.7.5	Pourcentage d'étudiants de 15 ans montrant une maîtrise des connaissances de la science environnementale et de la géoscience	Non	Oui. Il faudrait augmenter les données par les données d'évaluation et des enquêtes auprès des ménages ou sur le lieu de travail. Ce sera un domaine difficile à mesurer. L'UNESCO travaille sur des méthodes.		Presque certainement

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
<p>Cible 4.a : Faire construire des établissements scolaires qui soient adaptés aux enfants, aux personnes handicapées et aux deux sexes ou adapter les établissements existants à cette fin et fournir un cadre d'apprentissage effectif qui soit sûr, exempt de violence et accessible à tous</p>					
<p>4.a.1 – Proportion d'établissements scolaires ayant accès à : a) l'électricité ; b) Internet à des fins pédagogiques ; c) des ordinateurs à des fins pédagogiques ; d) des infrastructures et des matériels adaptés aux élèves handicapés ; e) une alimentation de base en eau potable ; f) des installations sanitaires de base séparées pour hommes et femmes ; et g) des équipements de base pour le lavage des mains</p>	<p>Oui, dans certains pays. Cela pourrait être ajouté à l'enquête administrative du SIGE, mais peut-être pas chaque année. On pourrait fournir une enquête cyclique sur les variables (de sorte que chaque variable serait traitée tous les trois ans, par exemple).</p>	<p>Oui, s'il y a eu une enquête distincte sur les installations scolaires, qui ne fait pas partie du SIGE traditionnel.</p>		<p>Presque certainement</p>	
<p>4.a.2</p>	<p>Pourcentage d'élèves victimes de brimades, de harcèlement, de violence, de discrimination ou d'abus sexuels</p>	<p>Oui, dans certains pays. Cela pourrait être ajouté à l'enquête administrative du SIGE, peut-être pas chaque année. On pourrait fournir une enquête cyclique sur les variables (de sorte que chaque variable serait traitée tous les trois ans, par exemple).</p>	<p>Oui, s'il y a eu une enquête distincte sur le climat dans les écoles qui ne fait pas partie du SIGE traditionnel. Ou cela pourrait faire partie d'une enquête auprès des ménages, comme MICS ou DHS ou d'autres.</p>		<p>Presque certainement</p>

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.a.3 Nombre d'attaques sur les élèves, le personnel et les établissements	Oui, dans certains pays. Cela pourrait être ajouté à l'enquête administrative du SIGE, peut-être pas chaque année. On pourrait fournir une enquête cyclique sur les variables (de sorte que chaque variable serait traitée tous les trois ans, par exemple).	Oui, s'il y a eu une enquête distincte sur le climat dans les écoles qui ne fait pas partie du SIGE traditionnel. Ou cela pourrait faire partie d'une enquête auprès des ménages comme MICS ou DHS ou d'autres.			Presque certainement
Cible 4.b : D'ici à 2020, augmenter considérablement à l'échelle mondiale le nombre de bourses d'études offertes aux pays en développement, en particulier aux pays les moins avancés, aux petits États insulaires en développement et aux pays d'Afrique, pour financer le suivi d'études supérieures, y compris la formation professionnelle, les cursus informatiques, techniques et scientifiques et les études d'ingénieur, dans des pays développés et d'autres pays en développement					
4.b.1 Volume de l'aide publique au développement consacrée aux bourses d'études, par secteur et type de formation	Non	Non	Il faudrait probablement collecter ces données manuellement auprès du ministère des Finances du pays ou des unités de coordination budgétaire ou des donateurs du ministère de l'Éducation.		Oui
4.b.2 Nombre de bourses d'études de l'enseignement supérieur accordées par pays bénéficiaire	Non	Non	Il faudrait probablement que ces données soient collectées par un bureau spécialisé au ministère de l'Éducation ou des ambassades locales d'autres pays.		Oui

