

What data is needed for improving automated translation?

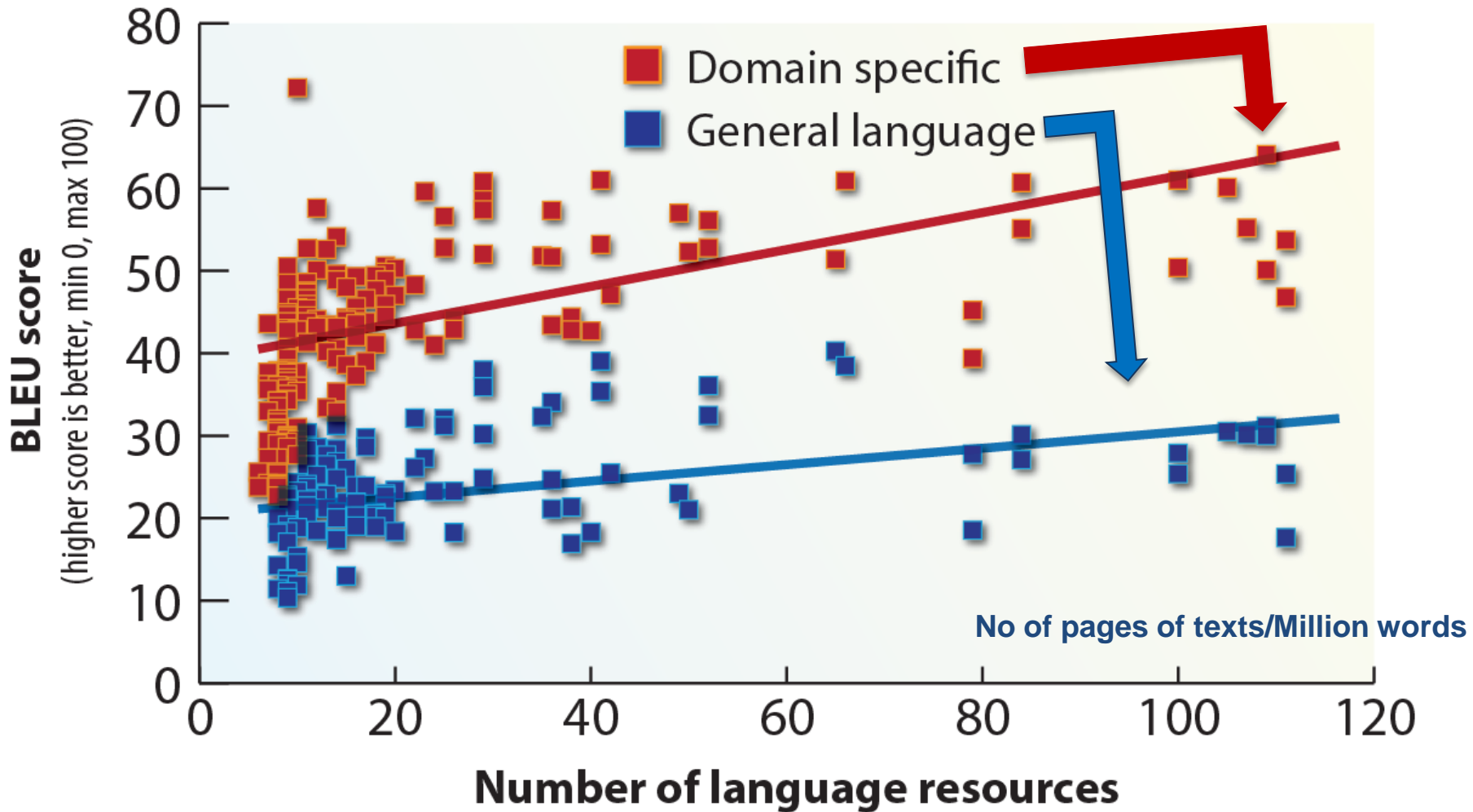
Dr. Andrejs Vasiļjevs, ELRC/Tilde

Credits:

