

# TOSHIBA

Leading Innovation >>>

---

## 可読性診断技術

東芝ソリューション株式会社  
IT技術研究所

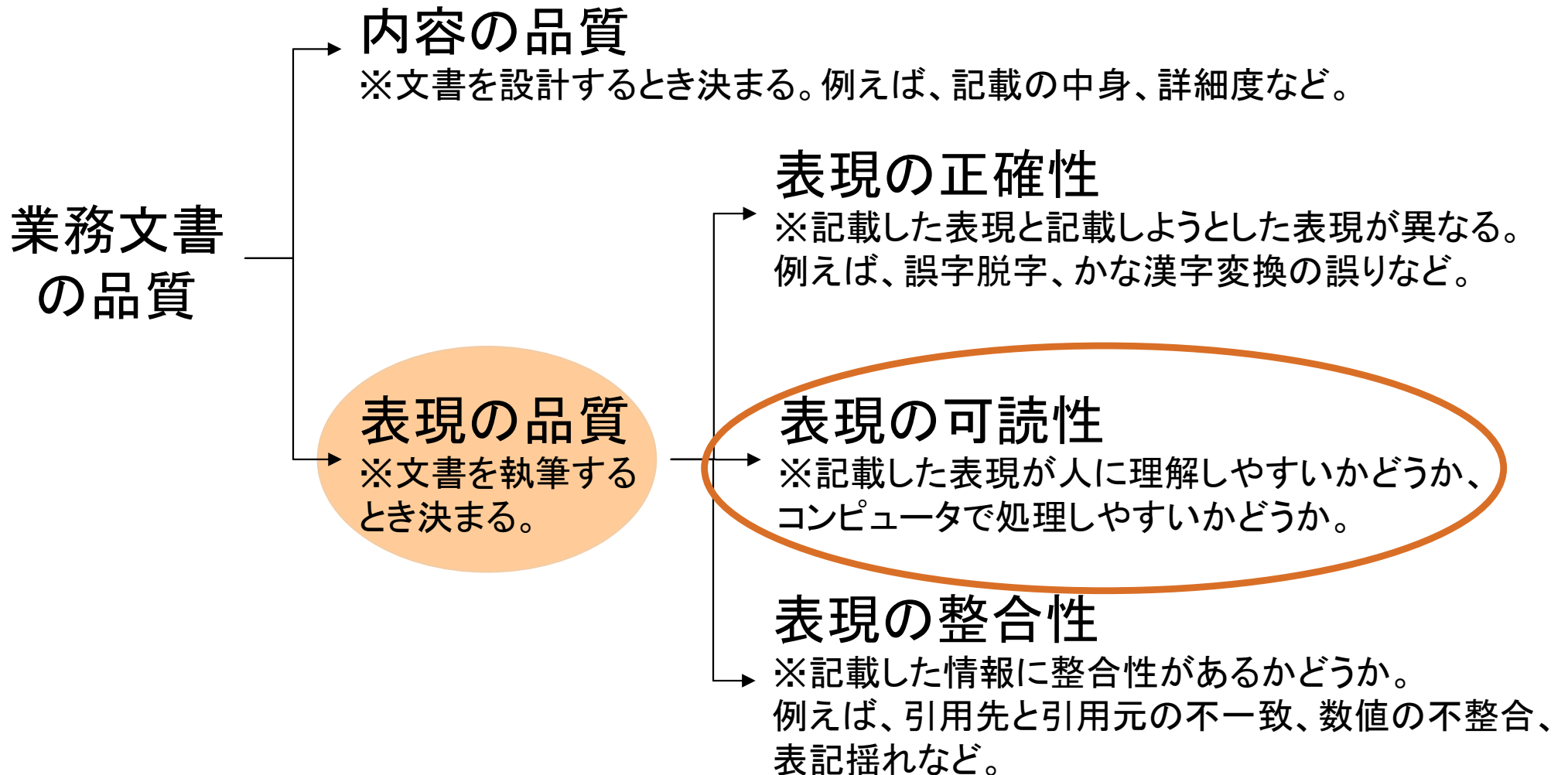
2012年2月29日

東芝ソリューション株式会社  
TOSHIBA SOLUTIONS CORPORATION

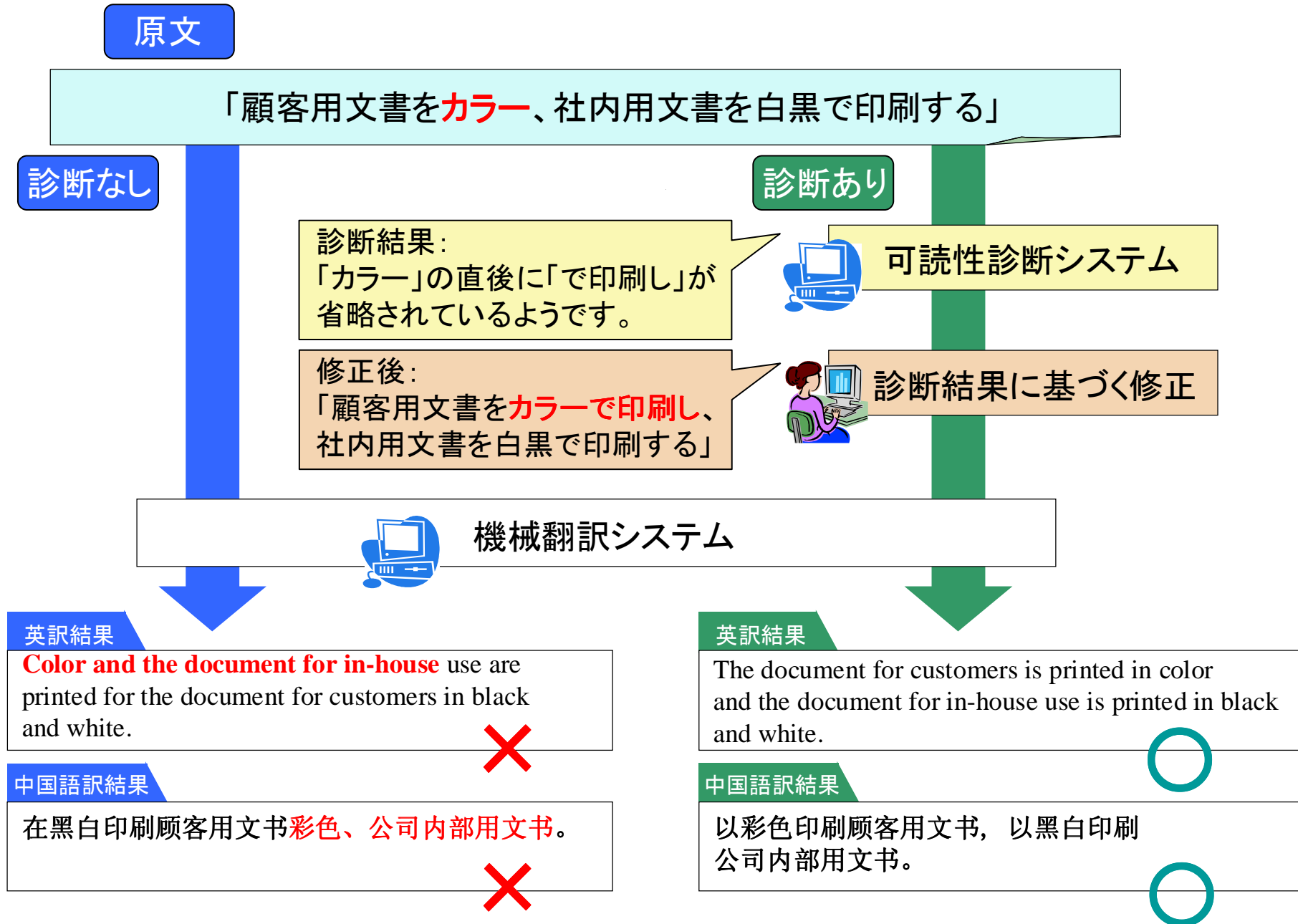


東芝グループは、持続可能な  
地球の未来に貢献します。

# 業務文書の品質



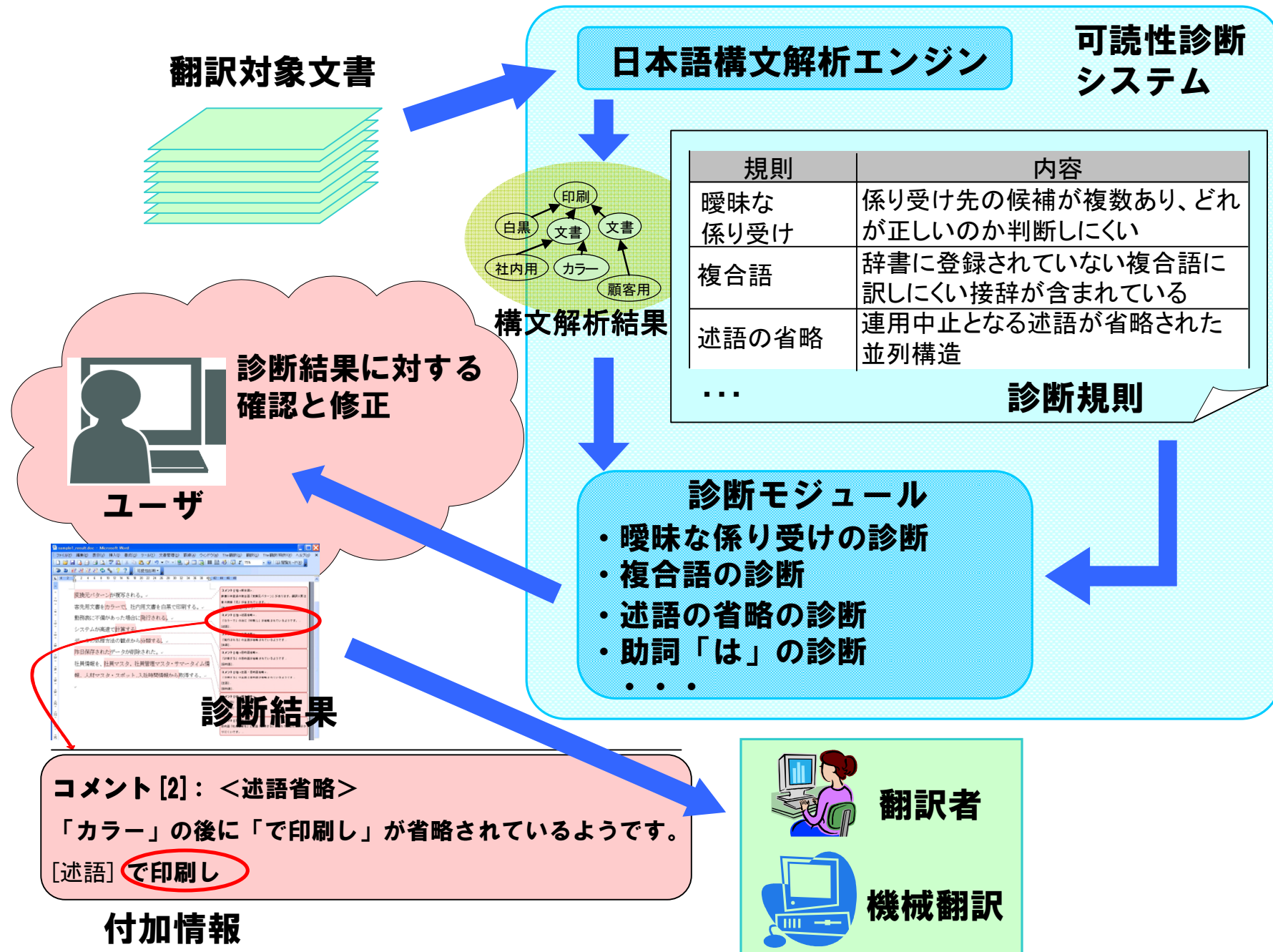
# 機械翻訳への活用例



# 可読性診断技術の活用シーン

- 機械翻訳の前処理
  - 曖昧性のある表現が指摘される
  - 原文を修正し、翻訳の精度を向上
- 翻訳コストの低減
  - 分かりやすい翻訳原文が作成できる
  - 翻訳者からの問合せ、翻訳者の誤解による誤訳が減少
- 人が理解しやすい文章の作成
  - 仕様書などの技術文書で、正しい情報伝達を支援
  - 理解度が増し、業務効率が向上