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
<p>Cible 4.c : D'ici à 2030, accroître considérablement le nombre d'enseignants qualifiés, notamment au moyen de la coopération internationale pour la formation d'enseignants dans les pays en développement, surtout dans les pays les moins avancés et les petits États insulaires en développement</p>					
<p>Proportion d'enseignants dans : a) le préscolaire ; b) le cycle primaire ; c) le premier cycle du secondaire ; et d) le deuxième cycle du secondaire qui ont suivi (avant leur entrée en fonction ou en cours d'activité) au moins les formations minimales organisées à leur intention (notamment les formations pédagogiques) qui sont requises pour pouvoir enseigner au niveau pertinent dans un pays donné</p> <p>4.c.1</p>	<p>Oui, dans de nombreux pays, mais pas tous.</p>	<p>Dans certains pays, les données de l'épine dorsale pourraient être complétées dans un entrepôt par des données provenant des systèmes RH ou de paie et placées dans un entrepôt de données.</p>	<p>Interopérabilité : Données du SIGE à lier aux données RH ou de paie au moyen d'un champ commun.</p>	<p>Presque certainement</p>	
<p>Ratio élèves/enseignant formé par niveau d'études par type de formation</p> <p>4.c.2</p>	<p>Oui, dans de nombreux pays, mais pas tous.</p>	<p>Dans certains pays, les données de l'épine dorsale pourraient être complétées dans un entrepôt par des données provenant des systèmes RH ou de paie et placées dans un entrepôt de données.</p>	<p>Interopérabilité : Données du SIGE à lier aux données RH ou de paie au moyen d'un champ commun.</p>	<p>Presque certainement</p>	
<p>Pourcentage d'enseignants qualifiés selon les normes nationales par niveau d'études et type d'établissement</p> <p>4.c.3</p>	<p>Oui, dans de nombreux pays, mais pas tous.</p>	<p>Dans certains pays, les données de l'épine dorsale pourraient être complétées dans un entrepôt par des données provenant des systèmes RH ou de paie et placées dans un entrepôt de données.</p>	<p>Interopérabilité : Données du SIGE à lier aux données RH ou de paie au moyen d'un champ commun.</p>	<p>Presque certainement</p>	

Indicateur	Épine dorsale centrale du SIGE	Ne figure pas habituellement dans l'épine dorsale du SIGE, mais plutôt dans un entrepôt de données lié au SIGE	Commentaires	Sans aucun rapport avec le SIGE ou même avec toute base de données ¹	Calculé à partir du SIGE et de données similaires, mais en tant que sortie manuelle
4.c.4 Ratio élèves/enseignant qualifié par niveau d'études	Oui, dans de nombreux pays, mais pas tous.	Dans certains pays, les données de l'épine dorsale pourraient être complétées dans un entrepôt par des données provenant des systèmes RH ou de paie et placées dans un entrepôt de données.	Interopérabilité : Données du SIGE à lier aux données RH ou de paie au moyen d'un champ commun.	Presque certainement	
4.c.5 Salaire moyen des enseignants par rapport aux autres professions exigeant un niveau comparable de qualification	Non	Non	Il s'agit d'une variable de politique complexe qui est généralement dérivée des enquêtes sur les revenus et les dépenses des ménages. Les échelles salariales peuvent également être utiles si des informations empiriques ne sont pas disponibles.	Oui	
4.c.6 Taux d'attrition des enseignants par niveau d'études	Pas habituellement car la plupart des SIGE ne suivent pas, jusqu'à présent, les enseignants individuels.	Peut-être. On pourrait avoir un entrepôt de données sur la paie ou les RH, permettant de calculer les différences annuelles.		Oui	
4.c.7 Pourcentage d'enseignants qui ont bénéficié de la formation continue au cours des 12 derniers mois, par type de formation	Pas habituellement demandé dans la plupart des SIGE, mais ce serait possible.	Peut-être. Cela pourrait être une base de données de la formation fournie par les services pertinents du ministère de l'Éducation ou d'un ministère de l'enseignement supérieur. Cependant, une grande partie de cette formation est dispensée par la société civile ou les enseignants s'y inscrivent simplement eux-mêmes. Ainsi, l'ajouter en tant que question dans l'épine dorsale du SIGE serait le plus logique.		Oui	

Annexe C. Exemple de calendrier de collecte SIGE

Le calendrier de collecte des données SIGE couvre toutes les étapes du processus de la collecte et il doit être diffusé à l'ensemble des acteurs pertinents. Voici un exemple de l'Ohio, Etats-Unis.¹

¹ Notez que ce calendrier SIGE est probablement beaucoup plus élaboré que ce qui serait adapté à des pays à revenu faible ou à revenu intermédiaire de la tranche inférieure. Extrait de : <http://education.ohio.gov/getattachment/Topics/Data/EMIS/Reporting-Responsibilities/EMIS-Data-Collection-Calendar/Processing-Schedule-FY19-v0-3-4-29-19.pdf.aspx?lang=en-US> ou <https://tinyurl.com/y9rqmfwf>.