Khalid Choukri, ELRA

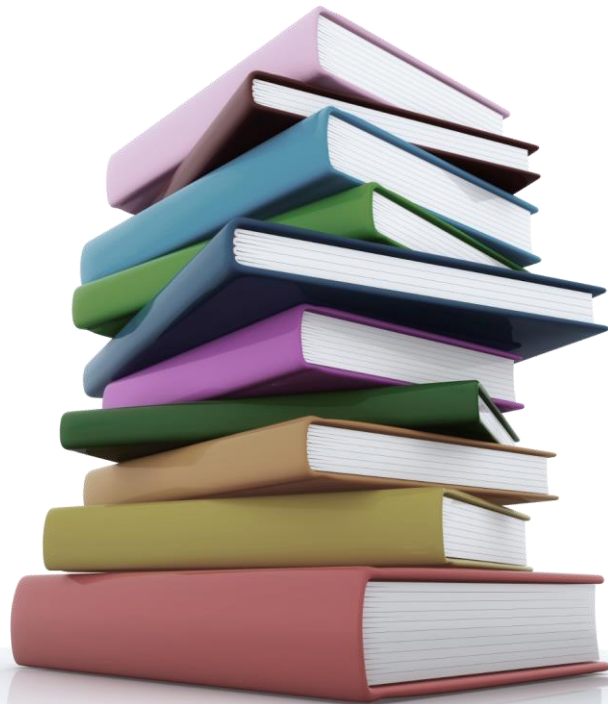
Prof. Dr. Josef van Genabith, DFKI

Impact of number of language resources on Statistical MT quality

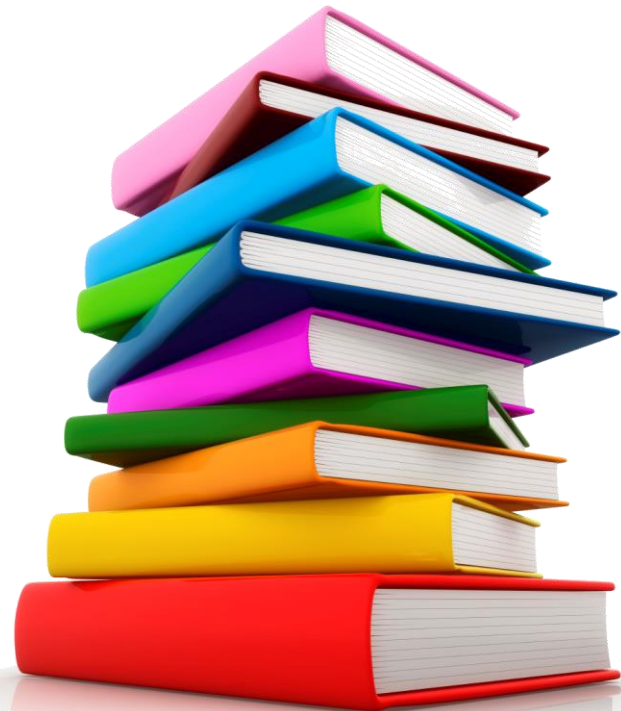


- Anything that contains “words”
- Preferences for “sentences”, even for sentences expressed in multiple languages
- Examples: reports, speeches, documents, web pages, brochures, etc.
- Bags of “words”, “sentences”, multiple bags

What types of data? “Aligned” Translation



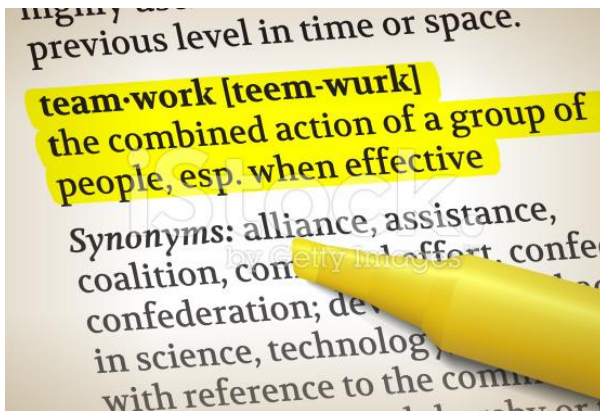
English



Danish



Digital textual data in various formats – valuable language resources



	ES	EL
	abandono escolar	διακοπή της σχολικής φοίτησης
	despojo	παραπροϊόντα σφαγίων
	sacrificio de animales	σφαγή ζώων
	derogación	κατάργηση
5075	Abruzzes	Αβρουζία
5339	absentéisme	συστηματική απουσία από την εργασία
5984	abstentionnisme	αποχή
2	abus de confiance	απιστία
	abus de droit	κατάχρηση δικαιώματος
	abus de pouvoir	κατάχρηση εξουσίας
	accès à l'éducation	πρόσβαση στην εκπαίδευση
	accès à l'emploi	πρόσβαση στην αγορά εργασίας



Comparable Collections



English

Telecommunication occurs when the exchange of information between two or more entities (communication) includes the use of technology.