# 可読性診断技術の構成



# 技術的特徴

- 翻訳しにくい文の特徴を考慮した診断規則に基づいて、文章の可読性を診断
- 日本語構文解析の結果を使用し、文の構造から誤訳しやすい箇所や曖昧な箇所を検出
- 翻訳の精度を向上するための付加情報や選択肢を生成

コメント [2] : <述語省略>

「カラー」の後に「で印刷し」が省略されているようです。

[述語] **で印刷し**

## 付加情報

昨日保存されたデータが削除された

コメント [3] : <曖昧な係り受け>

「昨日」 - 「保存」

「昨日」 - 「削除」

## 選択肢

# 可読性診断結果の出力例

sample\_一般例\_MTI向け\_診断結果.doc (プレビュー) - Microsoft Word

変換元パターンが複写される。

顧客用文書をカラー、社内用文書を白黒で印刷する。

昨日保存されたデータが削除された。

勤務表に不備があった場合に発行される。

システムが高速で計算する。

データの処理方法の観点から分類する。

社員情報を、社員マスタ、社員管理マスタ・サマータイム情報、人材マスタ・スポット、入社時間情報から取得する。

コメント [1]: <複合語>  
辞書に未登録の複合語「変換元パターン」があります。翻訳に要注意の接辞「元」が含まれています。  
[登録語]変換元パターン:

コメント [2]: <述語省略>  
「カラー」の後に「で印刷し」が省略されているようです。(目的語)  
[述語]で印刷し

コメント [3]: <主語省略>  
「印刷する」の主語が省略されているようです。  
[主語]

コメント [4]: <構文曖昧>  
0. 「昨日」 - 「保存さ」 (2)  
1. 「昨日」 - 「削除さ」 (11)  
[選択]

コメント [5]: <主語省略>  
「発行さ」の主語が省略されているようです。  
[主語]

コメント [6]: <目的語省略>  
「計算する」の目的語が省略されているようです。  
[目的語]

コメント [7]: <主語・目的語省略>  
「分類する」の主語と目的語が省略されているようです。  
[主語]  
[目的語]

コメント [8]: <主語省略>  
「取得する」の主語が省略されているようです。

東芝 TOSHIBA Corporation 7

# 産業日本語との関連

- 可読性診断の判定基準と産業日本語ライティングルールとの共通性
- ライティングの支援システムや支援環境としての活用



**ご清聴ありがとうございました**