Calendrier de collecte de données SIGE pour 2018-19			
Par grands groupes de données	Ensemble de données	Date d'ouverture	Date de clôture
Principales séries relatives aux élèves			
Table de référence des élèves (FY19) (2019SSCRS)	S	7/6/2018	7/17/2019
Rapport sur le taux de rétention, toutes les années (2019SRTNT)	S	7/27/2018	8/31/2018
SOES (Système d'inscription aux options scolaires), Collecte élèves, début d'année (FY19) (2019SAODE)	S	8/3/2018	12/20/2018
SOES, Collecte contact(s) élèves (FY19) (2019SSSDT)	S	8/3/2018	8/9/2019
Collecte élèves, début d'année (FY19) (2019S1TRD)	S	9/4/2018	12/20/2018
Collecte élèves, milieu d'année (FY19) (2019S2TRD)	S	1/4/2019	4/30/2019
SOES, Collecte élèves, fin d'année (FY19) (2019SBODE)	S	1/4/2019	7/17/2019
Collecte élèves, fin d'année (FY19) (2019S3TRD)	S	5/3/2019	7/17/2019
Collectes supplémentaires, élèves et personnel, Calendrier			
Calendrier de collecte - initial (FY19) (2019CINIT)	C	7/10/2018	9/28/2018
Calendrier de collecte - final (FY19) (2019CFINL)	C	10/2/2018	8/9/2019
Collecte initiale, personnel et cours (FY19) (2019L1STR)	L	10/2/2018	2/4/2019
Collecte de suivi, mars (FY19) (2019D0000)	D	2/1/2019	5/10/2019

Collecte finale, personnel et cours (FY19) (2019L2FNL)	L	2/5/2019	8/9/2019
Collecte, diplômés FY19 (2019G0000)	G	5/3/2019	10/18/2019
Collecte des évaluations			
Collecte, Évaluation fin de cursus été et automne (FY19) (2019AGEFL)	A	9/14/2018	3/29/2019
Collecte, Évaluation OGT (Ohio Graduation Test), été FY19 (2019AGXSM)	A	9/14/2018	11/30/2018
Collecte, Évaluation apprentissage de la petite enfance, automne FY19 (2019AGBFL)	A	9/28/2018	3/1/2019
Collecte, Évaluation niveau de préparation à la maternelle FY19 (2019AG0FL)	A	9/28/2018	2/1/2019
Collecte, Évaluation sommaire des résultats d'apprentissage des enfants FY19 (2019AGMFY)	A	10/15/2018	7/12/2019
Collecte, Évaluation lecture 3e année du primaire, automne FY19 (2019AGNFL)	A	10/23/2018	3/1/2019
Collecte, Évaluation DORP (Prévention du décrochage) automne FY19 (2019AGDFL)	A	10/26/2018	7/8/2019
Collecte, Évaluation OGT (Ohio Graduation Test) FY19 (2019AGXFY)	A	1/8/2019	8/9/2019
Autres évaluations de responsabilisation FY19 (2019ANACC)	A	2/5/2019	8/14/2019
Collecte, Évaluation de remplacement, printemps FY19 (2019AALTS)	A	2/5/2019	8/9/2019
Collecte, Évaluation DORP (Prévention du décrochage, printemps FY19) (2019AGDSP)	A	1/8/2019	7/8/2019
Collecte, Évaluation OELPA (évaluation de la maîtrise de l'anglais) FY19 (2019AGFSP)	A	2/22/2019	8/9/2019
Collecte, Évaluation au niveau de l'État, fin de cursus, printemps FY19 (2019AGESP)	A	3/1/2019	7/17/2019
Collecte, Évaluation, apprentissage de la petite enfance, printemps FY19 (2019AGBSP)	A	4/9/2019	8/9/2019
Collecte, Évaluation de responsabilisation ETP FY19 (2019ACTSP)	A	3/29/2019	10/11/2019
Collecte des données financières			
Collecte, données financières FY19 (2019H0000)	H	6/3/2019	8/30/2019
Collecte supplémentaire, données financières FY19 (2019H0000)	H	9/3/2019	9/27/2019
Prévisions à 5 ans - initial requis (FY19) (2019P0000)	P	10/1/2018	10/31/2018
Prévisions à 5 ans - initial optionnel (FY19) (2019P1OPT)	P	11/6/2018	3/29/2019