Communication technology uses channels to transmit information (as electrical signals), either over a physical medium (such as signal cables), or in the form of electromagnetic waves.

The word is often used in its plural form, telecommunications, because it involves many different technologies.

Greek

Με τον γενικό όρο τηλεπικοινωνίες, (telecommunications), χαρακτηρίζεται η κάθε μορφής ενσύρματη ή ασύρματη, ηλεκτρομαγνητική, ηλεκτρική, κ.λπ., ακουστική και οπτική επικοινωνία που πραγματοποιείται ανεξαρτήτως απόστασης.

Στους σύγχρονους καιρούς, αυτή η διαδικασία σχεδόν πάντα περιλαμβάνει την αποστολή ηλεκτρομαγνητικών κυμάτων ή ηλεκτρικών σημάτων από κατάλληλες ηλεκτρονικές συσκευές, όπως το τηλέφωνο ή ο ασύρματος, αλλά παλαιότερα περιελάμβανε τη χρήση ακουστικών σημάτων, όπως τυμπάνων, ή οπτικών, όπως ο σηματοφόρος καπνός ή η λάμψη της φωτιάς.

Spanish

Una telecomunicación es toda transmisión y recepción de señales de cualquier naturaleza, típicamente electromagnéticas, que contengan signos, sonidos, imágenes o, en definitiva, cualquier tipo de información que se desee comunicar a cierta distancia.

Por metonimia, también se denomina telecomunicación (o telecomunicaciones, indistintamente) a la disciplina que estudia, diseña, desarrolla y explota aquellos sistemas que permiten dichas comunicaciones; de forma análoga, la ingeniería de telecomunicaciones resuelve los problemas técnicos asociados a esta disciplina.

Source: First sentences of articles for Telecommunications in the English, Greek and Spanish Wikipedias

- Visible data e.g. [Web](#)
 - Public websites
 - leaflets, brochures, news etc.
- [Invisible Data](#): archives , hidden web, internal repositories
 - translated documents, reports, speeches, meeting minutes etc.
- Data from outsourced translations:
 - From [Language Service Providers](#) and freelance translators
 - Translation Memory as part of contract delivery
 - Respective provisions in the contracts for outsourced translations

An iceberg floating in the ocean. The small tip above the water represents the public web, while the much larger submerged part represents the deep web. The background is a blue gradient representing the sky and water.

4%

The Public
Web

96%

The Deep
Web

Multilingual databases
Public sector resources
Organization specific repositories
Legal documents
Scientific reports
Medical records
Etc.

*Information stored inside
institutions or online with
password protection*

Processing of Bilingual Documents for Machine Translation



Word docs from <http://www.diplomatie.gouv.fr/fr/photos-videos-publications/publications/enjeux-planetaires-cooperation/rapports/article/rapports-du-groupe-pilote>,
Financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition, Ministère des Affaires étrangères et du Développement international



Agriculture_GB_bd_cle03aa11_p7.docx - Word

EXECUTIVE SUMMARY

This report is the result of a collective work carried out by the high-level expert Committee and a writing team commissioned by the Task Force on Innovative Financing for agriculture, food security and nutrition created by the Leading Group on Innovative Financing for Development at its 9th plenary session in Mali (Bamako) in June 2011.

The report includes an analysis of the need for innovating financing dedicated to the agricultural, food security and nutrition sector, a critical review of existing and possible mechanisms and a proposed selection of avenues for the development of such mechanisms on the basis of the expertise of a high-level Committee of experts, literature review, meetings with relevant professional actors and an on-line consultation on the Global Forum on food security and nutrition (FSN Forum).

The setting up of the Task Force on Innovative Financing for agriculture, food security and nutrition responds to current and future crucial challenges faced by the international community regarding food insecurity and malnutrition and is related to the achievement of the first Millennium Development Goal (MDG 1) (reduction of extreme poverty and hunger by half by 2015).

With almost 870 million chronically undernourished people in 2010-2012, the number of hungry people in the world remains unacceptably high.

Given the current trends in world population growth, particularly in developing countries, the global demand for food is expected to increase by 60 percent by 2050¹.