Prévisions à 5 ans - mise à jour du requis, printemps (FY19) (2019P2MAY)	P	4/2/2019	5/31/2019
Prévisions à 5 ans - final optionnel (FY19) (2019P3OPT)	P	6/4/2019	8/9/2019

Les cellules ombrées indiquent un changement de date ou toute autre modification de la version antérieure du calendrier

Version 0.3

Dernière mise à jour 29/04/2019

Calendrier de collecte de données SIGE pour 2018-19			
Par date d'ouverture	Ensemble de données	Date d'ouverture	Date de clôture
Tableau de référence des élèves (FY19) (2019SSCRS)	S	7/6/2018	7/17/2019
Calendrier de collecte - initial (FY19) (2019CINIT)	C	7/10/2018	9/28/2018
Rapport sur le taux de rétention, toutes les années (2019SRTNT)	S	7/27/2018	8/31/2018
SOES (Système d'inscription aux options scolaires), Collecte élèves, début d'année (FY19) (2019SAODE)	S	8/3/2018	12/20/2018
SOES, Collecte contact(s) élèves (FY19) (2019SSSDT)	S	8/3/2018	8/9/2019
Collecte élèves, début d'année (FY19) (2019S1TRD)	S	9/4/2018	12/20/2018
Collecte, Évaluation fin de cursus été et automne (FY19) (2019AGEFL)	A	9/14/2018	3/29/2019
Collecte, Évaluation OGT (Ohio Graduation Test), été FY19 (2019AGXSM)	A	9/14/2018	11/30/2018
Collecte, Évaluation, apprentissage de la petite enfance, automne FY19 (2019AGBFL)	A	9/28/2018	3/1/2019
Collecte, Évaluation, niveau de préparation à la maternelle FY19 (2019AGOFL)	A	9/28/2018	2/1/2019
Prévisions à 5 ans - initial requis (FY19) (2019P0000)	P	10/1/2018	10/31/2018
Calendrier de collecte - final (FY19) (2019CFINL)	C	10/2/2018	8/9/2019
Collecte initiale, personnel et cours (FY19) (2019L1STR)	L	10/2/2018	2/4/2019
Collecte, Évaluation sommaire des résultats d'apprentissage des enfants FY19 (2019AGMFY)	A	10/15/2018	7/12/2019
Collecte, Évaluation lecture 3e année du primaire, automne FY19 (2019AGNFL)	A	10/23/2018	3/1/2019

Collecte, Évaluation DORP (Prévention du décrochage), automne FY19 (2019AGDFL)	A	10/26/2018	7/8/2019
Prévisions à 5 ans - initial optionnel (FY19) (2019P1OPT)	P	11/6/2018	3/29/2019
Collecte élèves, milieu d'année (FY19) (2019S2TRD)	S	1/4/2019	4/30/2019
SOES (Système d'inscription aux options scolaires), Collecte élèves, fin d'année (FY19) (2019SBODE)	S	1/4/2019	7/17/2019
Collecte, Évaluation OGT (Ohio Graduation Test) FY19 (2019AGXFY)	A	1/8/2019	8/9/2019
Collecte de suivi, mars (FY19) (2019D0000)	D	2/1/2019	5/10/2019
Collecte finale, personnel et cours (FY19) (2019L2FNL)	L	2/5/2019	8/9/2019
Autres évaluations de responsabilisation FY19 (2019ANACC)	A	2/5/2019	8/14/2019
Collecte, Évaluation de remplacement, printemps FY19 (2019AALTS)	Un	2/5/2019	8/9/2019
Collecte, Évaluation DORP (Prévention du décrochage), printemps FY19 (2019AGDSP)	A	2/8/2019	7/8/2019
Collecte, Évaluation OELPA (évaluation de la maîtrise de l'anglais) FY19 (2019AGFSP)	A	2/22/2019	8/9/2019
Collecte, Évaluation au niveau de l'État, fin de cursus, printemps FY19 (2019AGESP)	A	3/1/2019	7/17/2019
Collecte, Évaluations années 3-8 au niveau de l'État, FY19 (2019AGNSP)	A	3/1/2019	7/17/2019
Collecte, Évaluation de responsabilisation Career Tech (ETP) FY19 (2019ACTSP)	A	3/29/2019	10/11/2019
Prévisions à 5 ans - mise à jour du requis, printemps (FY19) (2019P2MAY)	P	4/2/2019	5/31/2019
Collecte, Évaluation, apprentissage de la petite enfance, printemps FY19 (2019AGBSP)	A	4/9/2019	8/9/2019
Collecte élèves, fin d'année (FY19) (2019S3TRD)	S	5/3/2019	7/17/2019
Collecte, diplômés FY19 (2019G0000)	G	5/3/2019	10/18/2019
Collecte, données financières FY19 (2019H0000)	H	6/3/2019	8/30/2019
Prévisions à 5 ans - final optionnel (FY19) (2019P3OPT)	P	6/4/2019	8/9/2019
Collecte supplémentaire, données financières FY19 (2019H0000)	H	9/3/2019	9/27/2019

Les cellules ombrées indiquent un changement de date ou toute autre modification de la version antérieure du calendrier

Version 0.3

Dernière mise à jour 29/04/2019

Calendrier de collecte de données SIGE pour 2018-19			
Par date de clôture	Ensemble de données	Date d'ouverture	Date de clôture
Rapport sur le taux de rétention, toutes les années (2019SRTNT)	S	7/27/2018	8/31/2018
Calendrier de collecte - initial (FY19) (2019CINIT)	C	7/10/2018	9/28/2018
Prévisions à 5 ans - initial requis (FY19) (2019P0000)	P	10/1/2018	10/31/2018
Collecte, Évaluation OGT (Ohio Graduation Test), été FY19 (2019AGXSM)	A	9/14/2018	11/30/2018
SOES (Système d'inscription aux options scolaires), Collecte élèves, début d'année (FY19) (2019SAODE)	S	8/3/2018	12/20/2018
Collecte élèves, début d'année (FY19) (2019S1TRD)	S	9/4/2018	12/20/2018
Collecte, Évaluation, niveau de préparation à la maternelle FY19 (2019AGOFL)	A	9/28/2018	2/1/2019
Collecte initiale, personnel et cours (FY19) (2019L1STR)	L	10/2/2018	2/4/2019
Collecte, Évaluation, apprentissage de la petite enfance, automne FY19 (2019AGBFL)	A	9/28/2018	3/1/2019
Collecte, Évaluation lecture 3e année du primaire, automne FY19 (2019AGNFL)	A	10/23/2018	3/1/2019
Collecte, Évaluation, fin de cursus été et automne (FY19) (2019AGEFL)	A	9/14/2018	3/29/2019
Prévisions à 5 ans - initial optionnel (FY19) (2019P1OPT)	P	11/6/2018	3/29/2019
Collecte élèves, milieu d'année (FY19) (2019S2TRD)	S	1/4/2019	4/30/2019
Collecte de suivi, mars (FY19) (2019D0000)	D	2/1/2019	5/10/2019
Prévisions à 5 ans - mise à jour du requis, printemps (FY19) (2019P2MAY)	P	4/2/2019	5/31/2019
Collecte, Évaluation DORP (prévention du décrochage) automne FY19 (2019AGDFL)	A	10/26/2018	7/8/2019
Collecte, Évaluation DORP, printemps FY19 (2019AGDSP)	A	2/8/2019	7/8/2019
Collecte, Évaluation sommaire des résultats d'apprentissage des enfants FY19 (2019AGMFY)	A	10/15/2018	7/12/2019
Tableau de référence des élèves (FY19) (2019SSCRS)	S	7/6/2018	7/17/2019
SOES, Collecte élèves, fin d'année (FY19) (2019SBODE)	S	1/4/2019	7/17/2019
Collecte, Évaluation au niveau de l'État, fin de cursus, printemps FY19 (2019AGESP)	A	3/1/2019	7/17/2019
Collecte, Évaluations années 3-8 au niveau de l'État, FY19 (2019AGNSP)	A	3/1/2019	7/17/2019
Collecte élèves, fin d'année (FY19) (2019S3TRD)	S	5/3/2019	7/17/2019