Meeting this challenge and allowing developing countries to ensure food security and nutrition imply an important investment in the agricultural sector (i.e. 50 percent more than the current level according to FAO²) and in the field of nutrition (i.e. additional annual investment of USD 11.8 billion is needed to implement at scale the priority nutrition-specific interventions identified by the Lancet series³ and endorsed by the Scaling Up Nutrition (SUN) movement⁴, according to the World Bank⁵). These investments need to be as well combined with efforts across cross-cutting sectors such as economic empowerment (especially women and young people), education, health, water sanitation and hygiene, etc.

Although they are progressing, budgets for food security, including agriculture and nutrition components, in developing countries are severely constrained. Regarding the agriculture side, the dedicated Official Development Assistance (ODA) has increased in recent years, with higher amounts allocated to emerging middle income countries compared to Sub-Saharan Africa and less developed countries. The challenge is all the more important in Sub-Saharan Africa, the continent where population growth will be the highest, where yields have tended to stagnate

1- FSN Forum : <http://www.fao.org/fsnforum/>
2- FAO, IFAD and WFP: The State of food insecurity in the world (SOFI) 2012. Economic growth is necessary but not sufficient to accelerate reduction of hunger and malnutrition. <http://www.fao.org/publications/sofi/>
3- Idem
4- The Lancet (janvier 2008) "Maternal and Child Undernutrition, Special Series"
5- SUN movement: <http://scalingupnutrition.org/>
6- Horton and al. (2010) "Scaling Up Nutrition. What will it cost?" Washington DC: World Bank

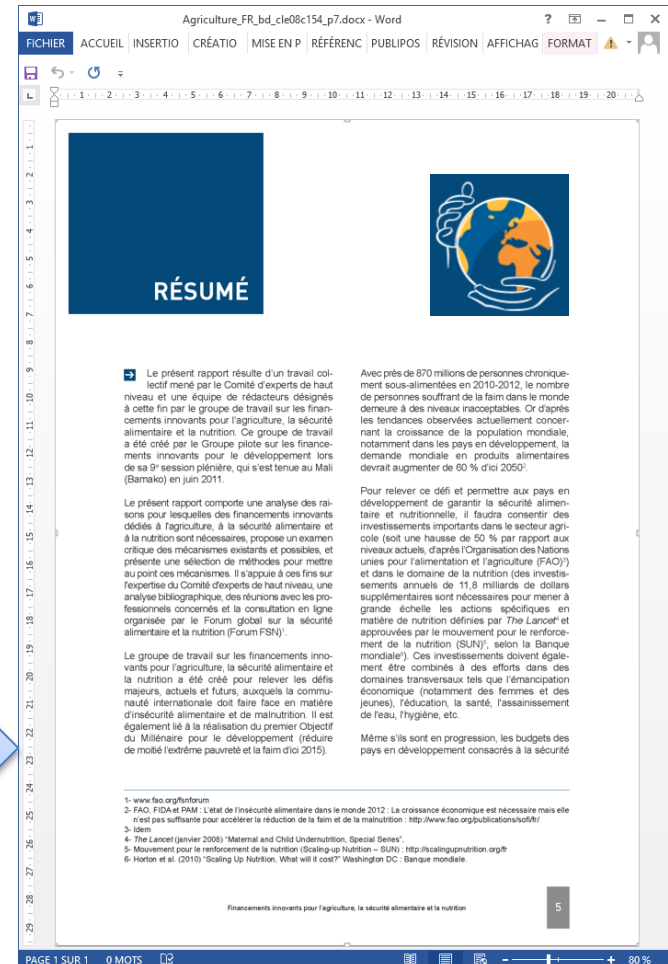
Innovative financing for agriculture, food security and nutrition

PAGE 1 SUR 1 0 MOTS

English
version



French
version



Agriculture_FR_bd_cle08c154_p7.docx - Word

RÉSUMÉ

Le présent rapport résulte d'un travail collectif mené par le Comité d'experts de haut niveau et une équipe de rédacteurs désignés à cette fin par le groupe de travail sur les financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition. Ce groupe de travail a été créé par le Groupe pilote sur les financements innovants pour le développement lors de sa 9^e session plénière, qui s'est tenue au Mali (Bamako) en juin 2011.