SOES, Collecte contact(s) élèves (FY19) (2019SSSDT)	S	8/3/2018	8/9/2019
Calendrier de collecte - Final (FY19) (2019CFINL)	C	10/2/2018	8/9/2019
Collecte, Évaluation OGT (Ohio Graduation Test) FY19 (2019AGXFY)	A	1/8/2019	8/9/2019
Collecte finale, personnel et cours (FY19) (2019L2FNL)	L	2/5/2019	8/9/2019
Collecte, Évaluation de remplacement, printemps FY19 (2019AALTS)	A	2/5/2019	8/9/2019
Collecte, Évaluation OELPA (évaluation de la maîtrise de l'anglais) FY19 (2019AGFSP)	A	2/22/2019	8/9/2019
Collecte, Évaluation, apprentissage de la petite enfance, printemps FY19 (2019AGBSP)	A	4/9/2019	8/9/2019
Prévisions à 5 ans - final optionnel (FY19) (2019P3OPT)	P	6/4/2019	8/9/2019
Autres évaluations de responsabilisation FY19 (2019ANACC)	A	2/5/2019	8/14/2019
Collecte, données financières FY19 (2019H0000)	H	6/3/2019	8/30/2019
Collecte supplémentaire, données financières FY19 (2019H0000)	H	9/3/2019	9/27/2019
Collecte, Évaluation de responsabilisation ETP (FY19) (2019ACTSP)	A	3/29/2019	10/11/2019
Collecte, diplômés FY19 (2019G0000)	G	5/3/2019	10/18/2019

Les cellules ombrées indiquent un changement de date ou toute autre modification de la version antérieure du calendrier

Version 0.3

Dernière mise à jour 29/04/2019

Systemes d'information pour la gestion de l'education

Efficience et efficacite

dans le choix et l'utilisation d'un SIGE

Les Recommandations concernant la gestion des donnees et les fonctionnalites des Systemes d'information pour la gestion de l'education (SIGE) ont ete redigees en partenariat avec l'Institut de statistique de l'UNESCO (ISU) et le Partenariat Mondial pour l'Education (PME).

Elles decoulent directement d'une demande de conseils exprimee par les pays en developpement sur les fonctions que devrait ideallement proposer un SIGE et la maniere de mieux les utiliser. Lors d'une reunion entre les representants des pays et l'UNESCO, le PME et d'autres partenaires ont estime que dans la mesure ou il n'existait aucun logiciel SIGE par defaut standard produit par la communaute internationale et qu'il n'etait sans doute pas souhaitable d'en elaborer un, on pourrait au moins fournir un guide sur les fonctions souhaitables. Les pays ont souvent congu leurs propres systemes ou ont adopte – partiellement ou en totalite – des systemes proposes par differents fournisseurs. De plus, en l'absence de normes ou de lignes directrices, les pays avaient le sentiment de ne pas savoir ce qu'ils devaient negocier ou discuter en matiere de soutien financier ou technique lorsqu'ils discutaient avec les organismes de developpement.

Les recommandations ont pour but d'aider les pays a choisir et a utiliser de maniere efficace une approche SIGE. Autrement dit, elles visent a s'assurer que de tels systemes soient aussi complets que possible, tout en contenant le moins d'elements redondants possible. En outre, elles insistent sur le fait que les systemes devraient faciliter la fourniture d'information a valeur ajoutee aux decideurs politiques et aux gestionnaires en maximisant l'interopabilite et l'interdependance entre les differentes bases de donnees gerees par le secteur educatif. Par exemple, permettre l'identification des ecoles susceptibles d'etre moins performantes au vu de leur contexte ou celles qui obtiendraient de meilleures performances que prevu, et fournir les enseignements tires. Enfin, les partenaires ont demande que les lignes directrices donnent une vue globale de ce que devrait etre un systeme SIGE qui fonctionne bien, au regard des correspondances avec les objectifs declares du secteur educatif ainsi qu'avec les flux et les liens entre les elements de donnees (inscriptions, cout, personnel, etc.) du secteur. Les lignes directrices s'inscrivent dans le contexte du quatrieme Objectif de developpement durable pour l'education (ODD 4), sans pour autant supposer que les rapports mondiaux devaient etre l'objectif principal des systemes SIGE. Certains pays ont egalement demande des informations et des lignes directrices sur la maniere de mettre en oeuvre des systemes d'identification des apprenants et des enseignants en vue d'ameliorer le suivi des apprenants qui abandonnent et donc l'achevement scolaire. Ces lignes directrices s'efforcent de remplir les missions ci-dessus mentionnees. L'ISU et le PME espent que ces lignes directrices seront utiles aux pays et aux partenaires et qu'elles repondront a leurs demandes.