Le présent rapport comporte une analyse des raisons pour lesquelles des financements innovants dédiés à l'agriculture, à la sécurité alimentaire et à la nutrition sont nécessaires, propose un examen critique des mécanismes existants et possibles, et présente une sélection de méthodes pour mettre au point ces mécanismes. Il s'appuie à ces fins sur l'expertise du Comité d'experts de haut niveau, une analyse bibliographique, des réunions avec les professionnels concernés et la consultation en ligne organisée par le Forum global sur la sécurité alimentaire et la nutrition (Forum FSN).

Le groupe de travail sur les financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition a été créé pour relever les défis majeurs, actuels et futurs, auxquels la communauté internationale doit faire face en matière d'insécurité alimentaire et de malnutrition. Il est également lié à la réalisation du premier Objectif de Millénaire pour le développement (réduire de moitié l'extrême pauvreté et la faim d'ici 2015).

Avec près de 870 millions de personnes chroniquement sous-alimentées en 2010-2012, le nombre de personnes souffrant de la faim dans le monde demeure à des niveaux inacceptables. Or, d'après les tendances observées actuellement concernant la croissance de la population mondiale, notamment dans les pays en développement, la demande mondiale en produits alimentaires devrait augmenter de 60 % d'ici 2050¹.

Pour relever ce défi et permettre aux pays en développement de garantir la sécurité alimentaire et nutritionnelle, il faudra consentir des investissements importants dans le secteur agricole (soit une hausse de 50 % par rapport aux niveaux actuels, d'après l'Organisation des Nations unies pour l'alimentation et l'agriculture (FAO)) et dans le domaine de la nutrition (des investissements annuels de 11,8 milliards de dollars supplémentaires sont nécessaires pour mener à grande échelle les actions spécifiques en matière de nutrition définies par The Lancet² et approuvées par le mouvement pour le renforcement de la nutrition (SUN)³, selon la Banque mondiale⁴). Ces investissements doivent également être combinés à des efforts dans des domaines transversaux tels que l'éducation économique (notamment des femmes et des jeunes), l'éducation, la santé, l'assainissement de l'eau, l'hygiène, etc.

Même s'ils sont en progression, les budgets des pays en développement consacrés à la sécurité

1- www.fao.org/fsnforum/
2- FAO, IFAD et PAM: L'état de l'insécurité alimentaire dans le monde 2012. La croissance économique est nécessaire mais elle n'est pas suffisante pour accélérer la réduction de la faim et de la malnutrition. <http://www.fao.org/publications/sofi/>
3- Idem
4- The Lancet (janvier 2008) "Maternal and Child Undernutrition, Special Series"
5- Mouvement pour le renforcement de la nutrition (Scaling Up Nutrition - SUN) : <http://scalingupnutrition.org/>
6- Horton et al. (2010) "Scaling Up Nutrition. What will it cost?" Washington DC : Banque mondiale.

Financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition

PAGE 1 SUR 1 0 MOTS

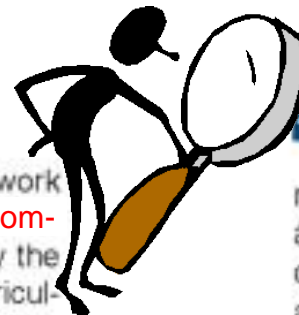
Management of Bilingual Data Example (2/4)



EXECUTIVE SUMMARY

→ This **report** is the result of a collective work carried out by the high-level **expert Committee** and a writing team commissioned by the Task Force on Innovative Financing for agriculture, food security and nutrition created by the **Leading Group on Innovative Financing for Development** at its 9th plenary session in **Mali (Bamako)** in June 2011.

The **report** includes an analysis of the need for innovating financing dedicated to the agricultural, food security and nutrition sector, a critical review of existing and possible mechanisms and a proposed selection of avenues for the development of such mechanisms on the basis of the



RÉSUMÉ

→ Le présent **rapport** résulte d'un travail collectif mené par le **Comité d'experts** de haut niveau et une équipe de rédacteurs désignés à cette fin par le groupe de travail sur les financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition. Ce groupe de travail a été créé par le **Groupe pilote sur les financements innovants pour le développement** lors de sa 9e session plénière, qui s'est tenue au **Mali (Bamako)** en juin 2011.

Le présent **rapport** comporte une analyse des raisons pour lesquelles des financements innovants dédiés à l'agriculture, à la sécurité alimentaire et à la nutrition sont nécessaires, propose un examen critique des mécanismes existants et possibles, et

Management of Bilingual Data Example (3/4)



English version – Raw text

Executive Summary

This report is the result of a collective work carried out by the high-level expert Committee and a writing team commissioned by the Task Force on Innovative Financing for agriculture, food security and nutrition created by the Leading Group on Innovative Financing for Development at its 9th plenary session in Mali (Bamako) in June 2011.

The report includes an analysis of the need for innovating financing dedicated to the agricultural, food security and nutrition sector, a critical review of existing and possible mechanisms and a proposed selection of avenues for the development of such mechanisms on the basis of the expertise of a high-level Committee of experts, literature review, meetings with relevant professional actors and an on-line consultation on the Global Forum on food security and nutrition (FSN Forum)¹.

The setting up of the Task Force on Innovative Financing for agriculture, food security and nutrition responds to current and future crucial challenges faced by the international community
[...]

French version – Raw text

Résumé

Le présent rapport résulte d'un travail collectif mené par le Comité d'experts de haut niveau et une équipe de rédacteurs désignés à cette fin par le groupe de travail sur les financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition. Ce groupe de travail a été créé par le Groupe pilote sur les financements innovants pour le développement lors de sa 9e session plénière, qui s'est tenue au Mali (Bamako) en juin 2011.

Le présent rapport comporte une analyse des raisons pour lesquelles des financements innovants dédiés à l'agriculture, à la sécurité alimentaire et à la nutrition sont nécessaires, propose un examen critique des mécanismes existants et possibles, et présente une sélection de méthodes pour mettre au point ces mécanismes. Il s'appuie à ces fins sur l'expertise du Comité d'experts de haut niveau, une analyse bibliographique, des réunions avec les professionnels concernés et la consultation en ligne organisée par le Forum global sur la sécurité alimentaire et la nutrition (Forum FSN)¹.

Le groupe de travail sur les financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition a été créé pour relever les défis majeurs, actuels et futurs, auxquels la communauté
[...]



Alignment of English and French versions

S1. Executive Summary

S2. This report is the result of a collective work carried out by the high-level expert Committee and a writing team commissioned by the Task Force on Innovative Financing for agriculture, food security and nutrition created by the Leading Group on Innovative Financing for Development at its 9th plenary session in Mali (Bamako) in June 2011.

S3. The report includes an analysis of the need for innovating financing dedicated to the agricultural, food security and nutrition sector, a critical review of existing and possible mechanisms and a proposed selection of avenues for the development of such mechanisms on the basis of the expertise of a high-level Committee of experts, literature review, meetings with relevant professional actors and an on-line consultation on the Global Forum on food security and nutrition (FSN Forum)1.

S4. The setting up of the Task Force on Innovative Financing for agriculture, food security and nutrition responds to current and future crucial challenges faced by the international community [...]

S1. Résumé

S2. Le présent rapport résulte d'un travail collectif mené par le Comité d'experts de haut niveau et une équipe de rédacteurs désignés à cette fin par le groupe de travail sur les financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition.

S3. Ce groupe de travail a été créé par le Groupe pilote sur les financements innovants pour le développement lors de sa 9e session plénière, qui s'est tenue au Mali (Bamako) en juin 2011.

S4. Le présent rapport comporte une analyse des raisons pour lesquelles des financements innovants dédiés à l'agriculture, à la sécurité alimentaire et à la nutrition sont nécessaires, propose un examen critique des mécanismes existants et possibles, et présente une sélection de méthodes pour mettre au point ces mécanismes.

S5. Il s'appuie à ces fins sur l'expertise du Comité d'experts de haut niveau, une analyse bibliographique, des réunions avec les professionnels concernés et la consultation en ligne organisée par le Forum global sur la sécurité alimentaire et la nutrition (Forum FSN)1.

S6. Le groupe de travail sur les financements innovants pour l'agriculture, la sécurité alimentaire et la nutrition a été créé pour relever les défis majeurs, actuels et futurs, auxquels la communauté [...]

We need your involvement



- You know your data
 - visible vs. invisible
- Access to archives, deep web, etc. is often not possible from the outside.
- Clearance and access to data that is not yet opened by PSI or a permissive license
- Access to source formats (e.g. Word) as derived forms (e.g. PDF) is less efficient than access to internal source content repositories.